

文章编号: 2095-2163(2021)08-0056-05

中图分类号: R338.63; TP181

文献标志码: A

# 基于支持向量机的睡眠监测系统设计

钟佳良, 易 钢

(湖南中医药大学 信息科学与工程学院, 长沙 410208)

**摘要:** 睡眠问题逐渐成为当今快节奏社会人们对生命健康的关注重点, 由此本文提出了一套基于支持向量机的睡眠监测系统。该系统利用生物传感器, 对心率和呼吸信号的实时监测, 通过 Wi-Fi 的方式将数据传送到上位机, 由设计好的机器学习算法, 对睡眠呼吸暂停低通气综合征等进行判断, 并搭配终端 APP 便于使用者跟踪病情。

**关键词:** 支持向量机; 睡眠监测; 机器学习

## Design of sleep monitoring system based on Support Vector Machine

ZHONG Jiali, YI Gang

(School of Informatics, Hunan University of Chinese Medicine, Changsha 410208, China)

**[Abstract]** Sleep problem has gradually become the focus of people's life and health in today's fast-paced society. Therefore, this paper proposes a sleep monitoring system based on Support Vector Machine. The system mainly uses biosensors to monitor heart rate and breathing signals in real time, and transmits data to the upper computer through Wi-Fi. The designed machine learning algorithm is used to judge sleep apnea-hypopnea syndrome. In order to facilitate the user to track the condition, it is also designed a terminal APP.

**[Key words]** support vector machine; sleep monitoring; machine learning

## 0 引言

睡眠是人们生活中不可缺少的重要生命活动, 但如今快节奏的生活, 以及各方面的压力, 使得人们出现了各种严重的睡眠障碍问题。例如: 失眠、呼吸暂停等, 会对人们次日的学习和工作造成负面影响。其中, 睡眠呼吸暂停低通气综合征 (sleep apnea-hypopnea syndrome, SAHS) 是其中较为典型的一种疾病。其临床表现为: 患者上呼吸道塌陷, 使得在睡眠时出现呼吸阻塞甚至呼吸暂停的现象。SAHS 不仅仅影响着人们日常的睡眠质量, 还带来了一系列的隐性并发症。如心律失常、心肌梗塞、高血压、中风等<sup>[1]</sup>。根据调查发现, 全球大约有 9%~38% 的人深受 SAHS 的困扰。而肥胖人口增加和老龄化程度加深的社会情况加剧之下, 该病的发病率也在逐年增高<sup>[2]</sup>。但是, 由于患者对 SAHS 了解不深, 该病往往得不到应有的重视, 使患者深受其并发症的困扰。因此, 对睡眠状态的监测和识别作为诊治 SAHS 的前提, 越来越多科研人员投入对其研究工作中。

现今临床上, 对 SAHS 进行分析诊断的最权威方

法是多导睡眠监测仪 (Polysomnography, PSG), PSG 对患者睡眠时的一系列相关生理数据进行采集 (如: 脑电、心电等), 再由专业的实验记录人员针对这些数据进行手动评估得出诊断结果<sup>[3]</sup>。虽然 PSG 的诊断具有高度的可靠性, 但不仅对操作人员专业要求高, 而且受试者需要佩戴一系列传感设备, 严重影响受试者的真实睡眠状态。因此, PSG 这类传统睡眠监测系统不仅不适宜作为对 SAHS 等睡眠疾病的初步筛查手段, 也难以布置在居家环境中使用。

为了使患者能够早期了解自身的患病情况, 本文基于物联网设计了一套适用于居家环境的无线睡眠监测系统。该系统通过采集患者睡眠时的呼吸和心率信号, 采用无线模块将得到的睡眠信号传输至上位机进行分析处理, 采用机器学习方法对 SAHS 进行判断, 最后将初步的诊断结果传送到手机终端, 方便患者了解自己的病情, 也能够给医生提供较为准确的诊断依据。

