

Vibration Event Recognition Algorithm of Φ -OTDR System Based on Two-Level Support Vector Machine *

ZHU Haiqiang¹, ZHANG Zhili^{2*}, GAO Huimin², MA Xiaoming²

(1.School of Electrical Engineering and Automation, Tianjin University of Technology, Tianjin 300384, China;

2.Intelligent Manufacturing College, Tianjin Sino-German University of Applied Sciences, Tianjin 300350, China)

Abstract: Phase sensitive optical time domain reflectometer (Φ -OTDR) can realize continuous and distributed measurement with high sensitivity for vibration signals, so it is often used for the identification and classification of vibration events. However, the current vibration event recognition algorithms often extract features from a single dimension, such as time domain or frequency domain, which can not achieve the fusion analysis of multi-dimensional and large sample feature parameters. At the same time, the existing algorithms generally use a simple unipolar recognition algorithm, which has a relatively simple structure, resulting in low accuracy and poor generalization ability of the final model recognition. To solve the above problems, multi parameter features of the experimental vibration signals are extracted and fused from the time domain, frequency domain and spatial domain. At the same time, for the specific vibration signal recognition problem, a two-level support vector machine (SVM) recognition algorithm is constructed to classify the vibration events, which can realize the accurate recognition of similar vibration events, and the recognition accuracy is more than 90%.

Key words: Φ -OTDR distributed measurement; vibration time identification; multi feature fusion; support vector machine

EEACC: 7220; 7320G

doi: 10.3969/j.issn.1005-9490.2023.03.034

基于两级 SVM 的 Φ -OTDR 系统振动事件识别算法 *

朱海强¹, 张志利^{2*}, 高慧敏², 马晓明²

(1.天津理工大学电气工程与自动化学院,天津 300384;2.天津中德应用技术大学智能制造学院,天津 300350)

摘 要: 相位敏感光时域反射计 (Φ -OTDR) 对于振动信号能够实现高灵敏度的连续分布式测量,目前的振动事件识别算法常从一个或者两个维度来提取特征,如时域或频域等,未能实现多维度大样本特征参量的融合分析;现有的算法一般采用简单的单级识别算法,结构比较简单,导致最终的模型识别准确率不高、泛化能力较差。针对上述问题,对实验采集的振动信号从时域、频域和空间域的多参量特征进行提取和融合,针对具体的振动信号识别问题,构建了一种两级支持向量机 (SVM) 识别算法,对振动事件进行两级分类,能够实现对相似振动事件的精确识别,识别准确率达 90% 以上。

关键词: Φ -OTDR 分布式测量;振动事件识别;多特征融合;支持向量机

中图分类号: O329; TP181

文献标识码: A

文章编号: 1005-9490(2023)03-0783-07

基于相位敏感光时域反射计^[1] (Phase-sensitive Optical Time Domain Reflectometer, Φ -OTDR) 的分布式光纤振动传感系统可以实现远距离、连续测量,具有监测盲区小、高分辨率和高灵敏度,且安装简单,通用性强,在石油管道监测、桥梁健康监测、周界安保监测等诸多领域具有广泛的应用。近几年, Φ -OTDR 系统在输水管线监测领域成为研究热点。

为了提高预警的准确性,降低误报率,对于振动信号模式识别的研究一直是输水管线监测领域的热点。2014 年,王兴奇等^[1]以信号强度为识别依据在埋地环境下对人工挖掘、车辆经过以及挖掘机挖掘

三种模式进行了识别研究。2015 年,张颜等^[2]以信号平方差、短时过电平率、短时傅里叶变换等特征参量为识别依据,对敲击、攀爬、浇水等模式进行识别,在时域和频域上进行分析,使用中识别正确率较低,且对相似度较高的振动识别困难。2018 年,彭宽等^[3]以平均片段间隔、片段长度和峰均比、频域能量等特征参量,对脚踩、自行车轧过、拍击等模式进行识别,识别正确率较高,但对相似事件的识别能力不够。2019 年, Wang 等^[4]提取了信号时域中最大值、平均值、均方差、脉冲因子等 10 个特征,创新性地 将随机森林分类器应用于 Φ -OTDR 系统的模式

项目来源:天津市教委科研项目(2018KJ256)

