

文章编号: 2095-2163(2021)05-0065-05

中图分类号: TN274

文献标志码: A

基于端点检测的光纤振动信号识别

徐方辰¹, 邢炜宁²

(1 国家石油天然气管网集团有限公司西气东输分公司, 上海 200126;

2 中国石油天然气管道通信电力工程有限公司, 河北 廊坊 065000)

摘要: 利用光纤分布式传感系统对入侵事件进行识别主要难点在于对入侵事件的识别准确率低, 为了提高对入侵事件的识别准确率, 本文提出一种基于端点检测与信号重组的光纤振动信号的识别方法。该方法首先使用基于谱质心与短时能量的端点检测算法对振动信号的振动部分进行检测, 然后将检测到的振动信号进行振动信号的重组, 最后使用一个多尺度卷积神经网络结合随机森林树对重组后的信号进行识别。实验证明该识别方式能快速完成对识别模型的训练, 并且能有效识别在实际环境中采集的入侵振动信号, 对入侵信号的识别准确率可达 97.4%。

关键词: 分布式光纤传感; 相敏光时域反射计(φ -OTDR); 端点检测; 多尺度卷积神经网络; 随机森林树

Fiber optic vibration signal identification based on endpoint detection

XU Fangchen¹, XING Weining²

(1 National Oil and natural gas pipeline Network Group Co., Ltd. West-East Gas Transmission Branch, Shanghai 200126, China;

2 China Oil and natural gas pipeline Communication and Power Engineering Limited Corporation, Langfang Hebei 065000, China)

[Abstract] The main difficulty in identifying intrusion events using fiber optic distributed sensing system is the low recognition accuracy of intrusion events. In order to improve the recognition accuracy of intrusion events, this paper proposes a fiber optic vibration signal recognition method based on endpoint detection and signal reorganization. The endpoint detection algorithm based on spectral center of mass and short-time energy is used to detect the vibration part of the vibration signal, then the detected vibration signal is reorganized, finally a multi-scale convolutional neural network combined with random forest tree is used to identify the reorganized signal. The experiment proves that the recognition method can quickly train the recognition model and effectively recognize the intrusion vibration signals collected in the actual environment, and the recognition accuracy of the intrusion signals can reach 97.4%.

[Key words] distributed fiber optic sensing; phase-sensitive optical time-domain reflectometer (φ -OTDR); endpoint detection; multiscale convolutional neural network; random forest tree

0 引言

目前, DAS 已被广泛用于管道泄露监测、围栏入侵、周界安防等诸多领域^[1-4]。DAS 具有不同的实现原理, 其中使用 Φ -OTDR 原理构建的 DAS 系统具有检测信号的信噪比高、检测范围广、成本相对较低、可实现远程入侵检测等优点, 具有很强的发展潜力。目前使用 DAS 系统对入侵事件的识别过程可分为 4 个步骤, 即: 先对入侵振动信号进行预处理, 然后进行特征提取, 再进行数据集的制作, 最后搭建识别模型, 对数据进行分类。在此过程中, 对数据的预处理主要包含对数据去噪, 进行端点检测^[5]等。端点检测的目的在于减小一段信号中存在的噪

音, 将只包含振动的振动片段提取出来, 减小噪音对模型的影响, 双门限端点检测就是时下常用的 VAD 算法。而在研究中常常见到的特征提取方式则有提取振动信号的傅里叶变换特征^[6]、梅尔倒频系数^[7]等, 或者使用 EMD^[8]、VMD^[9]将信号分解后提取相关的统计学特征。典型的分类算法主要有: 支持向量机(SVM)^[10]、随机森林树^[11]等。上述分类算法虽然已经在 DAS 振动信号的分类识别中取得了很好的效果, 但是由于需要进行人为的特征提取, 加大了 DAS 模式识别的实现难度。随着深度学习的发展以及大范围的应用, 深度学习也已开始更多地应用在 DAS 系统的模式识别上, 例如使用 BP 神经网络^[12]、径向基神经网络^[13]等神经网络算法对振动

