

基于深度卷积神经网络的铁路接触网 鸟窝检测方法研究

贺德强¹, 江 洲¹, 陈基永¹, 杨严杰¹, 姚晓阳²

(1. 广西大学 机械工程学院, 广西 南宁 530004;

2. 中车株洲电力机车研究所有限公司, 湖南 株洲 412001)

摘 要: 鸟类在铁路接触网筑巢一直是造成接触网故障的一个重要原因, 目前主要依靠人工巡检的方式确定是否存在鸟窝, 不仅工作量大、漏检率高, 而且效率低。因此提升接触网鸟窝的检测效率, 及时排除隐患, 对保障铁路安全运营具有重要的意义。针对此问题, 提出了一种基于深度卷积神经网络的 Faster R-CNN 模型用于接触网鸟窝的自动识别。通过自定义合适的网络结构和参数, 经过预训练、RPN 网络训练、Fast R-CNN 网络训练以及对 RPN 和 Fast R-CNN 的联合训练, 建立了适合鸟窝检测的 Faster R-CNN 模型, 实现对鸟窝的检测。经试验, Faster R-CNN 的准确率为 88.5%, 每张图片的识别速度为 79 ms, 通过传统的 HOG 方法、DPM 方法和卷积神经网络方法进行比较, 验证了深度卷积神经网络对铁路接触网鸟窝检测高效性。

关键词: 深度学习; Faster R-CNN; 接触网; 鸟窝检测; 卷积神经网络

中图分类号: U225; TP391.9

文献标识码: A

doi: 10.13890/j.issn.1000-128x.2019.04.027

Research on Detection of Bird Nests in Overhead Catenary Based on Deep Convolutional Neural Network

HE Deqiang¹, JIANG Zhou¹, CHEN Jiyong¹, YANG Yanjie¹, YAO Xiaoyang²

(1. College of Mechanical Engineering, Guangxi University, Nanning, Guangxi 530004, China;

2. CRRC Zhuzhou Institute Co., Ltd., Zhuzhou, Hunan 412001, China)

Abstract: Nesting on railway catenaries by birds has been an important cause of catenary failure. At present, the inspection for judging the existence of bird's nests is carried out artificially, which has much shortcomings in heavy working intensity, high missing-inspection rate and low detecting efficiency. So improving detecting efficiency and then removing potential hazards in time is significant to ensure the security operation of railway vehicles. Aiming at such shortcomings, a Faster R-CNN model based on deep convolution neural network was proposed to identify bird's nests on catenaries automatically. After the steps of customizing appropriate network structure and parameters and going on pre-training, RPN network training, Fast R-CNN network training and the joint training of RPN network and Fast R-CNN network, a Faster R-CNN model which could be implemented bird's nests detection was established. Experiment has proved that the recognizing accuracy of Faster R-CNN model is 88.5% and the recognizing time of a picture is 79 ms. Compared with the traditional HOG, DPM and convolutional neural network method, it was verify that the Faster R-CNN model based on deep convolutional neural network was efficient on detecting the existence of bird's nests on railway catenaries.

Keywords: deep learning; Faster R-CNN; catenary; nest detection; convolution neural network

收稿日期: 2018-09-05

基金项目: 国家自然科学基金项目(51765006); 广西科技攻关项目(桂科攻 1598009-6); 广西自然科学基金重点项目(2017GXNSFDA198012)

0 引言

接触网是沿铁路线上空架设, 为列车提供电能的特殊供电线路, 是与铁路安全运行直接相关的设施^[1]。每年3月至5月是鸟类筑巢的高峰时期, 鸟类经常在接触网隔离开关底座、硬横梁、钢柱等处筑巢, 筑巢用的树枝、铁线等物体容易短接绝缘子导致接触网设备短路跳闸, 影响列车的安全运行甚至引起机破, 这种情况在阴雨天尤为突出。目前, 铁路沿线检测采用专线接触网设备进行在线巡检, 由于存在隐患的接触网位置是分散的, 此方式需要记录海量的视频数据, 面对众多的视频数据, 若仅仅依靠人工判读的方法进行甄别, 不仅工作量极大、效率低, 而且可靠性和准确率也很难保证。因此, 对巡检图像进行准确、高效的智能分析, 减少工作人员的工作强度, 提升工作效率, 对保障铁路安全运营具有重要的意义。目前深度学习是图像识别领域的一个热门研究方向, 本文提出一种基于深度卷积神经网络的Faster R-CNN目标检测模型, 实现对铁路接触网鸟窝的自动识别。该模型将一般目标检测的特征提取、目标识别和检测过程统一到一个深度神经网络框架中, 可以在保证检测速度的同时不降低检测精度。

