

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

李凯彦 赵兴群 孙小菡 万遂人

A regular composite feature extraction method for vibration signal pattern recognition in optical fiber link system

Li Kai-Yan Zhao Xing-Qun Sun Xiao-Han Wan Sui-Ren

引用信息 Citation: [Acta Physica Sinica](#), 64, 054304 (2015) DOI: 10.7498/aps.64.054304

在线阅读 View online: <http://dx.doi.org/10.7498/aps.64.054304>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2015/V64/I5>

您可能感兴趣的其他文章

Articles you may be interested in

基于椭球封闭空气腔的光纤复合法布里-珀罗结构折射率传感特性研究

Refractive index sensing characteristic of a hybrid-Fabry-Pérot interferometer based on an in-fiber ellipsoidal cavity

物理学报.2014, 63(24): 240701 <http://dx.doi.org/10.7498/aps.63.240701>

基于 Voronoi 图盲区的无线传感器网络覆盖控制部署策略

A deployment strategy for coverage control in wireless sensor networks based on the blind-zone of Voronoi diagram

物理学报.2014, 63(22): 220701 <http://dx.doi.org/10.7498/aps.63.220701>

星载激光多普勒测风雷达鉴频系统仿真(I): 基于 Fizeau 干涉仪的 Mie 通道大气风速反演研究

Simulation of frequency discrimination for spaceborne Doppler wind lidar (I): Study on the retrieval of atmospheric wind speed for Mie channel based on Fizeau interferometer

物理学报.2014, 63(14): 140702 <http://dx.doi.org/10.7498/aps.63.140702>

星载激光多普勒测风雷达鉴频系统仿真(II): 基于 Fabry-Perot 标准具的 Rayleigh 通道大气风速反演研究

Simulation of frequency discrimination for spaceborne Doppler wind lidar (II): Study on the retrieval of atmospheric wind speed for Rayleigh channel based on Fabry-Perot interferometer

物理学报.2014, 63(14): 140703 <http://dx.doi.org/10.7498/aps.63.140703>

基于无线传感器网络的地震信号特征提取方法研究

A novel approach to research on feature extraction of seismic wave signal based on wireless sensor networks

物理学报.2013, 62(10): 104301 <http://dx.doi.org/10.7498/aps.62.104301>

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*

李凯彦 赵兴群 孙小菡 万遂人†

(光传感/通信综合网络国家地方联合工程研究中心, 东南大学生物科学与医学工程学院, 南京 210096)

(2014年7月21日收到; 2014年8月27日收到修改稿)

相位光时域反射链路监测系统是一种利用光纤作为传感介质的传感系统, 能够监测、定位、识别入侵信号. 模式识别模块是其重要组成部分, 实时智能区分安全扰动和危险入侵. 本文提出一种用于光纤链路振动信号模式识别的复合特征提取方法. 利用改进的双门限方法确定有效信号段的起止位置, 结合最大能量与最高信噪比挑选出采样周期内主要入侵扰动的特征段. 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, 使用支持向量机进行模式识别. 实验表明, 基于本文提出的规整化特征提取方法的模式识别准确率有了显著提高.

关键词: 光纤振动信号, 小波包, 复合特征提取, 模式识别

PACS: 43.60.Hj, 07.07.Df, 42.30.Sy

DOI: 10.7498/aps.64.054304

1 引言

分布式光纤传感技术^[1]使用长达几公里的光纤作为传感元件, 光纤上的每一个点同时作为信号的传感和传输单元^[2], 通过在重要设施周围铺设光纤, 实现远程实时监测外界入侵. 这一技术已经广泛应用在不同的监测环境中^[3]. 国外开展研究较早, 技术已经比较成熟, 国内的浙江大学和天津大学均研制了分布式光纤管道安全检测系统. 东南大学研制的光纤链路监测系统已经成功应用在了储油区、银行、动物园等重要场馆设施的周界监测, 系统框图如图1所示, 模式识别模块作为监测系统核心需要深入研究. 模式识别有遗传算法(GA), 模糊推理(FR), 线性判别分析(LDA)^[4], 人工神经网络(ANN)^[5], 支持向量机(SVM)^[6]以及相关向量机(RVM)^[7]等方法, 目前用于光纤信号分类的方法主要是ANN和SVM, 均能有效分类非线性问题^[8,9].

