

Article

Recognition of a Phase-Sensitivity OTDR Sensing System Based on Morphologic Feature Extraction

Qian Sun, Hao Feng *, Xueying Yan and Zhoumo Zeng

Tianjin University, State Key Laboratory of Precision Measurement Technology & Instruments, 92 Weijin Road, Nankai District, Tianjin 300072, China; E-Mails: sunqiansohu@126.com (Q.S.); fightingxy@126.com (X.Y.); zhonzeng@tju.edu.cn (Z.Z.)

* Author to whom correspondence should be addressed; E-Mail: fenghao@tju.edu.cn; Tel.: +86-22-2740-2366.

Academic Editor: Lorenzo Pavesi

Received: 8 May 2015 / Accepted: 19 June 2015 / Published: 29 June 2015

Abstract: This paper proposes a novel feature extraction method for intrusion event recognition within a phase-sensitive optical time-domain reflectometer (Φ -OTDR) sensing system. Feature extraction of time domain signals in these systems is time-consuming and may lead to inaccuracies due to noise disturbances. The recognition accuracy and speed of current systems cannot meet the requirements of Φ -OTDR online vibration monitoring systems. In the method proposed in this paper, the time-space domain signal is used for feature extraction instead of the time domain signal. Feature vectors are obtained from morphologic features of time-space domain signals. A scatter matrix is calculated for the feature selection. Experiments show that the feature extraction method proposed in this paper can greatly improve recognition accuracies, with a lower computation time than traditional methods, *i.e.*, a recognition accuracy of 97.8% can be achieved with a recognition time of below 1 s, making it is very suitable for Φ -OTDR system online vibration monitoring.

Keywords: Φ -OTDR; morphology; feature extraction; intrusion event recognition

1. Introduction

Optical fiber vibration sensing systems based on phase-sensitive optical time-domain reflectometer (Φ -OTDR) devices can detect and locate vibration signals by measuring the backscatter light scattered across the entire optical spectrum. A Φ -OTDR uses a single mode fiber for optical transmission and sensing, and can be used for real-time monitoring and accurate positioning across long distances [1,2]. It has been used for monitoring the health of engineering structures, optical fiber perimeter protection and gas or oil pipeline safety pre-warning systems [3–6]. Accurate recognition of different vibration signals is crucial in a Φ -OTDR pre-warning system. False positives will be an inefficient use of resources and more seriously, delays in processing time may threaten people's lives and property, therefore, there is a lot of attention is being given in Φ -OTDR research to techniques to accurately recognize the event type, ensure warnings are given in sufficient time and reduce the false positive rate.

Due to the nonlinear, dynamic nature of the signal acquired by a Φ -OTDR vibration sensing system, a location scheme based on the wavelet packet transform (WPT) is proposed to reduce the number of false alarms [7]. Previous studies have used the short-time Fourier transform (STFT) and the continuous wavelet transform (CWT) for recognition of Φ -OTDR systems [8]. These traditional methods focus on finding and locating the intrusion events. If the features of the time domain signal can be extracted after the event has been located, different event types can be classified based on these signal features. However, this method is time consuming, due to the requirement to firstly pinpoint the location in the recognition process. If multiple events occur simultaneously, the recognition time will increase significantly. The intrusion signal of an Φ -OTDR system is not a single point in the space domain; it occurs across a range, since the attenuation will continue for a period along the optic fiber. Within the attenuation range, each scattered light signal contains the vibration response, but since the initial phase of each interference signal is different, the amplitudes of the vibration responses are different within the attenuation range, as shown in Figure 1. The peaks are due to backscattering when an intrusion event occurs. However, the maximum point may not be the location of the intrusion. If the event location is identified as the point with the maximum light intensity, then an error in location will occur.

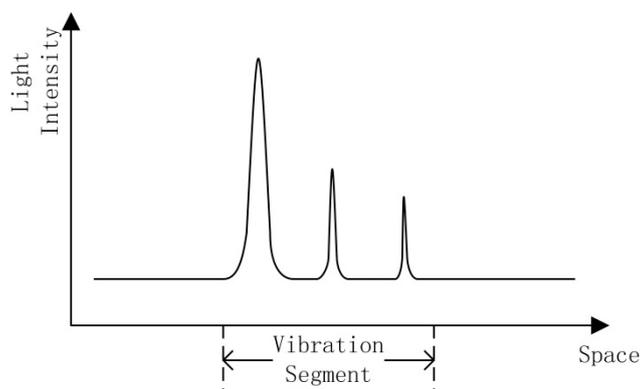


Figure 1. Scattering curve of intrusion signal.

Once there is an error in locating the event, the time domain signal of that location will not be the real signal corresponding to the intrusion event. Therefore the features of the event cannot be extracted

correctly, which will lead to a recognition error. This conventional method is inadequate for a Φ -OTDR system and thus a highly efficient and precise method is necessary for recognition of intrusion events. In this paper, after a brief presentation of the Φ -OTDR signal characteristics in Section 2, a method based on time-space domain signals is proposed. The algorithm is evaluated in this paper using the criteria of recognition accuracy and speed. The recognition accuracy is the percentage of correct recognition events of all events. Recognition speed is the time calculated from feature extraction to recognition. The method proposed in this paper improves both the speed and accuracy of the recognition method, and is therefore a very suitable method to recognize different events within a Φ -OTDR sensing system. In Section 3, a feature extraction method based on morphology is proposed. The classifier design and performance evaluation are demonstrated in Section 4, and the final section provides the conclusions.

2. Φ -OTDR Signal Characteristics

The Φ -OTDR system has unique signal characteristics compared with other optic fiber sensing systems. Rayleigh Scattering (RS) light traveling within a fiber is phase modulated by vibrations that are applied to the fiber to acquire a RS curve for the pulse duration. The time-space domain signal is acquired by clustering all RS curves. It can reflect the characteristics in the space and time domain simultaneously. The propagation distance of vibrations in the space domain reflects the energy of the intrusion signal and the signal characteristics in the time domain reflect the duration of the intrusion event. A simplified construction of an Φ -OTDR system is shown in Figure 2. It can be assumed that that $\epsilon(t)$ is the optical pulse and L is the length of optical fiber. The vibration occurs at position z_0 and the optical phase is therefore modulated by the vibration, with variation Φ .

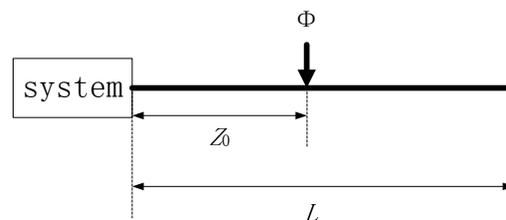


Figure 2. Simplified construction of a Φ -OTDR system.

The Rayleigh scattering field intensity can be expressed as follows:

$$E(\Phi, s, n) = E_1(s, n) + E_2(s, n) \exp(2i\Phi) \quad (1)$$

$$E_1 = \int_0^{z_0} \epsilon\left(s - \frac{2z}{c}\right) r(z) \exp(-2i\beta z) dz \quad (2)$$

$$E_2 = \int_{z_0}^L \epsilon\left(s - \frac{2z}{c}\right) r(z) \exp(-2i\beta z) dz \quad (3)$$

where n is the number of optical pulses. If the optical pulse period is T , $s \in (nT, nL/c + nT)$. s can also be regarded as the abscissa of any point in the optical fiber. c is the velocity of light, β is the transmission constant, $2z/c$ is the optical pulse delay and $2\beta z$ is the optical phase delay. E_1 and E_2 are the Rayleigh scattering field intensities before and after the intrusion event, respectively.