1 Faster R-CNN 检测模型

文献[2]提出Fast R-CNN的改进模型Faster R-CNN, 该模型引入一个区域生成网络RPN, 用来对目标物产生候选区域。Faster R-CNN可以看作是由产生目标候选区的RPN网络和利用这些候选区域进行预测分类的Fast R-CNN检测器组成的, 如图1所示。首先输入一张任意尺寸的图片, 向前传播到最后一个共享卷积层, 一方面将得到的特征映射传送到RPN网络, 另一方面继续向前传播产生更高维的特征映射。在RPN网络中, 经过一系列的处理, 产生多个候选区域和候选区域得分, 对候选区域框采用非极大值过滤^[3], 以减少候选区域数量。将大于设定阈值的候选区域和之前产生的高维特征映射一同输入到RoI池化层提取对应区域的特征, 最后将该区域特征和全连接层连接输出目标分类和得分, 以及bounding box目标定位框回归值。

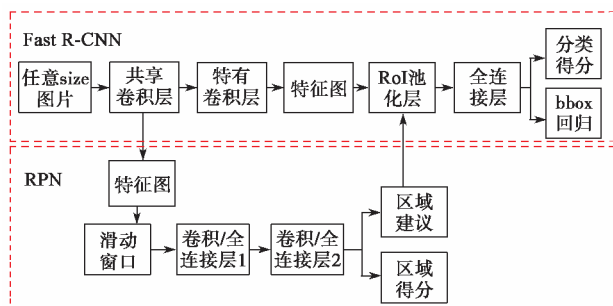


图1 Faster R-CNN 网络处理流程

Faster R-CNN已应用在行人检测、移动电话识别等众多领域。

2 基于Faster R-CNN模型的鸟窝识别检测

基于Faster R-CNN模型的鸟窝识别检测模型共分为4个阶段：第1阶段需要对整个网络进行初始化预训练；第2阶段训练RPN网络，用于对每张图片产生候选区域；第3阶段训练Fast R-CNN网络，用于从候选区域提取特征；第4阶段对已训练好的RPN网络和Fast R-CNN网络进行联合调优训练。Faster R-CNN训练过程如图2所示。

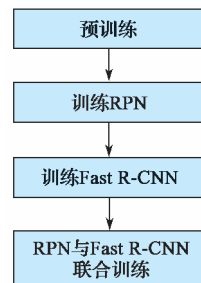


图2 训练流程

2.1 预训练

训练一个深度卷积神经网络, 其困难在于缺乏足够的数据对其进行充分训练, 容易造成过拟合问题。因此, 需要对网络进行预训练。通常用于预训练的网络模型有ZF-Net模型^[4]和VGG模型^[5]。预训练的数据通常采用ImageNet数据集^[6]。ImageNet是目前国际上图像识别领域最大的数据集, 在深度学习和图像领域有着广泛的应用。将已用ImageNet数据集训练好的通用模型对网络进行初始化即可完成预训练。采用这种方式不仅可以避免从零开始训练大型卷积神经网络容易导致过拟合的问题, 而且可以大幅减少训练时间。

2.2 RPN训练

RPN是一个全卷积网络, 该网络将一张任意大小的图片作为输入, 输出带有目标得分的候选区域集合。RPN设置一个尺寸为 3×3 的空间窗口, 在共享卷积层输出的卷积特征映射上滑动, 每滑动一次窗口都将窗口内的高维特征向量映射成低维向量, 接着进行 1×1 卷积运算并有2个输出支路: 支路1输出是目标或者非目标的概率, 支路2输出box定位框中心坐标 x, y 与长宽值 w, h , 如图3所示。

