

基于特征工程和支持向量机的甲烷预混火焰当量比测量

陈长友¹, 傅钰雯², 涂沛驰¹, 舒文¹, 杨健晟¹

(1. 贵州大学电气工程学院, 贵阳 550025; 2. 贵州交通职业技术学院物流工程系, 贵阳 550025)

摘要: 利用火焰颜色建模测量火焰当量比是燃烧诊断技术的一个新兴研究方向。目前的建模方法主要利用RGB(Red-green-blue)模型中蓝色/绿色特征(B/G)作为模型输入, 但通过单一颜色比值简单拟合得到的颜色-当量比模型存在较大的不确定性及测量误差, 因此本文提出利用多颜色模型下的多颜色特征参数作为模型输入。首先, 采用数字火焰颜色分布(Digital flame colour distribution, DFCD)技术对采集甲烷燃烧预混火焰图像进行处理并获取火焰图像目标区域(Region of interest, RoI)。其次, 综合分析火焰颜色特征变量构建特征工程, 设计并提取火焰目标区域的不同颜色模型下的多颜色特征, 共计36维火焰颜色特征, 利用Spearman秩相关性分析与随机森林(Random forest, RF)算法筛选出表征燃烧当量比更深层的颜色特征, 得到16维优质特征子集。最后, 通过优化支持向量机(Support vector machine, SVM)参数选择, 并采用网格搜索方法(Grid search method, GSM)寻求最优参数优化SVM, 进一步利用特征工程构建得到的特征子集训练SVM以建立预混火焰燃烧当量比较测量模型。将该算法与传统的BP神经网络和极限学习(Extreme learning machine, ELM)算法进行对比, 实验结果表明, 本文方法具有较好的回归预测效果, 均方误差(Mean square error, MSE)低至0.023。

关键词: 当量比; 特征工程; 支持向量机; Spearman秩相关; 随机森林; 网格搜索法

中图分类号: TP391.41 **文献标志码:** A

Methane Premixed Flame Equivalence Ratio Measurement Based on Feature Engineering and Support Vector Machine

CHEN Changyou¹, FU Yuwen², TU Peichi¹, SHU Wen¹, YANG Jiansheng¹

(1. College of Electrical Engineering, Guizhou University, Guiyang 550025, China; 2. Department of Logistics Engineering, Guizhou Communications Polytechnic, Guiyang 550025, China)

Abstract: Flame equivalence ratio measurement using flame color modeling method, is an emerging research direction in the combustion diagnosis technology. At present, the modeling methods mainly use the blue/green color features (B/G) in the RGB(Red-green-blue) model as the modeling input, however, the color equivalence ratio modeling by single color ratio fitting has large uncertainty and measurement errors. Therefore, this paper proposes to use the multi-color feature parameters under different-color models as the modeling inputs. Firstly, the digital flame color distribution (DFCD) technology is used to process the methane premixed flame image and obtain the region of interest (RoI) of flame images. Secondly, the flame color feature variables are comprehensively analyzed, and the multi-color features

under different color models are designed and extracted, which are 36 color features. Then, the Spearman rank correlation analysis and random forest (RF) algorithm are used to screen out the deeper color features, and 16 dimensional high-quality features are selected. At last, the optimal support vector machine (SVM) parameters are selected using the grid search method (GSM). Furthermore, the equivalence ratio measurement model of premixed methane flame is trained by SVM using the feature subset constructed. The algorithm is compared with the traditional BP neural network and the extreme learning machine (ELM) algorithm. Experimental results show that the algorithm has better regression prediction effect, and the mean square error (MSE) decreases to 0.023.

Key words: equivalence ratio; feature engineering; support vector machine (SVM); Spearman rank correlation; random forests; grid search method

引 言

化石燃料是我国最主要的能源形式。随着我国经济发展对能源需求与日俱增,在提高化石燃料燃烧效率和减少污染物排放方面都提出了更高的要求。在燃烧工业领域中,检测燃烧炉内火焰相关参数判断燃烧状态,有效控制燃料与空气的供给关系,是提高能源转换效率和减少燃烧污染物生成的有效手段。其中,燃空当量比是表征燃料与空气混合的一个重要参数,是确定系统性能的因素之一,也是影响燃烧污染物生成的重要因数^[1]。因此测量火焰当量比对节能减排具有重大的意义。

火焰是燃料和空气燃烧过程中释放的可见光和热辐射的形式,是燃烧过程的综合表征。对于碳氢火焰而言,火焰发出的蓝绿光是由于燃烧反应中产生的激发态自由基 CH^* (430 nm)和 C_2^* (516 nm)等回归基态时在可见光谱内释放相应光子引起的化学发光,而火焰的红黄光是不完全燃烧生成的碳黑在连续光谱段的黑体辐射^[2]。有研究表明,碳氢火焰自由基 CH^* 和 C_2^* 的化学发光与预混火焰燃烧当量比之间存在密切的关系,通过测量自由基 CH^* 与 C_2^* 浓度比值可以间接测量当量比^[2-4];Clark等^[5]研究发现自由基 C_2^* 与 CH^* 的比值可以用来近似表示当量比,Haber等^[6]也证实了自由基化学发光 CH^*/C_2^* 的比值与当量比具有相似的响应。由于颜色是光谱的表征形式,所以通过测量火焰颜色也可以间接获得自由基化学发光信息。2008年,Huang等^[7]发现在RGB(Red-green-blue)颜色模型中,预混甲烷火焰图像的蓝色和绿色通道的中值强度与 CH^* 和 C_2^* 化学发光强度非常匹配,以此为依据提出了首个基于火焰颜色测量火焰当量比的软测量模型,该模型以蓝色/绿色(B/G)的比值作为输入,以对应火焰当量比值作为模型输出,通过线性拟合建立B/G与当量比的函数关系。随后,Huang等利用该方法分别对甲烷和乙烯两种燃料在不同当量比下的化学发光进行测量,通过对比传统的光谱仪法和滤波片法,证实了在RGB模型下的B/G值可以用来测量燃空当量比^[8]。通过颜色建模测量当量比方法对成像设备要求低、操作简单、数据处理便捷,因此很快受到其他研究者的关注。2014年,Migliorini等^[9]利用光谱仪法、滤波片法和数字颜色建模法,对不同当量比下甲烷预混火焰中 CH^* 和 C_2^* 自由基化学发光进行了测量,实验结果验证了数字颜色建模法对燃空当量比测量的可行性。目前,基于颜色建模的火焰当量比检测方法已经应用于多类基础燃烧研究,并取得了一定的成果^[10-14]。