## 1 系统设计

### 1.1 整体设计方案

系统结构由睡眠监测终端、上位机服务器和移

**基金项目:** 湖南中医药大学电子科学与技术学科开放基金 (2018DK05)。

**作者简介:** 钟佳良 (1996-), 男, 硕士研究生, 主要研究方向: 物联网、机器学习; 易 钢 (1976-), 男, 硕士, 副教授, 主要研究方向: 中医药信息采集与传输技术研究。

**通讯作者:** 易 钢 Email: 6636244@qq.com

**收稿日期:** 2020-05-25

动端 APP 3 部分组成,如图 1 所示。

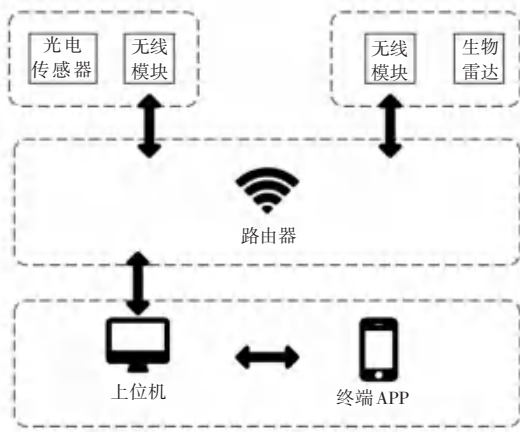


图 1 系统结构

Fig. 1 System Structure

其中,睡眠监测终端将生物传感器获取到的使用者的生理信息,通过无线模块将睡眠信号传输至上位机中对睡眠信号进行分析处理,最后获得使用者的诊断信息并使用户能够通过移动端 APP 查看报告。

### 1.2 系统硬件设计

系统硬件主要由心率监测模块、呼吸监测模块、无线模块组成:

(1) 心率监测模块: 心率监测模块采用手环的形式,供使用者在睡眠时佩戴,选用光电反射式传感器 PAH8011ET 作为监测使用者心率的芯片。该芯片是一款具有集成 DSP 的低功耗 CMOS 工艺光学心率监测芯片,与 3 个板载芯片 LED 封装在一起: 其中 3 个 LED 分别为两个绿色光源,一个红外光源。其基于光学传感技术,可捕获比传统光电二极管更高分辨率的图像,然后通过集成的 DSP 处理图像,获得 PPG(光电容积描记)数据,用来推断心率。

(2) 呼吸监测模块: 呼吸监测模块选用 Novalda 公司生产的 X4M03 型号的超宽带雷达模块作为核心器件,将其封装在一个床头摆件中充当手环的放置平台,该雷达是一种超宽带短程脉冲雷达收发 SoC(System on Chip),工作中心频率为 7.29 GHz,工作带宽为 1.5 GHz,有效范围在 10 m 左右。其主要的组件是一组发射-接收天线和微型控制模块。其基本工作流程为:先由发射天线将电磁波发射出去,再将人体反射回的电磁波由接收天线接收,最后由微型控制模块对信号进行初步的处理,并将其通过 SPI 接口传输至无线模块中。

(3) 无线模块: 无线模块主要是将心率监测和呼吸监测两大模块的数据,通过 Wi-Fi 方式传输至

上位机中。本系统选用乐鑫公司开发的 ESP8266 芯片,该芯片集成了完整且自成体系的 Wi-Fi 网络功能,集成的 32 位内核处理器和片上 SRAM,使其作为数据传输桥梁的同时还有较强的片上计算和存储能力。本系统中,ESP8266 主要承担数据通道的作用,并不涉及到数据的相关运算。

### 1.3 系统软件设计

本系统的软件设计主要分为两部分,一是系统功能的实现,二是移动终端 APP 的设计与制作,具体实现步骤如下:

(1) 系统的初始化。其中包括无线传输模块内部寄存器的初始化,芯片各种外设接口的初始化,两个监测模块的初始化,Wi-Fi 服务的初始化等。

(2) 两个监测模块分别实时采集使用者的心率和呼吸信号后,通过无线传输模块将以上数据通过 Wi-Fi 传输至上位机中,上位机通过设计好的机器学习算法根据使用者的睡眠信号进行分类,来判断是否出现 SAHS 的症状,并生成睡眠报告。

(3) 移动终端 APP 的工作主要是通过图表等报告形式,使得用户能够方便地了解自己的睡眠状况; 监控人员能够实时的查看使用者的心率和呼吸数据,在病情严重时能够利用这些数据供医生进行初步评估,方便患者进一步的治疗。