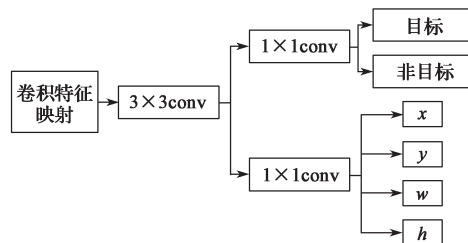


图3 RPN处理过程

整个RPN网络采用反向传播法^[7]和随机梯度下降法^[8]进行端到端的训练, 损失函数为分类误差和回归误差之和^[2]:

$$L(\{p_i\} + \{t_i\}) = \frac{1}{N_{\text{cls}}} \sum_i L_{\text{cls}}(p_i, p_i^*) + \lambda \frac{1}{N_{\text{reg}}} \sum_i p_i^* L_{\text{reg}}(t_i, t_i^*) \quad (1)$$

$$L_{cls}(p_i, p_i^*) = -\lg[p_i^* p_i + (1 - p_i^*)(1 - p_i)] \quad (2)$$

$$L_{cls}(t_i, t_i^*) = S_{L_i}(t_i - t_i^*) \quad (3)$$

$$S_{L_i}(x) = \begin{cases} 0.5x^2 & |x| < 1 \\ |x| - 0.5 & |x| \geq 1 \end{cases} \quad (4)$$

其中, i 指的是第 i 个 anchor 的检索下标; P_i 为第 i 个 anchor, 是目标物的预测概率。anchor 是 RPN 的核心部分, 通过提供一个基准边框, 然后按比例给出各种长宽不一的边框进行滑框映射。当 anchor 是正样本时, 真实值标签 p_i^* 取值为 1, 否则取值为 0。 t_i 表示预测边框坐标向量, 而 t_i^* 是与正 anchor 所对应的真实值边框的坐标向量。 L_{cls} 为分类误差损失函数, L_{reg} 为回归误差损失函数, N_{cls} 与 N_{reg} 是归一化参数, 由 λ 权衡, 这里设置 $\lambda=10$ 。

2.3 Fast R-CNN 网络训练

Fast R-CNN 网络首先用一个 5 层的卷积网络对图像进行卷积运算得到一个特征映射, 然后与候选区域组成特征图输入到 RoI 池化层进行深层特征提取, 由此生成规格化的特征向量。将规格化的向量输入到全连接层最终产生 2 个支路: 支路 1 利用 Softmax 预测类别概率, 支路 2 输出检测框坐标值。Fast R-CNN 同样采用反向传播法和随机梯度下降法进行训练, 通过分类误差和回归误差构成的损失函数进行联合训练^[9]:

$$L(p, u, t^u, v) = L_{cls}(p, u) + \lambda[u \geq 1] \times L_{loc}(t^u, v) \quad (5)$$

$$L_{cls}(p, u) = -\log(p_u) \quad (6)$$

$$L_{loc}(t^u, v) = \sum_{i \in \{x, t, w, h\}} S_{L_i}(t^u - v_i) \quad (7)$$

$$S_{L_i}(x) = \begin{cases} 0.5x^2 & |x| < 1 \\ |x| - 0.5 & |x| \geq 1 \end{cases} \quad (8)$$

式中: p 为类别概率值; u 为真实类别; t 为回归框参数; L_{cls} 为真实类别 u 的对数损失; L_{loc} 为回归损失。

与训练 RPN 网络一样, 在训练过程中, 由反向传播算法和随机梯度算法不断更新网络的权重和偏置。

2.4 Fast R-CNN 与 RPN 联合调优训练

前部分 Fast R-CNN 与 RPN 网络的训练都是单独训练, 最后需要对二者进行联合训练, 精调各参数。在此过程中 Fast R-CNN 和 RPN 共享一个卷积网络的特征参数。Faster R-CNN 有 3 种训练共享卷积网络特征参数的方法: 交替训练、迭代训练和非近似联合训练。本文采用交替训练的方式, 训练步骤为:

①对 RPN 网络进行预训练, 在初始化的模型上进行端到端微调;

②将步骤①产生的候选区域用于 Fast R-CNN 检测网络的端到端训练, 此时 2 个网络并无卷积共享;

③将步骤②训练的 Fast R-CNN 检测网络重新对 RPN 网络进行初始化, 此时将学习率设置为 0 固定共享层, 仅微调 RPN 各层参数;