收稿日期:2021-10-18 修改日期:2022-02-22

识别中,最终分类效果能够达到 90%,但这种方法仅仅提取了单一维度特征,算法鲁棒性不高。2019 年 Shi Y^[5] 等人采用深度学习的方法,将时-空数据矩阵输入至卷积神经网络中进行训练,针对 5 种不同的振动数据进行分类,平均识别精度能够达到 97%,识别精度较高,但是算法复杂度也较高,实时性不强。2020 年, Wang Z D 等^[6] 将多分支长短时记忆卷积神经网络 (Multi branch Long Short-term Memory Convolutional Neural Network, MLSTM-CNN) 应用于 Φ -OTDR 系统的模式识别中。实现了针对浇水、敲击、攀爬等扰动事件的分类。但这种算法的算法复杂度较高,实时性较差。

针对上述存在的问题,本文提出一种基于多特征融合的振动事件识别方法。该算法从时域、频域和空间域三个维度,提取信号振动持续时间、峰峰值、均值、有效值、峰值因子、峭度因子、重心频率、均方频率和频率方差等特征参量。采用了一种基于支持向量机的两级分类思想,在第一级分类中仅根据信号的振动持续时间将信号分为连续振动信号和间断振动信号;在第二级分类中结合了峰峰值、均值等时域特征对连续振动信号大类中的各个相似的小类进行精确分类。利用多种特征向量更加有效地提高识别准确性,对相似度较高的振动目标有较强的识别能力。

1 Φ -OTDR 实验原理及系统

Φ -OTDR 系统原理如图 1 所示,系统由窄线宽激光器、声光调制器、EDFA 放大器、环形器、采集卡、平衡探测器以及光纤构成。其原理与传统 OTDR 系统原理相近,首先由窄线宽激光器发出一束连续的激光,经由声光调制器调制为指定频率和脉宽的脉冲光,其中一束脉冲光作为参考光被平衡探测器接收,另外一束探测光经过 EDFA 放大器放大后,通过环形器进入到光纤中,光在光纤中发生背向瑞利散射,散射回的光经过 EDFA 放大器放大后进入平衡探测器,将包含待测信息的探测光和作为基准的参考光在光电探测器的光敏面上进行混频。采集卡接收信号后传输到上位机进行进一步的处理。

本实验系统采用市面上常用的 G652 光纤作为传感器,传感器伴随排水管道进行铺设。激光器发

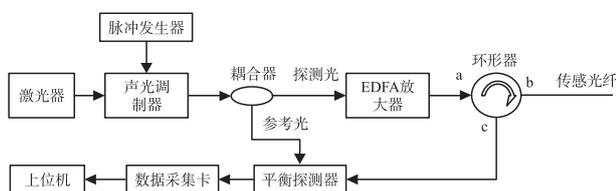


图 1 系统原理图

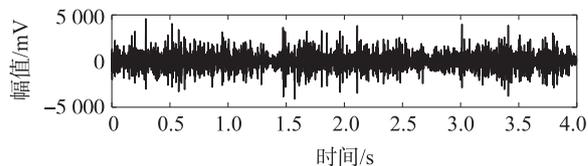
出一束脉冲光,并由声光调制器调制为 100 ns 的脉冲光,经过放大器进行放大后打入光纤,光纤总长 40 km,系统采集频率为 250 MHz,由光电探测器接收传感光纤传回的后向散射信号。探测器接收到的信号经过放大后进入信号处理模块进行处理,信号处理模块包括信号采集和信号处理两部分,信号采集部分接收探测器检测的信号,在上位计算机中显示时域和空间域信号。

2 多参量特征提取

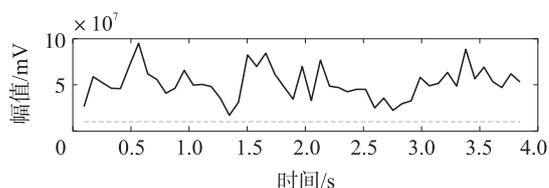
由于振动信号特征的多样性和易变性,因此用单一特征向量不足以准确鉴别出目标类别。本文采用了多特征融合的方法,充分并且有效地利用了振动信号的多维度特征信息。本文所采用的多角度多特征提取信号特征参量包括:振动持续时间、峰峰值、均值、有效值、峰值因子、峭度因子、重心频率、均方频率和频率方差等。这些指标从不同的角度突出了不同扰动事件的特征。