由于成像颜色会受到彩色图像传感器光谱响应函数差异的影响,2019年Yang等^[15]等通过标定相机传感器的光谱响应,改进了利用蓝色和绿色通道测量 CH^* 和 C_2^* 化学发光的方法,优化了基于

图像颜色与化学发光之间的数学模型。2020年龙凯等^[16]试图通过白平衡处理校正成像颜色来提高火焰当量比颜色模型的普适性。Trindade等^[17]也曾设计了4种不同的颜色输入构建模型,试图通过对比寻找最优的表征火焰当量比模型的数字颜色输入。由于成像过程影响因素较多,如光学镜头色偏、图像传感器滤波片光谱特性不一致、制造工艺差异以及插值算法区别等,使得这些方法仅通过单一颜色比值简单拟合得到的颜色与当量比软测量模型存在较大的不确定性及误差。近年来,随着机器学习智能算法的快速发展,为燃烧过程中生成的中态产物、终态产物以及相关参数的软测量建模提供了新的方法。以支持向量(Support vector machine, SVM)、BP神经网络和极限学习(Extreme learning machine, ELM)等为代表机器学习算法已经逐步应用于燃烧诊断领域。例如:通过提取研究对象时间和空间上的相关特征,包括颜色、纹理和运动等,进一步应用机器学习算法实现火焰、烟雾和燃料的识别以及燃烧污染物排放预测,并取得很好的识别率和回归预测效果^[16-21];通过融合火焰静态和动态特征,基于ELM算法对特征数据集进行训练与测试,进而实现火焰的检测,实验结果表明,该方法在环境干扰较小的情况下具有较高的时效性和检测精度;把目标集中在算法的改进与优化方面,通过优化算法来提高识别率和降低预测误差^[22]。目前这些研究主要基于特定数据集或特定应用取得了良好的表现性能,并未有通过机器学习火焰颜色特征检测火焰当量比的相关研究。

因此,本文基于火焰颜色-自由基化学发光-当量比之间的物理转换原理,通过特征工程构建火焰多维颜色特征数据,利用SVM机器学习算法实现甲烷预混火焰燃烧当量比的软测量。

1 实验方法

1.1 实验对象与平台

甲烷是天然气、煤层气等能源的主要成分,为此本文主要针对甲烷预混火焰燃烧当量比测量展开实验研究。实验平台包括3个部分。

(1) 燃烧实验平台

本文使用的燃烧实验平台为自制本生灯预混/扩散火焰基础燃烧装置,包括空气压缩机、甲烷(99.9%)气瓶、气体流量计、压力表、预混管以及燃烧器(本生灯)等装置组成,基本架构如图1所示。其中气体流量计MFC300最大量程为500 sccm,精度高达 $\pm 0.5\%$ FS,本生灯燃烧管口直径为6 mm。试验中通过Labview控制流量计使得甲烷(CH_4)与空气(Air)分别从不同的气路管道输送至预混管进行充分混合(为保证混合均匀,每组工况点燃前预混5 min),然后送至本生灯进行点燃燃烧。

(2) 成像实验平台

本文所采用的成像装置为WP-030 CMOS微型工业相机,5 Gb/s传输宽带,采集传输速度可达 2×10^9 像素/s,最高帧率达到815帧/s,最高像素500万,内置64 MB缓存,确保图像数据传输更加稳定,成像镜头光圈为F1.4、焦距 $f=15\sim 30$ mm。

(3) 数据样本处理及训练平台

数据样本处理及训练平台采用的是CPU为Intel酷睿I5-4210U、可超频至2.7 GHz的计算机、64位

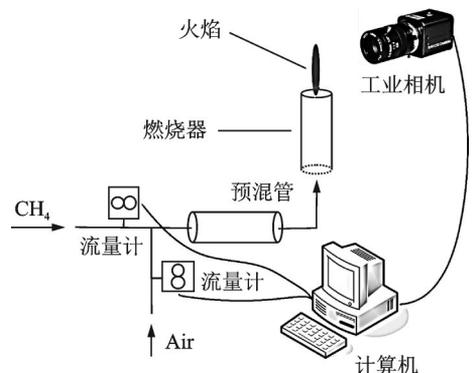


图1 燃烧平台基本示意图

Fig.1 Schematic of combustion platform

操作系统,编程软件为MATLAB。

1.2 技术路线

建模过程的技术路线大致可分为火焰图像数据采集、火焰颜色特征工程构建与SVM训练数据建模3个部分,如图2所示。火焰图像使用WP-030微型工业相机进行采集并将其送至计算机,在计算机上运用数字图像处理技术对采集的火焰图像进行预处理,以获取火焰图像目标区域,即感兴趣区域(Region of interest, RoI)。进一步在RoI提取不同颜色空间多颜色特征,以此设计并构建火焰颜色特征工程。最后,在构建的火焰颜色特征工程基础上结合改进的SVM算法训练得到甲烷预混火焰燃烧当量比测量模型。

1.3 样本数据采集

对于数据样本采集,采用WP-030工业相机采集了当量比 Φ 从0.7~1.5不同工况下的火焰图像,其中,甲烷流量恒定设置为100 sccm,根据理论空燃比、实际空燃比以及当量比之间的变换关系,例如 $\Phi=1.0$ 时, CH_4 为100 sccm,此时空气为171 sccm;通过改变空气流量值进而实现不燃烧工况;在稳定燃烧状态下以当量比间隔 $\Phi=0.05$ 为一组样本工况进行采集。成像快门时间设置为0.005 s、帧率为100 帧/s,每种当量比工况连续采集时间 $t=1$ s数据,共计100张火焰图像。所有工况下共采集1 800张火焰图像数据,不同工况火焰图像样本数据示例如图3所示。

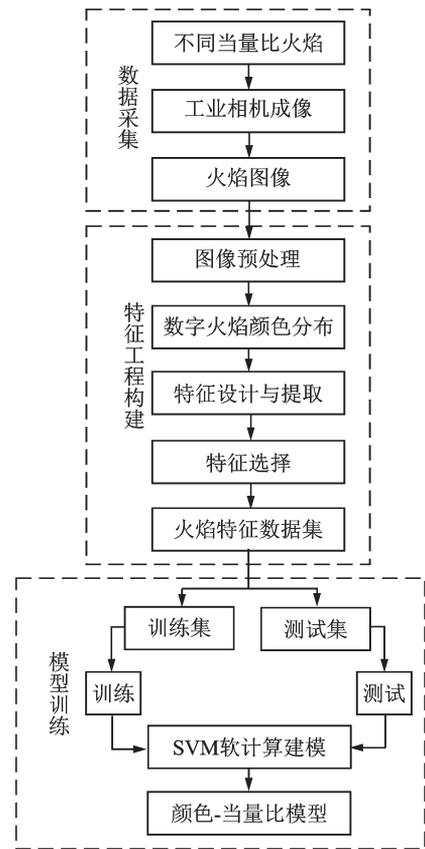


图2 本文提出的甲烷预混火焰当量比测量系统框图

Fig.2 Diagram of the proposed methane pre-mixed flame equivalent ratio measurement system