④固定步骤③的共享卷积层, 将步骤③得到的候

选区域用于对 Fast R-CNN 独有的层参数进行微调。

3 试验及结果分析

3.1 样本来源与数据处理

试验数据来源于 2 方面: 铁路巡检所拍摄图像; 部分鸟窝图片来源于网络。试验数据作为训练样本使用, 训练样本数总共 1 500 张, 测试样本总共 150 张, 为了防止过拟合, 增加整个网络的鲁棒性, 随机翻转每张图片, 将训练样本数量扩充至 3 000 张。将全体数据做成 PASCAL VOC 2007 格式, 选取训练样本中出现的完整且无遮挡、变形小、光线好的鸟窝图片予以标记, 如图 4 所示。在测试时, 设置当 IoU 阈值不小于 0.7 时, 则视为成功识别。



图 4 训练样本标记

3.2 试验结果分析

试验在基于深度学习框架 caffe 上完成, 动量设置为 0.9, 权值衰减设置为 0.000 5, batch size 设置为 1, 依次以 0.001 和 0.000 1 的学习速率分别进行 50 000 次迭代训练, 然后以 0.000 01 学习速率训练 20 000 次。部分测试结果如图 5 所示。被识别的鸟窝被标记在方框内, 方框上方标签显示是鸟窝的概率。

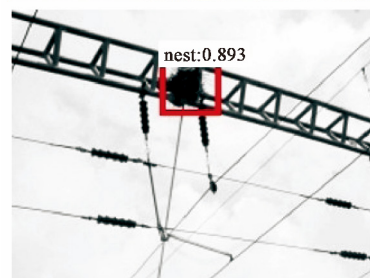
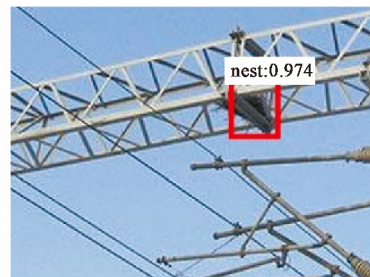
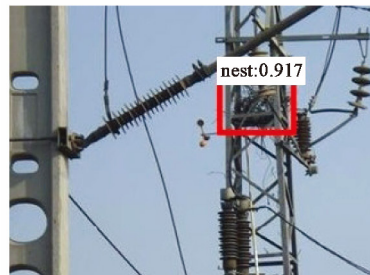


图 5 试验结果

RPN 网络中的 anchor, 提供一个基准边框, 然后按比例衍生出长宽比不一的边框进行滑框映射, 主要解决目标物多尺度的问题。默认基准边框有 128, 256, 512 共 3 个基本尺寸, 1:1, 2:1, 1:2 共 3 个长宽比, 考虑到鸟窝大小占整幅图片比例小, 本文通过将 512 大小的基准边框修改为 64, 增加一组基准边框设置进行对比。在训练过程中, 选择合适的 mini-batch size (即

最小批次和候选框数量)可以最大化整个网络的效率。不同参数的结果见表1。

表1 不同参数结果比较

mini-batch size	基准边框尺寸	候选框数量	准确率/%
32	64,128,256	300	80.7
64	64,128,256	300	82.8
128	64,128,256	300	83.3
256	128,256,512	300	86.1
256	64,128,256	300	88.5
256	64,128,256	100	87.5

从结果可以看出,随着 mini-batch size 的提高,准确率逐步升高。基准边框有 64,128,256,512 共 4 个比例,将 512 大小的基准边框修改为 64,准确率提高了 2.4%,这表明尺寸为 64 的基准边框大小更适合对鸟窝的检测。在其他条件相同的情况下,候选框数量为 300 的准确率与候选框数量为 100 的网络只相差 1.0% 左右。从整体看 mini-batch size 设置为 256,基准边框为 64,128,256,候选框数量为 300 是最合适的选择。

3.2.2 全卷积网络模型比较

RPN 产生候选区域所使用的全卷积网络模型有 ZF-Net 模型和 VGG 模型,将 2 个模型分别进行测试,加入对应模型与 Fast R-CNN 检测网络是否共享卷积特征的组合,以及与非全卷积网络模型 Selective Search 方法^[10]组合作为对比项,除 Selective Search 候选框数量为 2 000 外,其他模型均为 300,测试结果见表 2。

表2 不同模型比较

模型	候选框数量	准确率/%
SS	2 000	83.5
RPN+ZF unshared	300	86.4
RPN+ZF	300	88.5
RPN+VGG unshared	300	89.2
RPN+VGG	300	90.0

