Recognition of Pipeline Safety Events Applied to Optical Fiber Pre-warning System

Qian Sun, Hao Feng, Jian Li and Shijiu Jin
State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin, China

Keywords: Pipeline Safety, Recognition, Feature Extraction, Wavelet Energy Spectrum, Support Vector Machine.

Abstract: Recognition of pipeline safety events is a key problem in the research of the optical fiber pre-warning system. In this paper a feature extraction method combined with wavelet energy spectrum (WES) and wavelet information entropy (WIE) is proposed. In order to avoid kernel function being dominated by trivial relevant or irrelevant features, a support vector machine (SVM) approach is also put forward based on the feature weighting, i.e. Feature Weighted SVM (FWSVM). The experiment shows that the method proposed in this paper is effective for recognition of the pipeline safety events and can be applied in optical fiber pre-warning system.

1 INTRODUCTION

By virtue of a great many advantages, pipelines have become the principal means of oil and gas transportation. However, pipeline leakage takes place due to some natural or artificial damages. Leakage accidents may cause loss of life and properties along with environmental pollution (Yang et al., 2004).

Pre-warning technique based on distributed optical fiber for the long oil and gas pipeline can give an alarm when the pipeline is threatened; therefore it is an important means to reduce the economic expense and to ensure people's security. Optical fibers are used to compose distributed optical fiber vibration signal sensor based on Mach-Zehnder interferometer principle (Zhou et al., 2007). How to accurately recognize the types of pipeline safety events is a key problem in the research of the optical fiber pre-warning system. In (Qu et al., 2006), a distributed optical fiber alarming system for the safety of oil and gas pipeline has been developed. Unfortunately, the system, with low intelligence, cannot tell which kind of activity causes the leaking accident, and the man on duty can not take corresponding action at once. In order to overcome the defect a recognition method based on feature weighted support vector machine is purposed in this paper.

The remainder of this paper is organized as follows. Considering that the non-stationary and random characteristics of the pipeline safety detection signals, a feature extraction method combined with WES and WIE is proposed in section 2. In section 3 a support vector machine approach based on the feature weighting, i.e. Feature Weighted SVM (FWSVM) is put forward in order to avoid kernel function being dominated easily by trivial relevant or irrelevant features. The experiment in section 4 analyzes the performance of the method proposed in this paper. Finally the conclusions and future research are given in Section 5.

2 FEATURE EXTRACTION

2.1 Wavelet Energy Spectrum

According with energy mode, the result of wavelet decomposition is called wavelet energy spectrum (Qu et al., 2008). The energy of the signal at different scales can be arranged as a feature vector. In other words the characteristic bands of the signal are extracted and can be used for classifier. The original sequence $x(n)$ can be expressed as

$$x(n)=\sum_{j=1}^{J}D_j(n)+A_J(n)=\sum_{j=1}^{J+1}D_j(n) \qquad (1)$$

Sun Q., Feng H., Li J. and Jin S..
Recognition of Pipeline Safety Events Applied to Optical Fiber Pre-warning System .
DOI: 10.5220/0004277300730077
In *Proceedings of the International Conference on Photonics, Optics and Laser Technology* (PHOTOPTICS-2013), pages 73-77
ISBN: 978-989-8565-44-0

$D_j(n)$ is the component of the signal $x(n)$ at different scales.

The signal energy in a certain scale is the sum of squares of wavelet coefficients in the same scale. Definition is shown as

$$E_j = \sum_{k=1}^{N} \left| D_j(k) \right|^2, j = 1, 2, \cdots, J \tag{2}$$

In the formula (2), N is the length of the sample points, $D_j(k), k = 1, 2, \cdots, N$ is the reconstructed wavelet coefficient at scale J. $E = [E_1, E_2, \cdots, E_J]$ is the WES of signal $x(n)$ of J scales.

2.2 Wavelet Information Entropy

In the information theory, entropy is used to represent the average information of the source output. It can provide signal useful information of potential and dynamic process. Its value is the measure of signal average uncertainly and complexity (EI-Zonkoly et al., 2011). The traditional entropy can express the uncertainty of a signal in the whole period, but cannot analyze the local nonlinear characteristic of the non-stationary signal. Therefore an time window is defined to calculate the WIE value, and to observe the change of the WIE following the sliding window. In this paper signal is decomposed to J scales, and the discrete wavelet coefficient at scale j is $D_j(k)$. The length of the short-time window is L, the sliding step is δ, and the calculation of the signal energy within a certain time window at each scale is

$$E_j = \sum_{k=1}^{L} \left| D_j(k) \right|^2 \tag{3}$$

The total energy within the time window is the sum of the energy component at each scale.

$$E_{tot} = \sum_{j=1}^{J+1} E_j \tag{4}$$

The signal relative energy of each scale within the time window is

$$p_j = \frac{E_j}{E_{tot}} \tag{5}$$

p_j is the energy distribution of different scale. The WIE of the signal within time window is defined as

$$S_{WT} = -\sum_{j=1}^{J+1} p_j \log_2(p_j) \tag{6}$$

Therefore, the change law of WIE along with time window sliding can be obtained.