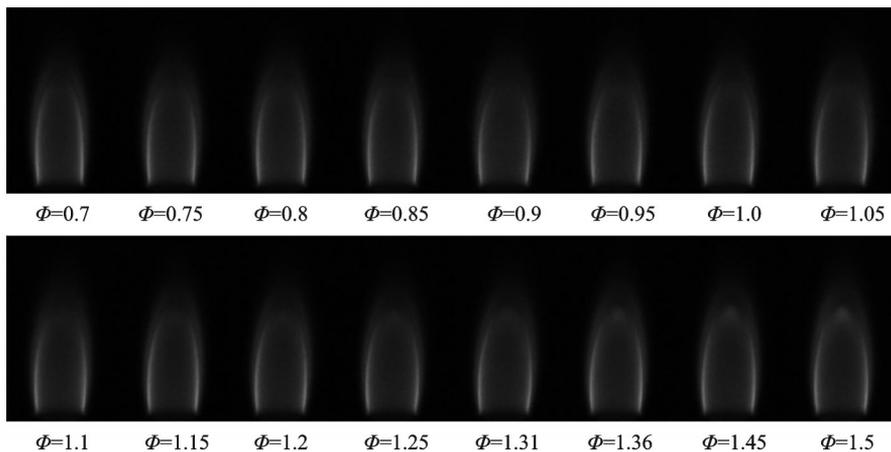


图3 不同工况火焰图像数据

Fig.3 Flame image data under different working conditions

2 特征工程构建

2.1 图像预处理

随着当量比值增加,从图3不同工况火焰图像数据可以明显看出,火焰图像在 $\Phi=1.15\sim 1.5$ 工况下火焰顶部存在黄色区域,由于黄色火焰与蓝色火焰产生机理完全不同,为了避免这种黄色火焰对当量比较测量实验结果造成影响,需要采用相关图像处理技术对火焰图像进行预处理。本文根据预混火焰与扩散火焰的数字火焰颜色分布(Digital flame colour distribution, DFCD)^[13]存在明显的分布差异,利用DFCD技术实现火焰区域的分割,并对分割的火焰图像进行滤波处理以滤除图像背景噪声的干扰。考虑到由于中值滤波算法在去除图像噪声同时也会对图像细节造成影响,因此对初步处理后RGB图像进行中值滤波后的区域图像不能视为目标区域,需要进一步将原始图像在RGB、HSV、YCbCr不同颜色空间模型分布下的各颜色通道值分别与DFCD技术处理之后得到的二值图像进行逻辑乘运算,通过逻辑运算最终得到火焰图像目标区域,详细的火焰图像目标区域获取算法流程如图4所示。其中,HSV是根据颜色的直观特性得到一种颜色空间模型,具体指色相(Hue)分量 H 、饱和度(Saturation)分量 S 和明度(Value)分量 V ^[23],而对于YCbCr颜色空间模型,其中 Y 指亮度分量, Cb 指蓝色色度分量, Cr 指红色色度分量,YCbCr模型来源于YUV色彩模型^[24]。根据算法流程使用MATLAB软件编程实现火焰图像RoI获取,如图5所示。其中图5(b,c)清晰地展示了采用DFCD技术可以很好地实现黄色火焰与蓝色火焰的分离以及火焰图像目标区域的获取。

2.2 特征设计与提取

对于火焰颜色特征提取,常用的颜色空间模型有RGB、HSV和YCbCr等,火焰图像在不同的颜色空间模型下具有不同的颜色分布,其中RGB应用最为广泛,并且HSV和YCbCr颜色空间模型可以由RGB颜色模型通过相应的线性变换或者非线性变换得到^[25-26]。

通过图像预处理阶段获取火焰图像ROI,设计并提取甲烷燃烧预混火焰图像特征:选取不同颜色空间模型下的基本颜色分量及相互间比值作为颜色特征,以RGB颜色空间模型为例,设计了 R 、 G 、 B 、 B/G 、 $R/(R+B+G)$ 等12维火焰颜色特征,其中 R 、 G 、 B 分别指RGB颜色模型中的红色分量、绿色分量和蓝色分量,类似地提取HSV和YCbCr颜色空间模型的颜色特征,设计并提取了36维火焰图像颜色特征,如表1所示。

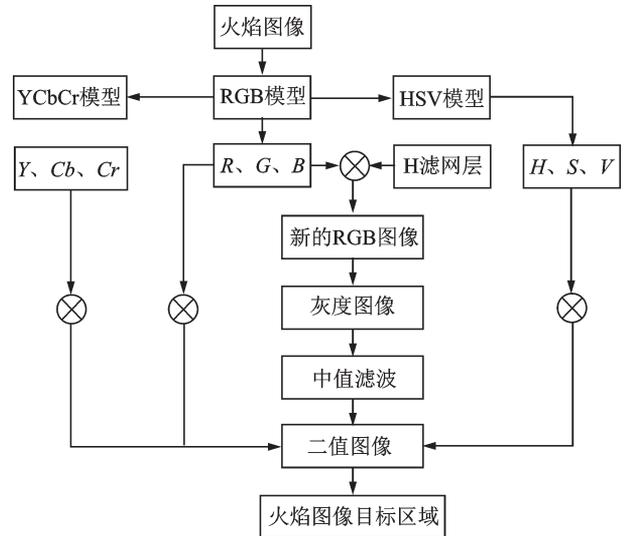


图4 基于DFCD技术的火焰图像目标区域获取算法流程

Fig.4 Diagram of flame RoI using DFCD

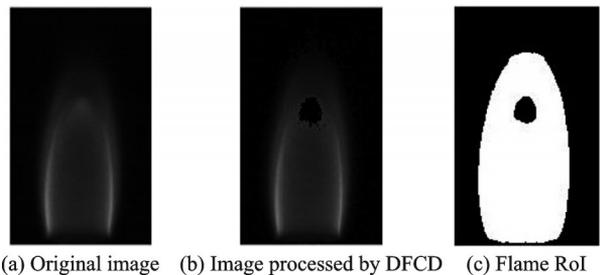


图5 基于DFCD技术的火焰图像分割结果

Fig.5 Flame image segmentation by DFCD

表1 火焰图像颜色特征
Table 1 Color features of flame image

颜色空间模型	颜色特征
RGB	$R, G, B, R/G, R/B, G/R, G/B, B/R, B/G, R/(R+G+B), G/(R+G+B), B/(R+G+B)$
HSV	$H, S, V, H/S, H/V, S/H, S/V, V/H, V/S, H/(H+S+V), S/(H+S+V), V/(H+S+V)$
YCbCr	$Y, Cb, Cr, Y/Cb, Y/Cr, Cb/Y, Cb/Cr, Cr/Y, Cr/Cb, Y/(Y+Cb+Cr), Cb/(Y+Cb+Cr), Cr/(Y+Cb+Cr)$

