

Article

Research of the Vibration Source Tracking in Phase-Sensitive Optical Time-Domain Reflectometry Signals Based by Image Processing Method

Yanzhu Hu, Song Wang * and Xinbo Ai 

Automation College, Beijing University of Posts and Telecommunications, Beijing 100876, China; yanzhu_hu@bupt.edu.cn (Y.H.); axb@bupt.edu.cn (X.A.)

* Correspondence: wongsanwongsang@163.com; Tel.: +86-138-1015-3283

Received: 8 July 2018; Accepted: 31 July 2018; Published: 2 August 2018



Abstract: This paper aims to improve the source tracking efficiency of distributed vibration signals generated by phase-sensitive optical time-domain reflectometry (Φ -OTDR). Considering the two dimensions (time and length) of Φ -OTDR signals, the authors saved and processed these signals as images after particle filtering. The filtering method could save 0.1% of hard drive space without sacrificing the original features of the signals. Then, an integrated feature extraction method was proposed to further process the generated image. The method combines three individual extraction methods, namely, texture feature extraction, shape feature extraction and intrinsic feature extraction. Subsequently, the signal of each frame image was recognized to track the vibration source. To verify the effect of the proposed method, several experiments were carried out to compare it with popular and traditional approaches. The results show that: Hard drive space is greatly conserved by saving the distributed vibration signals as images; the proposed particle filter is a desirable way to screen the vibration signals for monitoring; the integrated feature extraction outperforms the individual extraction methods for texture features, shape features and intrinsic features; the proposed method has a better effect than other popular integrated feature extraction methods; and, the signal source tracking method has little impact on the positioning accuracy of the vibration source. The research findings provide important insights into the source tracking of Φ -OTDR signals.

Keywords: phase-sensitive optical time-domain reflectometry (Φ -OTDR); image-based signal storage; integrated feature extraction; vibration source tracking

1. Introduction

Proposed by Taylor and Lee in 1993, phase-sensitive optical time-domain reflectometry (Φ -OTDR) is a typical monitoring technique for distributed vibrations [1]. Capable of positioning distributed signals, this technique has been widely applied to health monitoring of large buildings [2], perimeter security of important places [3], etc. Compared to traditional 1D monitoring of vibration signals [4–8], Φ -OTDR can realize long-term and high-accuracy monitoring. However, these advantages are achieved at the cost of a huge amount of distributed vibration data, which may lead to insufficient storage and inefficient data processing.

To solve this defect, this paper attempts to store and process the vibration signals as images according to existing image-based approaches for vibration signal processing [9–14]. Specifically, the vibration signals were converted into storable images, and then an integrated strategy was proposed to extract features of these images. Next, the proposed strategy was applied to analyze Φ -OTDR signals and track the vibration source. Overall, our research mainly tackles two issues, namely, signal storage

(i.e., the conversion of signals into images for storage) and signal analysis (i.e., the processing of the stored images by the proposed method).

Concerning signal storage, Han et al. [15] combined autoregressive-moving-average (ARMA) with swing door trending (SDT) to compress the vibration signals without sacrificing the key features. Malovic et al. [16] applied time delay estimation (TDE) in conjunction with differential pulse code modulation (DPCM) as the entropy coding of preprocessor, revealing that the integrated method can encode different types of aperiodic signals and compress vibration signals. Inspired by block compression, Huang et al. [17] put forward a lossless compression plan that draws on the merits of both lossy and lossless compressions. Guo et al. [18] developed a vibration signal compression technique called intrinsic mode function (IMF) based on ensemble empirical mode decomposition (EEMD), aiming to decompose the components of vibration signals in different frequency bands. To sum up, the above signal storage methods can be easily derived through analyzing and calculating the vibration signals. However, most of these methods require complex computation and do not apply to exceptional cases. By contrast, the Φ -OTDR technique can overcome these problems by collecting distributed vibrations signals with two dimensions: Time and length. Therefore, this paper aims to convert vibration signals directly into storable images after a few simple steps of preprocessing.

Concerning signal analysis, image target recognition has long been regarded as the key problem. The existing methods of image target recognition fall into five categories: Color feature extraction, texture feature extraction, shape feature extraction, intrinsic feature extraction, and spatial feature extraction. The color and spatial features are neglected here due to the lack of color and spatial information in the grayscale images generated from vibration signals. Because signal types of features cannot meet engineering requirements, many scholars have explored integrated feature extraction for image processing. For instance, Yang et al. [19] achieved high-speed tracking of image targets via hybrid rotation invariant description and skip search. Xia et al. [20] use color and edge feature distribution to build a mixture model to search for matching targets in the next frame image. Xiao et al. [21] combined the effective region index and multi-scale edge index for image processing. Considering the shape feature and other details of moving objects, Ren et al. [22] presented a robust visual tracking method called the SURF Mean Shift Deep Learning Tracker (SMS-DLT). Nevertheless, the above integrated methods are not comprehensive enough to process the images generated from distributed vibration signals. In these images, there is no complex background, light conversion or other factors common in traditional image processing. Hence, the texture, shape and intrinsic features should be taken into account.

In light of all three types of features, this paper adjusts the weight of each pixel in the original image by the speeded-up robust features (SURF) method and embeds the extraction methods of the three features in the particle filter. Based on the extraction of hybrid image features, a method was proposed to track the vibration source of Φ -OTDR signals. The steps of the proposed method are presented in Figure 1 below. First, the vibration signals of optical fiber in different sources were acquired by the Φ -OTDR technique, subjected to pre-processing, and stored as images to reduce storage space; then, three types of features (i.e., texture features, shape features and intrinsic features) were extracted from the images; finally, the effect of the proposed method was verified through experiments. The research findings shed new light on the tracking of vibration sources.

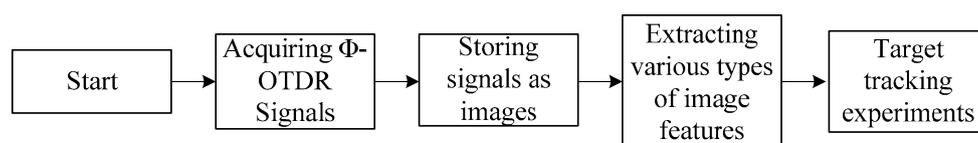


Figure 1. Steps of the proposed method.

2. Φ -OTDR Signal Storage Model

2.1. Φ -OTDR

Unlike traditional monitoring, the Φ -OTDR technique regards the optical fiber as an organic whole in the monitoring process. In other words, the optical fiber is considered as a single vibration signal appearing at multiple points on the same line [23–25]. Owing to the distributed feature, the vibration signal of the Φ -OTDR contains three kinds of information—amplitude, time and length—in which the amplitude varies with time and length.

The optical fiber takes the shape of a long line. The signal at each point on the length axis must be observable from the time axis, and the inverse is also true. The signal is the strongest at the point where the vibration source is vertical to the fiber. From this point, the signal strength gradually decays until reaching the two ends of the fiber. If the length axis is between 4.5 m and 5.9 m, then the strongest signal will appear at 5.2 m. Then, the signal amplitude between 4.4 m and 6.1 m can be measured by time and length (Figure 2).

