

文章编号:1007-5321(2017)06-0109-06

DOI:10.13190/j.jbupt.2017-030

Phi-OTDR 技术信号处理中的特征提取方法

胡燕祝¹, 王松¹, 李雷远^{1,2}

(1. 北京邮电大学 自动化学院, 北京 100876; 2. 南京大学 光通信工程研究中心, 南京 210009)

摘要: 在相位光时域反射(Phi-OTDR)技术分布式振动信号的处理中,为了保证时间轴上的信噪比和长度轴上的作用域同时得到提升,同时兼顾保留信号特征和时域快速处理 2 个条件,采用了“信噪分类、信号保留、噪声归零”的方式进行信号处理. 在信噪分类的过程中,首次提出利用改进的实时编码信号处理算法,在时域上对 Phi-OTDR 分布式振动信号数据进行时域编码和特征提取. 所提方法可以同时提高长度轴上的作用域和时间轴上作用域内各点的信噪比,运算速率也得到了一定的提升.

关键词: 相位光时域反射信号; 特征提取; 信号处理

中图分类号: TP751.2

文献标志码: A

Feature Extraction Research in Signal Process with Phi-OTDR Technology

HU Yan-zhu¹, WANG Song¹, LI Lei-yuan^{1,2}

(1. Automation College, Beijing University of Posts and Telecommunications, Beijing 100876, China;

2. Institute of Optical Communication Engineering, Nanjing University, Nanjing 210009, China)

Abstract: In the phase-sensitive optical time domain reflectometer (Phi-OTDR), noise-signal ratio on time axis and effective range on length axis are two important factors. Reservation of signal characteristics and rapid processing in time domain are also two important conditions. To optimize these factors and conditions, the signal by signal-noise classification, signal reservation and noise reduction was processed. In signal-noise classification, the time encode signal processing algorithm was improved for coding and extract feature of distributed vibration signal in time domain. The signal range on the length axis and the signal-to-noise ratio of each point on the time axis can be improved. At the same time, the calculation speed is also improved.

Key words: phase-sensitive optical time domain reflectometer signal; feature extraction; signal processing

伴随着光纤通信技术的快速发展,光纤传感技术近年来得到了深入研究,并已广泛应用于多个领域. 相位光时域反射(Phi-OTDR, phase-sensitive optical time domain reflectometer)技术^[1-2]是检测分布式振动信号的一种比较典型的方法,通过 Phi-OTDR 技术采集的信号而了解监测物的状态,成为了人们

关心的热点问题.

传统的振动信号处理方法主要分为 3 类:第 1 类是在时域上直接进行信号处理,如低通、高通、带通等;第 2 类是转换到频域或其他域进行处理,如快速傅里叶变换(FFT, fast Fourier transformation) $f-k$ 变换法、希尔伯特-黄变换(HHT, Hilbert-Huang

收稿日期:2017-03-22

基金项目: 国家自然科学基金资助项目(61627816,61503034); 国家重点研发计划项目(2015BAK40B03,2016YFC0701309); 北京市科技计划项目(D161100004916002)

作者简介: 胡燕祝(1970—),男,教授,博士生导师, E-mail: yzhu@263.net.

transform)、Karhunen-Loève 变换等;第3类是一些智能算法,如小波、模糊判别、奇异值分解等。在上述方法中,有的方法不能很好地表述信号的特征;有的方法增加了运算时间,降低了效率;还有些方法不能在扩展成分布式信号上使用。总结上述方法,为了避免上述3种情况的发生,兼顾保留特征、保证效率、适合 Phi-OTDR 技术特点3个情况,“信噪分类、信号保留、噪声归零”这种方法是目前最适合 Phi-OTDR 信号处理的方法。