2.3 特征选择

特征选择是从多种特征中选择出较为有效的特征,即数据挖掘,是特征工程构建中的关键环节之一。因为特征的选择对甲烷预混火焰燃烧当量比测量模型的表现性能具有重要的影响,主要体现在模型计算复杂度和泛化能力2个方面。为此,本文采用基于 Spearman^[27]相关性分析和随机森林算法(Random forest, RF)^[28-29]相结合手段实现特征的降维处理。

2.3.1 基于 Spearman 的特征选择

对于特征提取的36维火焰颜色特征,采用统计学理论中的 Spearman 相关性分析剔除冗余特征,即该类特征所包含的信息可以从其他特征中推演出来或者存在一些特征它们所包含的信息具有相似性, Spearman 相关系数法通过计算特征之间的相关系数进而筛选出相关性较小的特征,其基本原理为

$$\rho_s = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N(N^2 - 1)} \tag{1}$$

式中: d_i 表示秩序; N 表示样本容量;以RGB颜色空间模型的颜色特征为例,计算其12维颜色特征 Spearman 相关系数,得到火焰颜色特征之间相关系数矩阵如表2所示。

表2 RGB颜色空间特征之间的相关系数矩阵
Table 2 Correlation coefficient matrix between RGB color space features

特征值	R	G	B	R/G	R/B	G/R	G/B	B/R	B/G	R/S	G/S	B/S
R	1.00	0.96	0.94	-0.22	-0.51	0.48	0.48	-0.48	0.21	0.51	-0.17	-0.42
G	0.96	1.00	0.98	-0.22	-0.32	0.26	0.30	-0.26	0.21	0.30	-0.02	-0.30
B	0.94	0.98	1.00	-0.07	-0.25	0.26	0.22	-0.26	0.06	0.25	-0.13	-0.18
R/G	-0.22	-0.22	-0.07	1.00	0.53	-0.11	-0.56	0.07	-0.99	-0.41	-0.67	0.82
R/B	-0.51	-0.32	-0.25	0.53	1.00	-0.87	-0.98	0.85	-0.49	-0.98	0.19	0.89
G/R	0.48	0.26	0.26	-0.11	-0.87	1.00	0.85	-0.99	0.07	0.93	-0.61	-0.60
G/B	0.48	0.30	0.22	-0.56	-0.98	0.85	1.00	-0.82	0.53	0.97	-0.15	-0.91
B/R	-0.48	-0.26	-0.26	0.07	0.85	-0.99	-0.82	1.00	-0.04	-0.92	0.64	0.56
B/G	0.21	0.21	0.06	-0.99	-0.49	0.07	0.53	-0.04	1.00	0.38	0.70	-0.80
R/S	0.51	0.30	0.25	-0.41	-0.98	0.93	0.97	-0.92	0.38	1.00	-0.32	-0.82
G/S	-0.17	-0.02	-0.13	-0.67	0.19	-0.61	-0.15	0.64	0.70	-0.32	1.00	-0.20
B/S	-0.42	-0.30	-0.18	0.82	0.89	-0.60	-0.91	0.56	-0.80	-0.82	-0.20	1.00

注:S=R+G+B。

设置阈值为0.95,相关系数大于0.95表示特征之间相关性较高,从表2中可以看出特征 R 与 G 、 R 与 B 、 G/B 与 R/S 、 R/B 与 B/S 之间的相关性较高,因此剔除 G 、 B 、 R/B 、 G/B 颜色特征,最后得到8维的火焰颜色特征。类似地,计算HSV、YCbCr颜色空间的火焰颜色特征之间的相关系数,剔除相关性较大的特征,最终得到22维颜色特征子集。

2.3.2 基于RF的特征选择

RF算法是一个包含多棵决策树的集成学习方法,每棵决策树代表一个分类器。RF算法学习过程中会得到多种决策结果,然后根据决策结果评估最后的分类效果。因此,本文采用RF算法对基于Spearman算法相关性分析得到的22维火焰颜色特征进行选择,通过进一步降维处理,选择出相对较为重要的特征,即筛选与预混火焰燃烧当量比相关性较大的特征。其基本思想通过分析每个特征在随机森林中的每颗树上所做的贡献,然后求取平均值,对比特征之间的贡献大小,最后筛选出贡献较大的特征。这里采用基尼指数(Gini index, GI)指标进行衡量贡献,表达式为

$$GI = 1 - \sum_{k=1}^N P_{mk}^2 \quad (2)$$

式中: N 、 m 分别表示类别和树节点; P_{mk} 表示节点 m 中类别 k 所占的比例。在Python软件上实现RF算法,通过相关性分析得到火焰颜色特征属性为22,并以此建立决策树数目为22随机森林,进一步分析不同特征在每一颗决策树上的贡献。碳氢预混火焰燃烧当量比作为标签,通过学习得到火焰颜色特征与火焰燃烧当量比之间相关性排序,火焰颜色特征重要性如图6所示。图中展示每个火焰特征在每颗决策树上贡献大小,选择GI大于0.02的火焰颜色特征作为特征数据集,完成火焰颜色特征的选择,最终样本数据维度降至16维。

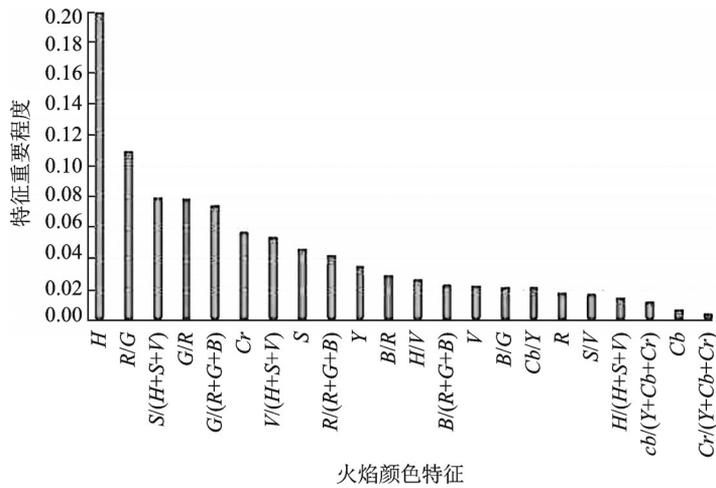


图6 基于RF算法火焰颜色特征重要性排序

Fig.6 Importance ranking of flame color features based on RF algorithm

3 基于SVM的燃烧当量比软测量模型

SVM是一种寻求全局最优解和实现结构风险最小的机器学习算法,常用于分类和回归预测^[30]。与当前发展较为迅速的深度学习算法^[31]相比,SVM可以通过核函数映射关系在高维空间实现低维空间的回归预测,而深度学习网络模型表现性能很大程度上取决于特征样本数据集容量,同时对硬件设备要求非常高,即成本较高。因此本文采用SVM实现碳氢预混火焰燃烧当量比软测量模型的构建。