基金项目: 基于数据融合的站场安全风险智能监控及分析平台研究项目; 国家石油天然气管网集团有限公司西气东输分公司项目 (KJ201901)。

作者简介: 徐方辰(1985-), 男, 学士, 工程师, 主要研究方向: 通信工程、安防; 邢炜宁(1985-), 男, 学士, 助理工程师, 主要研究方向: 通信工程、大数据。

收稿日期: 2021-03-30

信号进行识别,在文献[14]中即提出使用一维卷积神经网络对油气管道入侵事件进行识别,该次研究直接将经过预处理的振动信号输入一维卷积神经网络,达到了与二维卷积神经网络相近的准确率,降低了识别时间。

基于此,本文提出使用基于谱质心与短时能量的端点检测算法提取出振动信号中包含的振动片段,并对振动片段进行信号的重组,形成新的振动信号后,再使用一个多尺度的卷积神经网络结合随机森林树的模型对振动信号进行识别。

1 识别方式原理

1.1 基于 φ -OTDR 分布式光纤声传感原理

基于 Φ -OTDR 的 DAS 系统原理如图 1 所示。其中,激光源为超窄线宽激光。 φ -OTDR 的工作原理为:超窄线宽激光源发出连续光,该连续光经过声光调制器(AOM)后被调制为脉冲光,由于脉冲光的强度偏弱,因此再让该脉冲光通过掺铒光纤放大器(EDFA)进行放大,被放大后的脉冲光经过带通滤波器滤波(BPF)后由光循环器(OC)的 1 端口进入,从 OC 的 2 端口射入传感光纤内。在传感光纤中,产生的后向瑞利散射光反向传播到 OC 的 2 端口,由 OC 的 3 端口射出。由于后向瑞利散射光强度较低,因此使用 EDFA 进行放大,放大后使用带通滤波器滤除自发辐射噪声(AES),稍后使用光电探测器将脉冲光转换为电信号,再使用 DAQ 去采集电信号。由于不同类型的振动信号所包含的频率、持续时间等不尽相同,所以通过 DAS 采集的信号波形也不同,因此可以对信号进行直接识别。

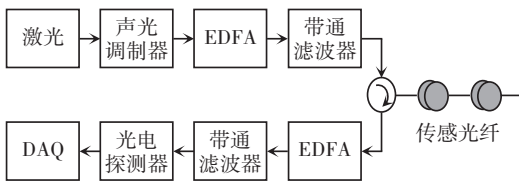


图 1 φ -OTDR 结构

Fig. 1 The construction of φ -OTDR

1.2 基于谱质心与端点检测算法原理

基于短时能量与谱质心的端点检测算法可以有效地提取出一段振动信号中的振动片段。短时能量能反映出信号时域的特征,谱质心能反映出信号频域的特征,结合两种特征可以很好地提取出信号中的振动部分。基于短时能量与谱质心的端点检测算法的原理为:首先对信号进行分帧处理,并且对每帧信号进行加窗,这样可以减小信号在分帧后的能量

泄露,然后计算出每帧信号的短时能量,短时能量的计算公式为:

$$E = \sum_{i=0}^{m-1} s^2(i). \quad (1)$$

其中, s 为帧信号, i 为第 i 帧。再计算出每帧信号的谱质心,谱质心计算公式为:

$$C = \frac{\sum_{i=0}^n (i+1)f(i)}{\sum_{i=0}^n f(i)}. \quad (2)$$

其中, f 为帧信号经过傅里叶变换后的频率值。

设置一个短时能量的阈值与一个谱质心的阈值。此后进行判断。如果帧信号的短时能量与谱质心都小于设置的阈值,认为该段信号没有振动;如果大于设置的阈值,认为存在振动,保留此时的帧数,根据帧与帧移的关系得出该段帧在原信号中的位置。

1.3 多尺度卷积神经网络原理

通常单个的神经网络只能提取出信号中部分的信息,不能提取到信号不同尺度的信息,所以通过使用多个不同尺度的卷积神经网络来更加全面的提取振动信号的特征。多尺度卷积神经网络的结构如图 2 所示。

