

文章编号: 2095-2163(2022)02-0191-05

中图分类号: TB535

文献标志码: A

基于 LS-SVR 的高速列车车内声品质主观评价

王增政, 王岩松, 郭辉, 袁涛, 郑立辉, 孙裴

(上海工程技术大学 机械与汽车工程学院, 上海 201620)

摘要: 针对传统主观评价方法用于高速列车车内噪声信号评价存在效率低、评价时间长、可重复性差等问题, 本文运用心理声学参数表征主观评价结果, 从而建立评价模型, 以探究两者之间的复杂非线性关系, 用最小二乘法对原始信号进行点拟合, 反向筛选短时噪声信号, 对筛选出的声音信号分别计算声压级、响度、粗糙度、尖锐度、抖动度和 AI 指数 6 种心理声学参数的算数平均值。以“烦躁度”作为主观评价指标, 用单位间隔的语义细分法作为主观评价方法来主观评价实验得到评价结果, 以样本集中样本的心理声学参数作为输入, 主观评价结果作为输出, 建立高速列车车内声品质的最小二乘法—支持向量机回归 (LS-SVR) 预测模型, 同时设置测试集对 LS-SVR 模型进行预测精度检验, 并与线性回归模型进行比较。检验结果发现, 在测试集中模型预测误差率较低, 比多元线性回归模型预测精度高, 说明 LS-SVR 模型针对高速列车车厢内声品质评价预测是有效适用的。

关键词: 最小二乘; 支持向量机; 高速列车; 声品质; 主观评价

Study on subjective evaluation interior sound quality of high-speed train based on LS-SVR

WANG Zengzheng, WANG Yansong, GUO Hui, YUAN Tao, ZHENG Lihui, SUN Pei

(School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

【Abstract】 Aimed at the problems of low efficiency, long evaluation time and poor repeatability when the traditional subjective evaluation method is used to evaluate the noise signal in high-speed train, psychoacoustic parameters are used to characterize the subjective evaluation results to establish an evaluation model to explore the complex nonlinear relationship. Point fitting of original signal is conducted with the least square method to reverse screen the short-term noise signal, and the average sound pressure level, loudness, roughness, sharpness of six psychoacoustic parameters of jitter and AI index are calculated. Taking irritability as the subjective evaluation index and the semantic segmentation method of unit interval as the subjective evaluation method, the subjective evaluation experiment is carried out to obtain the evaluation results. Taking the psychoacoustic parameters of the samples in the sample set as the input and the subjective evaluation results as the output, the LS-SVR prediction model of high-speed train interior sound quality is established. At the same time, the test set is used to test the prediction accuracy of LS-SVR model. Compared with the linear regression model, the test results show that the prediction error rate of the model in the test set is lower and the prediction accuracy is higher than that of the multiple linear regression model. It shows that LS-SVR model is effective and applicable for the evaluation and prediction of sound quality in high-speed train compartment.

【Key words】 least square (LS); support vector machine (SVM); high-speed train; sound quality; subjective evaluation

0 引言

目前, 针对高速列车车内声品质评价方式的研究, 大多数还是沿用传统的问卷调查和主观评价实验等评价形式。而高速列车车内噪声信号具有复杂性、多特征性, 这种简易的评价方式用于评价高速列车车内声品质的准确性和合理性有待商榷。因此, 通过结合信号本身的心理声学参数特性和人的主观评价结果, 利用多参数建立可以预测人的主观评价结果的评价模型是非常有必要的。

高速列车车内声品质主观评价研究中, 一般会根据评价对象和评价目标进行评价方法和评价模型的改进。Hardy^[1]采用了多种评价内部噪声环境的标准, 如: 噪声准则 (NC)、首选噪声准则 (PNC)、噪声等级 (NR)、房间准则 (RC), 研究轨道车辆噪声抑制过程中的固有频率, 确定了该噪声与乘客响应之间的关系; 周家中^[2]基于交通网络距离重新划分影响范围, 用分距离影响带的线型和指数型空间权重系数方程, 构建轨道列车交通客流的 LS-SVM 预测模型, 结果显示, 模型不仅可以简化数据的需求量,