针对 Phi-OTDR 振动信号的特征提取工作,近几年才开始有所涉及。在国外,Tan 等^[3]以信号强度为主要参考因素进行信号处理;Zhu 等^[4]则以短时过电频率为主要参考要素进行信号处理。国内方面,李凯彦等^[5]综合利用特征段时域持续时间和小波包能量谱提取复合特征向量的方法对 Phi-OTDR 数据进行处理;张颜等^[6]采用提取多特征参量进行信号处理。在上述的研究中,有的没有考虑 Phi-OTDR 信号包含长度、时间、幅度三维数据的情况^[7-8];有的没有考虑 Phi-OTDR 信号数据量大的情况;还有的方法在对特征提取的过程中,仅仅限于振幅等信号基础信息,特征表述过于简单,不能很好地表达信号的特点。

笔者主要通过“信噪分类、信号保留、噪声归零”的方式进行信号处理,针对 Phi-OTDR 技术采集的三维振动信号的特点,采用改进的 TESP 算法进行信噪分离,最终使 Phi-OTDR 信号在时间轴上和长度轴上的性能都得到了提升。

1 评估指标

基于 Phi-OTDR 监测到的是分布式振动数据,对其的评判指标是由信噪比和有效信噪比下的最大距离2个指标共同决定的。有效信噪比下的最大距离是分布式振动信号特有的指标,简称“作用域”。因此,下面结合 Phi-OTDR 技术的特点对这2个指标做具体明确的说明。

1.1 信噪比

信噪比是指一个电子设备或者电子系统中信号与噪声的比例。结合实际中 Phi-OTDR 的特点,监测到的振动信号的波形并不已知,因此采用把信号进行无监督分类的方式进行处理,具体步骤如下。

1) 对信号进行分窗处理,设窗宽为 n ,第 m 个采样点的窗口、 $n \sim n+m$ 窗口作为一个单元。

2) 对每个单元在时域上进行编码处理后,再进

行特征提取。

3) 对各单元的特征信息代入到经典的聚类算法中进行聚类处理。

4) 对聚类结果进行观察,属于信号类的单元保留原来数值,其余采样点的数值归为0。

1.2 作用域

对于光纤上的作用域是指可以检测到信号的光纤长度范围段,求解作用域的步骤如下。

1) 求解该函数的最高点,记作 i_1 。

2) 求解距离 i_1 左右最近与横轴的交点,记作 i_2 和 i_3 。

3) 作用域如式(1)所示。

$$L = (i_2 - i_3) \times 90\% \quad (1)$$

4) 作用域的边界值如式(2)和式(3)所示。

左边界:

$$i'_2 = i_2 + (i_2 - i_3) \times 5\% \quad (2)$$

右边界:

$$i'_3 = i_3 - (i_2 - i_3) \times 5\% \quad (3)$$

图1为信号计算后作用域示意图。

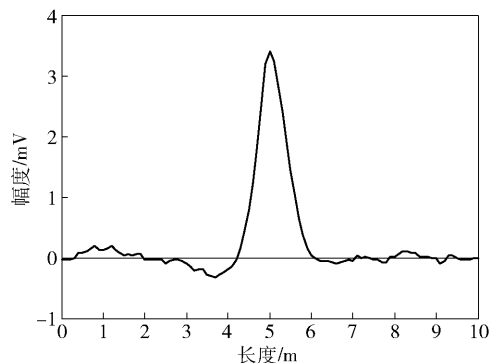


图1 作用域示意图

1.3 信噪比和作用域的关系

因为光纤是一个有机的整体,在长度轴上作用域内的点,在时间轴上是可以观察到信号的;同理,在时间轴上可以观察到信号的点,在长度轴上一定是在作用域之内的。振源与光纤的垂点为信号最强点,分布式信号会以最强点为中心,沿光纤向两端衰减直到消失。根据一个实例,长度轴上的作用域为4.5~5.9 m处,5.2 m信号最强。因此观察到4.4~6.1 m中信号的幅度随时间和长度变化而变化。表1是作用域内各点的信噪比数据。