光纤振动信号是典型的非稳定信号, 非稳定信号常规的特征提取方法有过零检测、经验模态分解法、频谱分析法^[10]等. 目前光纤振动信号主要的特征提取方法有小波包能量谱特征提取^[11]等方法. 在实际应用中, 基于小波包能量谱的特征提取方法能较好区分无害信号和入侵信号, 但入侵信号的具体类型识别率较差. 本文提出一种基于小波包分解的复合特征提取方法, 可以有效解决这一问题. 选取包含入侵信息最多的特征段, 规整化处理后将有效信号分段持续时间作为特征量加以利用, 与小波包能量谱整合, 共同构成最终的特征向量. 分别使用传统小波包能量谱和本文提出的复合特征提取方法得到的特征向量用于分类. 实验表明, 对5类光纤振动信号的4种分类情况进行模式识别, 基于本文特征提取方法的模式识别准确率比传统的小波包分解特征提取方法最高有17.55%的提高. 识别准确率最低为91.23%, 最高达100%, 平均识别准确率为95.998%.

* 江苏省产学研前瞻性项目(批准号: BY2009155, BY2013073)资助的课题.

† 通信作者. E-mail: srwan@seu.edu.cn

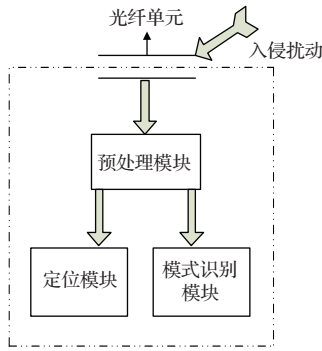


图1 OTDR 光纤链路系统框图

2 特征段选取

光纤对外界扰动十分敏感,有效入侵信号之间夹杂着噪声,对特征提取和模式识别都容易产生不良影响.从图2中可知,一个信号采样周期内可能并不包含任何入侵信息,也可能包含一至两段入侵信息,一段入侵信息可能分散在两个或者更多采样周期内.一个采样周期内如果出现多段光纤振动信号,表明光纤传感单元受到了多次入侵扰动.例如,一次剪网事件之后紧接着发生攀爬事件,在信号上表现为一段攀爬信号紧接着一段剪网信号.对包含不同类型外界入侵扰动的光纤振动信号进行特征提取,提取到的特征是多次扰动事件的综合.因此需要采取一定措施从中选取出表示主要入侵动作的有效特征段.

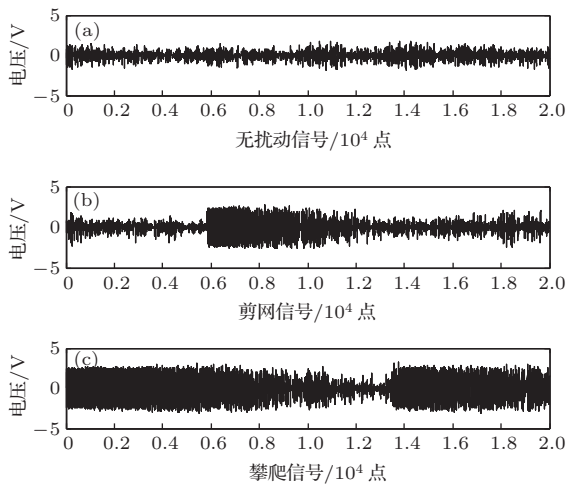


图2 三种不同类型光纤链路信号

经大量实验验证,本文采用改进的双门限判定法,即时间门限取700个采样点,阈值门限取峰值电位30%,确定有效信号段的起止位置.利用

Hilbert 变换对预处理后的信号进行包络提取并检查包络幅度,如果包络幅度在一段时间 t 内均大于阈值 $thresh$ 时,确定该段扰动信号有效,据此逐点检查,直至检测出所有有效信号段.其中 t 为门限1, $thresh$ 为门限2.信号其余小于阈值或者持续时间过短的部分视为干扰噪声段,均置零.而后分别计算各段有效信号的能量和信噪比,将能量最大和信噪比最高的两段信号分别标记出来,记为 S_1, S_2 .