2.1 一级特征提取

振动持续时间: Φ -OTDR 系统采集到的振动信号也是随时间变化的非平稳随机过程,因此对于该信号的分析也可以类比于语音信号,用短时分析来研究。图 2(a)和图 3(a)为典型的连续振动信号和非持续振动信号时域图,求得短时能量如图 2(b)和图 3(b)所示,信号有振动部分的短时能量相对较高,无振动部分的短时能量相对较低。在本文中,通过设置短时能量阈值的方式来求得振动的持续时间,计算的具体步骤如下:①将时间域信号碎片化:通过窗函数对信号进行加窗、分帧,将信号进行碎片化,本文所用加窗函数为海宁窗,每帧的长度为 200,帧移为 80。②设置振动持续时间阈值:将无振动部分的振动持续时间均值设定为阈值 T 。③振动片段提取:以帧为单位,求出过阈值的振动持续时间 t 。



(a) 典型持续振动时域波形



(b) 典型持续振动短时能量

图 2 典型连续振动信号

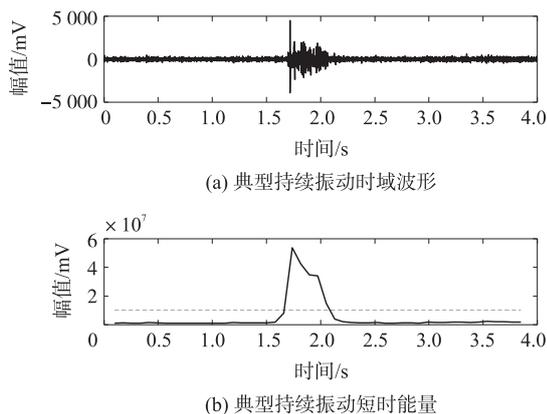


图3 典型非连续振动信号

2.2 二级特征提取

二级特征提取模型主要针对持续振动中的管道泄漏、机械挖掘、管道空腔中的水流流动信号,以及非持续振动事件中的人为触碰和人工挖掘信号进行进一步准确分类。针对这些信号的特征提取思路主要是从时域、频域、空间域对振动事件的特征进行指标的量化^[7]。

2.2.1 时域特征提取

①峰值:反映信号幅值的增大过程。

$$X_{\max} = \max(x) \quad (1)$$

②平均值^[8]:均值是反映信号中心趋势的一个指标,反映了信号中的静态部分;

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i \quad (2)$$

式中: \bar{X} 代表信号的平均值, N 为采样点数, X_i 为某一时间点的幅值。

③有效值(均方根值):反映信号的平均能量。

$$X_{\text{rms}} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} \quad (3)$$

④峰值因子^[9]:信号峰值与有效值的比值,代表的是峰值在波形中的极端程度。

⑤峭度因子:能够反映波形的陡度的指标,如果波形陡度很高,说明外界扰动变化剧烈。

$$C_k = K/\sigma^4 \quad (4)$$

2.2.2 频域特征提取

频域特征反映信号中的频率成份以及各频率成份的能量大小情况。其主要分为频谱分析、能量谱分析、功率谱分析和倒频谱分析等,其中以频谱分析最为常用也最为重要。

①重心频率:反映主频带位置的变化。

$$F_{\text{FC}} = \frac{\sum_{n=1}^N \bar{u}(n)u(n)}{2\pi \sum_{n=1}^N u(n)^2} \quad (5)$$

②均方频率:反映主频带位置的变化。

$$F_{\text{RMSF}} = \sqrt{\frac{\sum_{n=1}^N u(n)^2}{4\pi^2 u(n)^2}} \quad (6)$$

③频率方差:反映频谱的分散或集中程度。

$$S_4 = \frac{\sum_{k=1}^k (f_k - S_2)^2 S(k)}{\sum_{k=1}^k S(k)} \quad (7)$$

2.2.3 空间域特征提取

振动传导距离^[10]:当光纤上某一点位置上有振动事件发生时,光纤上一段距离的光强度都会受到影响,而不同振动事件在空间域所影响的空间长度不同。振动传导距离的计算思路如下:找到振动点的位置,并以此位置为中心向两侧发散,分别计算空间上两点信号的相关性,当信号相关性小于一定值并持续减小时,将这个距离作为振动的传导距离。信号相关性的计算公式如下:

$$r_{xy}[l] = \sum_{n=-\infty}^{\infty} x[n]y[n-l] \quad (8)$$

式中: $r_{xy}[l]$ 表示 $x[n]$ 和 $y[n-l]$ 之间的相关性大小,它的值越大代表两个序列之间的相关程度越高,即这两个序列越相似。通过分析振动中心点两侧信号的相关性,可以得到振动的传导距离。