系统软件流程如图 2 所示。

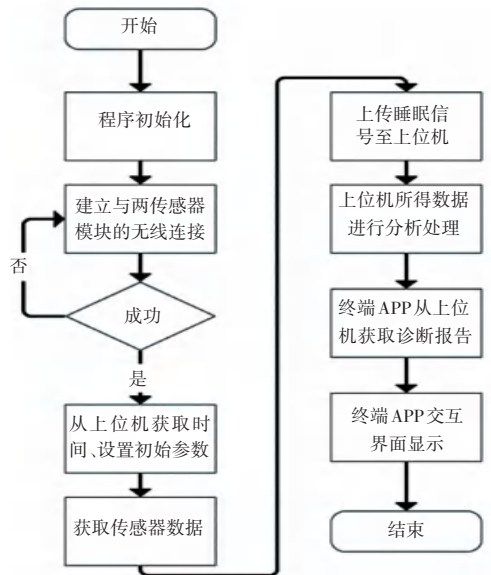


图 2 系统软件流程

Fig. 2 Process of System Software

## 2 SVM 分类算法

### 2.1 原理简述

现阶段性能优秀的分类算法有很多。例如:单

一的分类方法(决策树、K近邻),支持向量机(Support Vector Machine, SVM)和人工神经网络等,也有像随机森林和XGBoosting这样集成学习算法<sup>[4]</sup>。上述这些分类算法并不是照搬即可使用,而需要针对应用场景和识别目标对分类器进行改善。下面具体介绍SVM算法。

SVM是一种能够将线性数据进行分类的算法,由于其直观的原理和较优的性能,在许多领域得到了应用。其基本分类原理是:在给定的样本空间中寻找一个性能最佳的超平面,该平面能够将正负样本最大程度地分离。SVM的基本原理如图3所示。

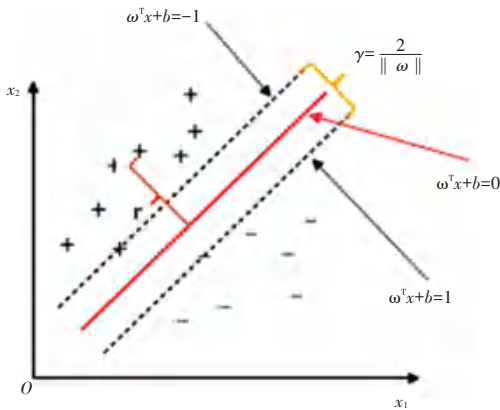


图3 SVM原理

Fig. 3 Principle of SVM

假设:已知的样本集为  $(x_1, x_2), i = 1, 2, \dots, n$ 。

其中,  $x_1$  为样本点的  $d$  维特征,  $x_2$  代表样本所属的类别,在图中分别用正负号表示。图中  $\omega^T x + b = 0$  即为划分两个样本类的超平面,  $\gamma = \frac{2}{\|\omega\|}$  表示

由两支持向量  $\omega^T x + b = -1$  和  $\omega^T x + b = 1$  隔开的分类间隔。而寻找最佳超平面,其实就是寻找能够使得  $\|\omega\|^2$  最小的超平面。SVM作为一种线性分类器,无法处理非线性的数据,所以为了达到分离非线性数据的目的,需要选择合适的核函数,将特征向量投射至高维空间中进行分类,这样做的同时还能降低运算的复杂度。常用的核函数见表1。

表1 常用核函数表

Tab. 1 Several Kernels of SVM

函数名称	公式
线性核	$k(x_i, x_j) = x_i^T x_j$
多项式核	$k(x_i, x_j) = (x_i^T x_j)^d$
高斯核	$k(x_i, x_j) = \exp\left(-\frac{\ x_i - x_j\ ^2}{2\sigma^2}\right)$
Sigmoid核	$k(x_i, x_j) = \tan h(\beta x_i^T x_j + \theta)$

## 2.2 特征提取

检测是否发生睡眠呼吸暂停这样的问题,实际上是一个分类问题,目的是将患有SAHS的患者的睡眠数据和正常的睡眠数据区分开。分类问题的解决,通常都需要进行特征提取和分类器设计两方面的工作。其中,特征提取的数据准备工作质量,对后续分类器得出的结果有很大的影响。