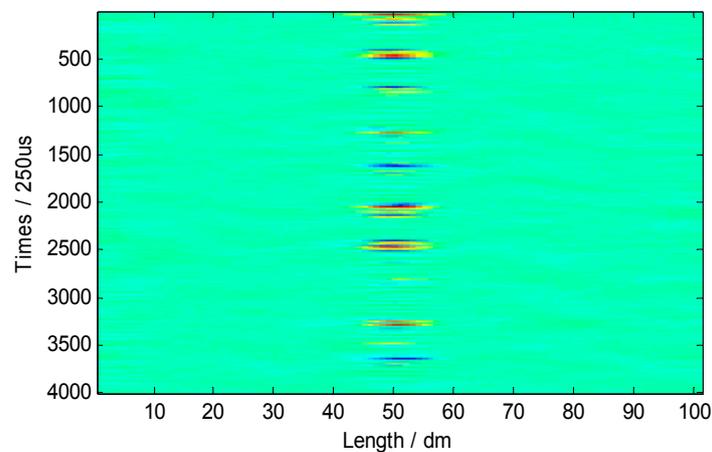


Figure 2. 2D diagram of distributed vibration signals.

Table 1 below provides a detailed explanation of Figure 2.

Table 1. Signal-to-noise ratio (SNR) at each point of the length axis.

No.	Length (m)	SNR (dB)	No.	Length (m)	SNR (dB)
1	4.4	0.5061	10	5.3	6.7691
2	4.5	1.7981	11	5.4	6.3194
3	4.6	3.5218	12	5.5	5.9771
4	4.7	5.0084	13	5.6	5.4368
5	4.8	5.2964	14	5.7	4.8608
6	4.9	6.0163	15	5.8	3.8066
7	5.0	6.4029	16	5.9	1.5836
8	5.1	6.8485	17	6.0	0.4238
9	5.2	7.9589	18	6.1	0

2.2. Signal Storage

In practice, the Φ -OTDR technique often results in a huge amount of vibration signals. Taking the NBX-S3000 distributed vibration monitoring device (Nebreux, Kobe, Japan; sampling rate, 4000 Hz; spatial resolution, 0.1 m; monitoring range, 10 m; format, double-precision floating-point) for example, the monitoring process generates 3.2 MB of data per second and 270 GB of data each day. The massive amount of data adds to the difficulty in data operations, such as storage and analysis. Similar to those

shown in Figure 2, Φ -OTDR signals are displayed regularly. This naturally associates the image-based approach with the data reduction of the Φ -OTDR technique. Clearly, the image-based approach only works if the hard drive space can be released substantially without losing a significant quantity of signal features.

Before generating images from Φ -OTDR signals, it is necessary to pre-process the original signals through the following steps:

- Step 1: During the acquisition process, the Φ -OTDR signals appear as a slow and low-amplitude sine wave due to the features of the acquisition card. Therefore, a high pass filter (threshold: 1 Hz) was added to filter the signals.
- Step 2: The signals were further screened by a sliding window filter (window: 10 Hz) to eliminate noises and possible error points.
- Step 3: The signals were magnified exponentially to obtain a better signal-to-noise ratio (SNR).

To save storage space, the pre-processed signals were saved in an image model through the following steps:

- Step 1: The distributed vibration signals were split into one-second segments.
- Step 2: Taking time as the horizontal axis and length as the vertical axis, the signal amplitude was normalized into the greyscale range between 0 and 1.
- Step 3: In the generated image, the number of pixels on the horizontal axis ($Pixel_x$) is the number of sampling points per second, also known as sampling rate (F_s) in Equation (1):

$$Pixel_x = F_s \quad (1)$$

Since the natural frequencies of large structures usually fall between 0 Hz and 60 Hz, the sampling rate should reach 120 Hz according to the Nyquist-Shannon sampling theorem. As for concrete structures, the frequency of vibrations signals varies from hundreds to thousands of hertz. Thus, the sampling frequency for concrete structures should fall between 1 kHz and 4 kHz.

- Step 4: In the generated image, the number of pixels on the vertical axis ($Pixel_y$) is the ratio of length (L) to the spatial sampling rate (R):

$$Pixel_y = \frac{L}{R} * 20 \quad (2)$$

- Step 5: The image was saved in the .jpg or .bmp format.

3. Vibration Source Tracking Based on Various Types of Image Features

In this research, the tracking target is the change in the image of vibration signals. To track the target, three types of features were extracted from the image: Texture features, intrinsic features and shape features. The shape features were obtained by the histogram of gradient directions, while the intrinsic features were acquired by GoogLeNet (Google, San Francisco, CA, USA). Then, the particle filter was adopted to track the target on the image based on the shape and intrinsic features. Subsequently, the greyscales of the original image pixels were adjusted by the SURF matching algorithm. The salient pixels were given relatively high greyscales. The adjustment helps improve the feature extraction results. The flow of the vibration source tracking is shown in Figure 3.

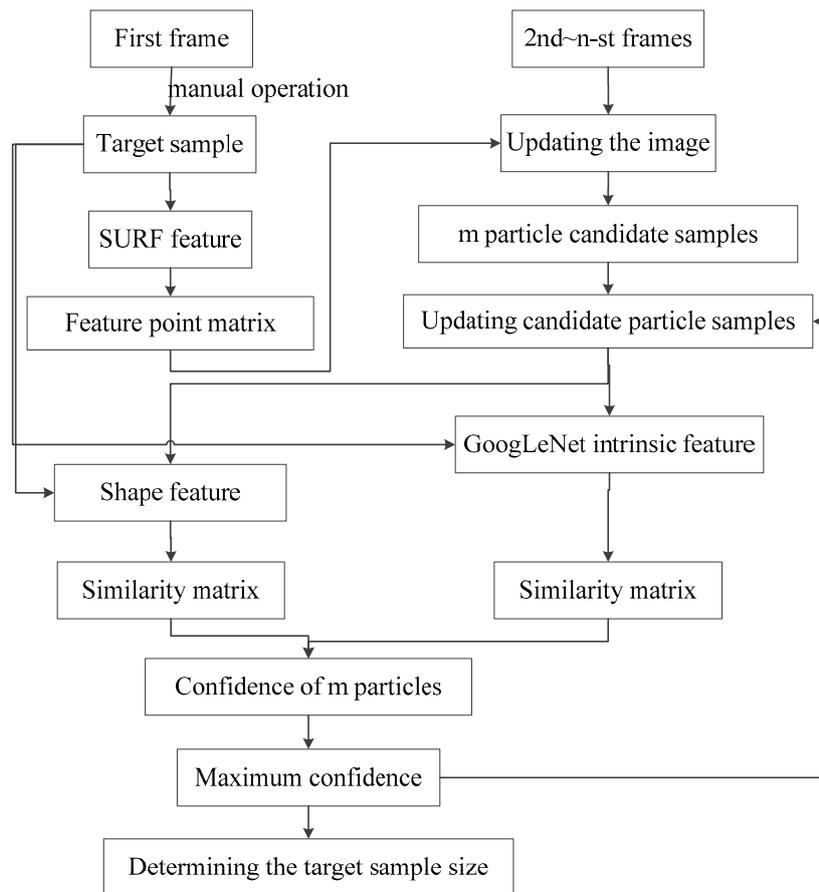


Figure 3. The flow diagram of algorithm.