The interference intensity can be expressed as:

$$I(\phi, s, n) = |E(\phi, s, n)|^2 = I_1(s, n) + I_2(s, n) + 2[I_1(s, n)I_2(s, n)]^{1/2} \cos(2\phi + \phi_0) \quad (4)$$

$$\phi_0 = \arg(E_2) - \arg(E_1) \quad (5)$$

$$I_i = |E_i|^2, i = 1, 2.$$

where ϕ_0 is the initial phase of the Φ -OTDR system. The value of ϕ in Equation (4) not only changes in the time domain but also in the spatial domain. The attenuation of ϕ in the spatial domain can be expressed as:

$$\phi(s, d) = \begin{cases} \Delta\delta e^{-\alpha(s_0-s)} & s < s_0 \\ \Delta\delta e^{-\alpha(s-(s_0+d))} & s > (s_0 + d) \\ \Delta\delta & s_0 \leq s \leq (s_0 + d) \end{cases} \quad (6)$$

In Equation (6), α is the phase attenuation coefficient, s_0 is the position of the intrusion event and d is the range of the intrusion force. $\Delta\delta$ is the phase variation in the time domain. The schematic diagram is shown in Figure 3.

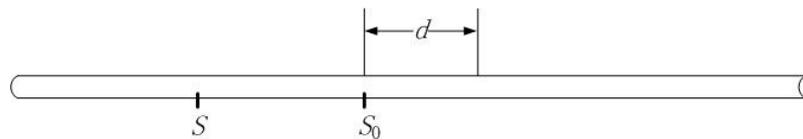


Figure 3. Parameters of Equation (6).

The length and refractive index of the optical fiber will be modified when there is an intrusion force on the optical fiber. $\Delta\delta$ can be calculated as follows:

$$\Delta\delta = \beta\Delta L + L\Delta\beta = \beta L \frac{\Delta L}{L} + L \frac{\partial\beta}{\partial n} \Delta n \quad (7)$$

where β is the propagation constant of the optical wave, L is the length of the optical fiber and n is the refractive index. The $\beta L(\Delta L/L)$ term is the effect due to strain, which can be regarded as being due to the phase variation along with the optical length variation. The $L(\partial\beta/\partial n)\Delta n$ term is the photoelastic effect, due to the phase variation along with the refractive index variation. ΔL and Δn are determined by the characteristics of the intrusion force and edatope, which are calculated in detail in the literature [9].

A Φ -OTDR pipeline safety pre-warning system was used as the case study in this paper. Common events which threaten pipeline safety are a vehicle passing over the pipeline, soil digging above the pipeline, and walking over the pipeline. Experiments were done using the Dagang-Zaozhuang product oil pipeline. The sensing cable used is a GYTA six-core single mode communication optical fiber, which was buried 30 cm above the pipeline and 1.5 m underground. Twenty km of optical fiber was used in the experiment. The laser pulse repetition frequency is 500 Hz and the sampling frequency is 50 MHz. Two seconds of data are used to constitute an image. For this particular application, the recognition accuracy must be as high as possible.

The signal calculated by Equation (4) and the experimentally measured signals of three pipeline safety events are shown in Figures 4–6. Due to the long distance of the Φ -OTDR system monitoring, the

vibration range is relatively small. An intrusion event is usually a fleck in the time-space domain image. The intrusion event occurs continuously in practice, so there are not only one event region in general in time-space signal image. Usually, walking frequency is three steps within 2 s, the interval between the wheels of the vehicle pressing the deceleration strip is about 0.25 s, and the interval due to digging is much larger than 2 s in the experiment. There are three event regions A1, A2, A3 in the experimental walking event, A2, A3 are repetition of A1. There are only one event region B1 in the experimental digging event, and three event regions C1, C2, C3 in the vehicle passing event. Therefore events are calculated only once for simulation signal, marked as region A, region B and region C in Figures 4–6. In Figures 4–6, the graphs in (a) are the simulation signals calculated by Equation (4) and the graphs in (b) are the experimentally measured signals. In order to illustrate the problem more clearly, the simulation signals are appropriately amplified. In the simulation process, for calculation convenience, many environmental factors were ignored. Soil is idealized as elastic half-space [10], and intrusion events are idealized as single frequency signals. Although there are some differences between the simulation and experimental signal images, they are basically the same.

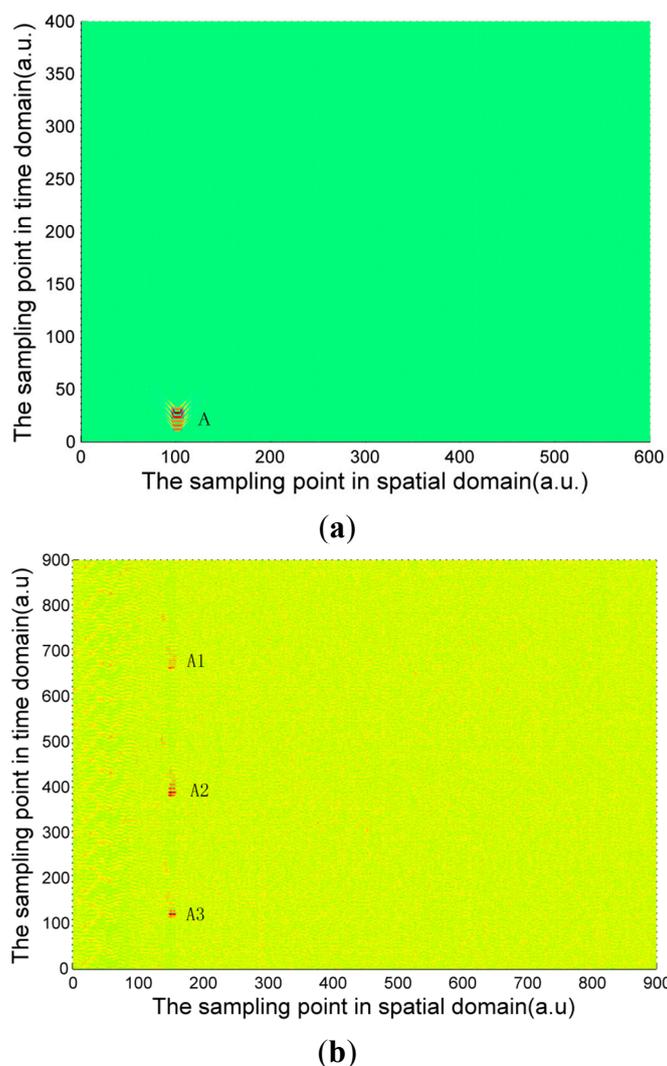


Figure 4. (a) Calculated walking signal; (b) Experimentally measured walking signal.