由表 2 可知,采用 2 000 个候选框数量的 Selective Search 模型准确率为 83.5%,准确率最低,unshared 表示卷积不共享,即 RPN 网络与 ZF 或者 VGG 网络单独训练。不共享卷积的 ZF 与 VGG 模型准确率较 SS 网络都有提升,VGG 模型达到 89.2%,甚至超过了与 RPN 网络共享卷积的 ZF 模型,共享卷积的 VGG 模型取得了 90.0% 的准确率,单从准确率上看 VGG 模型准确率最高。

模型的检测速度是衡量实时性的一个重要指标,将共享卷积的 VGG 与 ZF 模型、非全卷积网络模型 Selective Search 三者检测速度进行比较,结果见表 3。

表3 不同模型检测速度比较

模型	候选框数量	平均耗时/ms
Faster R-CNN(SS)	2 000	2 633
Faster R-CNN(ZF)	300	79
Faster R-CNN(VGG)	300	275

从结果看,在整个 Faster R-CNN 框架中,选择不同的模型,平均每张图片的测试速率相差很大,ZF 模型仅为 79 ms,VGG 模型为 275 ms,耗时约为 ZF 模型的 3.5 倍,SS 模型最慢,平均耗时为 2 633 ms,是 ZF 模型的约 33.3 倍。导致 SS 模型速度过慢的一个主要原因是 SS 模型无法利用 GPU 运算,只能通过 CPU 运算来产生候选区域,且不同于卷积神经网络直接对图像做卷积映射,SS 在生成候选区域时,为了定位到目标的具体位置,会把图像分成许多子块然后把子块作为输入,穷举所有的图块进行匹配检测。同属卷积神经网络的 ZF 模型比 VGG 模型快,主要是因 VGG 模型相较 ZF 模型网络规模更大,更为复杂,因此虽然 VGG 模型准确率比 ZF 模型有所提高,但耗时远比 ZF 模型长。

通过几个控制试验可知,选择 ZF 模型,mini-batch size 设置为 256,基准边框为 64,128,256,候选框数量为 300 可以使整个接触网鸟窝检测模型达到最优效果。

3.2.3 与其他方法比较

为了验证 Faster R-CNN 的有效性,将其与文献[11]中的 HOG 方法、文献[12]中的 DPM 方法及改进的基于 CNN 方法进行比较,采用每张图片的测试速度、准确率和召回率进行评价,其定义如下:

$$\text{准确率} = \frac{\text{正确检测出含有鸟窝图片的数量}}{\text{所有检测出含有鸟窝图片的总数}} \quad (9)$$

$$\text{准确率} = \frac{\text{正确检测出含有鸟窝图片的数量}}{\text{所有含有鸟窝图片的数量}} \quad (10)$$

其结果统计见表 4。

表4 不同方法比较

模型	准确率/%	召回率/%	平均耗时/ms
HOG	24.6	18.7	5 771
DPM	31.2	28.7	6 523
CNN	73.3	66.0	2 119
Faster R-CNN	88.5	82.0	79

从表 4 中可以看出,HOG 与 DPM 模型准确率、召回率远不如 Faster R-CNN 模型,这是因为 HOG 和 DPM 这类传统识别模型采用的是人工提取特征作为检测模板再进行滑框匹配检测的方式,这类方法很容易受到形状、纹理特征的影响,而鸟窝的形态差异较大又处于复杂的接触网环境中,很难提取出标准的检测特征,因此准确率和召回率不高。基于卷积神经网络的方法使用卷积核作为特征提取器,图像可以直接作为网络的输入,通过训练得到鸟窝特征,避免了传统识别算法中复杂的特征提取和数据重建过程,在准确率和召回率上有明显提升。Faster R-CNN 模型和 CNN 模型相比,Faster R-CNN 模型采用 RPN 网络生成候选区域的策略可以比 CNN 提取出精度更高的特征,因此,

准确率和召回率较高。从时间上看,传统识别模型平均耗时较多,这类模型在检测时,将特征模板在检测图像上滑框匹配,需要穷举所有的图块进行匹配检测,耗时过长。与传统模型相比卷积网络模型检测速度更快,CNN模型平均耗时只有2 119 ms,但CNN模型各层神经元之间都是全连接且各个识别过程是分开的,计算量大且存在重复计算,耗时依然较长,而Faster R-CNN是一个整体的模型,所有处理过程都在一个深度网络框架中完成,因此大幅减少了计算时间。综合比较,Faster R-CNN表现最好,在实时性和准确性方面都能得到令人满意的结果,可以满足实时性巡检的要求。