As shown in Figure 3, the vibration sources were tracked in the following steps:

- Step 1: The first frame (f_1) of the image was sampled and the target and background templates were obtained manually; the target template was obtained by shape feature extraction, intrinsic feature extraction, and SURF feature extraction.
- Step 2: Let f_i be the i -th frame of the image ($i = 2, 3, \dots, n$).
 - (1) The candidate sample sets were obtained through random generation of sampled particles.
 - (2) A set of SURF features (S_i) was established to reflect the target positions from f_i to f_{i-1} . After matching, the SURF feature point mapping matrix (W_{is}) was obtained. The grayscale of the original image was multiplied by 0.7 and then increased by 0.3 at the feature point position, forming the updated samples.
 - (3) The shape and intrinsic similarities (ρ_i) between each candidate sample and the target template were calculated, respectively.
 - (4) The confidence was obtained for each particle, and the particle with the highest confidence was determined as the target position of f_i .
- Step 3: If the update condition was satisfied, the target and background templates were resampled. If not, let $i = i + 1$ and return to Step 2.

3.1. Target Contour Feature Extraction

As mentioned previously, the shape features of the sample image were extracted by the histogram of gradient directions [26,27], and the similarity between the sample and the target template was

calculated, laying the basis for subsequent motion estimation. The first step is to determine the gradient direction (i.e., the angle between the x - and y -axis gradients of a pixel). Let $a \times b$ be the number of pixels of the greyscale image and $\{\delta_{ij}\}_{i \in [1,a], j \in [1,b]}$ be the gradient angles of these pixels. Then, we have:

$$\Phi_{ij} = \arctan\left(\frac{\partial(\text{gray}_{ij})}{\partial x} / \frac{\partial(\text{gray}_{ij})}{\partial y}\right) \quad (3)$$

where gray_{ij} is gray value of the point (i, j) , and $\partial(\text{gray}_{ij})/\partial x$ and $\partial(\text{gray}_{ij})/\partial y$ are the x -axis and y -axis gradients of point (i, j) , respectively.

Then, the histogram of gradient directions can be determined by dividing the gradient angle into different intervals:

$$H_k = \Phi_{ij} / \Delta\Phi \quad (4)$$

where $\{H_k\}_{k \in [1,m]}$ is the interval and $\Delta\phi$ is size of interval. The histogram of gradient directions is the probability of the encoded pixels in the image in each direction (H_k).

Next, the histogram of gradient directions was weighted to ensure the robustness of density estimation. Through the weighting process, the pixels were assigned their respective weights according to their proximity to the target center. In the weighted histogram, the probability of the k -th interval, p_k , can be expressed as:

$$p_k(y) = H_h \sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|\right) \delta[b(x_i) - k] \quad (5)$$

where y is the center of the sample; $\{x_i\}_{i \in [1, n_h]}$ is the position of each pixel in the sample; $k(x)$ is the kernel function; H is the window width of the kernel function; $b(x_i)$ is direction encoding index of pixel x_i ; and, δ is the Dirichlet function. The importance of each particle in each frame image was determined according to the particle's confidence. To obtain the confidence, the histogram of gradient directions was established for each candidate sample, and the similarity between each sample and the target template was computed at the same time. The similarity, $\rho(y)$, between the histogram of gradient directions, $p(y)$, of each candidate sample and $p(y_0)$ of the target template was measured by the Bhattacharyya distance:

$$\rho(y) = \sum_{k=1}^n \sqrt{p_k(y)p_k(y_0)} \quad (6)$$

3.2. GoogLeNet-Based Feature Extraction

At the frontier of machine learning, deep learning mimics the mechanism of the human brain to interpret such data as images, audio and text, and supports the automatic extraction of the intrinsic features of an image. A typical example of deep learning is GoogLeNet, a deep convolutional neural network designed by Google [28,29] for the Large Scale Visual Recognition Challenge 2014. As shown in Figure 4, the GoogLeNet consists of 22 layers and reflects the idea of sparse learning. The size of GoogLeNet can be expanded by adjusting the parameters of the sparse network.

The GoogLeNet adopts a modular structure that is easily addable or modifiable. The fully-connected layer is replaced by the average pooling, which improves the accuracy by 0.6%. Despite the removal of the fully-connected layer, the dropout concept is still used in the network. To prevent the vanishing gradient problem, two additional modules are added for the forward propagation of gradient. Here, the GoogLeNet is employed to extract feature samples from images, and the confidence of the target sample is discussed according to both intrinsic and shape features.

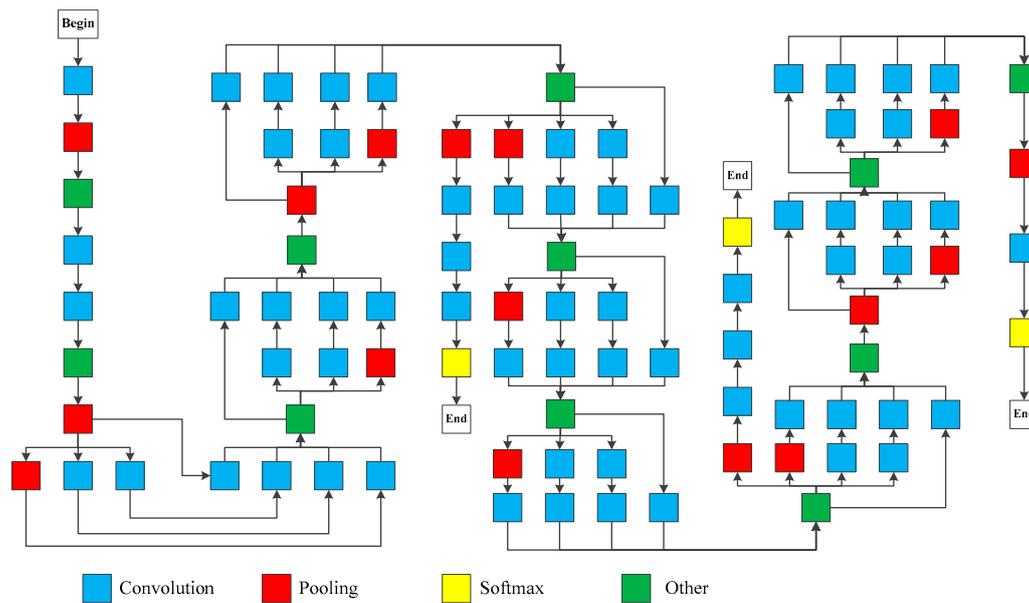


Figure 4. Structure of GoogLeNet.