3 FEATURE WEIGHTED SUPPORT VECTOR MACHINE

Support vector machine has excellent learning, classification and generalization abilities which use structural risk minimization instead of empirical risk minimization. The basic idea of SVM is to transform the input space into a high-dimensional feature space through non-liner transformation, and optimal separating hyperplane can be obtained in this feature space (Kurek et al., 2010). In a training sample set with classification mark $(x_i, y_i), x_i \in R^n$, $y_i \in \{+1, -1\}$, $i = 1, \cdots, l$. The optimal separating hyperplane can be created by the following optimization problem,

$$\left.\begin{aligned} &\max \sum_{i=1}^{l} a_i - \frac{1}{2} \sum_{i,j=1}^{l} a_i a_j y_i y_j k(x_i, x_j) \\ &s.t. \sum_{i=1}^{l} a_i y_i = 0, \\ &0 \le a_i \le C, i = 1, \cdots, l \end{aligned}\right\} \tag{7}$$

a_i is the Lagrange multiplier of x_i. The corresponding decision function is shown as

$$f(x) = sign\left(\sum_{i=1}^{l} a_i y_i k(x_i, x) + b \right) \tag{8}$$

$k(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$ is called kernel function which should be selected as integral operator of feature space.

SVM makes the nonlinear separable problem become linearly separable. According to the functional theory, as long as a kernel function satisfies mercer condition, it corresponds to a dot product in the transform space. Different algorithm formed by different kernel function. Common kernel functions are shown as follows

(1) Polynomial kernel function

$$k(x_i, x_j) = (x_i \cdot x_j + 1)^d, d = 1, 2, \cdots \quad (9)$$

(2) Gauss radial basis kernel function

$$k(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (10)$$

(3) Sigmoid kernel function

$$k(x_i, x_j) = \tanh(b(x_i \cdot x_j) + c), b > 0, c < 0 \quad (11)$$

According to some criteria, the features of the data set are given certain weights, this is called feature weighting. The key issue of feature weighting is to obtain the weight vector ω . Calculation of the weight vector is important for analyzing feature correlation. The basic idea of feature correlation analysis in classification learning process is to calculate a certain metric in order to quantify the correlation of feature and a given category. In this paper class separability criterion is used to compute weight vector. Calculating the divergence of the one-dimensional d_{ij} of each pair of classes, the criterion value of each feature is $c(k) = \min d_{ij}$. Characteristic which has bigger criterion value is well differentiable, in other words , that is a greater contribution to the classification. Assuming that each sample of data set is described by n features, the vector $C = (c(1), c(2), \cdots, c(n))$ describes the weight of each feature. The weight vector ω is constructed by vector C .

The support vector machine based on feature weighted kernel function is called feature weighted support vector machine. The calculation of the weighted kernel function can avoid kernel function being dominated easily by trivial relevant or irrelevant feature; therefore the better classification results can be obtained (Zhang et al., 2009). The weighted kernel function can be computed as

$$k_P(x_i, x_j) = k(x_i^T P, x_j^T P) \quad (12)$$

k is defined as the kernel function of $X \times X$, $X \subseteq R^n$, P is the transformation matrix, $P = diag(\omega)$, ω is the weight vector. Common weighted kernel functions are shown as follows

(1) Feature weighted polynomial kernel function

$$\begin{aligned} k_p(x_i, x_j) &= (x_i^T P \cdot x_j^T P + 1)^d \\ &= (x_i^T P P^T x_j + 1)^d \qquad d = 1, 2, \cdots \end{aligned} \quad (13)$$

(2) Feature weighted gauss radial basis kernel function

$$\begin{aligned} k_p(x_i, x_j) &= \exp(-\gamma \|x_i^T P - x_j^T P\|^2) \\ &= \exp(-\gamma ((x_i - x_j)^T P P^T (x_i - x_j))) \end{aligned} \quad (14)$$