选取部分上述二级特征提取的结果,由图4可以看出,不同类别的特征值具有一定的区分度和差异,因而可以利用 SVM 算法进一步进行近似信号的识别。

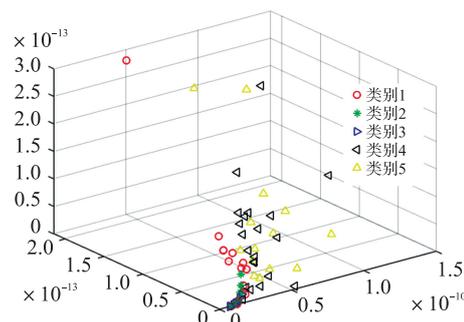


图4 二级特征提取结果

3 两级 SVM 分类算法

每种振动事件都有其独有的特性,这些特点主要反映在振动信号的波形中,不同振动事件波形的连续性,幅值、均值、频率等都会不同。因此,我们可以从时间域、频域等维度来求取一些指标来量化这些振动事件的特征。

在实际场景中不同振动事件的某些指标也会存

在相似的情况,例如过车的振动事件和管道泄漏的振动事件,虽然在频谱上有所差异,但是这两者在能量上的区别并不大,因此如果仅仅从某个指标上来区分不同的振动事件的话,最终的分类效果会很差,因此我们需要使用多个特征指标来描述不同的振动事件。这样能够在一定程度上降低事件的误识别率,但是仍然存在某两种相似振动事件难以区分的情况,无法满足精准识别的需求,除此之外,振动信号数量庞大,种类多,计算机在面对如此多的事件进行分类时,其压力无疑是巨大的,并且准确性和系统的实时性都无法得到保障。

为了解决上述问题,本文提出一种两级振动信号模式识别方法,算法的流程如图 5 所示:①为了消除不同指标的量纲对最终分类结果的影响,首先需要将原始信号进行归一化,接下来才能从不同的维度进行特征值的提取。②在第一级分类中,不考虑振动信号的具体类别,仅仅考虑信号的振动持续时间,将振动信号粗略地划分为连续信号和非连续信号两大类,在此过程中,可能会出现属于不同相似类的信号被划分到同一大类中去的情况。③在第二级分类中,主要从时域、频域、空域三个维度进行特征提取,针对持续振动以及非持续振动中的小类进行细分类。

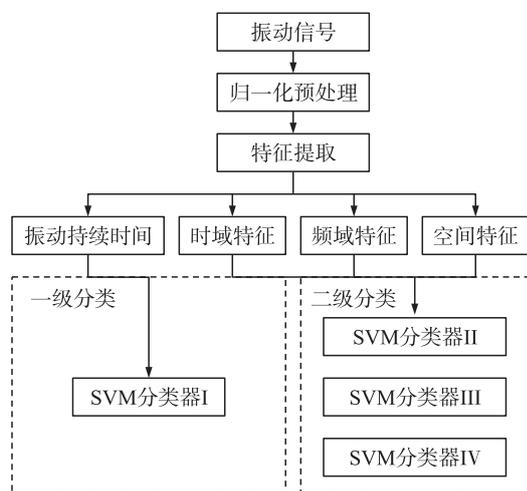


图 5 二级分类算法示意图

在两级分类模型中^[11],第一级分类相当于将外界的振动信号进行了粗分类,把外界所发生的振动事件区分为持续性振动事件和非持续性振动事件,信号在经过一级分类处理后可以识别到信号的大致类别。第一级分类器将海量的数据进行了预处理,起到了初步的过滤作用,同时避免了振动信号的频域、空域某些指标过于近似导致的误识别问题。

在第二级分类中,设计了多个二分类的分类器,

使多个分类器同时并行工作,这样在一定程度上提高了运算效率,缓解了计算机的计算压力,保障了系统处理数据的实时性,并且识别的准确率也有所提高。

本模型整体主要分两级来对外界的振动事件进行识别分类,并且在第二级分类模型中存在多个二分类器并行工作,因此模型的总体识别精度要从两级分类角度综合考虑,其计算方式如下:

$$\text{Accuracy} = A_f - \sum_i \frac{|c_i|}{|c|} (1 - A_i)$$

式中: A_f 代表了第一级分类器的识别精度, $|c_i|$ 为第 i 类的扰动事件的信号量, $|c|$ 表示需要处理的总体信号量, A_i 为第二级分类中第 i 种振动事件的识别精度。