### 2.2.1 心率信号

对心率信号的需要,主要是从其中得到两个R波之间的时间间隔,这个间隔也被称作RR间期。该参数是分析连续心跳之间瞬时心率差异的关键,该差异在医学上被冠名为心率变异性(Heart Rate Variability, HRV)。

临床研究上,对RR间期做HRV分析主要采用线性分析方法,分别对HRV进行时域和频域进行分析。现今在时域中,对于HRV的分析方法已经较为成熟,常常采用统计学的方法,根据RR间期的生物学特征,来制定相应的指标进行描述<sup>[5]</sup>。例如SDNN,意为RR间期的标准差,能够直观地表征心率变化的集中程度。而频域分析则根据人体正常的心率信号功率谱能量在0~0.4 Hz这一依据<sup>[6]</sup>,采用小波变换的方法分别对极低频段、低频段和高频段这3方面进行特征的提取,这3种频段的频率分别为0~0.04 Hz、0.04~0.15 Hz和0.15~0.4 Hz。

### 2.2.2 呼吸信号

根据PSG中对SAHS的判定标准,睡眠过程中口鼻气流停止,呼吸幅度较基准水平降低程度 $\geq 30\%$ ,且持续时间 $\geq 10$  s<sup>[7]</sup>,也能够利用统计学方法,对呼吸信号指定一些指标,对其进行时域分析。例如SDRP,意为呼吸波峰间隔标准差,表示呼吸变化的集中程度。对呼吸信号的频域分析与心率信号类似,同样采用小波变换对3个频段进行特征提取。

## 2.3 分类器设计

经过对心率信号和呼吸信号的特征提取后,就能够根据这些特征来对SAHS进行分类。但是,由于各项特征的量纲不同,直接将特征作为输入向量会导致分类效果较差,因此需要对这些特征做归一化处理,不仅能够降低运算的复杂度还能提高收敛的效率。SVM能够利用不同的核函数对非线性的数据进行分类,所以选择一个合适的核函数来寻找更优的超平面显得极为重要。

在对SVM进行训练的过程中,设计采用了交叉验证的方式,其工作思想是将训练数据按一定比例分为训练样本集与测试样本集。在训练过程中,只

使用拆分得到的训练样本集对 SVM 进行训练, 完成模型的训练后, 对训练得到的 SVM 使用测试样本集进行验证, 最后根据一些评判标准对分类结果进行评估, 再选取效果最优的 SVM 分类器<sup>[8]</sup>。

SVM 算法的具体流程如图 4 所示。

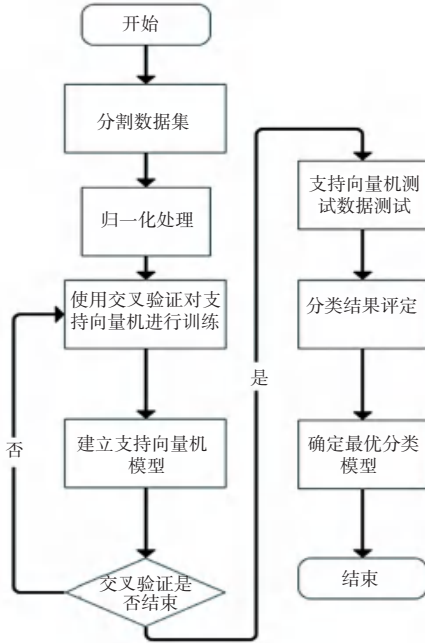


图 4 SVM 算法流程

Fig. 4 Process of SVM Algorithm

## 3 数据与验证

### 3.1 数据集获取

本研究中用来建立和测试模型的心率信号是来自 PhysioNet Apnea-ECG 数据集, 该数据集由 Phillips University 的 Dr. Tomas Penzel 提供。其中包含了一组开放和一组封闭的数据集合, 分别用作模型的训练和测试, 各集合中有 35 条从 PSG 记录中提取得到的 ECG 信号数据。这些 ECG 信号的分辨率为 16 bit, 采样频率为 100 Hz。医学专家结合 PSG 中的其它信号, 以分钟为单位, 对 ESG 信号进行标注, 并根据 AHI 的值, 将其分为 3 个等级。当记录的 AHI 低于 5 时, 该时间段被标注为 C 级, 意为正常的呼吸状况; 当记录的 AHI 高于 10 时, 该时间段被标注为 A, 意为出现呼吸暂停症状; 在当 AHI 值介于 5~10 之间, 表示临界状态, 该时间段被标注为 B。