设 E 表示能量, x_i 表示每个有效信号段中各点的幅值, $i = 1, 2, \dots, N, N$ 为有效信号段的长度.则 E 的计算公式为

$$E = \sum_{i=1}^N |x_i|^2, \quad i = 1, 2, \dots, N. \quad (1)$$

信噪比用分贝数表示

$$SNR = 10 \log_{10} \frac{power_{signal}}{power_{noise}}, \quad (2)$$

其中, $power_{signal}$ 为去噪之后的信号功率, $power_{noise}$ 为噪声功率,定义为

$$power_{signal} = \frac{1}{n} \sum_n f^2(n), \quad (3)$$

$$power_{noise} = \frac{1}{n} \sum_n (f(n) - \tilde{f}(n))^2, \quad (4)$$

其中 $f(n)$ 为原始带噪信号, $\tilde{f}(n)$ 为去噪之后的信号.设 S_1 的最大能量为 $\max E_1$,最高信噪比为 $\max SNR_1, S_2$ 的最大能量为 $\max E_2$,最高信噪比为 $\max SNR_2$.若

$$\begin{aligned} & \left| \frac{\max E_1 - \max E_2}{\max E_1} \right| \\ & \geq \left| \frac{\max SNR_1 - \max SNR_2}{\max SNR_1} \right|, \quad (5) \end{aligned}$$

则选择 S_1 作为有效特征段保留,否则,选择 S_2 作为有效特征段.

3 特征提取

3.1 小波包分解

傅里叶变换广泛用于稳定信号分析中,但分布式光纤传感器俘获的振动信号是非稳定信号,此时傅里叶变换不再适用.小波的多分辨率分析克服了FFT的缺陷,在时域和频域都具有较强的局部分析能力^[12].小波包通过以二叉树的形式同时分解信号的高频和低频部分^[13],能保留高频部分更多的信号细节.Mallat使用多分辨率分析提出了

Mallat 快速小波分解和重构算法^[14]. 将尺度子空间 V_j 和小波子空间 W_j 用一个新的子空间 U^n 统一表征. 令

$$U_j^0 = V_j, \quad U_j^1 = W_j, \quad j \in Z. \quad (6)$$

定义子空间 U_j^n 是函数 $w_n(t)$ 的闭包空间, 而 U_j^{2n} 是函数 $w_{2n}(t)$ 的闭包空间, 并令 w_n 满足如下双尺度方程:

$$\begin{aligned} w_{2n}(t) &= \sqrt{2} \sum_k h(k)w_n(2t - k), \\ w_{2n+1}(t) &= \sqrt{2} \sum_k g(k)w_n(2t - k), \end{aligned} \quad (7)$$

其中, $h(k)$ 和 $g(k)$ 为正交滤波器系数, Mallat 快速分解算法为

$$\begin{aligned} c_{j+1,k} &= \sum_{n \in k} h(2n - k)c_{j,n}, \\ d_{j+1,k} &= \sum_{n \in k} g(2n - k)d_{j,n}, \end{aligned} \quad (8)$$

其中, c_j 和 d_j 分别为不同分辨率下的低频系数和高频系数.

3.2 特征段规整

不同类别的光纤振动信号其能量谱各不相同, 利用小波包分解计算能量谱^[11]的方法来构建特征向量被广大研究者广泛采用. 小波包分析包含时域和频域两部分的信息. 提取位于不同时间段的同一信号的特征向量, 结果如图 3 所示.