3.3. Feature Extraction Based on SURF Method

The SURF is a simple and fast algorithm to extract interest points and describe eigenvectors [30,31]. The classical SURF uses the difference of Gaussians (DoG) operator, which is inspired by the Laplacian of Gaussian (LoG) operator in scale invariant feature transform (SIFT). In general, the SURF contains five steps: Constructing the Hessian matrix; calculating eigenvalue; constructing Gaussian pyramid; determining the principal direction of feature point and locating feature points; and, constructing feature descriptors.

The box filter plays an important role in these steps: It can simplify and approximate the Hessian matrix, making it possible to segment the second-order Gaussian template. With three values (i.e., 1-white, 0-gray and -1-black), the traditional box filter approximates white and light white regions as white regions, and black and light black regions as black regions. In this way, the speed is increased but the accuracy is not preserved. This gives rise to the improved box filter that has five values: 1, 0.5, 0, -0.5 and -1. The improved box filter (Figure 5) ensures that the regional size increases consistently in the SURF.

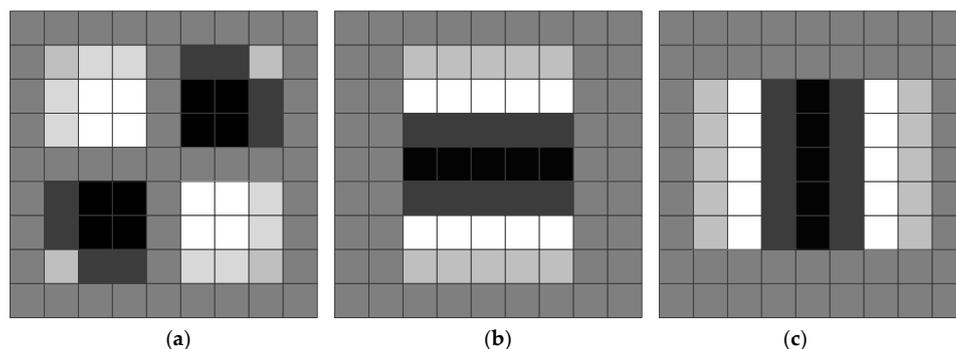


Figure 5. Improved box filter. (a) Box filter of D_{xy} (Two order derivatives for X and Y axes) (b) Box filter of D_{yy} (Two order derivatives for Y axis) (c) Box filter of D_{xx} (Two order derivatives for X axis).

In this paper, the SURF was adopted to extract the set of feature points from each frame image to form a new grayscale matrix of the same size as that of the original image. In the new matrix, the feature points were in black, and the other points were in white. Then, the original image was

generated from the new matrix. The grayscale matrix weights of the original image and the new matrix were 0.7 and 0.3, respectively. The two values were added together to derive the grayscale matrix of the updated image:

$$\hat{M}_{Image} = 0.7 \cdot M_{Image} + 0.3 \cdot M_{Surf} \quad (7)$$

where \hat{M}_{Image} is the gray scale matrix of updated image; M_{Image} is the gray scale matrix of the original image; and, M_{Surf} is the black-white matrix of SURF feature points.

3.4. Particle Filter Tracking Algorithm

In particle filtering, the particle states are described by affine transformation parameters, each of which is a six-dimensional vector. For each particle, the variables are distributed randomly and obey a probability distribution in the state space. The next most probable state is estimated by probability calculation according to the previous state:

$$M = [x \quad y \quad s_c \quad r_o \quad r_a \quad s_a] \quad (8)$$

where x and y are the abscissa and ordinate of the center of the particle sample; s_c is the length-width ratio of the sample; r_o is the rotation angle of the particle sample; r_a is the height-width ratio of the particle sample; and, s_a is the gradient of the tracking window.

Particle filtering is an important resampling process that places a number of particles, by certain rules, in the current frame. According to the placement rules, the particles are either placed evenly or denser near the target. The similarity between the particles and the target template is measured by particle weights. For simplicity, the weights should be normalized so that the sum of weights of all particles equals 1.

In this paper, the particle filter tracking algorithm is implemented as follows. First, n particle samples were obtained by random sampling in the initial frame. The weights of the particles were set to $1/n$. Let $\{s_{t-1}^i\}_{i \in [1,n]}$ be the state of the n particles at time $t - 1$ and $\{w_{t-1}^i\}_{i \in [1,n]}$ be the weights of these particles. Then, n particle samples were selected from the particle set based on the weights. The normalized weight probability set $\{C_{t-1}^i\}_{i \in [1,n]}$ can be expressed as:

$$C_{t-1}^i = \sum_{k=1}^i w_{t-1}^k / \sum_{k=1}^n w_{t-1}^k \quad (9)$$

The n sets of variables evenly distributed between 0 and 1 were randomly generated and denoted as $\{r^i\}_{i \in [1,n]}$. Then, $\{C_{t-1}^i\}_{i \in [1,n]}$, the set of n minimum indices $\{Idx_{t-1}^i\}_{i \in [1,n]}$, was established ($\{C_{t-1}^{Idx_{t-1}^i}\}_{i \in [1,n]} \geq 1$). Then, s_{t-1}^i was instated to the $s_{t-1}^{Idx_{t-1}^i}$, marking the end of the resampling process.

The updated particle set was transmitted via the system state-change equation. When a new frame arrives, the state of the particle state can be obtained as:

$$s_t = A \cdot s_{t-1} + v_{t-1} \quad (10)$$

where A is the state transition matrix and v_{t-1} are the multivariate Gaussian variables randomly generated by affine transformation parameters. The confidence of each particle can be obtained as:

$$E(y) = \rho(y)w(y) \quad (11)$$

where $\rho(y)$ is the similarity between the histogram of gradient directions of the candidate sample and that of the target template and $w(y)$ is the particle weight obtained by GoogLeNet. The maximum confidence particle was considered as the final estimation of the output frame. In this way, the texture features were combined with the shape features, and the sample with maximum similarity was identified to enhance the tracking accuracy.

4. Experiments

The NBX-S3000 (Figure 6) distributed vibration monitoring device was adopted for our experiments.

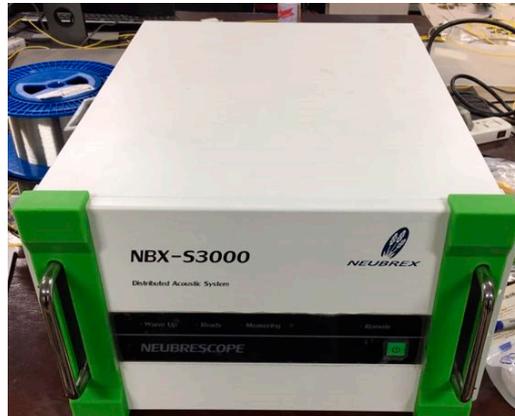


Figure 6. Image of Φ -OTDR instrument.