2 特征提取

为了在充分保留信号特征的基础上,进行快速

表 1 作用域内各长度点原始信号信噪比

序号	长度点/ m	信噪比/ dB	序号	长度点/ m	信噪比/ dB
1	4. 4	0. 506 1	10	5. 3	6. 769 1
2	4. 5	1. 798 1	11	5. 4	6. 319 4
3	4. 6	3. 521 8	12	5. 5	5. 977 1
4	4. 7	5. 008 4	13	5. 6	5. 436 8
5	4. 8	5. 296 4	14	5. 7	4. 860 8
6	4. 9	6. 016 3	15	5. 8	3. 806 6
7	5. 0	6. 402 9	16	5. 9	1. 583 6
8	5. 1	6. 848 5	17	6. 0	0. 423 8
9	5. 2	7. 958 9	18	6. 1	0

高效的滤波处理,并且还可以保证信噪比和作用域的同时提高,选取改进的实时编码信号处理(TESP, time encode signal processing)方法对信号进行特征提取.

2.1 TESP 算法

TESP 算法^[9-10]的实现步骤如下.

1) 在每个长度节点上,把划分好的窗口数据切割成若干时间段,切割原则寻找所有过零点,2 个相邻的过零点中间为一个时间段,按照 TESP 算法的算法习惯每个时间段称为元.

2) 在每个元指定 2 个指标. 一个是持续时间,一般用 D 表示;另一个是信号形态,一般用 S 表示. 同时,根据这 2 个指标得到如下信息.

① 在每个元内存在多少个采样点,也就是持续时间多长.

② 在每个元的采样点内,进行一次求导,得到每个元内的极值个数情况.

3) 以 D 和 S 作为 2 个维度,构建矩阵,并把每个矩阵中的相应元素进行编码处理.

4) 统计矩阵中每个编码所出现的概率,最后把概率分布作为特征代入到分类或者聚类器中.

2.2 基于改进 TESP 算法的特征提取

改进的 TESP 算法中,首先把 D - S 矩阵扩展成 D - $S1$ 和 D - $S2$ 矩阵;其次对每个矩阵的编码原则进行拓展,由 29 个拓展成 43 个;最后再把 D - $S1$ 和 D - $S2$ 矩阵中的编码进行联合概率分布统计,构成 A 矩阵.

2.2.1 扩展 TESP 的符号表

传统的 TESP 算法表中, S 只表示一阶求导的极值,但是把 S 是极值点的矩阵记作 D - $S1$,把 S 是拐点的矩阵记作 D - $S2$. 同时,因为 Phi-OTDR 技术的

特点,2 个指标 D 和 S 都会变大,所以采用 43 个字符的表示形式. 表 2 和表 3 分别为 D - $S1$ 矩阵标准的 29 字符符号表和扩展的 43 字符符号表. 表 4 和表 5 分别为 D - $S2$ 矩阵标准的 29 字符符号表和扩展的 43 字符符号表.

表 2 D - $S1$ 矩阵标准的 29 字符符号表

D	S					
	1	2	3	4	5	>5
1	1					
2	2	2				
3	3	3	3			
4	4	4	4	4		
5	5	5	5	5	5	
6	6	6	6	6	6	6
⋮	⋮	⋮	⋮	⋮	⋮	⋮
34	24	25	26	27	28	29
35	24	25	26	27	28	29

表 3 D - $S1$ 矩阵扩展的 43 字符符号表

D	S					
	1	2	3	4	5	>5
1	1					
2	1	1				
3	2	2	2			
4	3	3	3	4		
5	4	4	5	5	5	
6	5	5	6	6	6	6
⋮	⋮	⋮	⋮	⋮	⋮	⋮
34	37	38	39	40	41	42
35	38	39	40	41	42	43