4 结 语

本文将基于深度卷积神经网络的Faster R-CNN模型应用于接触网鸟窝检测。通过一系列控制试验确定合适的网络结构和参数,经过预训练、RPN网络训练、Fast R-CNN网络训练以及对RPN和Fast R-CNN的联合训练,建立了适合鸟窝检测的Faster R-CNN模型。该模型将目标检测的特征提取、目标识别和检测过程融入到一个深度神经网络框架中,可以在保证检测速度的同时不降低检测精度。试验结果表明检测平均处理时间为79 ms,正确率为88.5%,召回率为82.0%,经过与已有的方法进行对比,验证了该方法的准确性和快速性。该模型具有很好的检测效果,能够快速准确地对铁路接触网鸟窝进行识别,有利于提升巡检工作效率,保障铁路安全运营。

参考文献:

- [1] 曾怡达,杨金辉,谭金练,等.基于特征匹配的接触网定位支座检测[J].机车电传动,2018(3):100-103.
- [2] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [3] NEUBECK A, VAN GOOL L. Efficient non-maximum suppression[C]//IEEE. 18th International Conference on Pattern Recognition: Volume 3. Hong Kong: IEEE, 2006: 850-855.
- [4] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks[C]//European conference on computer vision. Computer Vision – ECCV 2014. Berlin: Springer, 2014: 818-833.
- [5] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J/OL]. Computer Science. arXiv: 1409.1556, 2014 [2018-09-01]. <https://arxiv.org/abs/1409.1556>.
- [6] RUSSAKOVSKY O, DENG J, SU H, et al. Imagenet large scale visual recognition challenge[J]. International Journal of Computer Vision, 2015, 15(3): 211-252.
- [7] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors[J]. Nature, 1986(323): 533-536.
- [8] BOTTOU L. Stochastic gradient descent tricks[C]//MONTAVON G, ORR G B, MÜLLER R KR. Neural networks: Tricks of the trade. Berlin: Springer, 2012: 421-436.
- [9] GIRSHICK R. Fast R-CNN[J/OL]. Computer Science, arXiv:1504.08083 [2018-09-01]. <https://arxiv.org/abs/1504.08083>. DOI: 10.1109/ICCV.2015.169.
- [10] UIJLINGS J R R, SANDE K E A, GEVERS T, et al. Selective search for object recognition[J]. International Journal of computer vision, 2013, 104(2): 154-171.
- [11] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//IEEE. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition: Volume 1. San Diego: IEEE, 2005: 886-893.
- [12] FELZENSZWALB P, MALLESTER D, RAMANAN D. A discriminatively trained, multiscale, deformable part model[C/OL]//IEEE. 2008 IEEE Conference on Computer Vision and Pattern Recognition. Alaska: IEEE, 2008 [2018-09-01]. <https://ieeexplore.ieee.org/document/4587597>. DOI: 10.1109/CVPR.2008.4587597.
- [13] 曾怡达,杨金辉,谭金练,等.基于特征匹配的接触网定位支座检测[J].机车电传动,2018(3):100-103.
- [14] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

作者简介:贺德强(1973—),男,教授,博士生导师,主要从事机车车辆、故障诊断与智能维护、网络化制造等技术的研究工作。

(上接第76页)

- [10] ÖHRTRÖM E, SKÅNBERGA.-B. A field survey on effects of exposure to noise and vibration from railway traffic, Part 1: Annoyance and activity disturbance effects[J]. Journal of Sound and Vibration, 1996, 193(1): 39-47.
- [11] 夏禾. 交通环境振动工程[M]. 北京: 科学出版社, 2010.
- [12] 常乐,任琨,闫维明,等. 城市公路与高架路交通诱发建筑振动实测与分析[J]. 北京工业大学学报, 2008, 34(10): 1053-1058.
- [13] Mechanical vibration and shock-Evaluation of human exposure to

whole-body vibration-Part 1: General requirements: ISO 2631-1—1997[S].

- [14] 日本工業規格. 全身振動—第2部: 測定方法及び評価に関する基本的要求: JIS B 7760-2: 2004[S]. 2004.

作者简介:宋波(1962—),男,博士,教授,博导,主要研究方向为城市基础设施(桥梁、港湾、电力设施等)的抗震及减灾。