4 实验结果与分析

4.1 实验结果

实验中基于 Φ -OTDR 光纤振动传感系统,采集了山东某地野外排水管线上几种频繁发生的振动事件进行分析,主要有:①人为触碰:为了方便对系统进行维护,需要每隔一段距离预留一部分光纤备用,假如光纤发生断裂时能够及时进行融接,这部分预留的光纤可能会被人为触碰到,从而引起系统的误报警。②人工挖掘:夜间可能会出现人为挖盗输水管线的情况存在,当外界有偷盗管道事件发生时,系统需要能够及时进行报警。③管道空腔的水流流动:由于下雨、河水流动等因素,管道下方经常会由于水流的冲击从而形成空腔,空腔中的水流流动会引起振动。当出现管道空腔时,需要及时对空腔进行填补,避免出现管道的断裂。④机械挖掘:为了对管道进行维护,现场经常会有挖掘机对管道进行挖掘。⑤管道泄漏^[12-13]:管道在长时间野外使用中会受到土壤中各种腐蚀性物质的侵蚀,很容易导致管道产生断裂,时常会有泄露产生,带来严重的安全隐患和巨大的财产损失。

实验中采集了人为触碰、人工挖掘、管道空腔内的水流流动、机械挖掘和管道泄漏这 5 种振动情形的实验数据,然后,每组实验进行 100 次,共计 500 次实验,得到 500 个样本。实验中采集了农机作业、机械挖掘、管道泄露、水流流动和人为触碰这 5 种振动情形的实验数据,然后,每组实验进行 100 次,共计 500 次实验,得到 500 个样本,并采用移动平均法^[14]对信号进行处理,去除信号中的直流分量。

图 6(a)~图 6(e)为不同振动事件的时域波形,几种振动事件分别为:人为触碰、人工挖掘、管道空

腔中的水流流动、机械挖掘和管道泄漏。从图中可以明显看出,后三种事件的持续事件较长,因此信号能量分布比较均匀,而人为触碰和人工挖掘事件的持续时间短,因此信号能量较集中。