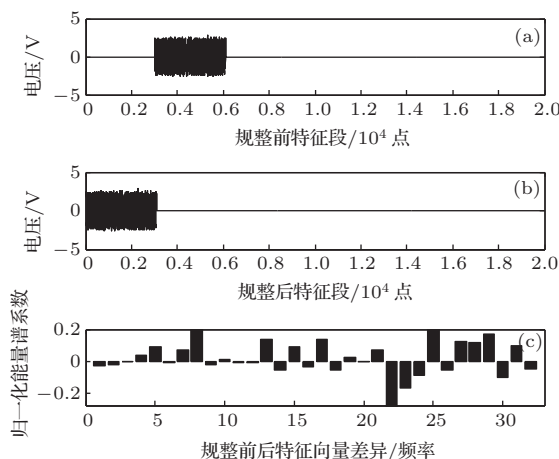


图 3 (a) 规整前特征段; (b) 规整后特征段; (c) 规整前后特征向量差异

图 3(a) 是选取到的特征段, 对其做 5 层小波包分解得到 32 维能量谱特征向量. 而后将特征段平

移到 (b) 所示位置, 再次计算 32 维特征向量. 前后两次计算得到的特征向量的差异如 (c) 所示, 两者之差中最大一维的绝对值为 0.3627, 而特征向量中每一维的最大值为 1, 时域平移前后特征量改变最高达 36.27%. 可见, 对于同一特征段, 处于时域中不同位置, 小波包能量谱特征向量有很大的区别, 且此区别不容忽视. 因此, 本文规定, 在进行特征提取之前, 将选取到的特征段统一规整到信号初始位置, 减少时域定位差异.

3.3 基于小波包能量谱特征提取

小波包系数的能量较幅度更能体现出小波包系数之间的差异, 按照能量方式表示的小波包分解结果称为小波包能量谱. 大量实验表明, 识别准确率随着分解层数增加而增加, 但是带来不利因素是数据量的增大导致的训练时间加长, 达到一定分解层数后 (比如 4 层), 识别准确率趋于稳值, 这时增加层数就没有意义了. 综合训练时间和识别准确率, 本系统的分解层数定为 4 层. 小波包能量谱的特征量定义为

$$E_{4j} = \int |S_{4j}(t)|^2 dt = \sum_{k=1}^n |x_{jk}|^2, \quad (9)$$

其中, S_{4j} 是分解第 4 层各频带的重构信号, x_{jk} 是 S_{4j} 各离散点的幅值. 令 E 为特征量总和, 特征向量 T 为

$$T = [E_{40}, E_{41}, \dots, E_{414}, E_{415}] / E. \quad (10)$$

3.4 复合特征提取

本文处理的光纤链路振动信号分为无扰动信号、大雨信号、大风信号、剪网信号和攀爬信号 5 类. 其中无扰动信号, 大雨信号, 大风信号属于无害信号; 剪网信号和攀爬信号属于入侵信号. 无害信号和入侵信号的特征向量区别明显, 利用支持向量机等模式识别方法能得到理想的分类效果. 无扰动信号, 大雨信号和大风信号之间的差异也明显, 分类准确率理想. 但剪网信号和攀爬信号特征向量类似, 当训练样本数较少时, 两类信号识别率较低. 注意到攀爬信号较剪网信号振动持续时间长且衰减速率慢, 时域能量分布更均匀. 因此考虑将有效信号持续时间和信号衰减率作为新的特征整合到能量谱特征向量中, 形成新的 $2^N + m + n$ 维特征向量, 其中 N 为分解层数, m 为新增加的持续时间特

度量, n 为新增加的信号衰减率特征量. 具体做法如下:

- 1) 计算特征段的总能量;
- 2) 对信号分段, 总长 20000 个信号点, 均分成 20 个子段, 每段 1000 个信号点;
- 3) 从第一段开始计算子段的能量, 依次相加, 当相加子段的能量和分别大于总能量的 20%, 40%, 60% 和 80% 时, 记录对应子段序号 $m_i, i = 1, 2, 3, 4$;
- 4) 将子段序号 m_i 除以 20 得到归一化的持续时间特征向量 $m = [m_1 \ m_2 \ m_3 \ m_4]/20$;
- 5) 求 4 段能量相对衰减率, $n_i = (m_{i+1} - m_i), i = 1, 2, 3$;
- 6) 将新特征量 m 和 n 整合进原 2^N 维特征向量中形成 $2^N + m + n$ 维复合特征向量.