(3) Feature weighted sigmoid kernel function

$$\begin{aligned} k_p(x_i, x_j) &= \tanh(b(x_i^T P \cdot x_j^T P) + c) \\ &= \tanh(b(x_i^T P P^T x_j) + c), b > 0, c < 0 \end{aligned} \quad (15)$$

The construction steps of feature weighted support vector machine shown as follows

Step1. Extract signal characteristics by WES and WIE described above. The normalized eigenvector is

$$T = [T_1, T_2, \cdots, T_n] \quad (16)$$

Step2. Calculate criteria value $c(k)$ of each characteristic. The weight vector ω and linear transformation matrix p are constructed as

$$\omega = \sqrt{C} = (\sqrt{c(1)}, \cdots, \sqrt{c(n)}), \quad (17)$$

$$p = diag(\omega) \quad (18)$$

Step3. Constitute the weighted kernel function with matrix p by formula (13) to (15) to replace standard SVM kernel function. In this paper FWSVM is constructed by the sigmoid weighted kernel function.

Step4. Recognize the pipeline safety events by FWSVM, and evaluate the performance of the classifier.

4 EXPERIMENTS

Three pipeline safety events including truck passing, excavator and digging were experimented based on optical fiber pre-warning system.

Optical fiber sensor was buried above oil products pipeline. The distance between pipeline and optical fiber is 500mm. The single mode optical fiber and semiconductor laser source were used in this system. The wavelength of laser source is 1550nm, and the power is 1mw. The optical interference signal is converted into electric signal by photodetector. Datas of these three events were collected through data acquisition module. Figure 2 shows the normalized voltage valve of three cases.

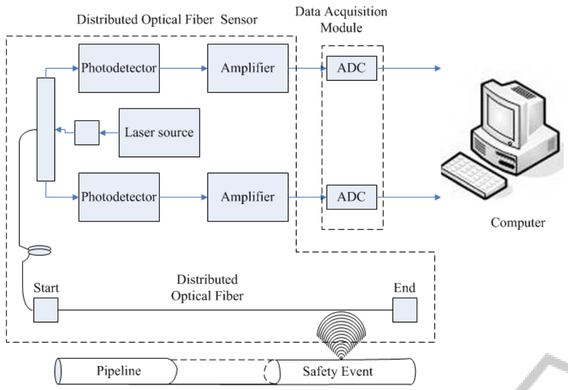

Figure 1: Optical fiber pre-warning system.

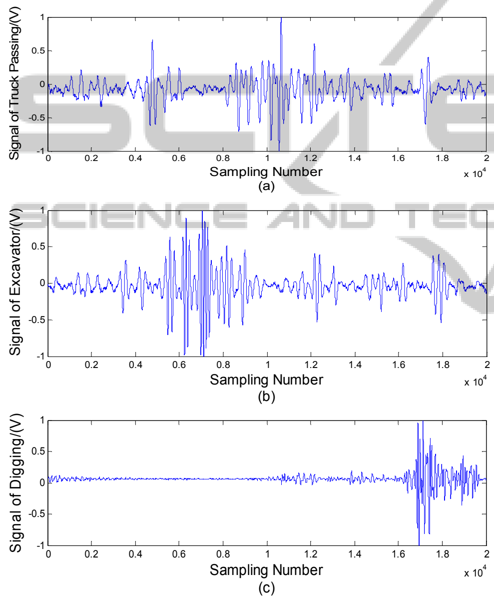

Figure 2: Signals of three cases.

Figure (a) is the signal generated by truck passing at the surface of the soil within 50cm of the pipeline on both sides. Figure (b) and figure (c) are the signal generated respectively by excavator and digging at the surface of soil just above the pipeline.

4.1 Signal Feature Extraction

In this paper, db6 wavelet function is used to decompose signals for 7 layers, and the wavelet energy spectrum of vibration signals caused by the three safety events is obtained by formula (2). The normalized energy of 8 frequency bands is shown in figure 3. Different energy on different frequency band provides a basis for the identification of the signal.

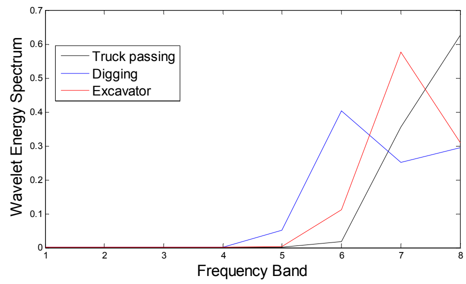

Figure 3: WES of three cases.