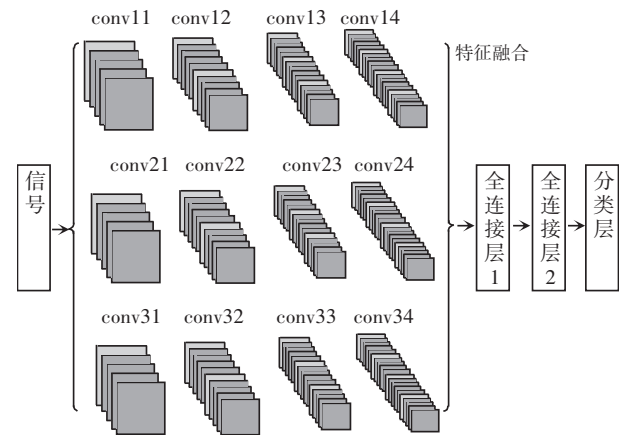


图 2 多尺度卷积神经网络

Fig. 2 Multi-scale convolutional neural networks

由图 2 可知,在多尺度卷积神经网络中使用了三层卷积神经网络,每个卷积层后都是使用了有一个最大池化层。通过对图 2 中的 3 个卷积神经网络使用不同的卷积核大小,从而使 3 个卷积神经网络能提取到信号不同尺度的特征。

1.4 随机森林树

构建随机森林树实际就是同时构建出多个决策树,所使用的特征是随机选择的,从不同的特征组合

中选择出最佳的特征组合。在随机森林树中设置的主要参数有:用于判断节点不纯度指标的计算方式、gini 系数与 entropy 系数;用于设置最大决策树棵树的参数,一般情况下该参数设置得越大,判断效果越好;用于决定树最大深度的参数,多数情况下设置为 5 层,越大则拟合效果越好,但也越耗时;以及一些剪枝参数,主要是用于降低模型的过拟合现象。使用随机森林树进行分类可以有效降低计算时间,因为相较于支持向量机来说,随机森林树中不存在复杂的计算。

2 实验验证与结果分析

研究中使用的 Φ -OTDR 的实验装置参如图 1

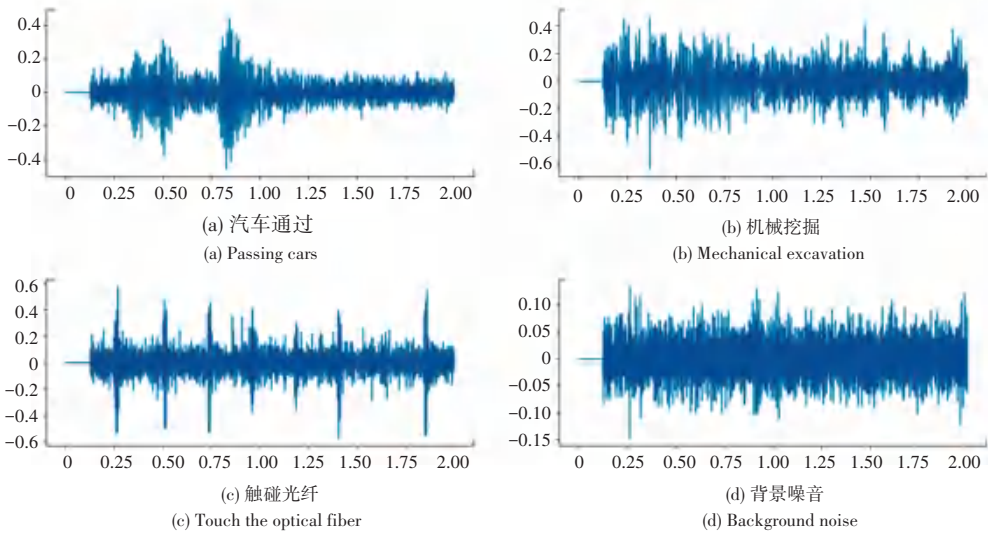


图 3 振动信号

Fig. 3 Vibration signal

考虑到是在实际的施工现场来采集振动信号,所以采集的振动信号中存在有大量的噪音。如果不去除信号中存在的噪音,就会对信号的识别造成严重的影响,因此需对信号进行小波阈值去噪处理。未去噪与去噪后的振动信号如图 4 所示。由图 4 (a)可以看出,原始未去噪的振动信号中存在有大量的噪音;由图 4(b)中可以看出,使用了小波阈值去噪后,明显减小了噪音,振动部分的信号变得清晰可见。