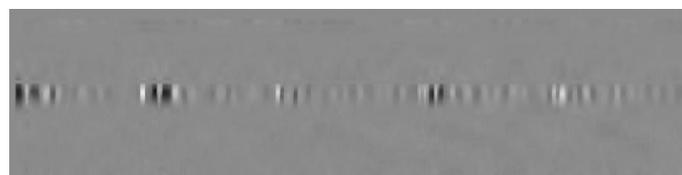
A standard five-hammer vibration device was taken as the vibration source (Figure 7). The experiments were carried out in an anechoic room to minimize the environmental noises and ensure the vibration effect.



Figure 7. Image of the standard five-hammer vibration device.

4.1. Signal Storage by Image Style

The vibration signal acquired by Φ -OTDR is stored by the image style. The specific image of each step is shown in Figure 8.



(a)



(b)

Figure 8. Cont.

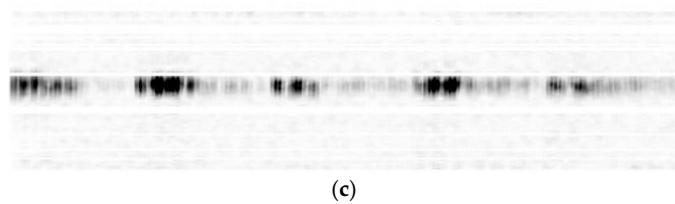


Figure 8. Images of the signals of each step: (a) Image of the original signals; (b) image of the filtered signals; and, (c) image of the enhanced signals.

According to Figure 8, it is obvious that the image of the original signals was not very clear, that of the filtered signals was clearer, with the noise becoming a white background, and that of the enhanced signals was clear and distinguishable.

Then, the proposed filtering method was contrasted with the Kalman filter and other mainstream methods in terms of the SNR and efficiency. The results are recorded in Table 2 below.

Table 2. Comparison of filtering effects among various methods.

No.	Name	Average SNR (dB)	Time (s)
1	The method in the paper	5.0422	6.78
2	Wiener filter	4.6231	12.71
3	Kalman filter	4.8672	11.18
4	Adaptive filter	4.4517	14.31
5	Wavelet filter	4.9951	16.14

Table 2 shows that the proposed method outperformed the other approaches in both SNR and efficiency. The advantage was particularly obvious in efficiency, as the proposed method consumed 45% less time than the Kalman filter. This is attributable to two main reasons: First, the vibration signals in our research are mechanically damped and thus easy to handle; second, it is very time-consuming to process distributed vibration signals on each length point. Meanwhile, the SNR of the proposed method was not notably stable because the strong and weak signals had not been fitted by advanced methods at the same time.

Next, several experiments were carried out between data files in different formats to see if the strategy of saving as images could save disk drive space. The experimental results are shown in Table 3.

Table 3. Comparison of data file sizes in different formats.

No.	Data Format	Time	Number of Files	Size of Single File	Total	Compress by RAR (Roshal ARchive)
1	.mat	10 s	1	604 MB	604 MB	312 MB
2	.csv	10 s	1	451 MB	451 MB	229 MB
3	.bin	10 s	1	367 MB	367 MB	190 MB
4	.jpg	10 s	20	16.6 KB	352 KB	189 KB

The results in Table 3 reveal that saving signals as images could greatly reduce the storage space. For instance, an image file occupied 1000 times less space than a binary file. In general, the image-based method allows the hard disk space to be searched about 2000 times. This is because the data in images are stored as integers while those in .csv and .mat files are saved as double-precision floating-points, and the image files are smaller than the other files. Thus, the image-based method is a desirable way to save the massive amount of data generated by the Φ -OTDR technique.

4.2. Image Feature Extraction

4.2.1. Effect Analysis of Target Tracking

Taking a vibration source image as the object, the features were extracted by the SURF (Figure 9) and expressed as center point error (vector 1) and success rate (vector 2). The center point error refers to the Euclidean distance between the center of the target frame and the real target frame. The mean error of the centers is the sum of center errors divided by the total number of frames. If more than 50% of the target frame and the real target frame is overlapped, the frame is considered as being correctly tracked. The success rate stands for the ratio of correctly tracked frames to the total number of frames.

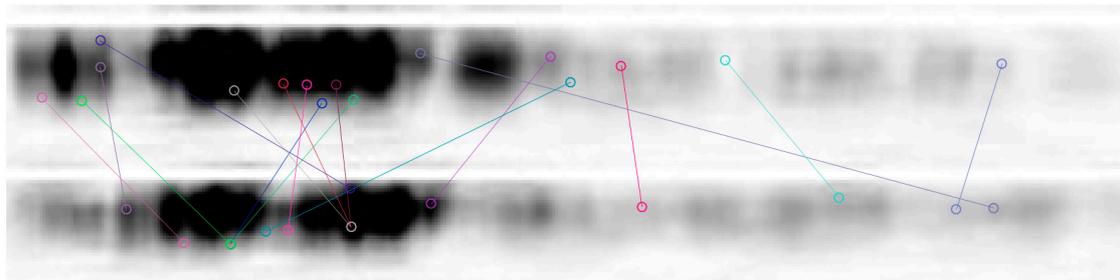


Figure 9. Comparison of two indexes under different weight.

After SURF extraction, a total of 4096 vectors were obtained:

- $Vector1_{(4096)} = [0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 2.453, 0.0, 5.042, 0.0, 1.899, \dots, 0.0, 0.0, 0.0, 0.6, 2.869, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0];$
- $Vector2_{(4096)} = [0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 1.485, 0.0, 0.941, 4.506, 2.171, \dots, 0.0, 0.0, 0.0, 0.57, 2.997, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0];$

Table 4 compares the results of the proposed integrated extraction method with those of several traditional feature extraction methods.

Table 4. Mean error of the center points (pixel) and the success rate (%).

No.	Name	Hammer No. 1	Hammer No. 2	Hammer No. 3	Hammer No. 4	Hammer No. 5
1	GoogLeNet	25.3 (90.1)	24.3 (90.6)	25.1 (89.9)	25.8 (91.1)	25.0 (90.4)
2	SURF	26.7 (89.7)	26.9 (89.9)	26.3 (89.1)	26.1 (90.4)	25.1 (90.8)
3	Target contour	27.4 (89.3)	27.8 (88.1)	27.2 (87.2)	28.9 (88.9)	27.9 (89.1)
4	Multiply-features	24.5 (91.1)	24.1 (92.3)	24.7 (90.7)	24.9 (93.6)	23.9 (91.5)

Note: the format in the table is: error (rate).

It can be seen that the integrated extraction method had a better effect than the traditional extraction methods for texture, shape and intrinsic features, respectively. In terms of mean center error, the proposed method surpassed the shape extraction method by 8%, the texture extraction method by 7% and the intrinsic extraction method by 4%. In terms of success rate, the proposed method outperformed the shape extraction method by 1.7%, the texture extraction method by 1.4% and the intrinsic extraction method by 1%.

In this paper, the effect of the integrated extraction method hinges on the modification of the original image according to the points detected by the SURF. Therefore, repeated experiments were conducted using the modified weights, and the results were plotted as Figure 10.