Pipeline safety events first make incursions into the soil near pipeline, and cause the vibration of soil. The wavelet information entropy increases when the optical fiber feels the vibration, and if the vibration is more intense the entropy will be greater. In this paper the length of time window $L = 100$, sliding step $\delta = 1$. The WIE is obtained by formula (6). Through experimental analysis the maximum WIE of the three pipeline safety events are significantly different, thus the maximum WIE can be also used as pattern recognition feature. The maximum WIE of three conditions are shown as table 1.

Table 1: Maximum WIE of pipeline safety events.

Pipeline events	Maximum WIE
Truck passing	0.1969
Digging	0.3112
Excavator	0.2230

4.2 Recognition of Pipeline Safety Events by FWSVM

WES is used as characteristic of each safety event. The energy of 8 frequency bands is marked as feature 1 to 8. The maximum of WIE is marked as feature 9. The three pipeline safety events truck passing, digging and excavator are labeled respectively as class 1, class 2 and class 3. FWSVM based on weighted sigmoid kernel function is used for recognizing the three events. 30 sets training sample of each event are obtained from field experiments for FWSVM classifier; besides, randomly select 20 sets sample of each event to test the classifier. Recognition results are shown in the following table 2 and table 3. In table 2 the accuracy is the percentage of the recognition correct number of each event. In table 3 the accuracy is the percentage of the total recognition correct number of three safety events.

Table 2: Recognition accuracy of different pipeline safety events.

Type of samples	Truck passing	Digging	Excavator
Number of samples	20	20	20
Truck passing	20	0	0
Digging	0	18	1
Excavator	0	2	19
Accuracy	100%	90%	95%

Table 3: Recognition accuracy of different method.

Pattern recognition method	Accuracy
WES-SVM	81.7%
WES-WIE-SVM	85%
WES-WIE-FWSVM	95%

As can be seen from table 2 and table 3, the FWSVM proposed in this paper is very effective for recognizing pipeline safety events. The recognition accuracy can reach 95%. Traditional SVM cannot satisfy the accuracy requirement because of its bad stability. Therefore FWSVM can be applied to the pattern recognition module of optical fiber pre-warning system.

5 CONCLUSIONS

This paper studies the recognition of pipeline safety events applied to optical fiber pre-warning system. In order to solve the typical recognition problem of pipeline safety events, the FWSVM is used in this paper. Firstly, wavelet energy spectrum and wavelet information entropy are used to extract features of signals, then the FWSVM is used for recognizing the three typical safety events. Through field experiment, the results show that FWSVM has high identification accuracy. The accuracy 95% is much higher than traditional SVM. The calculation of feature weighted kernel function can avoid kernel function being dominated by trivial relevant or irrelevant features. Therefore this method can satisfy the requirement of optical fiber pre-warning system. In the future work, recognition of more types of pipeline safety events still need further research and fieldexperiment.

REFERENCES

Yang, J., Wang, G. Z., 2004. Leak detection and location methods for gas transport pipeline, *Instruments in Chemical Industry*, Vol.31, No.2, pp.1-3.

Zhou, Y., Jin, S. J., Feng, H., 2007. Study on oil and gas pipeline leakage real-time inspection system based on distributed optical fiber, *Conference Committee of the 8th International Symposium On Measurement Technology and Intelligent Instruments*, pp.507-510.

Qu, Z. G., Jin, S. J., Zhou, Y., 2006. Study on the Distributed Optical Fiber Pre-warning System for the Safety of Oil and Gas Pipeline, *Piezoelectectrics &Acoustooptics*, Vol.28, No.6, pp.1-3.

Qu, W., Jia, X., Pei, S. B., Wu, J., 2008. Non-stationary signal noise suppression based on wavelet analysis, *Congress on Image and Signal Processing*, pp.303-306.

EI-Zonkoly, A. M., Desouki, H., 2011. Wavelet entropy based algorithm for fault detection and classification in FACTS compensated transmission line, *International Journal of Electrical Power & Energy Systems*, Vol.33, No.8, pp.1368-1374.

Kurek, J., Osowski, S., 2010. Support vector machine for fault diagnosis of the broken rotor bars of squirrel-cage induction motor, *Neral Computing & Applications,* Vol.19, No.4, pp.557-564.

Zhang, Y., Liu, X.D., Xie, F. D., Li, K. Q., 2009. Fault classifier of rotating machinery based on weighted support vector data description, *Expert Systems with Applications*, Vol.36, No.4, pp.7928-7932.