基金项目: 国家自然科学基金 (52172371)。

作者简介: 王增政 (1996-), 男, 硕士研究生, 主要研究方向: 高速列车车内声品质评价研究。

通讯作者: 王岩松 Email: jzwbt@163.com

收稿日期: 2021-08-17

还能明显提高客流量的预测精度,但受地域因素影响较大;孟凡雨^[3]以高速列车车内噪声品质的评价参量和主观评价结果分别作为输入和输出,确定网络的结构和参数,建立具有预测主观评价结果功能的BP神经网络模型,研究声音的物理属性和心理声学属性之间的关系;郑德署^[4]开发了一个基于NET平台的噪声预测系统,通过图形化的形式展现,实现了对噪声的计算分析功能;Park等^[5]利用具有不同语音噪声比和背景噪声水平的噪声源与隔间噪声进行了两次实验,结果显示语音隐私和烦恼受声噪比的影响显著。

而对于高速列车车内声音信号的声音特性的筛选,以期用有效心理声学参数量化表征主观评价结果的研究还是相对较少,大部分还是用线性回归模型进行分析,难以很好地表达复杂的数据,而且对于具有相关性的特征数据难以建模。在此基础上,区别于最小二乘支持向量机,本文中的最小二乘法—支持向量机回归(LS-SVR)首先利用最小二乘法(LS)筛选得到样本,再通过支持向量机回归(SVR),建立高速列车车内声品质预测模型。

1 高速列车车厢内噪声信号采集及筛选

1.1 车厢内噪声信号采集

车厢噪声信号由数字人工头记录采集,并由LMS Test.lab软件同步记录噪声数据。每个测点记录4种不同工况类型的信号,根据ISO 3381:2005并结合实际情况选择5个测点,人工头在车厢内安装位置如图1和图2所示。图1中人工头的站姿高度为1.6 m,图2中人工头的坐姿高度为1.2 m。采集前用吊锤校准人耳与车厢地面的垂直度。



图1 人工头站姿

Fig. 1 Standing pose



图2 人工头坐姿

Fig. 2 Sitting pose

1.2 样本信号的筛选

列车不同工况平稳运行时,原始信号被记录,每次记录时长17 s。根据实验需要和标准要求,截取样本时长为5 s,目的是挑选合适的短时信号,用于评价实验。本着信号不重叠的原则,先由原始信号截取出每种工况下的3个样本,再根据每个样本的

A计权值进行拟合,常用的拟合方法有插值拟合和曲线拟合,插值拟合适合于理想测量情况(没有测量误差)的数据拟合,曲线插值则允许误差存在。而在实际测量中有不可避免的系统误差,所以选用曲线拟合。曲线拟合的原理是最小二乘法原理,即根据样本数据点拟合后,再反向挑选出与拟合线误差平方最小的样本点所在的样本,即为实验听音样本。这里对加速和减速工况下的原始样本进行截取时,虽然速度区间有差异,但是本质都是同一原始信号截取的,速度作为一个维度可以进行比较。而且5 s内速度变化引起的变化量可以忽略,如匀速240 km/h情况下,原始信号的3个不同时间段内的样本点拟合,如图3所示。

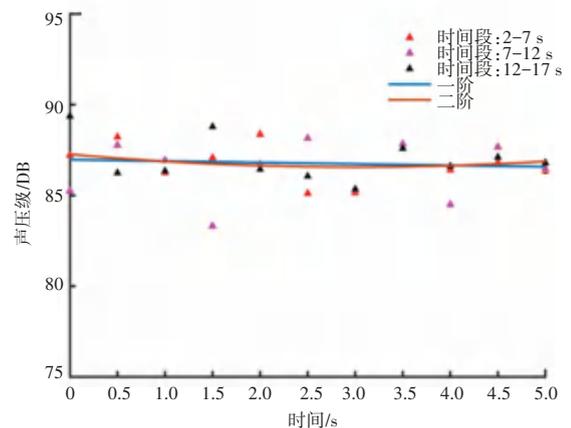


图3 同源信号不同时间段拟合

Fig. 3 Fitting of homologous signals of different time periods

目的是明确一阶(直线)拟合和二阶(曲线)拟合哪种拟合方式得到的拟合残差平方和更低,即确定样本中的时间和声压级之间的函数关系。分别计算了拟合后各样本点与拟合点间的残差平方和,见表1。由表1可知,同源不同时间段的一阶拟合残差平方和最低。故选择原始信号的2~7 s时间段作为实验样本。同理,按照此方法依次选取30个实验样本,得到样本后分别计算样本的声压级、粗糙度、尖锐度、抖动度和AI指数等心理声学参数值,见表2。