SVM核心思想是寻求最大分类间隔,对于回归问题,其思想和分类非常类似;区别在于不同的维度空间。分类问题是在二维空间中进行,而回归通常是在高维空间中寻求最优超平面以实现低维空间的不可分问题。对于预混火焰燃烧当量比软测量建模,该数学模型为典型的多输入单输出结构,给定训练数据集 $\{(x_i, y_i), i = 1, 2, \dots, n\}$,其回归估计预测模型^[31-33]表达式为

$$f(x) = w * \phi(x) + b \tag{3}$$

式中: $x_i \in R^l$ 为输入变量, l 表示维度; w 表示权重向量; $\phi(x)$ 为训练样本的高维非线性映射; b 表示偏置。

约束条件为

$$\begin{cases} (w * \phi(x_i)) + b \leq \epsilon + \xi_i \\ y_i - (w * \phi(x_i)) - b \leq \epsilon + \xi_i^* \end{cases} \quad \xi_i \geq 0, \xi_i^* \geq 0; i = 1, 2, \dots, n \tag{4}$$

优化目标为

$$\min_{w, b, \xi_i, \xi_i^*} = \left\{ \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \right\} \tag{5}$$

式中: ξ_i, ξ_i^* 为松弛因子; ϵ 为损失函数; $C(C > 0)$ 为规则化参数,也称惩罚因子。

综上所述,约束条件与优化目标构成了原所求的回归估计问题,根据式(4,5)可构建拉格朗日函数对该问题进行求解,分别对 w, b, ξ_i, ξ_i^* 求偏导数并令其等于零,结合支持向量机拉格朗日乘子与不等式约束的乘积^[34]条件最终求解出回归估计数学模型为

$$f(x) = \sum_{i=1, j=1}^n (a_i^* - a_i) \phi(x_i) \phi(x_j) + b \tag{6}$$

式中: a_i^*, a_i 为拉格朗日乘子向量,根据 Mercer 条件,特征空间中的内积 $[\phi(x_i) \phi(x_j)]$ 可以用核函数 $k(x_i, x_j)$ 代替,SVM的表现性能很大程度依赖核函数的选择,基本的核函数有线性核函数、多项式核函数、径向基核函数(Radical basis function, RBF)和 Sigmoid 核函数。其中 RBF 核函数应用较为广泛,所需确定的参数较少,计算相对简单、复杂度小,计算公式如式(7)所示。因此本文选 RBF 核函数代替特征空间的内积,则式(6)可改写为式(8)。

$$k(x_i, x_j) = - \exp(-\|x_i - x_j\|^2 / \sigma^2) \tag{7}$$

$$f(x) = \sum_{i=1, j=1}^n (a_i^* - a_i) k(x_i, x_j) + b \tag{8}$$

式(7)中 σ 为设定的参数。针对多输入多输出回归问题,只需运用上述求解方法分别对多个多输入单输出系统进行求解便可得到,对于火焰燃烧当量比软测量模型而言,某一种工况对应是多输入但输出系统,但是多种工况就形成了多输入多输出回归估计问题,因此不同工况下的火焰燃烧当量比软测量模型具体描述为

$$F(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_i(x) \end{bmatrix} = \begin{bmatrix} w_1 \phi(x) + b_1 \\ w_2 \phi(x) + b_2 \\ \vdots \\ w_i \phi(x) + b_i \end{bmatrix} \tag{9}$$

式中: i 表示工况情况,即不同当量比的情况。

4 试验结果与分析

试验中将火焰不同燃烧当量比作为 SVM 机器学习算法训练预测标签,每种工况下选择 100 张火焰图片,结合特征工程构建得到 16 维火焰颜色特征,用于学习 SVM。学习前,需要将特征数据集进行简

单的归一化处理,以避免较大的火焰特征数据掩盖较小的数据,使得最终的训练结果不佳^[19]。学习中80%特征数据作为训练集,20%特征数据作为测试集,通过训练SVM得到基于火焰颜色特征的甲烷预混火焰燃烧当量比软测量模型,其回归预测模型评价指标采用均方误差(Mean square error, MSE)和平方相关系数 R^2 (Squared correlation coefficient)进行评判,表达式为

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (\mathbf{y}_i - f_i(\mathbf{x}))^2 \quad (10)$$

$$R^2 = 1 - \frac{\sum_i (\mathbf{y}_i - f_i(\mathbf{x}))^2}{\sum_i (\mathbf{y}_i - \bar{\mathbf{y}}_i)^2} \quad (11)$$

式中: n 表示样本容量; \mathbf{y}_i 表示真实值, $\bar{\mathbf{y}}_i$ 表示 \mathbf{y}_i 的均值; $f_i(\mathbf{x})$ 表示预测值。

本文选择RBF径向基核函数进行预测,其惩罚参数 c 与 g 核函数参数对SVM算法的表现性能起着决定性的作用。这里采用网格搜索法^[35]寻求最优参数优化SVM算法,基本步骤如下:

(1)粗略设置网格搜索法搜索初始范围, c 与 g 的初始范围设置为 $[2^{-10}, 2^{10}]$,根据参数结果选择的高线图进一步改进 c 与 g 的搜寻范围。

(2)采用 k 折交叉验证方法对训练集进行测试,这里 k 是指将原始数据分成 k 组,常采用默认值 $k=3$ 。

(3)根据测试结果确定 c 与 g 搜索范围分别为 $[2^{-4}, 2^4]$ 和 $[2^{-4}, 2^6]$,搜索步长设置为0.5, $k=3$,寻得最优参数 $c=16$ 、 $g=64$;并将该最优参数优化SVM。通常,算法的效率主要表现为模型的训练与测试整个过程,而参数的选择对模型的表现性能与时效性具有至关重要的作用。本文选择GSM算法优化SVM参数选择,并不断改进参数搜寻范围,其目的是为降低时间消耗成本,同时确保模型的回归预测精度。因此,优化后的SVM算法具有很好的时效性和表现性能,其寻优结果如表3所示。