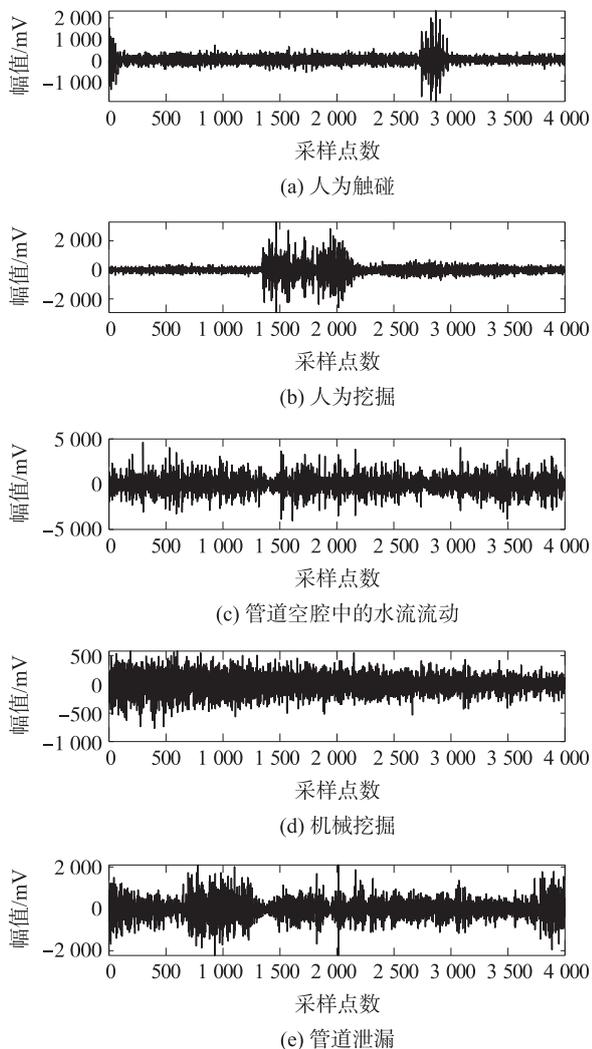


图 6 时域波形图

本文提取了 Φ -OTDR 系统中的 5 种外界振动事件的信号,并采用多特征融合的方法提取了振动事件的特征并构建了两级 SVM 分类器。首先,根据提取的振动持续事件,通过第一级分类器来区分非持续振动信号事件和其他连续振动事件;然后通过二级分类的分类器来针对非持续振动事件和持续振动事件的具体种类做详细的划分,该方法能够在一定程度上克服不同种类信号过于近似而难以识别的缺点,提高了振动源识别的准确性。

使用一级分类器的分类效果如图 7 所示,本文将人为触碰事件定义为非持续振动事件,将管道泄漏、农机作业等事件定义为持续振动事件,结果表明:这种模型能够很好地将持续振动与非持续振动区分开,识别精度能够达到 100%,且具有很好的实时性。

真实的类	1	19	3	1	82.6%	17.4%
	2	1	18		94.7%	5.3%
	3			18	100.0%	
		95.0%	85.7%	94.7%		
		5.0%	14.3%	5.3%		
		1	2	3		
		预测类				

图 7 二级分类器 I 混淆矩阵

表 1 一级分类器分类结果

事件分类	识别精度/%
持续振动	100
非持续振动	100

在一级分类器的基础上,进一步引入二级分类器,对近似的信号作进一步分类识别,识别的结果如表 2、表 3 所示。

表 2 主要针对持续振动做进一步的分类,管道泄漏的识别精度为 82.6%,管道空腔内的水流流动信号识别精度为 94.7%,机械挖掘信号的识别精度为 100%。整体的测试准确率在 92.43%,准确性优良。

表 2 二级分类器 I 分类结果

序号	振动事件	识别精度/%
1	管道泄漏	82.6
2	管道空腔内的水流流动	94.7
3	机械挖掘	100

表 3 为针对非持续振动事件中的人工挖掘和人为触碰事件进行进一步的分类,人为触碰的识别精度为 96%,人工挖掘的识别精度为 98%。能够有效区分不同的振动信号。

表 3 二级分类器 II 分类结果

序号	振动事件	识别精度/%
1	人为触碰	96
2	人工挖掘	98

4.2 实验结果分析

4.2.1 系统识别精度分析

使用混淆矩阵对最终的分类结果加以分析说明,图 7 为二级分类器 I 的混淆矩阵,事件 1—3 分别为:管道泄漏、管道空腔内的水流流动和机械挖

掘。在最终的分类结果中,有 4 组数据实际上是管道泄漏事件,然而却被识别为水流流动事件以及机械挖掘,有 1 组数据实际上是管道空腔中的水流流动事件,却被误识别成管道泄漏,其他事件均能准确识别。平均识别精度能够达到 92.43%。事件 1 管道泄漏的识别效率较低,影响了系统整体的平均识别效率,其主要原因是管道泄漏的流量大小比较复杂,容易与其他振动事件造成混淆。

图 8 为二级分类器 II 的分类结果,主要是针对非持续振动事件进行精确分类,事件 4 为人为触碰,事件 5 为人工挖掘,在测试集中仅有一组人为触碰数据被误识别为人工挖掘,二级分类器 II 的平均识别精度为 97.6%。

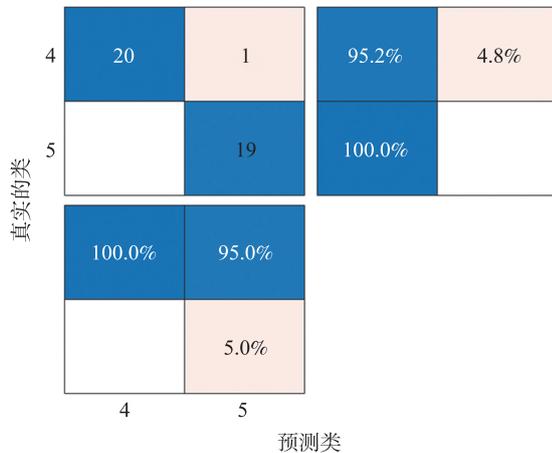


图 8 二级分类器 II 的混淆矩阵

4.2.2 系统实时性分析

使用训练好的模型来识别采集到的振动数据,进而验证算法的实时性。如表 4 所示,系统每次采集数据的时长为 2 s,实验中对 100 组数据的识别时间进行分析,得到了每种振动事件的平均识别时间,验证了基于本文所提的两级振动事件模式识别方案是可行的,算法具有良好的实时性。

