Research on Noise Reduction of Φ-OTDR Signal Based on Blind Source Separation Algorithm

Yue Zhang*, Song Wang* and Yanzhu Hu
School of Automation, Beijing University of Post and Telecommunications, Beijing, China

*Corresponding author e-mail: wongsangwongsang@163.com, zhanguye519@bupt.edu.cn, yzhu263.net

Abstract. Optical fiber sensing has become an important means of health and safety monitoring of large buildings and facilities due to its better sensing characteristics. Undesirable noise signals are inevitably generated during actual sensor data acquisition. This paper focuses on the noise reduction processing of optical fiber Φ-OTDR vibration signal. Firstly, the acquired signal is preprocessed with data, then the data is independently correlated, and finally the FastICA algorithm in blind source separation is used to filter the noise signal. In this paper, the traditional wavelet transform algorithm is compared to reduce noise. The experimental results show that the blind source separation algorithm has better separation effect on fiber vibration signals.

1. Introduction
As a high-precision, small-sized, fast-sensing sensing technology, fiber-optic sensing technology has received increasing attention and is becoming more and more widely used in the global field, such as structural safety real-time monitoring of large-scale building infrastructure. In the monitoring of vibration phenomena, the Φ-OTDR (Phase-sensitive OTDR) technology in fiber sensing monitoring is mostly used to monitor. However, due to the complex diversity of facilities or environments in the actual application process, the vibration source signal is often accompanied by many noise interference signals, so that the collected signals cannot extract the original signals better. The traditional noise reduction technology is not very effective in the application of fiber vibration signals. It is meaningful to propose an algorithm which uses blind source separation for the extraction of Φ-OTDR vibration signals.

There have been some studies that have applied blind source separation algorithms to noise separation. Sun YJ et al. applies blind source separation algorithm to different signal sources in multi-input and multi-output systems, and achieves good results [1]. Grotas, S applies blind source separation algorithm to solve the problem of blind estimation of states and topology in power systems [2]. Lei P et al. applies blind source separation technology to multiple interference signal denoising problems in driving voice command recognition scenarios, and has achieved good results [3]. Meschede, M et al. uses the blind source separation algorithm to separate the signals in the seismic waves, and achieves good results [4]. Zhou, YY et al. uses the blind source separation algorithm to separate the ground magnetic field pulsation signal and noise signal, and has achieved certain results [5]. Feng, FC proposed the blind source separation problem of undetermined mixed convolution and optimized the separation algorithm. The experimental results prove that the optimization effect is better [6]. Wang, LY optimized
the algorithm of compressed sensing theory (CS) and applied it to the underdetermined blind source separation algorithm, which achieved good results [7]. He, PJ et al. proposed an improved algorithm FastICA for multi-channel blind source separation is improved by using FSS-Kernel (Finite Support Samples Kernel) and separate two or more mixed signals [8]. He, YZ proposed a blind source separation algorithm to solve the thermal imaging problem of eddy current pulses [9]. He, CB applies blind source separation algorithm to separate pressure pulsation (PP) signal and improved machine fault detection efficiency [10].

However, none of the above studies applied the blind source separation algorithm to the noise separation of fiber vibration signals. In this paper, the blind source separation algorithm is applied to the separation of source and noise signals in fiber vibration.

2. Blind Source Separation Algorithm for Denoising Optical Fiber ϕ-OTDR Signal Signals

The theory of blind source separation is a new research field developed in the late 20th century. Its theory mainly refers to separating each signal from several observed mixed signals. The fiber vibration signal should take into account the vibration source that needs to be collected in the actual application process, and some undesired noise signals are collected. It is proposed to use the independent component analysis method in blind source separation to vibrate the fiber. The mixed signal is separated.

![Figure 1. Mathematical model of Blind source separation](image)

2.1. Feature correlation processing

When using independent component analysis, the features must be independently correlated, so that multiple features are independent in both linear and nonlinear cases. Considering that there may be some correlation between the noise of the vibration signal and the source vibration signal, it intends to use the PCA method to carry out the feature. The PCA principal component analysis is essentially a dimension reduction processing method, which is reduced by PCA principal component analysis. The redundancy of the input data and the correlation between the features is low. By calculating the eigenvalues and eigenvectors of the covariance matrix, the importance of the features is selected according to the eigenvalues and vectors. The ultimate goal is to reduce the correlation between input features and to satisfy the principle of using ICA algorithms in blind source separation.