表 4 D - $S2$ 矩阵标准的 29 字符符号表

D	S					
	1	2	3	4	5	>5
3	1					
4	2	2				
5	3	3	3			
6	4	4	4	4		
7	5	5	5	5	5	
8	6	6	6	6	6	6
⋮	⋮	⋮	⋮	⋮	⋮	⋮
34	24	25	26	27	28	29
35	24	25	26	27	28	29

表 5 D -S2 矩阵扩展的 43 字符符号表						
D	S					
	1	2	3	4	5	>5
3	1					
4	1	1				
5	2	2	2			
6	3	3	3	4		
7	4	4	5	5	5	
8	5	5	6	6	6	6
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
34	37	38	39	40	41	42
35	38	39	40	41	42	43

从表 2 ~ 表 5 的对比中可以看出,随着表示符号的增多,表示的特征区分得更为细致. 对于采集到的振动信号来说,使用扩展后的 D -S1 和 D -S2 矩阵 43 符号表示会更具有代表性,同时也会保留更多的信号特征信息.

2. 2. 2 二维 A 矩阵 TESP 算法

把原算法扩展到 43 字符后,除了用原来的概率密度进行表示外,进一步将算法的特征表示为二维的 A 矩阵. A 矩阵的 2 个维度分别是拓展 D -S1 和拓展 D -S2 矩阵中的 43 个元素, A 矩阵的数值是拓展 D -S1 和拓展 D -S2 矩阵的联合分布概率.

选取一个窗口的信号,按照改进的 TESP 算法进行编码,图 2 和图 3 分别为 D -S1 和 D -S2 的概率分布直方图. 图 4 为 D -S1 和 D -S2 联合分布概率 A 矩阵直方图.

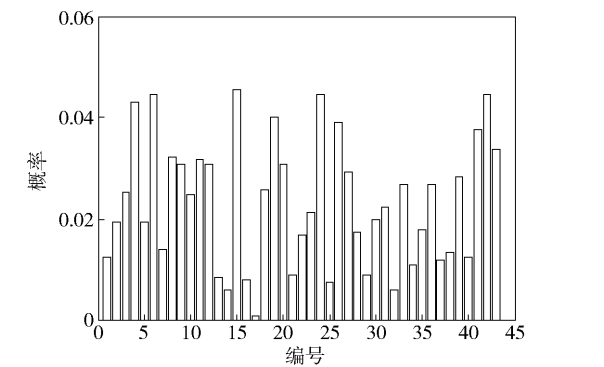


图 2 某信号 D -S1 编码的概率分布直方图

2. 3 算法分析

针对 TESP 的算法改进主要有如下几个方面的考虑.

1) 因为作用域的计算中,需要计算各采样点之

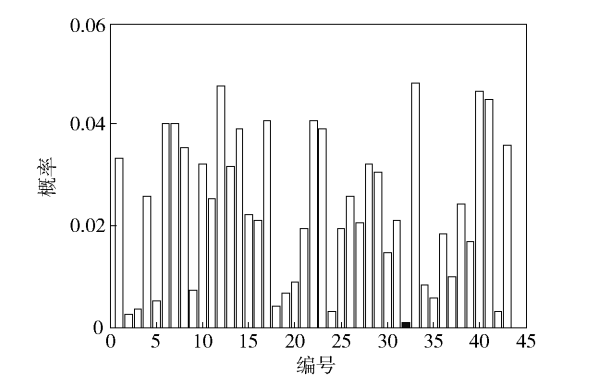


图 3 某信号 D -S2 编码的概率分布直方图

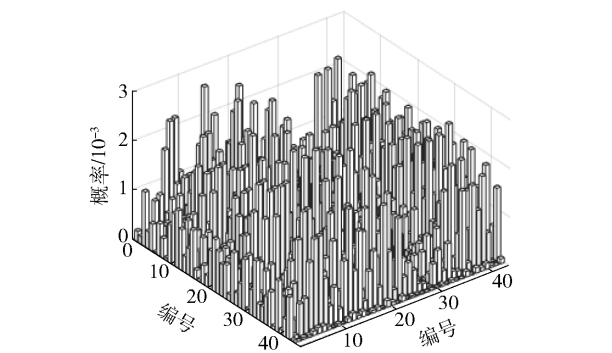


图 4 某信号 A 矩阵编码的概率分布直方图

和,在噪声部分的计算可能会出现互相抵消的现象,影响作用域的计算,所以选取改进的 TESP 滤波方式直接对信号进行保留,对噪声进行归零.