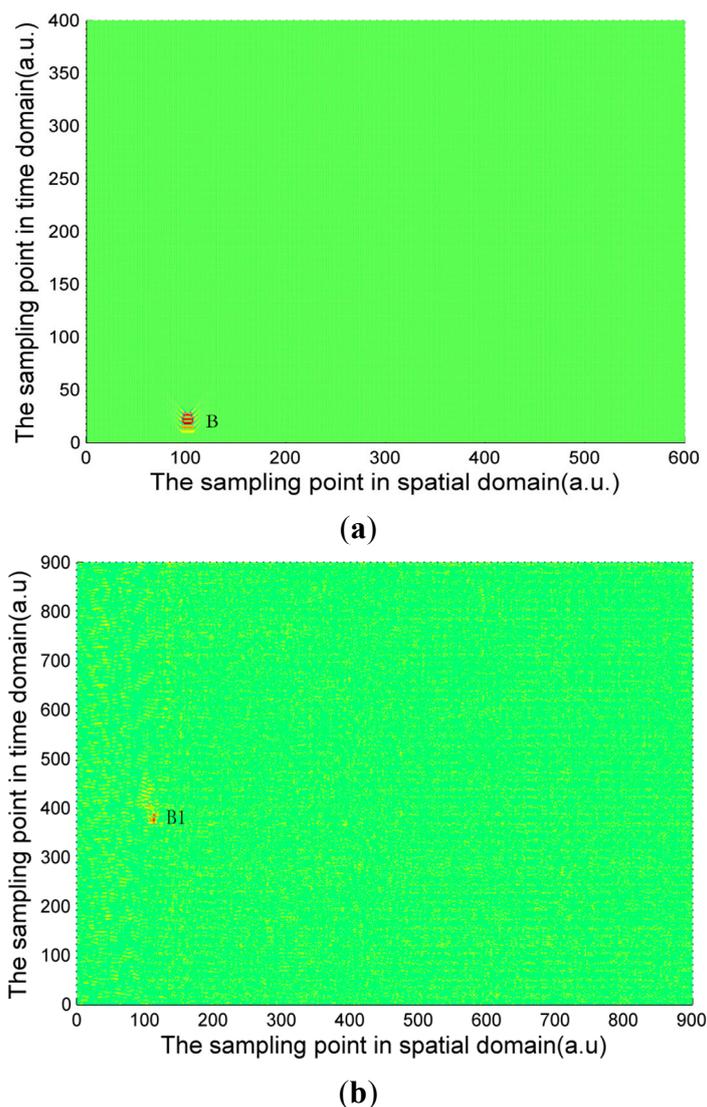


Figure 5. (a) Calculated digging signal (b) Experimentally measured digging signal.

The characteristic differences of the three types of events are obvious both in the simulation and experimental signals images. When a vehicle passes over the optical fiber, a vibration signal is produced by the wheels pressing on the deceleration strip on the ground. The duration of this event is very short but the energy is large, so the force can be considered ideally as an instantaneous force. Since the propagation distance is relatively high, a “V”-like shape will be formed. The forces due to walking and digging change relatively slowly with time and energy; the propagation distance is short. The energy and force variation of walking and digging are both different, so the size and the shape of the regions in the image are different. Based on the different image characteristics of each of the three different events, image processing technology can be used for image segmentation, and then a feature extraction method based on morphology proposed in this paper is used for event classification.

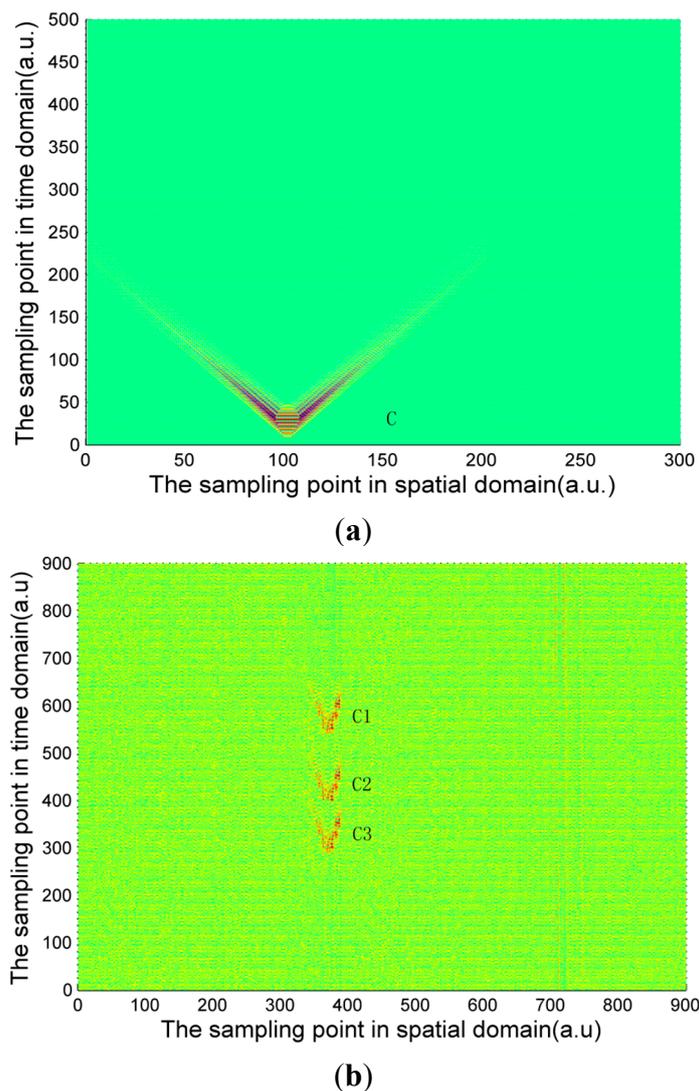


Figure 6. (a) Calculated signal of vehicle passing; (b) Experimentally measured signal of vehicle passing.

3. Feature Extraction Based on Morphology

3.1. Preprocessing

Before feature extraction, segmentation is required to separate the event region from the background. The grey level histograms of the three types of signals are shown in Figure 7. As can be seen from the grey level histograms, there is a large difference between the event region and the background, so threshold segmentation is an appropriate method [11].