4 实验结果与分析

4.1 特征段选取规整结果与分析

本文系统采样频率为 10 kHz, 处理的对象是监测区域受到入侵扰动的位置点上采集到的以 2 s 为采样周期, 随时间变化的光纤振动信号. 每段光纤振动信号含有 20000 个采样点, 每个采样点间隔 0.1 ms. 利用本文提出的特征段选取方法选取到的有效扰动段如图 4(b) 所示. 双门限判定法的时间门限和阈值门限分别取经验值 700 个采样点和 1.5 V. 结合最大能量和最高信噪比从中选取出特征段并做规整, 得到只包含主要入侵信息的特征段, 如图 4(c). 特征段规整选取过程消除了由于时

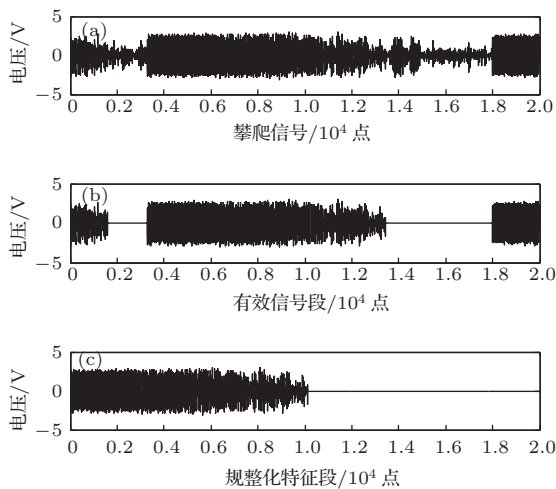


图 4 (a) 攀爬信号; (b) 有效信号段; (c) 规整化的特征段

域位置不同而导致的特征量差异, 减少了特征提取的运算数据量, 同时提高了分类准确率. 当信号为无扰动信号或者大风信号等无害信号时, 经过规整化特征段选取之后信号为零, 提取到的相应特征向量也为零向量, 据此可以对信号做一个预判.

4.2 复合特征选取结果与分析

分析图 5(a) 攀爬信号, 能量从 0 增加到 20% 耗时最短, 能量最集中, 几乎无衰减; 能量从 20% 增加到 60% 时, 耗时增加, 能量开始分散, 衰减速度加快; 能量增加到 80% 时, 耗时最长, 能量最分散, 衰减最快. 分析图 6(a) 剪网信号, 外界仅短暂触动光纤, 当扰动源脱离光纤时振动信号已经开始以相对稳定的速度衰减. 取总能量 80% 处作为特征段持续时间结束点是因为光纤振动信号尾端严重衰减, 含噪量大, 剩余的 20% 不作特征信息考虑.

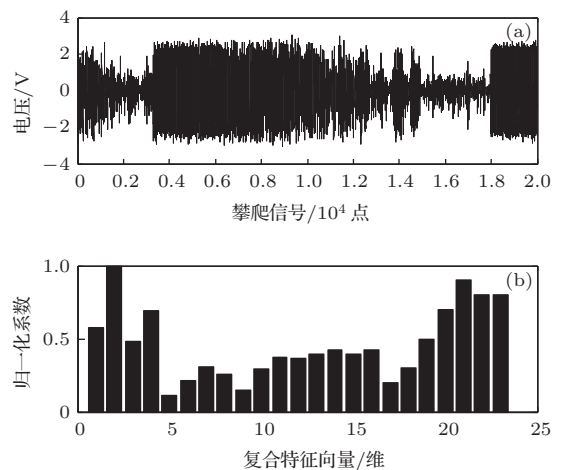


图 5 (a) 攀爬信号; (b) 复合 23 维特征向量

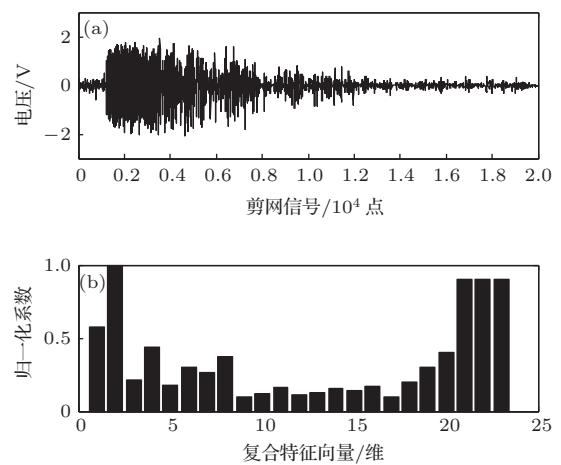


图 6 (a) 剪网信号; (b) 复合 23 维特征向量

综合考虑特征向量训练时间和分类准确率, 对光纤链路振动信号特征段作4层小波包分解, 计算16维能量谱. 而后按照本文提出的复合特征提取方法得到归一化的特征段持续时间特征量 m 和相对衰减率 n , 得到最终的23维特征向量.