2) 拓展字符后,43 字符肯定比 29 字符对信号的表达更为细致和精确. 43 字符是经过反复实验得到的结果,在 K -means、层次聚类 and 谱聚类中,43 字符的效果是最好的.

3) 因为 Phi-OTDR 技术存在三维数据,所以对信号的表达不仅用极值点,还采用了拐点,后续的实验结果证明,同时考虑极值点和拐点的情况表达效果会更好.

3 实验验证

3. 1 实验环境

采用的 Phi-OTDR 是日本光纳株式会社生产的 NBX-S3000 仪器设备,设备的实际参数如下:采样率为 4 kHz,记录时间为 1 s,空间分辨率为 0.1 m,记录距离为 10 m. 下面主要分为 5 种情况进行实验讨论,振动源采用标准的敲击设备,敲击周期为 0.1 s. 为了抑制环境噪声,保证振动效果的良好性,特意选取消声室进行实验,其中光纤紧贴地面,仪器放在了减振台上.

3.2 具体实验

首先针对同一信号,进行不同方法的特征提取,分别采用了 TESP 编码、改进 TESP 编码、FFT、HHT4 种方法,比较 4 种方法的运算时间,如图 5 所示.

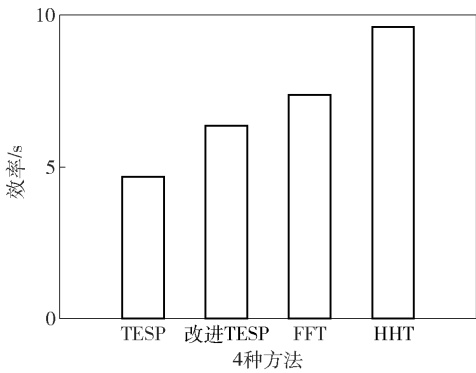


图 5 不同编码方法的运算时间对比

由图 5 可以很明显地看出,TESP 方法的运算时间最短,而改进 TESP 方法虽然比传统 TESP 方法的运算时间略有加长,但是相较于比较流行的 FFT 和 HHT 具有更加良好的效果.

在利用所提出的方法进行编码并提取特征后,把其分别代入 *K*-means 聚类算法、层次聚类 and 谱聚类中,对比极值 29 字符、极值 43 字符、拐点 29 字符、拐点 43 字符、极值拐点矩阵 5 种情况,各情况的聚类准确率如表 6 所示.

表 6 不同 TESP 编码之间的正确率比较

序号	名称	<i>K</i> -means/%	层次聚类/%	谱聚类/%
1	D-S1(29)	84.76	85.97	88.12
2	D-S1(43)	87.67	88.21	90.34
3	D-S2(29)	80.31	82.41	85.01
4	D-S2(43)	81.74	83.92	86.34
5	A	90.12	93.12	94.42

由表 6 可以看出,改进的 TESP 算法对于 Phi-OTDR 数据的特征提取具有良好的效果,使准确率平均提高了 6% 左右. 在上述实验中,需要强调两点:① 采用的数据已经提前经过了人为的标注,无监督学习的分类结果,会和标注好的结果进行对比,从而计算正确率;② 聚类方法采用 Matlab 软件中的工具包进行计算,所有参数均为默认参数. 为了验证所提方法对于信号信噪比的影响,观察使用所提方法后,4.4 ~ 6.1 m 长度段的数据,得到结果如表 7 所示.