通过测试得到预混火焰燃烧当量比预测误差结果如图7所示。从图7中可以清晰看出,基于SVM预混火焰燃烧当量比预测取得很好的预测结果,相对误差都集中在 $-0.3 \sim 0.5$ 之间,通过计算其均方根误差 $\text{MSE}=0.023$, $R^2=81.31\%$ 。

为了进一步验证该数学模型的效果,本文将该算法与传统BP神经网络^[36]、ELM算法^[37]进行对比分析。在同一特征数据集下,BP神经网络和ELM极限学习网络模型的隐含层统一设置为25,其中BP神经网络是一种由输入层、隐含层和输出层组成误差反向传播的前馈神经网络;本文构建的网络结构为16-25-1,网络最大训练次数为1 000、训练要求误差为 $1e-3$ 、学习率设置为0.1;而ELM算

表3 参数寻优结果对比

Table 3 Comparison of parameter optimization results

参数	第1组	第2组	第3组	第4组
c 范围	$(2^{-10}, 2^{10})$	$(2^{-8}, 2^{10})$	$(2^{-4}, 2^4)$	$(2^{-4}, 2^4)$
g 范围	$(2^{-10}, 2^{10})$	$(2^{-4}, 2^{-6})$	$(2^{-4}, 2^6)$	$(2^{-4}, 2^6)$
步长	0.5	0.5	0.5	0.5
k	3	3	3	5
最优 c	181	256	128	16
最优 g	90	64	64	64
$R^2/\%$	72.75	77.67	81.31	81.31
时间/s	323.45	156.75	20.22	42.17

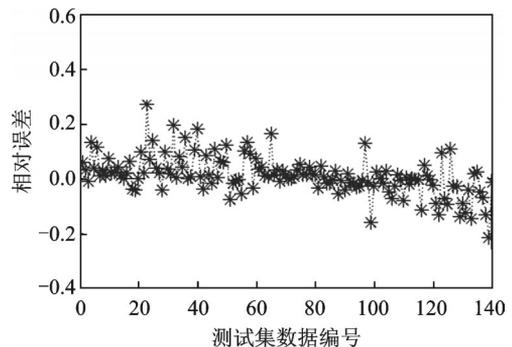


图7 基于SVM预混火焰燃烧当量比预测结果相对误差

Fig.7 MSE of prediction result based on SVM pre-mixed flame combustion equivalent ratio

法仅设置隐含层节点数目就可以获得唯一最优解,其训练速度较快。通过训练与测试得到基于不同算法的预混火焰燃烧当量比的预测结果,如图8所示。由于采集的火焰图片数据间隔 $\Phi=0.05$ 采集一次,采集间隔较小,因此相邻2组当量比工况下的火焰图像数据分布非常相近。在试验中,为了使试验结果便于分析与讨论,选择间隔 $\Phi=0.15$ 的火焰图片数据作为本文实验数据,共有7种不同工况,分别为当量比 $\Phi=0.7$ 、 $\Phi=0.85$ 、 $\Phi=1$ 、 $\Phi=1.25$ 、 $\Phi=1.31$ 、 $\Phi=1.4$ 、 $\Phi=1.5$ 。20%的火焰数据作为测试集,并且每20组测试数据对应同一个工况,通过对比不同算法的预测结果发现,SVM表现性能优于BP神经网络与ELM极限学习,基于SVM算法的预测结果相对于目标值其相关性高于其余2种算法。

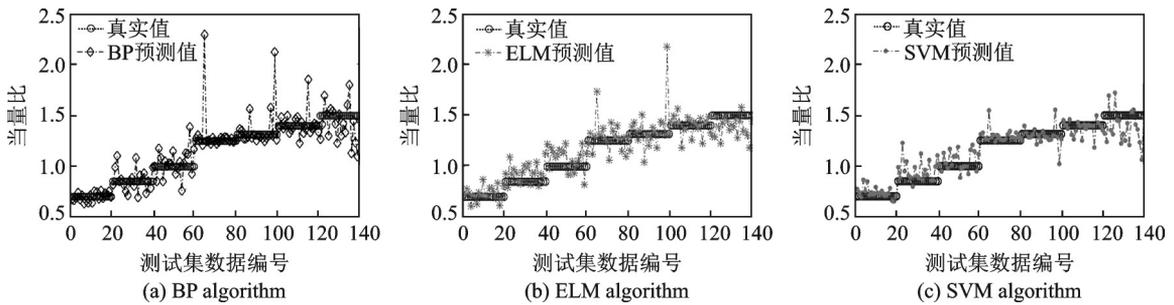


图8 基于不同算法的预混火焰燃烧当量比预测结果对比

Fig.8 Combustion equivalence ratio prediction result comparison of premixed flame based on different algorithms

从图8中可以看出,在当量比 $\Phi=1.25$ 与 $\Phi=1.31$ 时,基于BP神经网络和ELM算法预测值与参考值波动较大,BP神经网络易陷入局部最优,同时收敛速度较慢。ELM是针对传统神经网络训练速度较慢和易陷入局部最优的一种改进网络学习算法,通常其网络隐含层数目很难确定,相比较而言,基于GSM算法优化SVM实现甲烷预混火焰燃烧当量比预测具有一定的优越性,寻求的是全局最优解,总体预测效果较好。进一步计算基于不同算法下的火焰燃烧当量比预测绝对误差(Absolute error, AE),AE结果如图9所示。图9中展示了3种不同算法对预混火焰燃烧当量比预测误差结果图,对于同一个数据集,基于3种不同算法的最大AE分别为1.14、0.9、0.42,AE指标很好地反应了预测结果的可信度。表4展示了基于不同算法表现性能结果对比。通过对比不难发现SVM算法总体回归预测较好,结果表明本文模型具有较好的回归预测效果,同时,其决定系数优于其余2种算法。

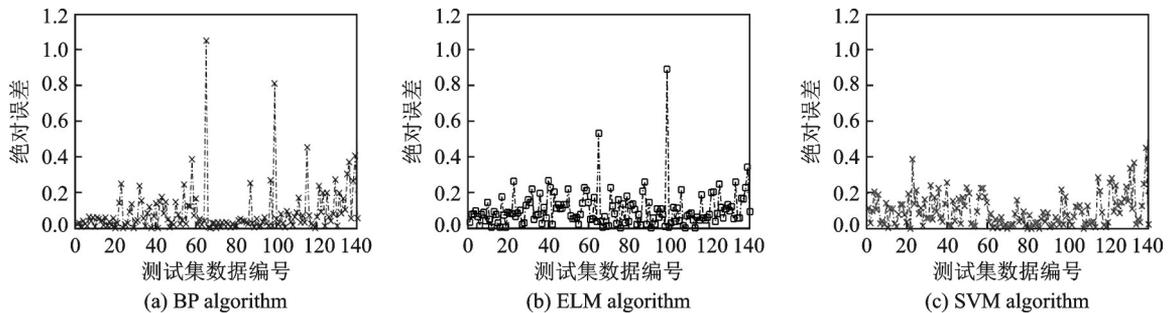


图9 基于不同算法预混火焰燃烧当量比回归预测绝对误差结果对比

Fig.9 AE comparison of combustion equivalence ratio regression prediction for premixed flames based on different algorithms