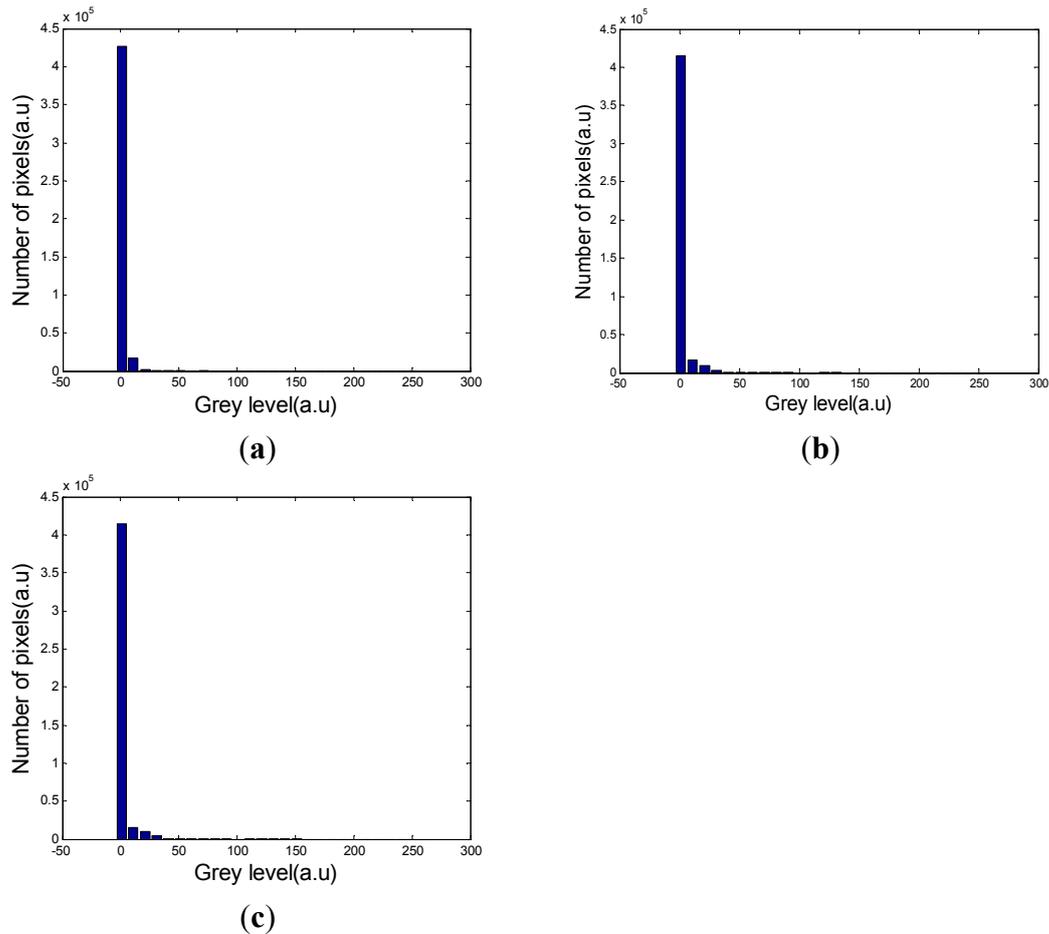


Figure 7. (a) Grey level histogram of vehicle passing; (b) Grey level histogram of digging. (c) Grey level histogram of walking.

An artificial threshold will lead to errors, because a high threshold will contain noise and a low threshold will discard information of the characteristics in the region. The Otsu threshold is calculated by the grey level histogram [12]. The main principle of this method is to find the maximum between-class variance threshold. The Otsu threshold regards the event region and the background as two classes. Firstly an initial threshold is set, and classification between the two classes can be calculated using Equation (8). Then a threshold is found which will maximize this formula:

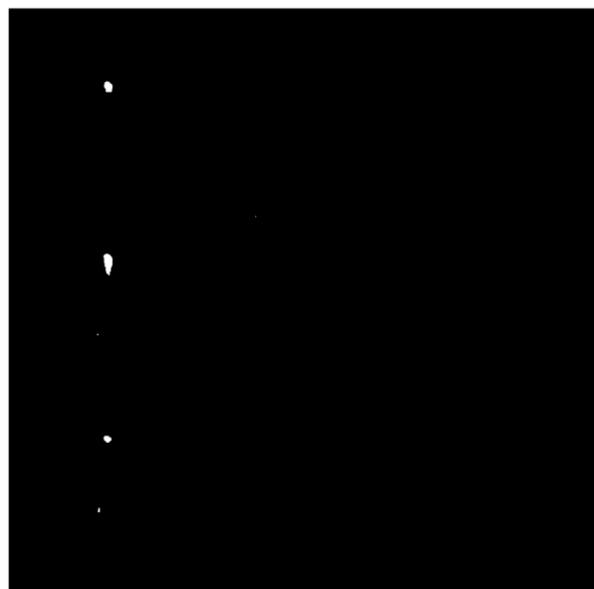
$$\eta = \frac{\sigma_B^2}{\sigma_G^2} \quad (8)$$

where σ_B^2 is the between-class variance and σ_G^2 is the global variance, which is a constant. The best threshold will maximize σ_B^2 :

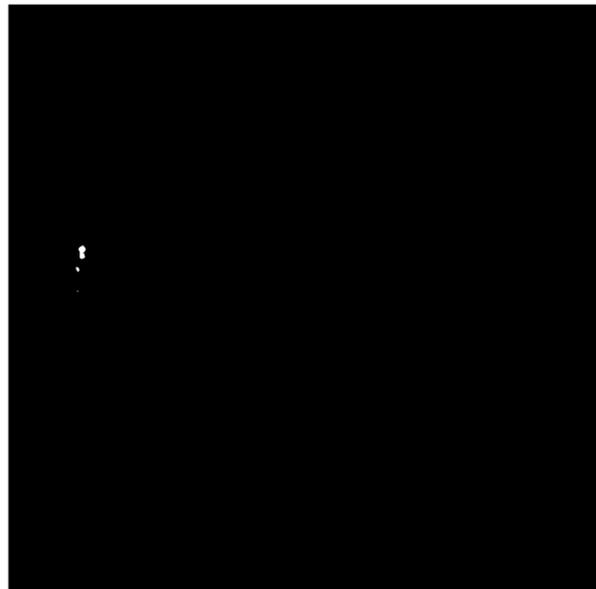
$$\sigma_B^2(t_{best}) = \max \sigma_B^2(t) \quad (9)$$

t_{best} is the optimal threshold. As can be seen from Figure 8, the segmentation method is effective and satisfactory for our purposes. There are some noise points appearing in the images in Figure 8, which are one order of magnitude smaller than the event region, and are widely dispersed. Median filtering is an effective way to eliminate these noise points [13]. Also, there are some holes in the images that

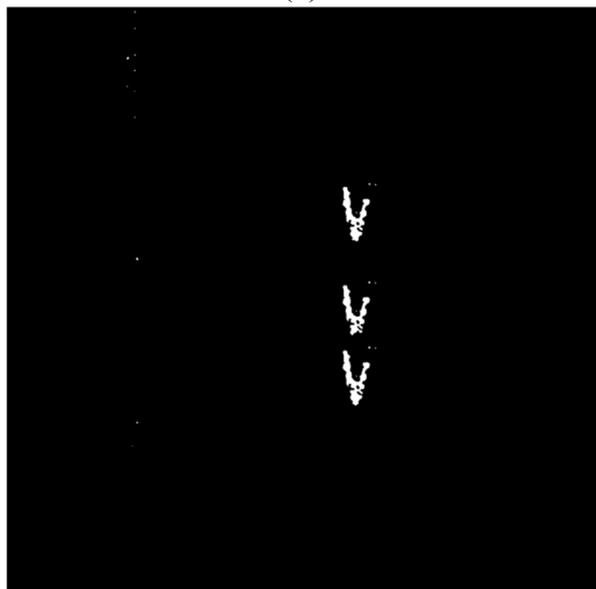
have the potential to seriously impact the effects of feature extraction, so an image dilation strategy was used for filling these holes.



(a)



(b)



(c)

Figure 8. (a) Image of walking after threshold segmentation; (b) Image of digging after threshold segmentation; (c) Image of vehicle passing after threshold segmentation.