对比表 7 和表 1 的数据可以看出,信号的信噪比有明显的提升,平均提升了 1 dB 左右. 再观察信

表 7 采用所提方法后作用域内各长度点原始信号信噪比

序号	长度/ m	信噪比/ dB	序号	长度/ m	信噪比/ dB
1	4.4	1.789 5	10	5.3	8.435 5
2	4.5	3.439 0	11	5.4	7.865 0
3	4.6	4.949 2	12	5.5	7.485 7
4	4.7	5.601 1	13	5.6	6.206 5
5	4.8	6.313 0	14	5.7	5.112 7
6	4.9	6.340 4	15	5.8	4.760 7
7	5.0	7.727 8	16	5.9	3.000 5
8	5.1	8.543 5	17	6.0	2.289 0
9	5.2	8.971 5	18	6.1	1.866 7

号的作用域,对比使用该方法前后作用域的情况,如图 6 所示.

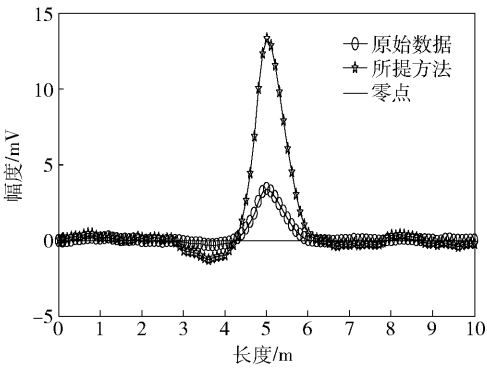


图 6 4 000 个采样点的平均值“长度-幅度”振动波形

由图 6 可以看出,信号的作用域在左侧边界没有变化,但是在右侧边界却有明显提升,这也恰恰对应了表 1 和表 7 中 6.1 m 处的信噪比数据的对比.

为了和其他信号处理方法有一个综合的对比,选取了一些比较流行的信号处理方法,比较各方法之间的最终效果,结果如表 8 所示. 在这里需要指

表 8 各种信号处理方法之间的比较

序号	名称	18 个长度点的信噪比平均值/dB	作用域/ m	运行时间/ s
1	原始信号	4.210 3	4.5 ~ 6.0	—
2	所提方法	5.036 2	4.5 ~ 6.1	10.78
3	卡尔曼	4.623 1	4.5 ~ 6.0	12.71
4	维纳	4.867 2	4.5 ~ 6.0	11.18
5	自适应	4.451 7	4.5 ~ 6.0	14.31
6	小波(6 层)	4.995 1	4.5 ~ 6.0	16.14

出的是,类似于低通、高通等方法过于简单且效果极为不佳,因此在表8中就不一一列出了。

由表8可以看出,在平均信噪比方面,所提方法比原始信号提高了18%,比最好的小波方法提高了0.2%;在作用域方面,只有所提方法使作用域范围得到了有效地提升;在运行时间方面,所提方法比用时最短的维纳滤波器快了3.8%。总体来说,所提方法是效率最高的,可以同时提高信噪比和作用域,且对信噪比和作用域的提高效果是最好的。