表1 残差平方和

Tab. 1 Sum of squares of residuals

时间段/s	残差平方和	
	一阶	二阶
2~7	20.26	22.38
7~12	26.97	28.63
12~17	30.74	35.67

表 2 样本的心理声学参数值

表 3 主观评价表

Tab. 2 Psychoacoustic parameter values of samples

Tab. 3 Subjective evaluation form

编号	声压级/响度(Z)/ DB	粗糙度/ sone	尖锐度/ asper	抖动度/ acum	抖动度/ vacil	AI/ %
1	85.72	24.56	0.257 0	0.673 7	0.237 3	0.95
2	82.63	21.74	0.264 1	0.689 7	0.260 7	0.91
3	80.74	20.36	0.303 2	0.736 1	0.276 6	0.92
4	88.12	26.74	0.347 7	0.891 5	0.363 4	0.61
5	90.78	25.97	0.321 9	0.863 0	0.338 9	0.85
6	91.87	28.13	0.271 9	0.675 9	0.197 3	0.74
7	78.97	21.03	0.335 0	0.771 7	0.347 1	0.75
8	73.68	19.74	0.262 0	0.740 3	0.247 4	0.85
9	74.63	19.43	0.294 6	0.825 2	0.279 6	0.92
10	75.95	19.97	0.285 9	0.891 5	0.377 9	0.83
11	73.74	19.46	0.214 5	0.826 1	0.286 3	0.99
12	76.37	20.17	0.292 1	0.848 3	0.296 3	0.96
...
...
28	85.23	24.98	0.253 4	0.834 8	0.257 6	0.97
29	86.71	25.96	0.287 1	0.673 7	0.313 4	0.95
30	84.96	24.63	0.345 1	0.753 6	0.287 3	0.61

极度	非常	比较	不烦躁	可接受	有点	比较	非常	极度
1	2	3	4	5	6	7	8	9

以 30 个样本作为横向维度, 21 名评价者作为纵向维度, 并由这 21 名评价者根据表 3 对样本进行打分, 最终得到 30×21 维评价结果数值矩阵, 对每个样本的横向主观得分求和, 并计算其算术平均值。30 个样本的烦躁度得分均值, 如图 4 所示。

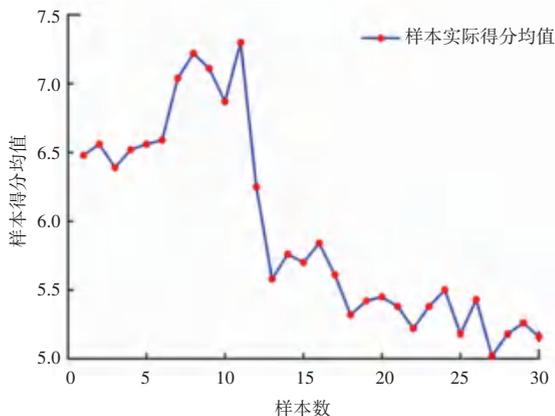


图 4 样本的实际得分均值

Fig. 4 Actual mean scores of the samples

2 主观评价实验

首先, 设计 9 级语义评分表, 结合数字等级评分法, 赋予形容词数字属性, 评分间隔为单位间隔。在区分对声音的“情感反应”和“情感评价”的词汇中, 常用的有尖锐度、愉悦度、烦恼度。尖锐度是通过人的分析能力判断得出的, 也就是对声音情感上的评价; 而烦恼度在情感表达上则较复杂, 是对声音情感上的反应, 这两个都是主观感觉的量化体现^[6]。但是尖锐度在主观的基础上略微偏客观。主观评价指标选择“烦躁度”, 相比于其他噪声评价指标, 高铁内部复杂噪声给人带来的烦躁程度更明显, 同时还设置具有“中性感觉”的样本作为参照, 避免评分差距过大, 评分等级为“5”, 主观评价见表 3。参考声样本由课题组声学专家在听完所有样本后根据声学经验选出, 选出的所有参照样本被试听完后都认同这种“中性感觉”。30 个声音打乱顺序随机呈现, 听者独自坐在隔音室, 待听者情绪平静后, 被指示用具有 9 个等级的评分表对样本进行打分, 从非常嘈杂到非常安静。30 个样本由 21 名评价者评价打分, 21 名评价者中有 14 位男性, 7 位女性, 每位评价者均没有听力障碍。评价实验后, 最终得到 30×21 维数值矩阵, 进而将抽象复杂的主观感受量化成数值, 对声音定量判断的评价结果通常可以作为一个区间尺度被认可。评价实验前, 对所有被试进行听音训练, 训练的声音在实验中不使用, 被试会被告知这些声音来自高铁车厢噪声。