Figure 9 shows each event region labeled in different colors. If there is more than one region in an image, the pixels of different regions will be labeled with different values; therefore different colors can be seen within the image.

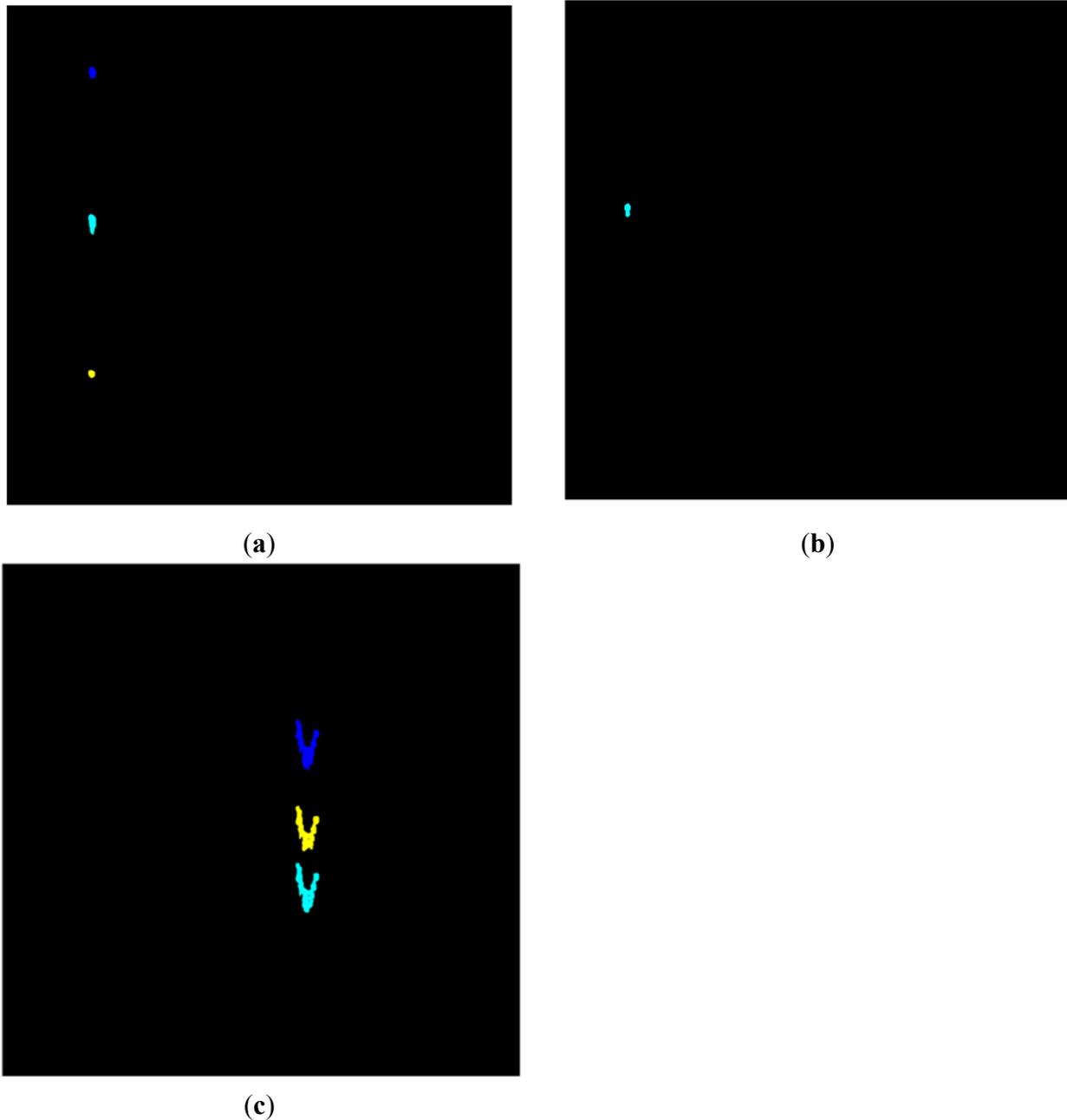


Figure 9. (a) Image of walking showing events which are labeled in different colors; (b) Image of digging showing events which are labeled in different colors; (c) Image of vehicle passing showing events which are labeled in different colors.

3.2. Feature Extraction

In Figure 9, every marked area is an event region. One region means an event happening once. In our experiment, there are three walking event regions, three vehicle passing regions and one digging event region every two seconds. If the signal is acquired for only one second, there may be only one walking event region, two vehicle passing regions or one digging region. There will be additional event regions within each image when the signal is acquired for a longer time. In order to uniformly process the image, every region in each image is used as a research object instead of the whole image.

Firstly, different events have different amplitudes because they exert different forces on the ground, so the amplitude “Amp” of the original signal in a different labeled region is used as a classification feature. Secondly, at the same frequency, the interval of two regions for one event in one image is different from the other events. If one region has two neighboring regions, the shortest interval is chosen as the feature value. The intervals between regions are shown in Figure 10.

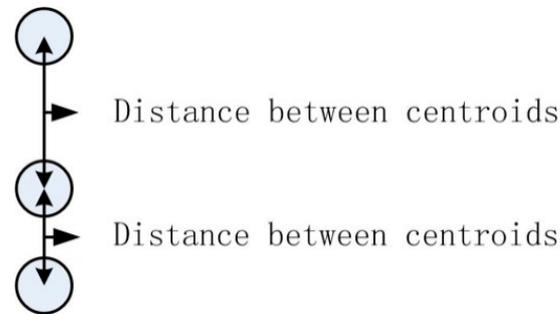


Figure 10. Intervals of centroids.

The centroid of a different region is first located as $C = [c_1, c_2, \dots, c_n]$, where n is the number of regions in the image. In order to avoid feature extraction errors when multiple events are happening at the same time, the abscissas of the centroids of regions are firstly compared. Two regions are classed as belonging to different events if their centroid abscissas have a sufficient distance from each other. When calculating the intervals between regions, the method will skip a region if its abscissa is far from the previous region. The distance threshold for the abscissa is set to be the minimum distance of abscissas between non-overlapping regions from different events. The distance from the centroid of one region to the centroid of its neighbor in the same event is then calculated and the shortest distance is found using Equation (10):

$$d_{\min 1} = \min(|c_i - c_{i-1}|, |c_i - c_{i+1}|) \quad i \in (2, \dots, n-1) \quad (10)$$

If there is only one region in an image, the interval feature is set to be a large value. Different events also have different region shapes, for reasons that were described in Section 1. The region where a vehicle is passing is similar to the letter “V”, due to its large energy but short duration. Other events have a lower energy and a relatively long duration time, so many “V” shapes are observed in the region in the time domain and a round shape is formed in the image. The size and roundness of the event regions for walking and digging are both different due to their different energies. The roundness of the region can be used as the shape feature. The process of obtaining the shape of the region is shown in Figure 11. Firstly, the boundary of the region b_i must be acquired. In this paper, the boundary can be found using a boundary tracking method [14] which is a simple and efficient method. The distance from the centroid to every point of the boundary is calculated using Equation (11):

$$D_{ik} = |b_{ik} - c_i|, \quad k \in (1, \dots, K) \quad (11)$$

In Equation (11), K is the number of boundary points used. The maximum distance $D_{i\max}$ and the minimum distance $D_{i\min}$ can be easily obtained. The shape feature is calculated using Equation (12):

$$S_i = |D_{i\max} - D_{i\min}| \quad (12)$$

A comprehensive description of the region is necessary for recognition; some common region descriptors which can describe differences between events are also used as features [15]. Seven region descriptors are chosen as features in this paper. All features are listed in Table 1.