表 4 系统实时性分析

振动事件		平均识别时间/s
非持续性 振动	人为触碰	0.32
	人工挖掘	0.27
持续性 振动	管道泄漏	0.36
	管道空腔中的水流流动	0.38
	农机作业	0.29

5 结论

本文搭建了 Φ -OTDR 分布式光纤实时监测系统,以识别多种振动事件为目的,使用多特征参数融合的二级 SVM 算法,能有效识别到多种振动事件。

该方法突破了单一级别振动事件识别中,不同振动事件相似性较高而难以识别的缺陷。采用了两级 SVM 分类模型,能够有效针对多种振动事件进行识别,并且算法的复杂度适中,系统具有很好的实时性,能够满足工程项目中的应用。

参考文献:

- [1] 王兴奇,朱瑞,李欣城,等. 基于 φ -OTDR 光纤振动传感系统模式识别方法[J]. 吉林大学学报:理学版, 2020,58(5):1243-1248.
- [2] 张颜,娄淑琴,梁生,等. 基于多特征参量的 φ -OTDR 分布式光纤振动传感系统模式识别研究[J]. 中国激光,2015,42(11):134-142.
- [3] 彭宽,冯诚,艾凡,等. 基于时/频域综合特征提取的分布式光纤入侵监测系统事件识别方法[D]. 武汉:华中科技大学,2017.
- [4] Wang X, Liu Y, Liang S, et al. Event Identification Based on Random Forest Classifier for Phase-Sensitive OTDR Fiber-Optic Distributed Disturbance Sensor[J]. Infrared Physics & Technology, 2019, 97: 319-325.
- [5] Shi Y, Wang Y, Zhao L, et al. An Event Recognition Method for Φ -OTDR Sensing System Based on Deep Learning[J]. Sensors, 2019, 19(15):3421.
- [6] Wang Z D, Lou S Q, Wang X, et al. Multi-Branch Long Short-Time Memory Convolution Neural Network for Event Identification in Fiber-Optic Distributed Disturbance Sensor Based on φ -OTDR[J]. Infrared Physics & Technology, 2020, 109: 103414.
- [7] Shi C, Zhao X, Wang L. A Multi-Branch Feature Fusion Strategy Based on an Attention Mechanism for Remote Sensing Image Scene Classification[J]. Remote Sensing, 2021, 13(10):1950.
- [8] 徐铖晋. 分布式光纤传感系统的信号处理技术研究[D]. 杭州:浙江大学,2017.
- [9] 张俊楠. 基于 SVM 的 φ -OTDR 分布式光纤振动传感系统模式识别研究[D]. 北京:北京交通大学,2017.
- [10] 孙茜. φ -OTDR 光纤预警系统模式识别方法研究[D]. 天津:天津大学,2015.
- [11] Qiu B, Xu C S, Tian Q. Two-Stage SVM for Medical Image Annotation [C]//2006 Conference and Labs of the Evaluation Forum (CLEF 2006), Alicante, Spain, 2006: qi-uCLEF2006.
- [12] 崔光磊,衣文索,牛卫丛,等. 基于分布式光纤传感器的管涌监测系统[J]. 长春理工大学学报(自然科学版),2018,41(1):39-43.
- [13] 张晖,吕宏伟,冯进良,等. 基于光纤振动传感器与 FPGA 的石油管道安防系统[J]. 长春理工大学学报

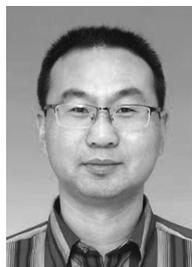
(自然科学版), 2016, 39(3): 116-119.

[14] Lu Y, Zhu T, Chen L, et al. Distributed Vibration Sensor

Based on Coherent Detection of Phase-OTDR[J]. Journal of Lightwave Technology, 2010, 28(22): 3243-3249.



朱海强(1994—),男,满族,河北人,硕士研究生,研究方向:数字信号处理,模式识别;



张志利(1978—),男,汉族,河北人,硕士生,教授,研究方向:工业自动化及智能控制。