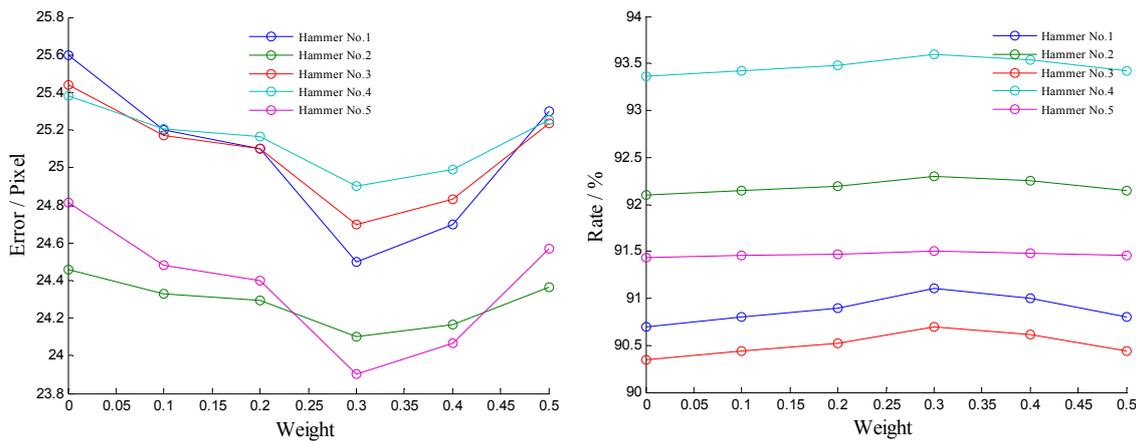


Figure 10. Comparison of two indexes under different weight.

As shown in Figure 10, with the increase of weight, the success rate remained essentially the same, while the center error exhibited a decreasing trend. This means the best grayscale ratio between the defect point and the original image is 3:7. Overall, the integrated extraction method did better than the signal method, and the effect of GoogLeNet was improved by superimposing SURF feature points onto the original image.

4.2.2. Comparison between Different Integrated Feature Extraction Methods

The proposed method was also compared against several popular integrated feature extraction methods, including discrete Fourier transform (DFT), incremental visual tracking (IVT), compressive tracking (CT), direct linear transform (DLT), etc. The mean center errors and success rates of these methods are listed in Table 5.

Table 5. Mean center errors (pixel) and the success rates (%) of different methods.

No.	Name	Hammer No. 1	Hammer No. 2	Hammer No. 3	Hammer No. 4	Hammer No. 5
1	DFT	14.1 (88.1)	12.3 (87.6)	15.1 (89.9)	13.8 (87.1)	12.0 (89.4)
2	IVT	5.7 (89.1)	6.9 (89.5)	6.3 (89.1)	5.9 (90.4)	5.1 (89.5)
3	CT	4.9 (89.2)	4.9 (89.1)	5.1 (89.2)	5.7 (92.2)	4.9 (89.8)
4	DLT	4.7 (89.6)	4.8 (90.8)	5.2 (89.7)	5.3 (92.9)	4.4 (90.7)
5	Mixed-features	4.5 (91.1)	4.1 (92.3)	4.7 (90.7)	4.9 (93.6)	3.9 (91.5)

As can be seen from Table 5, the proposed integrated feature extraction method outshined the other popular methods. In terms of mean center error, the proposed method achieved an accuracy 58.2% higher than the DFT, 17.3% higher than the IVT, 7.1% higher than the CT, and 3.2% than the DLT. In terms of success rate, the proposed method surpassed the DFT, IVT, CT and DLT by 2.5%, 1.7%, 0.9% and 0.4%, respectively. The across-the-board advantages arise from the complementary effect between shape and intrinsic features, the highlighting of target feature points by the SURF, and the edge of GoogLeNet over the other deep convolutional neural networks (DCNNs).

Taking one vibration source as an example, each frame image was analyzed by the proposed method, the DFT, the IVT, the CT and the DLT. According to the results in Figure 11, the proposed method lined out the range of the target source perfectly, while the other four methods each had its biases. Thus, the proposed method is an ideal tool to extract features from the images generated from Φ -OTDR signals.

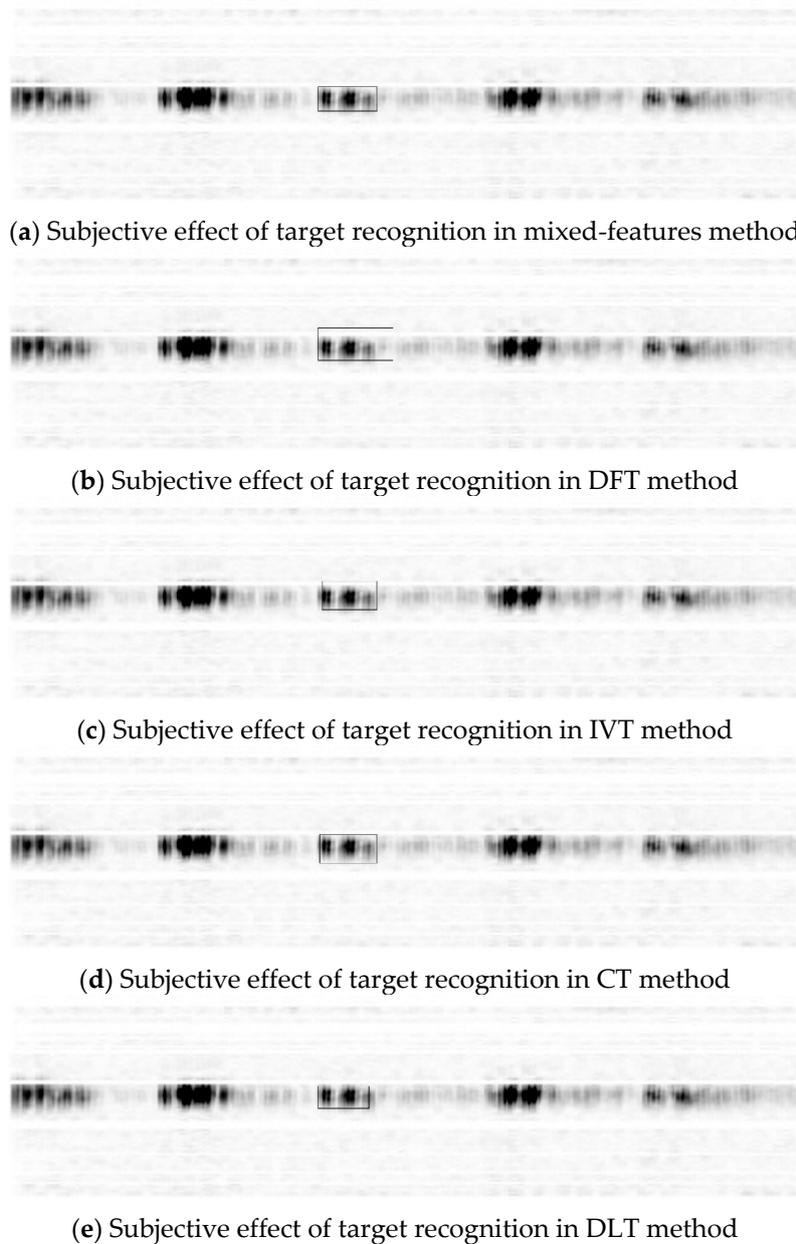


Figure 11. Subjective effect of target recognition in different methods.