3 高速列车车内声品质 LS-SVR 预测模型的建立

支持向量机 (SVM) 是一类按监督学习方式对数据进行二元分类的广义线性分类器, 是建立在统计学习理论 VC (Vapnik-Cher-vonenkis) 维理论和结构风险最小原理的基础上的^[7]。通过寻求最小结构化风险来提高学习机泛化能力, 实现经验风险和置信范围的最小化, 从而达到在统计样本较少的情况下, 亦能获得良好统计规律的目的, SVR 在解决小样本、非线性及高维模式识别中表现出许多独有的优势^[8]。高速列车本身的特殊构造决定列车内部噪声特性的复杂性, 在对其做评价研究过程中, 精确的选取噪声数据样本对分析心理声学参量与评价结果之间的关系尤为重要, 本文用最小二乘法对截取的车厢噪声样本进行精确筛选后, 再通过 SVR 预测分析, 建立高速列车车内声品质预测的最小二乘法——支持向量机回归 LS-SVR 评价模型。

将 SVM 由分类问题推广至回归问题可以得到支持向量回归 (SVR)^[9]。

由最小二乘法反向筛选出的样本计算的客观参数和样本的得分均值组成训练样本 $D = \{x_i, y_i\}, i = 1, 2, 3, \dots, n, x_i \in R^d$, 且 $y \in R$, 对于传统回归模型

通常直接基于模型输出 $f(x)$ 与真实输出 y 之间的差别来计算损失,当且仅当 $f(x)$ 与 y 完全相同时,损失才为 0。而 SVR 假设 $f(x)$ 与 y 之间最多有 ε 的偏差,即仅当两者之间的差别绝对值大于 ε 时才计算损失,相当于以 $f(x)$ 为中心,构建了一个宽度为 2ε 的间隔带,于是 SVR 问题可形式化,如式(1)所示:

$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n l_\varepsilon(f(x_i) - y_i) \quad (1)$$

其中, $\omega = (\omega_1, \omega_2, \omega_3, \dots, \omega_d)$ 为法向量; C 为正则化常数; l_ε 为不敏感损失函数,如式(2)所示:

$$l_\varepsilon(z) = \begin{cases} 0 & |z| \leq \varepsilon \\ |z| - \varepsilon & |z| > \varepsilon \end{cases} \quad (2)$$

其中, z 为自变量, $l_\varepsilon(z)$ 为因变量。

求解式(1)的优化问题,通过引入拉格朗日乘子 $u_i \geq 0, \hat{\mu}_i \geq 0, \alpha_i \geq 0, \hat{\alpha}_i \geq 0$, 把约束优化问题转化为无约束优化问题,并引入松弛变量 ξ_i 和 $\hat{\xi}_i$, 由拉格朗日乘子法可建立拉格朗日等式,如式(3)所示:

$$L(\omega, b, \alpha, \hat{\alpha}, \xi, \hat{\xi}, \mu, \hat{\mu}) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \hat{\xi}_i) - \sum_{i=1}^n \mu_i \xi_i - \sum_{i=1}^n \hat{\mu}_i \hat{\xi}_i + \sum_{i=1}^n \alpha_i (f(x_i) - y_i - \varepsilon - \xi_i) + \sum_{i=1}^n \hat{\alpha}_i (y_i - f(x_i) - \varepsilon - \hat{\xi}_i) \quad (3)$$