4 结束语

针对Phi-OTDR技术采集的信号进行研究,采用的特征提取方法通过理论分析和实验证明得到如下结论。

1) 所提出的编码及求特征值的方法,运算效率远远高于变换成频域或者其他域的方法。

2) 把TESP算法,从29字符拓展成43字符,一阶拓展成二阶,一阶二阶求联合概率,其特征表示效果得到了极大地提升,聚类准确率平均提高了6%。

3) 采用所提方法对信号进行处理后,可以达到长度轴上作用域提升,作用域内各点时间轴上的信噪比同时提升。

4) 从信噪比、作用域、运算速率3方面考虑,所提出的方法和目前比较流行的信号处理方法比较,3方面的效果均是最佳的。

综上所述,所提出的方法对Phi-OTDR技术信号的处理是快速准确的,但在本研究的同时也引发了一些思考。第一,Phi-OTDR技术数据量过于庞大,即使所提方法较传统的频域方法有所提高,但还是可能会导致运算资源不足的问题;第二,本研究主要是针对机械阻尼振动信号进行研究,但是实际中的信号类型会有很多种,所提出的方法是否也适用于其他类型的信号,这是一个未来有待探索的问题;第三,TESP算法实质上是在时域上编码的过程,是否可以实现时间轴和长度轴同时编码,这样可以更好地展现信号分布式的特点。希望这些问题,可以在今后的研究中逐渐得到解决。

参考文献:

- [1] Wang Zinan, Zhang Li, Wang Song, et al. Coherent Φ -OTDR based on I/Q demodulation and homodyne detection[J]. Optics Express, 2016, 24(2): 853-858.
- [2] Zhang Xuping, Guo Zheng, Shan Yuanyuan, et al. Enhanced Φ -OTDR system for quantitative strain measure-

ment based on ultra-weak fiber Bragg grating array[EB/OL]. Bellingham: SPIE, 2016[2017-01-02]. <http://spie.org/publications/Journal/10.1117/1.OE.55.5.054103>.

- [3] Tan Dongjie, Tian Xiaozhong, Sun Wei, et al. An oil and gas pipeline pre-warning system based on Φ -OTDR[EB/OL]. Santander: SPIE, 2014[2017-01-02]. <http://spie.org/Publications/Proceedings/Paper/10.1117/12.2054698>.
- [4] Zhu Hui, Pan Chao, Sun Xiaohan. Vibration pattern recognition and classification in Phi-OTDR based distributed optical-fiber vibration sensing system [C]//Smart Structures and Materials Nondestructive Evaluation and Health Monitoring. San Diego: SPIE, 2014: 3814-3818.
- [5] 李凯彦, 赵兴群, 孙小菡, 等. 一种用于光纤链路振动信号模式识别的规整化复合特征提取方法[J]. 物理学报, 2015, 64(5): 243-249.
Li Kaiyan, Zhao Xingqun, Sun Xiaohan, et al. A regular composite feature extraction method for vibration signal pattern recognition in optical fiber link system[J]. Acta Physica Sinica, 2015, 64(5): 243-249.
- [6] 张颜, 娄淑琴, 梁生, 等. 基于多特征参量的 Φ -OTDR 分布式光纤扰动传感系统模式识别研究[J]. 中国激光, 2015(11): 134-142.
Zhang Yan, Lou Shuqin, Liang Sheng, et al. Study of pattern recognition based on multi-characteristic parameters Φ -OTDR distributed optical fiber sensing system[J]. Chinese Journal of Lasers, 2015(11): 134-142.
- [7] Xu Haining, Zhang Yinxian, Han Lei, et al. Research on method to prevent external damage of submarine cable by distributed optical fiber sensing system with Φ -OTDR[J]. Electrical Engineering, 2016, 23(2): 46-48.
- [8] Wang Yu, Jin Baoquan, Wang Yuncai, et al. Real-time distributed vibration monitoring system using Φ -OTDR[J]. IEEE Sensors Journal, 2017, 17(5): 1333-1341.
- [9] 王艳, 李文藻, 张意, 等. 基于改进 TESP 算法的边防车辆类型声音识别[J]. 四川大学学报(工程科学版), 2014(S2): 122-127.
Wang Yan, Li Wenzhao, Zhang Yi, et al. Vehicle sound signal recognition in border surveillance based on improved TESP algorithm[J]. Journal of Sichuan University (Engineering Science Edition), 2014(S2): 122-127.
- [10] Sher M, Ahmad N, Sher M. TESP feature based isolated word speaker recognition system[C]//2012 18th International Conference on International Conference on Automation and Computing. Loughborough: IEEE, 2012: 1-4.