Signal damage is an unavoidable phenomenon in any data compression image. The proposed image-based approach is no exception. In light of this, several experiments were performed to measure the success rates of the proposed method and several traditional signal processing methods, and the results are shown in Table 6.

Table 6. The success rate (%) in different methods.

No.	Name	Hammer No. 1	Hammer No. 2	Hammer No. 3	Hammer No. 4	Hammer No. 5
1	FFT (Fast Fourier Transform)	95.8	96.5	95.2	96.7	96.6
2	HHT (Hilbert-Huang Transform)	96.3	96.7	97.0	97.1	97.0
3	WT (Wavelet Transform)	96.9	97.2	97.1	97.2	97.1

The data in Table 6 demonstrate that some traditional methods could achieve over 95% success rates in source tracking. However, the success rate of the proposed method was considerably lower because our method involves not only data compression, but also feature retention. In summary, the proposed image-based approach is improved on other image-based methods, but poorer than the traditional signal processing methods.

4.2.3. Tracking Effect Analysis of Vibration Source

As 3D data, the Φ -OTDR signals must be expressed from the length axis and the time axis. Figure 12 presents the effect of single point knocking on the optical fiber. It can be seen that the peak vibration occurred at the vertical intersection point between the vibration source and the optical fiber, and the signal gradually weakened from the intersection to each end.

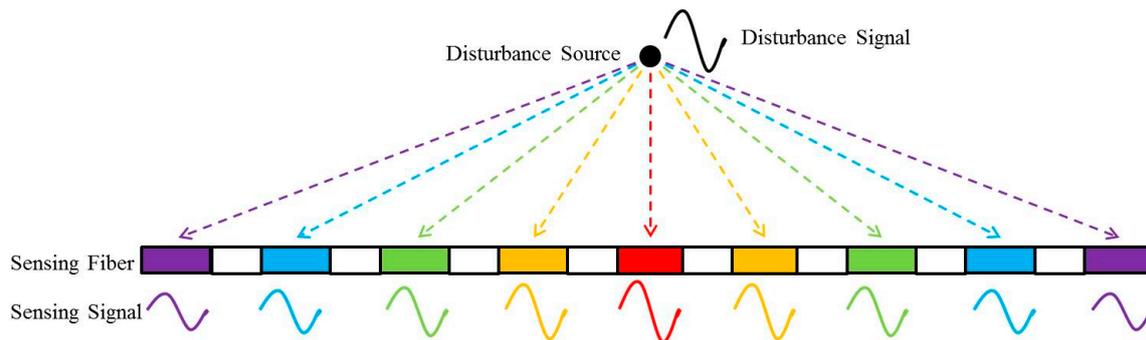


Figure 12. Perceptual model on length axis.

The five percussion hammers were equally distributed on the knocker. The distance between two strikes was 10 cm. Therefore, the four distances between the centers of the five target signals can be obtained by the proposed method. Our method was then compared with traditional methods, such as the fast Fourier transform (FFT), Hilbert-Huang transform (HHT) and wavelet transform (WT). The distances of each interval are presented in Table 7 below.

Table 7. Tracing effect of vibration sources between different methods (cm).

No.	Name	Interval No. 1	Interval No. 2	Interval No. 3	Interval No. 4
1	Mixed Feature	10.53	10.41	10.84	9.41
2	FFT	10.48	10.39	10.81	9.48
3	HHT	10.45	10.35	10.74	9.51
4	WT	10.46	10.37	10.78	9.52

As shown in Table 7, the proposed method had a positioning error of approximately 5.1%, 1.8% lower than the traditional methods. The accuracy is so small as to be negligible in actual practice. Further, the peak vibration occurred at the vertical intersection point between the vibration source and the optical fiber, and the signal gradually weakened from the intersection to each end, where it eventually disappeared. During the signal recognition on one frame, there were a number of errors in the recognition of the target signal, but this error could be corrected over time. To sum up, the vibration source was effectively tracked although the signals were saved as images.

4.2.4. Performance Difference between Image Formats

The .jpg image format requires an advanced compression technique. This format could maintain rich information at a high compression rate. To verify the feasibility of this format in our research, the above experiments were repeated on four different image formats. The results are shown in Table 8.

Table 8. Performance of different image formats (cm).

No.	Format	Interval No. 1	Interval No. 2	Interval No. 3	Interval No. 4
1	.jpg	10.53	10.41	10.84	9.41
2	.bpm	10.51	10.42	10.80	9.40
3	.png	10.53	10.52	10.85	9.43
4	.tiff	10.54	10.51	10.84	9.41

The data in Table 8 show that the four different formats had very similar performance, indicating that the .jpg format did not dampen the effect of the proposed method.

5. Conclusions

This paper proposes an integrated image feature extraction method for vibration source tracking of Φ -OTDR signals and compares the method with other popular approaches via experiments. According to the experimental results, it can be concluded that: Hard drive space is greatly conserved by saving the distributed vibration signals as images; the proposed particle filter is a desirable way to screen the vibration signals for monitoring; the integrated feature extraction outperforms the individual extraction methods for texture features, shape features and intrinsic features; the proposed method has a better effect than other popular integrated feature extraction methods; and, the signal source tracking method has little impact on the positioning accuracy of the vibration source.

Through our research, a simple, fast and lightweight source tracking method has been developed for Φ -OTDR signals. Considering the complexity of actual conditions and the fast development of deep learning networks and image processing methods, future research will improve the proposed method to suit other types of Φ -OTDR signals, such as non-damped leakage signals, and to reflect the latest techniques in image processing, such as the parallel use of multiple methods.

Author Contributions: Y.H. conceived and designed the experiments; S.W. performed the experiments, analyzed the data, wrote the paper; X.A. offered a lot of suggestions for paper writing.