图5(a)是攀爬信号, 5(b)是其经过规整化的特征段复合23维特征向量, 其中前16维是4层小波包能量谱, 17—20维是特征子段持续时间 m_i , 21—23维是相对衰减率 $(m_{i+1} - m_i)$. 后三维数值较小, 为了便于观察分析, 将 $1 - (m_{i+1} - m_i)$ 作为相对衰减率特征量. 从图5(b)中可以看到, 第17维与18维差值不大, 18维和19, 19维和20维差值变大, 即能量增加速度放慢, 衰减加快. 对应着21维比22, 23维数值都大. 同理, 图6是剪网信号及其对应的复合特征向量. 图6(b)的17—20维均匀增加, 衰减速度恒定. 比较图5(b)和图6(b)不难发现, 新增加的17—23维特征量在幅度上有明显差异. 经过实验验证, 当训练样本数为50时, 利用复合23维特征的分类准确率最高比原16维特征的分类准确率高出17.55%.

4.3 二叉树支持向量机分类结果

SVM是针对二分类问题提出的, 无法解决多类问题, 本文采用文献[15]提出的支持多分类的二叉树SVM. 首先使用SVM₁区分无害信号和入侵信号; 然后使用SVM₂进一步细分无害信号中的无扰动信号、大雨信号和大风信号; 最后使用SVM₃

区分入侵信号中的剪网信号和攀爬信号. 选取五类信号各50个分别使用16维特征和23维复合特征进行训练, 支持向量占向量总数百分比见表1. 可知, 处于分类边界的复合特征向量明显比小波包能量谱特征向量少, 绝大部分复合特征向量位于类内部, 使得分类准确率高.

表1 支持向量百分比(%)

	识别1		识别2		识别3	
	SVM ₁	SVM ₃	SVM ₁	SVM ₃	SVM ₁	SVM ₃
16维	58	34	38	34	11	34
23维	27.3	23.5	28.5	23.5	9.5	23.5

取剩余的150个无扰动信号, 50个大雨信号, 78个大风信号, 46个攀爬信号和57个剪网信号作分类测试. 分类结果如表2所示.

其中识别率1和识别率2分别是各类训练样本数为50时使用16维小波包能量谱特征向量和复合23维特征向量得到的识别准确率. 识别率3是各类训练样本数为100时使用复合23维特征向量得到的识别准确率. 由以上表格可知, 使用复合23维特征向量区分剪网信号和攀爬信号较使用16维小波包能量谱特征向量的准确率有不同程度的提高, 而对属于无害信号的无扰动信号、大雨信号和大风信号识别率无不良影响. 当训练样本增加到100个时, 各类平均识别率高达99.5967%. 因此, 本文提出的规整化特征段选取和复合特征提取方法切实可行.

表2 无害信号和有害信号识别结果(%)

信号类型	识别1			识别2			识别3			识别4		
	无扰动	剪网	攀爬	大雨	剪网	攀爬	大风	剪网	攀爬	无扰动	大雨	大风
识别率1	92	78.95	91.3	100	77.19	86.96	100	87.72	89.13	98	100	100
识别率2	92	91.23	95.65	100	94.74	93.48	100	91.23	95.65	98	100	100
识别率3	100	100	100	100	100	97.83	100	100	100	97.33	100	100

5 结论

本文提出的规整化特征段选取方法去除了主要扰动信号之间的干扰信息. 消除了由于时域位置不同而导致的小波包能量谱特征向量差异, 减少了特征提取的运算数据量同时提高了分类准确率. 本文提出的复合特征提取方法充分挖掘光纤链路振

动信号的特征, 综合利用了有效信号的持续时间和信号衰减率, 将其与小波包能量谱结合, 不但保证无害信号和入侵信号良好的区分度, 还能有效识别入侵信号的具体类型. 实验结果证明, 基于本文提出的特征段选取方法和复合特征提取的二叉树支持向量机能够很好地识别五类光纤链路振动信号, 并且在样本量较小的情况下也具有较高的分类准确率, 具有良好的实用价值和应用前景.