### 3.2 RR 间期特征提取

在实验过程中选择了 6 项时域特征和 3 项频域特征, 作为 SVM 分类器的分类特征。其中 6 项时域特征分别为:  $MRR$  (RR 间期平均值)、 $MHR$  (心率平均值)、 $RMSSD$  (相邻窦性间期的均方根)、

$SDNN$ 、 $NN50$  (相邻窦性间期之差超过 50 ms 的个数) 和  $pNN50$  ( $NN50$  的占比数)<sup>[9]</sup>。主要特征计算如式(1)~式(3)。

$$MRR = \widetilde{RR} = \frac{\sum_{n=1}^N RR(n)}{N}, \quad (1)$$

$$RMSSD = \sqrt{\frac{\sum_{n=2}^N (RR(n) - RR(n-1))^2}{N-1}}, \quad (2)$$

$$SDNN = \sqrt{\frac{\sum_{n=1}^N (RR(n) - \widetilde{RR})^2}{N-1}}. \quad (3)$$

其中,  $RR(n)$  表示  $RR$  间期信号,  $N$  为  $RR$  间期的个数。

### 3.3 性能评估

模型的性能评估根据准确率 (*Accuracy*)、精确率 (*Precision*) 和召回率 (*Recall*) 这 3 个指标来进行评定。其计算公式如下:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad (4)$$

$$Precision = \frac{TN}{TP + FP}, \quad (5)$$

$$Recall = \frac{TP}{TP + FN}. \quad (6)$$

其中,  $TP$  和  $TN$  分别表示被分类器正确分为阳性样例和阴性样例的数量,  $FP$  和  $FN$  分别表示被分类器错分为阳性样例和阴性样例的数量<sup>[10]</sup>。

### 3.4 结果与分析

首先, 将预处理后的数据输入采用不同核函数进行非线性变换的 SVM 分类器, 通过对模型的训练和验证, 其结果见表 2。

表 2 不同核函数 SVM 性能比较

Tab. 2 Comparison of SVM Using Different Kernels

核函数	Accuracy/%	Precision/%	Recall/%
线性核	85.3	80.4	88.4
多项式核	84.8	74.5	91.2
高斯核	86.8	80.7	90.7
Sigmoid 核	78.3	76.0	79.7

根据表 2 数据可以看出, SVM 选择不同核函数对非线性问题进行分类, 确实会得到性能不同的 SVM 分类器。在对本次研究中处理过后的数据进行分类可知, 选取高斯核作为 SVM 分类器的核函数, 能够得到较为优秀的超平面实现对 SAHS 的检测。

为了验证本研究中所使用的 SVM 算法的有效

性,选取了一些传统的机器学习算法对预处理后的数据进行对比实验,其比较的结果见表3。

表3 模型性能比较

Tab. 3 Comparison of Several Models

方法	Accuracy/%	Precision/%	Recall/%
LR	81.5	72.0	87.4
LDA	81.8	70.9	88.4
SVM	86.8	80.7	90.7

由表3可以看出,SVM算法在根据预处理后的心率信号进行SAHS检测具有一定的优越性,但由于数据集的限制,无法将呼吸信号也加入分类器中作为特征输入。因此,当本设计收集完初期数据后,将心率信号和呼吸信号分别建立SVM模型,最后将两者的性能进行对比验证,能够得到性能更加优越的SVM模型,为SAHS患者进行初步诊断提供了较为可靠的支持。

## 4 结束语

睡眠问题在现今社会人们越来越重视生命健康和生活质量的背景下,逐渐获得了更多人的重视,但由于医院专业设备的繁重复杂和紧缺,使得很多有睡眠问题的患者并不能及时得到诊疗。本文利用物联网技术将无线终端及睡眠监测相结合,提出的睡眠监测系统能够使得患者在居家环境中对自己的病情有