2.2. Maximum Likelihood Estimation Method for Estimating Independent Components

The input signal source data needs to be evaluated for the independence of the transformation matrix according to the steps of the ICA algorithm after being processed by the independent PCA. The maximum likelihood estimation comes from the use of probability theory in statistics. It is a method that must give an observation data to evaluate the model parameters. The principle of maximum likelihood estimation is to assume that the probability distribution is a hypothetical distribution rate, represents the sample, and represents the parameter that needs to be estimated, then represents how much the probability of occurs when the parameter is. After obtaining the estimation result of the transformation matrix independence, the feature set extracted by the independent component can be easily obtained. It should be noted that the features of the feature set at this time are independent of each other.

2.3. Blind Source Separation FastICA Algorithm and Wavelet Transform Algorithm

The FastICA algorithm is based on the fixed-point recursive algorithm. It is a fast optimization iterative algorithm. It substitutes a large number of samples and data into each operation, thus the batch operation
is different from ordinary neural networks. Therefore, it is a commonly used network algorithms, while maintaining the parallel, distributed and simple calculation. According to information theory, it can be known that among the same variance characteristics, the Gaussian distribution has the largest entropy. It can be obtained that when the entropy is minimum, it means that the non-Gaussian property of the variable is larger. In other words, when the non-Gaussian property of the variable reaches the maximum, the smaller the entropy, the smaller the relationship between the variables. Thus, the purpose of separation is achieved. The FastICA algorithm can directly find independent variables of non-Gaussian distribution through a nonlinear function. Compared with the ordinary ICA algorithm, the FastICA algorithm has a fast convergence speed, the step size is not required, and can separate each independent component to the maximum extent.

The wavelet transform is derived from Fourier, but unlike the Fourier transform, the wavelet transform is not based on an orthogonal basis space. Wavelet transform is a small region, a finite length, and a mean value of 0. The method of analysing the local time-frequency domain is used to gradually multi-scale the signal through the telescopic translation operation, which can adapt to the signal analysis requirements in the time domain and the frequency domain. Certain attenuation and volatility can be focused on any detail of the signal, and can be used to fully highlight some aspects of the problem. After the wavelet function is stretched and translated, the wavelet sequence is obtained as follows:

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right)$$  \hspace{1cm} (1)

Where a is the expansion factor and b is the translation factor. The effect of the translation factor is to change the position of the wavelet transform window in the plane time axis, while the scaling factor can change the size and shape of the window. Wavelet analysis can analyse and process local high frequency and low frequency signals of mixed signals.

3. Experiments
This section first describes the data collection process, from data collection to data cleaning. The data is then centralized, that is, correlated, and the independent components in the data are estimated to serve as a criterion for the ICA algorithm. Then, the FastICA algorithm is used to perform blind source separation and noise reduction on the collected signal data, and the effects before and after use are compared. At the same time, it also compares the denoising method of wavelet analysis and draws the experimental conclusion.

3.1. The introduction of Data
The signal data in the experiment is mainly from the fiber vibration collecting equipment. It uses a continuous acquisition device for fiber vibration signals. 10 is selected in the pulse width, and the pulse period is 81920, where the units are ns. In order to do the contrast experiment, it collected two sets of signals. One is the signal that the tapping signal is used to continuously tap the fiber in the anechoic chamber for a period of time. The tapping signal source is the standard signal source, and this signal has no noise interference. The other set of signals selects the outdoor tapping signal, and the standard source is still selected and the tapping time is the same as the previous section in the anechoic chamber. The data obtained is cleaned to ensure the integrity of the data during the analysis.