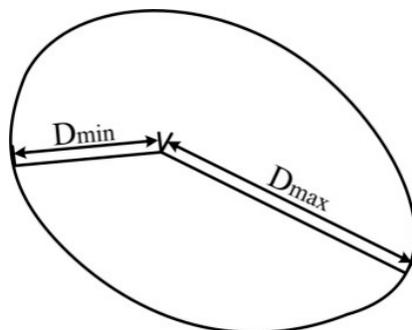


Figure 11. Shape of region.

Table 1. Features of an event region.

Feature	Definition
Amp	Amplitude of time-space domain signal
Cen	Minimum interval between regions
Shape	Roundness of the region
Con	Pixel number of the convex hull
Area	Pixel number of the region
Ecc	Eccentricity of ellipse which has the same second moment as the event region
Maj	Length of the long axis of the ellipse which has the same second moment as the event region
Min	Length of the short axis of the ellipse which has the same second moment as the event region
Equ	Diameter of a circle which has the same total area as the event region
Eul	Remaining number of objects, excluding holes in the image region

Due to limiting conditions, a large number of samples of the intrusion event cannot be acquired in a relatively short period of time. Therefore the Φ -OTDR pre-warning system recognition process is a small sample size problem. Fifty samples of each type event were acquired for a limited time from the field environment. One sample represents one event region. There are 150 samples of the three events used for recognition. Ten features of each region mentioned in Table 1 were extracted, and the mean value of each feature is shown in Table 2. There is only one digging region; and the region interval between the other two events is much less than digging. In order to ensure strong separability of features, the centroid interval feature “Cen” of digging is set to 1000. The distance threshold of the centroid abscissas is set by analysis of the region area. In this paper, the distance threshold is set to 25 sample points. If the abscissa distance from the centroid of one region to its neighbor is farther than the distance threshold, the regions are determined to belong to different events. The feature extraction method of centroid interval will therefore skip the region and continue the calculations on the next region.

Table 2. Mean value of 10 features of three events.

Event Type	Con	Cen	Area	Amp	Ecc
Walking	279	271.373	245	10.2871	0.7574
Digging	224	1000	205	12.6634	0.8751
Vehicle passing	2413	118.0326	1461	8.5084	0.9118
	Shape	Maj	Min	Equ	Eul
Walking	8.1041	24.4843	12.9812	17.3192	1
Digging	14.6010	25.9145	10.6407	16.1559	1
Vehicle passing	46.7452	84.0348	40.5031	43.1287	-1

3.3. Feature Selection

Comprehensive descriptions can sometimes use too many features, which not only makes the feature extraction method more complex, but also increases correlation between features, resulting in classification errors. A feature selection method to reduce the number of dimensions will increase the recognition speed and accuracy. A scatter matrix calculation method which can evaluate the discrimination of every feature is used in this paper for feature selection. This method will not change the feature attributes and has a very short calculation time. The main concept of the scatter matrix calculation method is to obtain the relationship of the eigenvector distribution in one-dimensional space and select the features with the smallest variance in class and largest distance between classes [16,17]. Firstly, the scatter matrixes of each class and between classes of 10 features of three events are calculated, and then the criterion value of every feature is computed using Equation (13):

$$J = tr\{S_{\omega}^{-1}S_m\} \quad (13)$$

where S_{ω} is the scatter matrix in the class and S_m is the hybrid scatter matrix calculated using Equation (14):

$$S_m = S_{\omega} + S_b \quad (14)$$

where S_b is the scatter matrix between classes.

If samples of every class are clustered around the mean value and there is complete separation between classes, the criterion value is highly suitable. The criterion value of every feature is calculated using Equation (13) and features with a larger value are selected as recognition samples. The result of feature selection is shown in Figure 12.

In Figure 12, 10 features are numbered according to Table 2. The first four largest value features are “Cen”, “Shape”, “Amp” and “Area”. The relevance of features is calculated in Equation (15). x_{nk} is the number k feature of number n sample. $n = 1, 2, \dots, N$, $k = 1, 2, \dots, m$. In this paper $N = 150$:

$$\rho_{ij} = \frac{\sum_{n=1}^N x_{ni}x_{nj}}{\sqrt{\sum_{n=1}^N x_{ni}^2 \sum_{n=1}^N x_{nj}^2}} \quad (15)$$

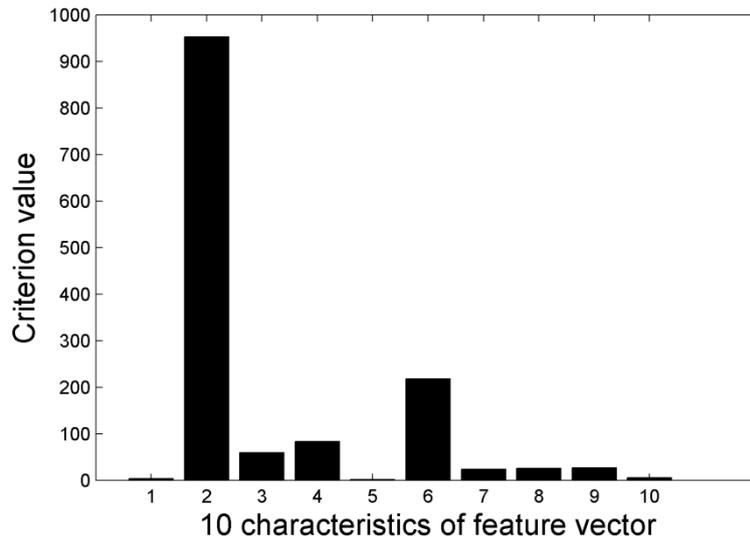


Figure 12. Criterion values of 10 characteristics.

First the feature with maximal criterion value “Cen” is selected as x_{i_1} . Second, the relevance ρ_{ij} between x_{i_1} and other $m-1$ features is calculated by Equation (15). Finally, the feature which satisfy the following equation will be selected as x_{i_2} :

$$i_2 = \arg \max \{ \alpha_1 J(j) - \alpha_2 |\rho_{i_1 j}| \}, \quad j \neq i_1 \quad (16)$$

α_1 α_2 are the weighting coefficients. In this paper, α_1 α_2 are set to 0.5 respectively. J is the criterion value. The other features x_{i_k} , $k = 3, \dots, l$, can be calculated by Equation (17):

$$i_k = \arg \max \{ \alpha_1 J(j) - \frac{\alpha_2}{k-1} \sum_{r=1}^{k-1} |\rho_{i_r j}| \}, \quad j \neq i_r, r = 1, 2, \dots, k-1 \quad (17)$$

The features relevance calculated by Equation (17) is shown in Table 3. x_{i_1} is the feature “Cen”.