表4 不同算法表现性能结果对比

Table 4 Comparison of performance results of different algorithms

算法	MSE	$R^2/\%$
BP	0.030	77.33
ELM	0.031	75.51
SVM	0.023	81.31

5 结束语

本文提出一种基于火焰图像颜色特征工程的甲烷预混火焰燃烧当量比软测量模型构建方法。相比传统单一颜色特征作为模型输入,本文在RGB、HSV和YCbCr不同颜色空间模型下,设计了多种颜色分量以及颜色分量之间的比值共计36维数据作为颜色特征,进而应用Spearman相关性分析和RF算法实现火焰颜色特征选择进行降维处理,最终筛选出表征燃烧当量比更深层次的16维火焰颜色特征作为模型输入,构建了甲烷预混火焰颜色特征工程。同时,相对传统简单的拟合数据建模方法,本文提出了通过改进GSM寻求最优参数优化SVM算法建模,结果显示在构建的火焰颜色特征工程数据集上表现出较好的回归预测结果。最后,通过与传统的BP神经网络和ELM算法对比发现,SVM算法回归预测效果在构建的特征数据集上优于BP与ELM。

尽管本文提出的甲烷预混火焰燃烧当量比测量方法取得了较好的回归预测结果,但是仍然存在一些不足之处:(1)针对基于Spearman相关性分析和RF算法特征选择阈值设置主要参考相关研究选择与对比,并没有深入研究与系统分析;(2)模型评估仅采用均方误差和平方相关系数指标进行评价。下一步工作将把本文方法应用到不同燃料燃烧当量比检测,同时对本文存在的不足作进一步的研究,包括多元颜色模型选择(例如引入与设备无关的CIE 1931颜色模型等)、参数选择以及引入多评价指标的模型评估方法。

参考文献:

- [1] 李鹏飞,王飞飞,米建春,等.当量比和初始混合模式对无焰燃烧的影响[J].工程热物理学报,2011,32(9):1592-1596.
LI Pengfei, WANG Feifei, MI Jianchun, et al. Effects of equivalence ratio and initial mixing mode on flameless combustion[J]. *Journal of Engineering Thermophysics*, 2011, 32(9): 1592-1596.
- [2] 王伟豪.基于化学发光特性的火焰动力学实验方法研究[D].杭州:浙江工业大学,2018.
WANG Weihao. Experimental method of flame dynamics based on chemiluminescence characteristics [D]. Hangzhou: Zhejiang University of Technology, 2018.
- [3] KOJIMA J, IKEDA Y, NAKAJIMAT. Spatially resolved measurement of OH^* , CH^* , and C_2^* chemiluminescence in the reaction zone of laminar methane/air premixed flames[J]. *Proceedings of the Combustion Institute*, 2000, 28(2): 1757-1764.
- [4] 张浪,谭建国,刘瑶.甲烷/空气同轴射流火焰反应热与 OH^*/CH^* 定量表征特性研究[J].光谱学与光谱分析,2020,40(6):1703-1709.
ZHANG Lang, TAN Jianguo, LIU Yao. Study on reaction heat and OH^*/CH^* quantitative characterization of methane/air coaxial jet flame [J]. *Spectroscopy and Spectral Analysis*, 2020, 40(6): 1703-1709.
- [5] CLARK T P. Studies of OH, CO, CH and C_2 radiation from laminar and turbulent propane-air and ethylene-air flames: 19930085300[R]. USA: NTRS-NASA, 1958.
- [6] HABER L C. An investigation into the origin, measurement and application of chemiluminescent light emissions from premixed flames[D]. Blacksburg, USA: Virginia Polytechnic Institute and State University, 2000.
- [7] HUANG H W, ZHANG Y. Flame colour characterization in the visible and infrared spectrum using a digital camera and image processing[J]. *Measurement Science & Technology*, 2008, 19(8): 085406.