参考文献

- [1] Li C Y, Teng Y T, Li X J 2012 *Journal of Earthquake Engineering and Engineering Vibratio* **32** 16 (in Chinese) [李彩云, 滕云田, 李小军 2012 地震工程与工程振动 **32** 16]
- [2] Juarez J C, Maier E W, Choi K N, Taylor H F 2005 *Journal of Lightwave Technology* **23** 2081
- [3] Turitsyn S K, Boscolo S 2009 *Transparent Optical Network Azores*, June 28–July 2 2009 p1
- [4] Gu Yu, Li Qiang 2014 *Chin. Phys. B* **23** 044213
- [5] Zhao J, Ding J, Wan S R, Sun X H 2011 *Journal of Southeast University: Natural Science Edition* **41** 41 (in Chinese [赵杰, 丁吉, 万遂人, 孙小菡 2011 东南大学学报: 自然科学版 **41** 41]
- [6] Zhang X G 2000 *Acta Automatic Sinica* **26** 32 (in Chinese) [张学工 2000 自动化学报 **26** 32]
- [7] Wang F F, Zhang Y R 2012 *Chin. Phys. B* **21** 050204
- [8] Qu Z G, Feng H, Zeng Z M, Zhuge J C, Jin S J 2010 *Measurement* **43** 513
- [9] Faiz Bukhari Mohd Suah, Musa Ahmad, Mohd Nasir Taib 2003 *Sensors and Actuators B: Chemical* **90** 182
- [10] Qi H, Wang F B, Deng H 2013 *Acta Phys. Sin.* **62** 104301 (in Chinese) [祁浩, 王福豹, 邓宏 2013 物理学报 **62** 104301]
- [11] Xu B G, Song A G, Fei S M 2011 *Acta Electronica Sinica* **38** 1025 (in Chinese) [徐宝国, 宋爱国, 费树岷 2011 电子学报 **38** 1025]
- [12] Yang G A, Zhong B L, Huang R, Jia M P, Xu Y F 2001 *Journal of Vibration and Shock* **20** 25 (in Chinese) [杨国安, 钟秉林, 黄仁, 贾民平, 许云飞 2001 振动与冲击 **20** 25]
- [13] Gao Q, Yi S H, Jiang Z F, Zhao Y X, Xu Y F 2012 *Chin. Phys. B* **21** 064701
- [14] Mallat S 1989 *IEEE Trans. Pattern Analysis and machine Intelligence* **11** 674
- [15] Xia S Y, Pan H, Jin L Z 2009 *Computer Engineering and Applications* **45** 167 (in Chinese) [夏思宇, 潘泓, 金立左 2009 计算机工程与应用 **45** 167]

A regular composite feature extraction method for vibration signal pattern recognition in optical fiber link system*

Li Kai-Yan Zhao Xing-Qun Sun Xiao-Han Wan Sui-Ren[†]

(National Research Center for Optical Sensing/Communications Integrated Networking, Biological Science & medical Engineering, Southeast University, Nanjing 210096, China)

(Received 21 July 2014; revised manuscript received 27 August 2014)

Abstract

Phase optical time-domain reflectometer link monitoring system is a kind of sensor system which uses optical fiber as sensing medium. It can detect, recognize and locate invasive signals. The module of pattern recognition, which is one of the important parts of the system, can intelligently and instantly distinguish dangerous intrusions from safe disturbances. This paper proposes a regular composite feature extraction method which can be used for vibrational signal pattern recognition in an optical fiber monitoring system. This method applies an improved double-threshold method to detect the start-stop positions of the valid signal segments, and then combines the maximum energy and maximum signal noise ratio to select the main invasion feature segment within a sampling period. The composite eigenvectors, which are extracted by using the last time of the feature segment and wavelet packet energy spectrum, can be used by the support vector machine to recognize the signal patterns. Experiment results show that the accuracy rate of the pattern recognition has been improved significantly based on the proposed method in this paper.

Keywords: optical fiber vibration signal, wavelet package, composite feature extraction, pattern recognition

PACS: 43.60.Hj, 07.07.Df, 42.30.Sy

DOI: 10.7498/aps.64.054304

* Project support by the Frontier Research and Development Projects of Jiangsu Province, China (Grant Nos. BY2009155, BY2013073).

[†] Corresponding author. E-mail: srwan@seu.edu.cn