Table 3. Relevance of features.

x_{i_k}	Feature	Value of i_k
x_{i_2}	Shape	159.8173
x_{i_3}	Area	71.8518
x_{i_4}	Amp	71.8435
x_{i_5}	Maj	11.8
x_{i_6}	Min	11.7677
x_{i_7}	Equ	11.7333
x_{i_8}	Eul	7.7
x_{i_9}	Con	7.6
$x_{i_{10}}$	Ecc	3.6

It can be seen from Table 3 the minimum relevance four features just are the features which have max scatter matrix value J . Therefore the first four features “Cen”, “Shape”, “Area”, “Amp” are finally selected as the feature vector. The features distribution figure of the three intrusion events is shown in Figure 13.

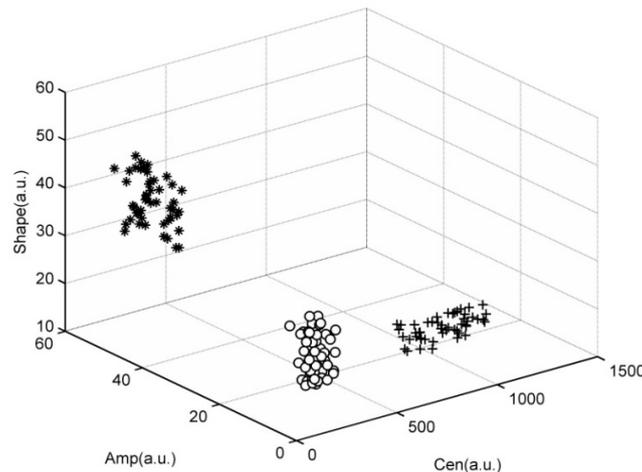


Figure 13. Feature distribution of three events.

In Figure 13, “o” represents the walking intrusion events, “+” represent the digging events, and “*” represents events due to vehicles passing. As we can see from Figure 13, there is a high dispersity between different intrusion events, and sample points from each type of event are located closely together. Features distribution can verify the separability of features extracted by morphologic method. The effective features can not only improve the recognition accuracy, but also can simplify the design of the classifier.

4. Experiments

The event samples acquired using the Φ -OTDR pre-warning system are multiclass and the number of samples is small, so the Relevance Vector Machine (RVM) technique is used in this paper. RVM is a machine learning method based on the Bayesian framework. It is sparser than the Support Vector Machine (SVM) technique; hence it has a shorter recognition time and a higher accuracy [18]. This makes it more suitable for use for recognition for an optical fiber pre-warning system [19–21]. The Gauss kernel function is used in this paper because of its widely usability and excellent performance. The kernel parameter is usually set between 0 and 1 [22]. Through experimental analysis, it was found that the highest accuracy was obtained when the parameter of the kernel function was set to 0.6. The RVM technique was designed for two-class classification problems; therefore a one-to-one multi-category method is used for recognition of the three events [23]. Each classifier recognizes two classes, so there are three classifiers for recognition of the three intrusion events. Each classifier is trained with two events and the training process is shown in Figure 14.

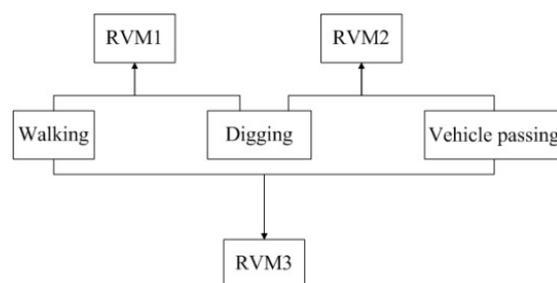


Figure 14. Training process of the three classifiers.

During the recognition process, an unknown event is recognized by the three classifiers. If two classifiers have the same results, the event will belong to that class. The recognition process is shown in Figure 15.

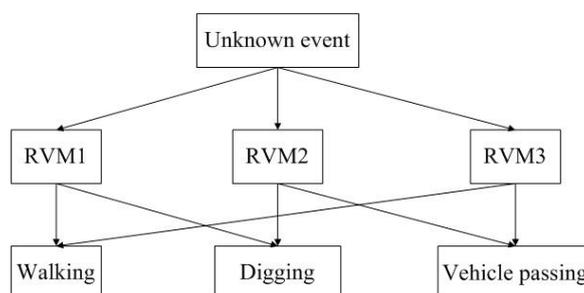


Figure 15. Recognition process of three classifiers.

The performance of the classifiers must be evaluated before recognition. Classifiers are evaluated by a five-fold cross validation. 20 samples from each set of 100 samples of each event are selected for training the classifiers. The accuracy of five-fold cross validation is shown in Figure 16.

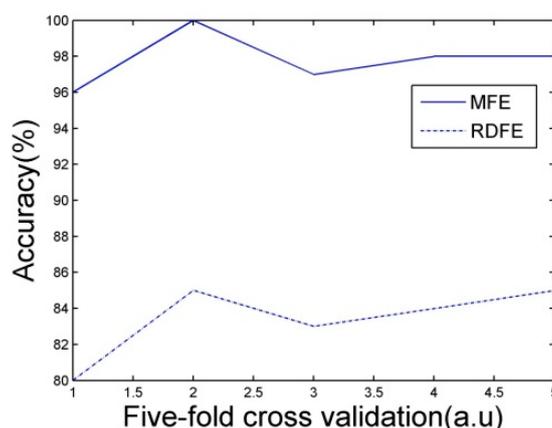


Figure 16. Accuracy of five-fold cross validation.

In Figure 16, the recognition accuracy of Morphological Feature Extraction (MFE) method with feature selection proposed in this paper is much higher than Region Descriptor Feature Extraction (RDFE) without feature selection. The recognition accuracy can reach over 95%. A single cross validation result is shown in Table 4.

Table 4. A single cross validation result.

Sample Model	Walking	Digging	Vehicle Passing
Sample size	20	20	20
Walking	20	1	0
Digging	0	19	0
Vehicle passing	0	0	20
Accuracy (%)	100	95	100

As seen from Table 4, the recognition accuracy of each event also reaches over 95%. This result contains the case where multiple events happen at the same time. There is a recognition error for the digging since the walking signal is mixed with the digging signal due to environmental disturbance. The category of the test sample is known in the test processing, so the classifiers sort the signal into the wrong class. This mistake can be avoided when this method is used in online monitoring, since the class that the signal belongs to is unknown. Traditionally, an event was located in the space domain by a moving average method, and then the event signal was extracted in the time domain. The wavelet method was used for extracting the features [24]. RVM was also used as classifier, and five-fold cross validation was used for evaluating the recognition accuracy. The results of the one dimension signal were shown in Table 4. Identification of the location of the event in the time domain is time-consuming, and since the position accuracy is easily affected by noises, the recognition performance is unsatisfactory. The mean accuracy and recognition time of different methods are shown also in Table 5. When Φ -OTDR is used for safety online monitoring, the algorithm computation time is the recognition time *i.e.*, the time calculated from feature extraction to recognition. When an event occurs, the real-time monitoring system firstly acquires the event signal, then extracts the features proposed in this paper, and finally recognizes the event using the RVM classifiers. The training process was performed before real-time monitoring. In this paper, the recognition time must be within one second.