Acknowledgments: This work was partially supported by the National Key Technology R&D Program of China (No. 2016YFC0701309), National Natural Science Foundation of China (No. 61627816 & No. 61503034).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Taylor, H.F.; Lee, C.E. Apparatus and Method for Fiber Optic Intrusion Sensing. U.S. Patent 5,194,847, 16 March 1993.
2. Bucaro, J.A.; Carome, E.F. Single fiber interferometric acoustic sensor. *Appl. Opt.* **1978**, *17*, 330–331. [[CrossRef](#)] [[PubMed](#)]
3. Juarez, J.C.; Maier, E.W.; Choi, K.N.; Taylor, H.F. Distributed fiber-optic intrusion sensor system. *J. Light. Technol.* **2005**, *23*, 2081–2087. [[CrossRef](#)]
4. Glowacz, A. Acoustic based fault diagnosis of three-phase induction motor. *Appl. Acoust.* **2018**, *137*, 82–89. [[CrossRef](#)]
5. Merizalde, Y.; Hernández-Callejo, L.; Duque-Perez, O. State of the art and trends in the monitoring, detection and diagnosis of failures in electric induction motors. *Energies* **2017**, *10*, 1056. [[CrossRef](#)]
6. Huang, J.; Tian, C.; Ren, J.; Bian, Z. Study on Impact Acoustic—Visual Sensor-Based Sorting of ELV Plastic Materials. *Sensors* **2017**, *17*, 1325. [[CrossRef](#)] [[PubMed](#)]
7. Sawczuk, W. The application of vibration accelerations in the assessment of average friction coefficient of a railway brake disc. *Meas. Sci. Rev.* **2017**, *17*, 125–134. [[CrossRef](#)]
8. Glowacz, A.; Glowacz, Z. Diagnosis of stator faults of the single-phase induction motor using acoustic signals. *Appl. Acoust.* **2017**, *117*, 20–27. [[CrossRef](#)]
9. Glowacz, A.; Glowacz, Z. Diagnosis of the three-phase induction motor using thermal imaging. *Infrared Phys. Technol.* **2017**, *81*, 7–16. [[CrossRef](#)]

10. Józwik, J.; Wac-Włodarczyk, A.; Michałowska, J.; Kłoczko, M. Monitoring of the Noise Emitted by Machine Tools in Industrial Conditions. *J. Ecol. Eng.* **2018**, *19*, 1. [[CrossRef](#)]
11. Xi, W.; Li, Z.; Tian, Z.; Duan, Z. A feature extraction and visualization method for fault detection of marine diesel engines. *Measurement* **2018**, *116*, 429–437. [[CrossRef](#)]
12. Józwik, J. Identification and monitoring of noise sources of CNC machine tools by acoustic holography methods. *Adv. Sci. Technol. Res. J.* **2016**, *10*. [[CrossRef](#)]
13. Praveenkumar, T.; Sabhrish, B.; Saimurugan, M.; Ramachandran, K.I. Pattern recognition based on-line vibration monitoring system for fault diagnosis of automobile gearbox. *Measurement* **2018**, *114*, 233–242. [[CrossRef](#)]
14. Glowacz, A.; Glowacz, Z. Diagnostics of stator faults of the single-phase induction motor using thermal images, MoASoS and selected classifiers. *Measurement* **2016**, *93*, 86–93. [[CrossRef](#)]
15. Han, S.; Liu, X.; Chen, J.; Wu, J.; Ruan, X. A real-time data compression algorithm for gear fault signals. *Measurement* **2016**, *88*, 165–175. [[CrossRef](#)]
16. Malovic, M.; Brajovic, L.; Sekara, T.; Miskovic, Z. Lossless Compression of Vibration Signals on an Embedded Device Using a TDE Based Predictor. *Elektronika Ir Elektrotehnika* **2016**, *22*, 21–26. [[CrossRef](#)]
17. Huang, Q.; Tang, B.; Deng, L.; Wang, J. A divide-and-compress lossless compression scheme for bearing vibration signals in wireless sensor networks. *Measurement* **2015**, *67*, 51–60. [[CrossRef](#)]
18. Guo, W.; Tse, P.W. A novel signal compression method based on optimal ensemble empirical mode decomposition for bearing vibration signals. *J. Sound Vib.* **2013**, *332*, 423–441. [[CrossRef](#)]
19. Yang, Y.; Yang, J.; Zhang, Z.; Liu, L.; Wu, N. High-speed visual target tracking with mixed rotation invariant description and skipping searching. *Sci. China* **2017**, *60*, 062401. [[CrossRef](#)]
20. Xia, J.; Rao, W.; Huang, W.; Lu, Z. Automatic multi-vehicle tracking using video cameras: An improved CAMShift approach. *KSCE J. Civ. Eng.* **2013**, *17*, 1462–1470. [[CrossRef](#)]
21. Xiao, C.; Gan, J.; Hu, X. Fast level set image and video segmentation using new evolution indicator operators. *Vis. Comput.* **2013**, *29*, 27–39. [[CrossRef](#)]
22. Ren, N.; Du, J.; Zhu, S.; Li, L.; Fan, D.; Lee, J. Robust visual tracking based on scale invariance and deep learning. *Front. Comput. Sci. China* **2017**, *11*, 230–242. [[CrossRef](#)]
23. Hu, Y.; Wang, S.; Wang, Z.; Zhang, Y. The Research on Information Representation of Φ -OTDR Distributed Vibration Signals. *J. Sens.* **2017**, *2017*, 6020645. [[CrossRef](#)]
24. Xu, C.; Guan, J.; Bao, M.; Lu, J.; Ye, W. Pattern recognition based on enhanced multifeature parameters for vibration events in ϕ -OTDR distributed optical fiber sensing system. *Microw. Opt. Technol. Lett.* **2017**, *59*, 3134–3141. [[CrossRef](#)]
25. Zhang, X.; Sun, Z.; Shan, Y.; Li, Y.; Wang, F.; Zeng, J.; Zhang, Y. A high performance distributed optical fiber sensor based on Φ -OTDR for dynamic strain measurement. *IEEE Photonics J.* **2017**, *9*, 1–2. [[CrossRef](#)]
26. Krieger, E.W.; Sidike, P.; Aspiras, T.; Asari, V.K. Deterministic object tracking using Gaussian ringlet and directional edge features. *Opt. Laser Technol.* **2017**, *95*, 133–146. [[CrossRef](#)]
27. Ahonen, T.; Matas, J.; He, C.; Pietikäinen, M. Rotation Invariant Image Description with Local Binary Pattern Histogram Fourier Features. In Proceedings of the Scandinavian Conference on Image Analysis (SCIA), Oslo, Norway, 15–18 June 2009; pp. 61–70.
28. Scott, G.J.; England, M.R.; Starms, W.A.; Marcum, R.A.; Davis, C.H. Training Deep Convolutional Neural Networks for Land-Cover Classification of High-Resolution Imagery. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 549–553. [[CrossRef](#)]
29. Zhao, G.; Ahonen, T.; Matas, J.; Pietikäinen, M. Rotation-invariant image and video description with local binary pattern features. *IEEE Trans. Image Process.* **2012**, *21*, 1465–1477. [[CrossRef](#)] [[PubMed](#)]
30. Zhao, X.; Dawson, D.; Sarasua, W.A.; Birchfield, S.T. Automated Traffic Surveillance System with Aerial Camera Arrays Imagery: Macroscopic Data Collection with Vehicle Tracking. *J. Comput. Civ. Eng.* **2017**, *31*, 04016072. [[CrossRef](#)]
31. Hu, Y.; Li, L. 3D Reconstruction of End-Effector in Autonomous Positioning Process Using Depth Imaging Device. *Math. Probl. Eng.* **2016**, *2016*, 8972764. [[CrossRef](#)]