3.2. Correlation processing and independence analysis of signal data
After the signal data is obtained, it needs to be correlated to ensure that each component meets the input criteria of the ICA algorithm. The PCA algorithm is used to first calculate the covariance of different dimensions between the input components, and at the same time obtain the eigenvalues and eigenvectors of the covariance matrix. The importance of the features will be judged according to the obtained eigenvalues. Then the transform matrix is processed according to the maximum likelihood estimation,
that is, the maximum likelihood function is solved. Under the condition that the maximum likelihood function is slightly continuous, the parameters are deductively guided and made equal to 0 and obtained. Estimated value.

3.3. Comparison of denoising effects of blind source separation algorithm
In this experiment, two sets of signal sources collected from different scenes are selected as experimental signals. The first group is the signal that is tapped by the standard signal source in the anechoic chamber. This signal ensures that the signal has no other noise interference, and this signal can be used as the standard control group; the other group of signals is the signal collected under the same conditions under outdoor conditions, and the two groups have the same signal length. Figure 2 is a source of noise-free interference collected in the anechoic chamber.

Figure 2. Vibration signal collected in the anechoic chamber

The vibration signals collected by the outdoor scenes of the second group inevitably have noise signals appearing outdoors. Using the FastICA algorithm to separate the second group of vibration source signals as shown in Figure 3.

Figure 3. Vibration signal separated by FastICA algorithm

The vibration source signal separated by wavelet transform is shown in Figure 4 below.

Figure 4. Vibration signal separated by wavelet transform algorithm
4. Conclusion
In this paper, the FastICA algorithm is used to separate the fiber vibration signal. The separation effect is compared with the first group of signals. The vibration source signal obtained by wavelet transform is also similar, which is a good result. However, comparing FastICA to wavelet transform, the former separates the signal from the vibration source signal closer to the first group of source signals than the latter. At the same time, there are some improvements in this paper. For example, the improvement of the FastICA algorithm itself is worthy of further study.

Acknowledgments
This work was partially supported by Beijing Natural Science Foundation (No.4192042), National Natural Science Foundation of China (No.61627816), Beijing Science and Technology Planning Project (No. D161100004916002).

References
[1] Sun Y, Zhu L, Li D, et al. Convex Combination of SISO Equalization and Blind Source Separation for MIMO Blind Equalization [J]. Wireless Personal Communications, 2019: 1 - 13.
[2] Grotas S, Yakoby Y, Gera I, et al. Power Systems Topology and State Estimation by Graph Blind Source Separation [J]. IEEE Transactions on Signal Processing, 2019.
[3] Lei P, Chen M, Wang J. Speech enhancement for in-vehicle voice control systems using wavelet analysis and blind source separation [J]. IET Intelligent Transport Systems, 2018.
[4] Meschede M, Stutzmann E, Schimmel M. Blind source separation of temporally independent microseisms [J]. Geophysical Journal International, 2018, 216 (2): 1260 - 1275.
[5] YuanYuan Z, Ying C, Hao C, et al. Application of reference-based blind source separation method in the reduction of near-field noise of geomagnetic measurements [J]. CHINESE JOURNAL OF GEOPHYSICS-CHINESE EDITION, 2019, 62 (2): 572 - 586.
[6] Feng F, Kowalski M. Underdetermined Reverberant Blind Source Separation: Sparse Approaches for Multiplicative and Convulsive Narrowband Approximation [J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2019, 27 (2): 442 - 456.
[7] Wang L, Yin X, Yue H, et al. A Regularized Weighted Smoothed L0 Norm Minimization Method for Underdetermined Blind Source Separation [J]. Sensors, 2018, 18 (12): 4260.
[8] He P, She T, Li W, et al. Single channel blind source separation on the instantaneous mixed signal of multiple dynamic sources [J]. Mechanical Systems and Signal Processing, 2018, 113: 22 - 35.
[9] He Y, Yang R, Wu X, et al. Dynamic scanning electromagnetic infrared thermographic analysis based on blind source separation for industrial metallic damage evaluation [J]. IEEE Transactions on Industrial Informatics, 2018, 14 (12): 5610 - 5619.
[10] He C, Li H, Zhao X. Weak characteristic determination for blade crack of centrifugal compressors based on underdetermined blind source separation [J]. Measurement, 2018, 128: 545 - 557.