初步的了解,为进一步寻求专业医生提供了帮助。该系统的推出,不仅对提高人们的睡眠质量有很大的帮助,同时也能够大大减轻公共医用资源的压力。

## 参考文献

- [1] 孙东升,徐少坤,张丽,等. 阻塞性睡眠呼吸暂停低通气综合征相关性高血压研究进展[J]. 中国医药,2020,15(10):1642-1645.
- [2] Akhter Shahin, Abeyratne Udantha R, Swarnkar Vinayak, et al. Snore Sound Analysis Can Detect the Presence of Obstructive Sleep Apnea Specific to NREM or REM Sleep.[J]. Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine,2018,14(6).
- [3] 许益彬. 基于非接触式睡眠监测系统的呼吸暂停综合征检测研究[D]. 北京:北京邮电大学,2019.
- [4] 柴建朴. 基于支持向量机的睡眠分期算法研究及应用实现[D]. 成都:电子科技大学,2018.
- [5] 马帅,侯世科,樊毫军,等. 智能心电信号监测设备研究现状[J]. 医疗卫生装备,2020,41(11):95-99.
- [6] 张硕. 可穿戴式睡眠呼吸暂停综合症检测系统研究[D]. 长春:长春理工大学,2020.
- [7] 张令晖,冀琳琳,于旭红,等. 阻塞性睡眠呼吸暂停低通气综合征患者睡眠参数间的关系[J]. 现代养生,2020,20(Z6):30-31.
- [8] 翁勇强. 支持向量机在医学数据中的应用研究[D]. 北京:北京邮电大学,2017.
- [9] 武悦. 基于心电信号对睡眠呼吸暂停综合征判别算法的研究[D]. 长春:吉林大学,2018.
- [10] 余辉,王硕,李心蕊,等. 基于LSTM-CNN的睡眠呼吸暂停与低通气事件实时检测算法研究[J]. 中国生物医学工程学报,2020,39(3):303-310.
- [1] 王新雨,汪驰升. 基于深度学习的密集人群安全监测系统[J]. 物联网技术,2019,9(11):8-12,17.
- [2] 梁静,杨君涛. 大型社会活动安全风险指标体系构建的思考[J]. 现代职业安全,2015(4):71-73.
- [3] 郭晟楠,林友芳,金文蔚,等. 基于时空循环卷积网络的城市区域人流量预测[J]. 计算机科学,2019,46(S1):385-391.
- [4] ZHANG J, ZHENG Y, QI D. Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction [J]. The Association for the Advance of Artificial Intelligence (AAAI), 2017:1655-1661.
- [5] JUNG Minju, LEE Haanvid, TANI Jun. Adaptive detrending to accelerate convolutional gated recurrent unit training for contextual video recognition[J]. Neural networks : the official journal of the International Neural Network Society,2018: 356-370.
- [6] LIN Z, FENG J, LU Z, et al. DeepSTN+: Context-Aware Spatial-Temporal Neural Network for Crowd Flow Prediction in Metropolis[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2019:1020-1027.
- [7] SUN J, ZHANG J, LI Q, et al. Predicting Citywide Crowd Flows in Irregular Regions Using Multi-View Graph Convolutional Networks [J]. IEEE Transactions on Knowledge and Data Engineering, 2020:1-1.
- [8] 重点区域人群密度预测[EB/OL]. <https://www.datafountain.cn/competitions/428>. 2020-02-22/2020-05-30.
- [9] 2020-CCF-Crowd-FlowPrediction[EB/OL].<https://github.com/agave233/2020-CCF-Crowd-Flow-Prediction>, 2020-06-27/2020-08-01.
- [10] ZHOU Jie, CUI Ganqu, HU Shengding, et al. Graph neural networks: A review of methods and applications[J]. AI ,2020:57-81
- [11] MACHADO M R, KARRAY S, DE SOUSA I T. LightGBM: an Effective Decision Tree Gradient Boosting Method to Predict Customer Loyalty in the Finance Industry [J]. International Conference on Computer Science & Education (ICSE), 2019: 1111-1116 .
- [12] CHEN T, Carlos Guestrin. XGBoost: A Scalable Tree Boosting System[C]//Proceedings of the 22<sup>nd</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,2016:785-794.
- [13] ZHAO Zheng, CHEN Weihai, WU Xingming, et al. LSTM network: a deep learning approach for short-term traffic forecast [J]. IET Intelligent Transport Systems, 2017:68-75.
- [14] WULFF S S. Time series analysis: Forecasting and control, 5<sup>th</sup> edition[J]. Journal of Quality Technology, 2017:418-419.

(上接第55页)

## 参考文献