Table 5. Performance comparison of different methods.

Method	Average Precision (%)	Recognition Time (s)
WFE-RVM	80	10.526
RDFE-RVM	85.4	2.169
MFE-RVM	97.8	0.7028

In Table 5, Wavelet Feature Extraction RVM (WFE-RVM) is the traditional recognition method. Region Descriptor Feature Extraction RVM (RDFE-RVM) is the recognition method without feature selection. Morphological Feature Extraction RVM (MFE-RVM) is the recognition method proposed in this paper. Compared with WFE and RDFE, the MFE method proposed in this paper can greatly improve the recognition accuracy. There are only four features after feature selection; therefore the speed of the algorithm also can be increased significantly. The recognition accuracy can reach 97.87%, and the recognition time is within one second, so it meets the requirement of Φ -OTDR online monitoring.

5. Conclusions

A time-space signal is regarded as an image in this paper, and a recognition method based on morphological feature extraction is proposed. Image processing technology is used for signal pretreatment, and a scatter matrix is calculated for feature selection. Recognition of pipeline safety events is used to test the performance of the algorithm. The results show that the accuracy and speed are both excellent compared with time domain signal recognition. The proposed method can completely meet the recognition requirements of a Φ -OTDR pre-warning system.

Acknowledgments

This work has been supported by National Natural Science Foundation of China (Grant No. 61304244), Tianjin Natural Science Foundation of China (Grant No. 14JCQNJC04900), Ministry of Education Foundation of China (Grant No. 20130032120066).

Author Contributions

Qian Sun made the algorithm of recognition and wrote the paper, Hao Feng and Zhoumo Zeng designed the overall system and revised the paper, Xueying Yan helped to collect our database and to measure the recognition accuracies.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Juarez, J.C.; Taylor, H.F. Field test of a distributed fiber-optic intrusion sensor system for long perimeters. *Appl. Opt.* **2007**, *46*, 1968–1971.
2. Zhong, X.; Zhang, C.X.; Li, L.J.; Liang, S.; Li, Q.; Lu, Q.Y.; Ding, X.X.; Cao, Q.Y. Influence of laser source on phase-sensitivity optical time-domain reflectometer-based distributed intrusion sensor. *Appl. Opt.* **2014**, *53*, 4645–4650.
3. Kuang, K.S.C.; Cantwell, W.J.; Thomas, C. Crack detection and vertical deflection monitoring in concrete beams, using plastic optical fibre sensors. *Meas. Sci. Technol.* **2013**, *14*, 205–216.
4. Lopez-Higuera, J.M.; Rodriguez Cobo, L.; Quintela Incera, A.; Cobo, A. Fiber Optic Sensors in Structural Health Monitoring. *J. Lightw. Technol.* **2011**, *29*, 587–608.
5. Hao, J.Z.; Dong, B.; Varghese, P. An Armored-cable-based Fiber Bragg Grating Sensor Array for Perimeter Fence Intrusion Detection. *Proc. SPIE* **2011**, 8332, doi:10.1117/12.918980.
6. Tanimola, F.; Hill, D. Distributed fibre optic sensors for pipeline protection. *J. Nat. Gas Sci. Eng.* **2009**, *1*, 134–143.
7. Lin, Q.; Zhang, C.X.; Li, C.S. Fiber-optic distributed sensor based on phase-sensitive OTDR and wavelet packet transform for multiple disturbances location. *Optik* **2014**, *125*, 7235–7238.
8. Qin, Z.; Chen, L.; Bao, X. Continuous wavelet transform for non-stationary vibration detection with phase-OTDR. *Opt. Exp.* **2012**, *20*, 20459–20465.
9. Anyaegbunam, A.J. Complete stresses and displacements in a cross-anisotropic half-space caused by a surface vertical point load. *Int. J. Geomech.* **2014**, *14*, 171–181.
10. Cai, Y.Q.; Sun, H.L.; Xu, C.J. Response of railway track system on poroelastic half-space soil medium subjected to a moving train load. *Int. J. Solids Struct.* **2008**, *45*, 5051–5034.
11. Yamamoto, K.; Guo, W.; Yoshioka, Y.; Ninomiya, S. On plant detection of infact tomato fruits using image analysis and machine learning methods. *Sensors* **2014**, *14*, 12191–12206.
12. Zhu, Q.D.; Jing, L.Q.; Bi, R.S. Exploration and Improvement of Ostu Threshold Segmentation Algorithm. In Proceedings of the 8th World Congress on Intelligent Control and Automation (WCICA), Jinan, China, 7–9 July 2010; IEEE: Piscataway, NJ, USA, 2010; pp. 6183–6188.

13. Samuel, M.; Valentin, G. Robustifying vector median filter. *Sensors* **2011**, *11*, 8115–8126.
14. Yuan, K.H.; Duan, C.J.; Han, W. Roundness curve of shape and application. *Int. J. Innov. Comput. Inf. Control* **2010**, *6*, 1915–1923.
15. Fowers, S.G.; Lee, D.J.; Ventura, D.A. The nature-inspired basis feature descriptor for UAV imagery and its hardware implementation. *IEEE Trans. Circuits Syst. Video Technol.* **2013**, *23*, 756–768.
16. Mahanta, M.S.; Aghaei, A.S.; Plataniotis, K.N. Regularized lda based on separable scatter matrices for classification of spatio-spectral EEG patterns. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Vancouver, BC, Canada, 26–31 May 2013; pp. 1237–1241.
17. Bianco, A.; Boente, G.; Pires, A.M.; Rodrigues, I.M. Robust discrimination under a hierarchy on the scatter matrices. *J. Multivar. Anal.* **2008**, *99*, 1332–1357.
18. Hu, J.F.; Peter, W.T. A relevance vector machine-based approach with application to oil sand pump prognostics. *Sensors* **2013**, *13*, 12663–12686.
19. Tipping, M.E. Sparse Bayesian Learning and th relevance vector machine. *J. Mach. Learn. Res.* **2001**, *1*, 211–244.
20. Zio, E.; Maio, F.D. Fatigue crack growth estimation by relevance vector machine. *Expert Syst. Appl.* **2012**, *39*, 10681–10692.
21. Maio, F.D.; Kwok, L.T.; Zio, E. Combining relevance vector machines and exponential regression for bearing residual life estimation. *Mech. Syst. Signal Process.* **2012**, *31*, 405–127.
22. Zhong, S.P.; Chen, D.Y.; Xu, Q.F. Optimizing the Gaussian kernel function with the formulated kernel target alignment criterion for two-class pattern classification. *Pattern Recognit.* **2013**, *46*, 2045–2054.
23. Schwenker, F. Hierarchical support vector machines for multi-class pattern recognition. In Proceedings of the Fourth International Conference on Knowledge-Based Intelligent Engineering Systems & Allied Technologies, Brighton, UK, 30 August–1 September 2000; Volume 2, pp. 561–565.
24. Feng, X.; Zeng, Z.M.; Feng, H.; An, Y. A method for evaluating the disturbance in distributed vibration sensor based on wavelet information entropy. *Acta Opt. Sin.* **2013**, *33*, 116005:1–1106005:7.