一段振动信号通常包括振动部分信号与噪音段信号。为了提取出振动信号中的振动部分,需要对振动信号使用 VAD 算法以提取出信号中的振动部分,从而排除非振动部分对识别模型的影响。使用基于短时能量与谱质心的 VAD 算法所提取的振动部分信号如图 5 所示。图 5(a)中的红色部分就是通过 VAD 算法提取出来的振动部分信号,通过对比图 5(b)中

所示。图 1 中,激光器为输出功率 10 mW、线宽为 3 KHz 的分布式反馈激光器 (DFB_LD),AOM 的带宽为 100 MHz,EDFA 的放大增益为 27 dB,PD 的带宽为 200 MHz,DAQ 的采样速率为 200 MHz/s、采样位数为 12 位。实验装置的传感光纤部分被埋设在中国上海市的文翔东路和沪松路的施工现场来采集振动信号。拟有针对性地采集如下振动信号:汽车通过路面或者井盖时引发的振动、施工机械工作时引发的振动、手触碰光纤时引发的振动以及背景噪声。采集的信号主要是单个事件工作时的振动信号,目前对混合型振动信号的识别依旧是一个难点。汽车通过、机械挖掘、触碰光纤、背景噪音的振动信号如图 3 所示。

的原始信号后可以看出,基于短时能量与谱质心的 VAD 算法很好地提取出了信号中的振动部分。

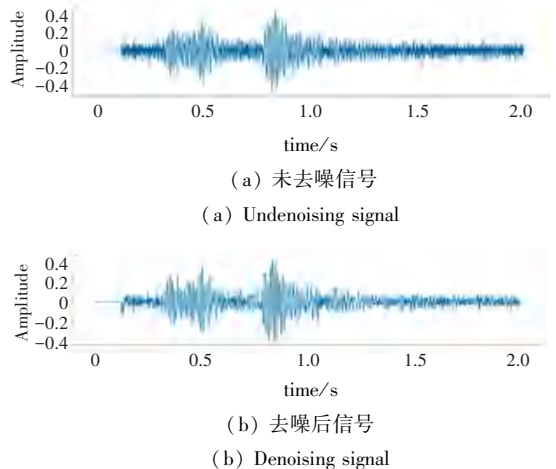
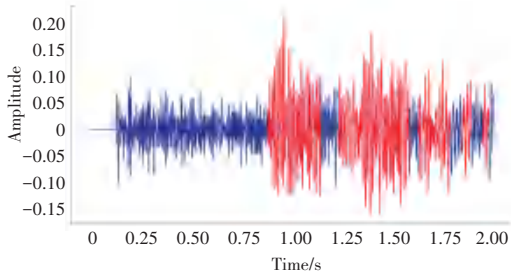


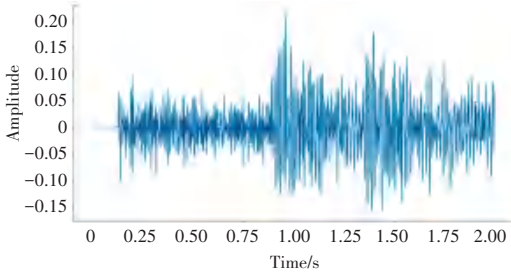
图 4 未去噪与去噪后的振动信号

Fig. 4 Undenoising and denoising signal



(a) 振动信号的端点检测

(a) End-point detection of the vibration signal



(b) 原始振动信号

(b) Original vibration signal

图 5 振动信号的检测

Fig. 5 Detection of the vibration signal

使用 VAD 算法提取出信号中的振动段后,就要对振动段进行重组,这是因为,每种振动的持续时间不同,导致提取出的振动段长度不同,为了使各种振动信号的振动段长度相同,需要对振动段进行信号的重组,所使用的方式为对所提取出的振动信号段进行重复,以此来重组出一个只包含振动片段的信号。重组信号如图 6 所示。