- [8] HUANG H W, ZHANG Y. Digital colour image processing based measurement of premixed CH_4 +air and C_2H_4 +air flame chemiluminescence[J]. *Fuel*, 2011, 90: 48-53.
- [9] MIGLIORINI F S, MAFFI S, IULIIS S, et al. Analysis of chemiluminescence measurements by grey-scale ICCD and colour digital cameras[J]. *Measurement Science and Technology*, 2014, 25 (5): 055202.
- [10] HUANG H W, ZHANG Y. Dynamic application of digital image and colour processing in characterizing flame radiation features [J]. *Measurement Science and Technology*, 2010, 21: 085202.
- [11] HUANG H W, WANG Q T, HAO J T, et al. Characterisation of external acoustic excitation on diffusion flames using digital color image processing[J]. *Fuel*, 2012, 94: 102-109.
- [12] HUANG H W, ZHANG Y. Imaging based chemiluminescence characterisation of partially premixed syngas flames through DFCD technique[J]. *International Journal of Hydrogen Energy*, 2013, 38: 4839-4847.
- [13] HUANG H W, YANG J, WANG Q, et al. Variation of hydrocarbon compositions and ignition locations on the radiative flame initiation characteristics through multi-dimensional DFCD incorporated image analysis[J]. *Fuel*, 2013, 103: 334-346.
- [14] YANG J, MOSSA F, HUANG H W, et al. Oscillating flames in open tubes[J]. *Proceedings of the Combustion Institute*, 2015, 35: 2075-2082.
- [15] YANG J, MA Z, ZHANG Y. Improved colour-modelled CH^* and C_2^* measurement using a digital colour camera[J]. *Measurement*, 2019, 141: 235-240.
- [16] 龙凯,杨健晟.基于成像颜色校正的机器视觉测量甲烷预混火焰当量比的方法研究[J].*智能计算机与应用*,2020,10(3): 288-293, 298.
LONG Kai, YANG Jiansheng. Measurement of equivalence ratio of methane premixed flame using machine vision based on image color correction[J]. *Intelligent Computer and Applications*, 2020, 10(3): 288-293, 298.
- [17] TRINDADE T, FERREIRA A, FERNANDES E. Characterization of combustion chemiluminescence: An image processing approach[J]. *Procedia Technology*, 2014, 14: 194-201.
- [18] 严云洋,唐岩岩,刘以安,等.使用多尺度LBP特征和SVM的火焰识别算法[J].*山东大学学报(工学版)*,2012,42(5): 47-52, 58.
YAN Yunyang, TANG Yanyan, LIU Yian, et al. Flame recognition algorithm using multi-scale LBP features and SVM [J]. *Journal of Shandong University (Engineering Science)*, 2012, 42 (5): 47-52, 58.
- [19] 李新利,李玲,卢钢,等.基于火焰自由基成像和支持向量机的燃烧过程 NO_x 排放预测[J].*中国电机工程学报*,2015,35(6): 1413-1419.
LI Xinli, LI Ling, LU Gang, et al. Prediction of NO_x emission in combustion process based on flame free radical imaging and support vector machine[J]. *Proceedings of the CSEE*, 2015, 35(6): 1413-1419.
- [20] 丁绪达,刘潇,金秀章.基于压缩感知最小二乘支持向量机的 NO_x 软测量模型[J].*热力发电*,2018,47(3): 76-81.
DING Xuda, LIU Xiao, JIN xiuzhang. NO_x soft sensing model based on compressed sensing least squares support vector machine[J]. *Thermal Power Generation*, 2018, 47(3): 76-81.
- [21] APPANA D K, ISLAM R, KHAN S A, et al. A video-based smoke detection using smoke flow pattern and spatial-temporal energy analyses for alarm systems[J]. *Information Sciences*, 2017, 418/419: 91-101.
- [22] 李新利,李楠,孙愉佳,等.火焰自由基成像和极限学习机在 NO_x 排放预测中的研究[J].*系统仿真学报*,2016,28(5):1179-1185.
LI Xinli, LI Nan, SUN Yujia, et al. Flame free radical imaging and extreme learning machine in NO_x emission prediction[J]. *Journal of System Simulation*, 2016, 28(5): 1179-1185.
- [23] 曹江涛,秦跃雁,姬晓飞.基于视频的火焰检测算法综述[J].*数据采集与处理*,2020,35(1): 35-52.
CAO Jiangtao, QIN Yueyan, JI Xiaofei. Survey of video based flame detection algorithms[J]. *Data Acquisition and Processing*, 2020, 35(1): 35-52.
- [24] LIU Z G, YANG Y, JI X H. Flame detection algorithm based on a saliency detection technique and the uniform local binary pattern in the YCbCr color space[J]. *Signal Image & Video Processing*, 2016, 10(2): 277-284.
- [25] CHEN Z Z, TIAN H, DING H J, et al. Video flame recognition based on feature fusion and extreme learning machine[C]// *Proceedings of 5th International Conference on Automation, Control and Robotics Engineering (CACRE 2020)*. Dalian, China:

[s.n.], 2020.

- [26] 耿庆田, 于繁华, 赵宏伟, 等. 基于颜色特征的火焰检测新算法[J]. 吉林大学学报(工学版), 2014, 44(6): 1787-1792.
GENG Qingtian, YU Fanhua, ZHAO Hongwei, et al. A new flame detection algorithm based on color feature [J]. Journal of Jilin University (Engineering Science), 2014, 44 (6): 1787-1792.
- [27] ZHANG W Y, WEI Z W, WANG B H, et al. Measuring mixing patterns in complex networks by Spearman rank correlation coefficient[J]. Physica A Statistical Mechanics & Its Applications, 2016, 451: 440-450.
- [28] GENUER R, POGGI J M, TULEAU-MALOT C. Variable selection using random forests[M]. USA: Elsevier Science Inc., 2010.
- [29] STATNIKOV A, WANG L, ALIFERIS C F, et al. A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification [J]. BMC Bioinformatics, 2008. DOI: 10.1186/1471-2105-9-319.
- [30] 郭建民, 刘石, 姜凡, 等. 基于SVM的对冲燃煤锅炉NO_x排放特性[J]. 燃烧科学与技术, 2006(3): 243-247.
GUO Jianmin, LIU Shi, JIANG Fan, et al. Characteristics of NO_x emission from opposed firing boilers based on SVM [J]. Combustion Science and Technology, 2006 (3): 243-247
- [31] HAO X, ZHANG G, MA S. Deep learning[J]. International Journal of Semantic Computing, 2016, 10(3): 417-439.
- [32] LLOYD B G R. Support vector machines for classification and regression[J]. Analyst, 2010, 135(2): 230-267.
- [33] 边靖伟, 寇立伟, 项基. 应用 PSO 和 SVM 的水下航行器黑箱建模[J]. 哈尔滨工业大学学报, 2019, 51(10): 55-60, 82.
BIAN Jingwei, KOU Liwei, XIANG Ji. Black box modeling of underwater vehicle based on PSO and SVM [J]. Journal of Harbin Institute of Technology, 2019, 51(10): 55-60, 82.
- [34] SMOLA A J, SCHLAKOPF B. A tutorial on support vector regression[J]. Statistics and Computing, 2004, 14(3): 199-222.
- [35] 李新利, 李一娇, 卢钢, 等. 基于火焰光谱和特征工程的生物质燃料识别[J]. 中国电机工程学报, 2018, 38(15): 4474-4481.
LI Xinli, LI Yijiao, LU Gang, et al. Biomass fuel identification based on flame spectrum and feature engineering [J]. Proceedings of the CSEE, 2018, 38 (15): 4474-4481.
- [36] XIAO Z, YE S J, ZHAGN B, et al. BP neural network with rough set for short term load forecasting[J]. Expert Systems with Applications, 2009, 36(1): 273-279.
- [37] HUANG G B, ZHU Q Y, SIEW C K. Extreme learning machine: Theory and applications[J]. Neurocomputing, 2006, 70(1/2/3): 489-501.

作者简介:



陈长友(1993-), 男, 硕士研究生, 研究方向: 机器视觉燃烧诊断技术, E-mail: 1753100818@qq.com。



傅钰雯(1988-), 女, 硕士, 讲师, 研究方向: 大数据分析, E-mail: 553408481@qq.com。



涂沛驰(1995-), 男, 硕士研究生, 研究方向: 机器视觉燃烧诊断技术, E-mail: 809888743@qq.com。



舒文(1996-), 男, 硕士研究生, 研究方向: 机器视觉燃烧诊断技术, E-mail: 771964326@qq.com。



杨健晨(1987-), 通信作者, 男, 博士, 副教授, 硕士生导师, 研究方向: 机器视觉燃烧诊断技术方面相关研究, E-mail: jsyang3@gzu.edu.cn。

(编辑: 张黄群)