再由式(3)分别对 ω, b, ξ_i 和 $\hat{\xi}_i$ 求偏导为 0, 可得 SVR 的对偶问题,如式(4)所示:

$$\begin{aligned} & \max_{\alpha, \hat{\alpha}} \sum_{i=1}^n y_i (\hat{\alpha}_i - \alpha_i) - \varepsilon (\hat{\alpha}_i + \alpha_i) - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\hat{\alpha}_i - \alpha_i) (\hat{\alpha}_j - \alpha_j) x_i^T x_j \\ & \text{s.t.} \sum_{i=1}^n (\hat{\alpha}_i - \alpha_i) = 0, 0 \leq \alpha_i, \hat{\alpha}_i \leq C \end{aligned} \quad (4)$$

对于上述不等式约束优化问题,需要满足前提 KKT 条件, KKT 条件是解决最优化问题时用到的约束方法,其一是对拉格朗日函数取极值时的必要条件;其二是对拉格朗日系数的约束优化; KKT 条件公式表达,如式(5)所示:

$$\begin{cases} \dagger \alpha_i (f(x_i) - y_i - \varepsilon - \xi_i) = 0 \\ \dagger \hat{\alpha}_i (y_i - f(x_i) - \varepsilon - \hat{\xi}_i) = 0 \\ \dagger \alpha_i \hat{\alpha}_i = 0, \quad \xi_i \hat{\xi}_i = 0 \\ \dagger (C - \alpha_i) \xi_i = 0, \quad (C - \hat{\alpha}_i) = 0 \end{cases} \quad (5)$$

在满足 KKT 条件下,将式(2)对 ω 的偏导为 0

带入到式(6)可得 SVR 解,如式(7)所示:

$$f(x) = \omega^T x + b \quad (6)$$

$$f(x) = \sum_{i=1}^n (\hat{\alpha}_i - \alpha_i) x_i^T x + b \quad (7)$$

由于高铁噪声中复杂的声学特性,输入的心理声学参数和输出的主观评价结果之间存在复杂的非线性关系,所以引入核函数将原始空间映射到一个更高维的非线性特征空间进行分析,常用的核函数有线性核、多项式核、高斯核等,由于本文中的样本数据点具有多维度特征,而高斯核函数具有优越的可分性和局域性,因此本文选用更稳定且泛化能力更强的高斯核函数,如式(8)所示:

$$\kappa(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right), \quad \sigma > 0 \quad (8)$$

则最后得到 SVR 的非线性高维解,如下所示:

$$f(x) = \sum_{i=1}^n (\hat{\alpha}_i - \alpha_i) \kappa(x, x_i) + b \quad (9)$$

以上过程在代码实现前,首先要设置好参数,其中,种群数量 v 设置为 15;交叉概率 P 设置为 0.8;迭代次数 eps 设置为 100;初始的 c, g, p 由交叉验证后直接得到默认值。对原始样本进行多次训练预测,最终的训练集预测结果如图 5 所示。

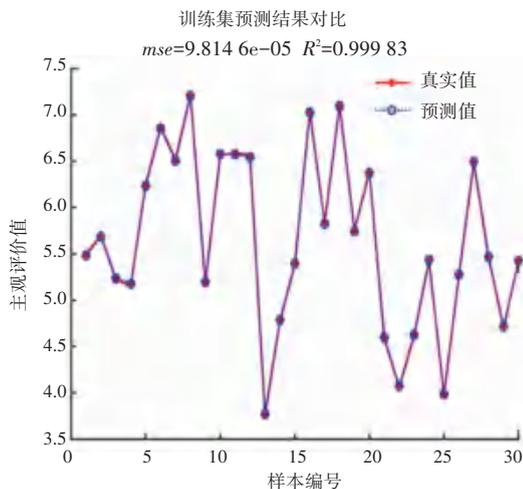


图 5 训练集样本预测结果

Fig. 5 Training sample prediction results

4 LS-SVR 预测模型的检验

为了进一步验证 LS-SVM 模型预测结果的准确性,再随机选取 12 个噪声样本作为测试集,并与传统多元回归模型进行比较。12 个样本都来自同一数据源,测试集中的数据点与训练集样本中数据点的获取方法相同。测试集的预测值与实际值之间的预测结果对比如图 6 所示。图 6 中测试集预测的