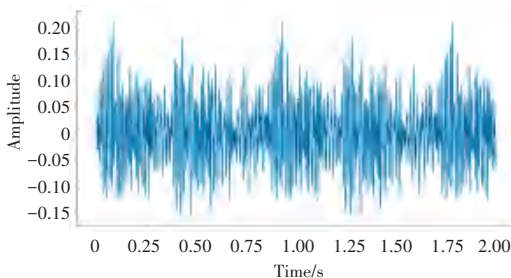


图 6 振动片段的重组

Fig. 6 Recombination of the vibration fragments

为验证本次研究提出的模型能准确识别出汽车通过、机械工作、触碰、噪音这 4 种振动事件,接着又进行了对比实验。研究中,为对比使用重组信号进行识别的效果,使用重组信号与原始信号进行对比,使用 LeNet-5 网络结构作为识别模型。两种信号的识别准确率如图 7 所示。

从图 7 中可以看出,通过使用重组信号,模型的识别率明显上升,这主要是因为使用 VAD 算法排除了信号中噪音的干扰。

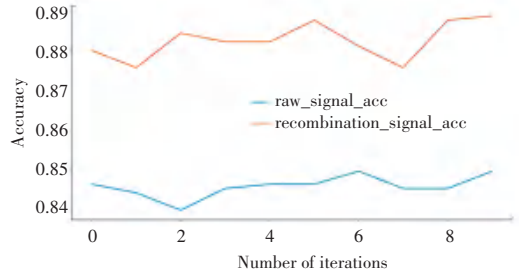
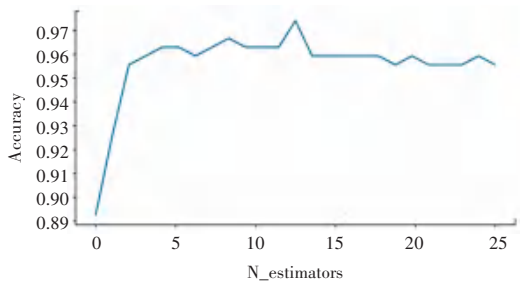


图 7 原始信号与重组信号识别率

Fig. 7 Original signal and recombination signal recognition rate

通过多尺度卷积神经网络提取重组信号的特征,再使用随机森林树对所提取特征进行识别,为了得到最优的随机森林树参数,设置不同的参数,根据学习率曲线来得出最优参数。 $N_estimators$ 的学习率曲线如图 8 所示。由图 8 中可以看出,当 $N_estimators$ 设为 13 时,识别准确率最高,可以达到 97.4%。此时使用的判断不纯度指标为 entropy 系数,最大深度为 5。

图 8 $N_estimators$ 学习曲线Fig. 8 $N_estimators$ learning curve

通过上述实验得出随机森林树最优参数后,开展模型的对比实验。数据集使用重组信号,模型则分别使用多尺度卷积神经网络结合随机森林树与 LeNet-5 模型进行识别。两种模型的识别准确率如图 9 所示。

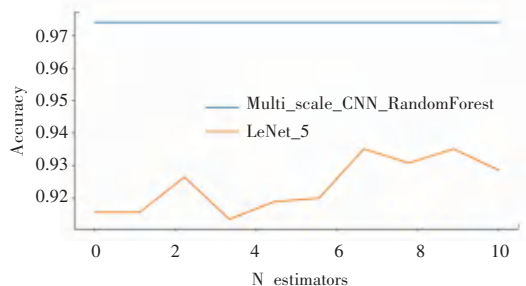


图 9 LeNet-5 与多尺度卷积神经网络结合随机森林树的识别准确率

由图 9 中可以看出,使用多尺度卷积神经网络对重组信号进行识别的准确率明显高于 LeNet-5 模型的识别准确率。(下转第 76 页)