Study on Extracting Pipeline Leak Eigenvector Based on Wavelet Packet

Xiaoyun Liu, Xiangfen Zhang*, Yan Ma and Yanqin Yang

The College of Information, Mechanical and Electrical Engineer, Shanghai Normal University, Shanghai, 200234, China

Abstract: Pipeline leak detection is an important part of pipeline safety, which is usually carried out by extracting feature vectors of leakage signal. However, the complexity of the leakage acoustic emission signal makes the extraction of feature vectors very difficult. To solve this problem, the authors propose an improved wavelet packet algorithm to extract the feature vectors which are constituted by five time-frequency domain parameters: time-domain energy, frequency-domain energy, frequency-domain peak, kurtosis coefficient and variance. Many experiments have been performed to extract feature vectors based on the proposed algorithm, with the results showing the proposed algorithm to be efficient enough to overcome the mixing effects caused by traditional wavelet packet when reconstructing the single sub-band signal. Thus, the proposed algorithm can accurately extract the feature vectors. The study of this article provides a good foundation for the subsequent work such as pipeline leak detection and positioning analysis.

Keywords: Acoustic emission, feature vectors, pipeline leakage, wavelet packet.

1. INTRODUCTION

Pipeline plays a very important role in petroleum, chemical, natural gas, urban construction and other industries. Pipeline leak can not only badly affect the normal transportation, but also cause the environmental pollution. To make matters even worse, it can sometimes lead to serious safety accidents [1].

As a non-destructive “testing method”, “acoustic emission” has “no damage to pipeline equipment”, but “high sensitivity, fast detection speed”, and low intensity of work, make it stand out from many other leakage testing methods [2, 3]. To detect the pipeline leakage condition, we usually get some time-frequency characteristic parameters from the acoustic emission signals collected by the acoustic emission sensors. Then, these parameters are adopted to construct feature vectors to judge whether the pipeline leakage has occurred.

Wavelet packet algorithms are usually adopted to extract feature vectors. “Wavelet packet transform has the ability of denoting local signal characteristics in time and frequency domain” [4] and it is suitable for the extraction of weak signal and analysis of non-stationary signal. The authors in literature [5] decomposed the acoustic emission signal of pipeline leakage with wavelet packet and reconstructed the single sub-band signal, with the energy of each frequency band as the feature vectors, which were then inputted to the BP neural network to complete pipeline leakage diagnosis. The authors in literature [6] decomposed the acoustic emission signal with wavelet packet and reconstructed the single sub-band signal, and then used the variance, kurtosis and energy from the single sub-band reconstructed signal to form the feature vectors, which were identified by the BP neural network to distinguish the causes of the failures. The authors in literature [7] decomposed the acoustic emission signal of pipeline leakage with wavelet packet and “extracted the energy of each frequency band” to form the feature vectors, and then used the feature vectors to represent the state of pipeline leakage.

All these references report the use of traditional wavelet packet to reconstruct single sub-band signal, and then extract the feature vectors from the reconstructed signal. However, because of the non-ideal characteristics of the traditional wavelet filter, the spectrum expansion and compression brought by the up-sample and down-sample, the single sub-band reconstructed signal we get might contain unwanted frequency components, and the acquired feature vectors may not be accurate.

To overcome the shortcomings of the traditional wavelet packet algorithm, this article puts forward a new wavelet packet algorithm. We select five time-frequency domain parameters to form the feature vectors to detect the pipeline leakage condition. We have performed many experiments based on real pipeline leakage data using the proposed method to extract the leakage feature vectors. That is, we decompose the experimental data and reconstruct the single sub-band signal with the new algorithm, and finally, the feature vectors are calculated from the signal to demonstrate whether the leakage has occurred or not. We also compared the results of the proposed algorithm with that of the traditional wavelet packet.
2. WAVELET PACKET ANALYSIS AND THE IMPROVED ALGORITHM

Acoustic emission signal is a kind of non-linear and non-stationary signal, with the statistical properties changing with time. Wavelet packet is an elaborate analysis method. The “decomposition of wavelet packet analysis on each level is conducted not only on the parts of low frequency, but also on those of high frequency” [8]. In this way, wavelet packet analysis can increase the time-frequency resolution. Therefore, it has widely been used [9].

2.1. The Traditional “Wavelet Packet Decomposition and Reconstruction” Algorithm

In engineering, “wavelet packet decomposition and reconstruction” are often accomplished through a wavelet filter bank [10], as shown in Fig. (1).

In Fig. (1), $h_0$ is the “decomposition low-pass filter”, $g_0$ is the “decomposition high-pass filter”, $h_1$ is the “reconstruction low-pass filter” [11], “ $g_1$ is the reconstruction high-pass filter” [12], $\downarrow 2$ and $\uparrow 2$ are respectively the down-sample and up-sample. $w_0, w_1, w_2, w_3, x_0, x_1, x_2, x_3$ are the results of each operation in the Fig. (1).

$f(n)$ is the time series signal, the length of signal is $N$, $S_j^i(n)$ is the wavelet packet coefficient of the $j$ layer on the node $i$, and $f(n) = S_0^0(n)$. The binary wavelet packet decomposition algorithm is as follows:

$$
\begin{align*}
S_j^{i+1}(n) &= \sum_k h_k(n - 2k)S_j^i(n) \\
S_j^{i+1}(n) &= \sum_k g_k(n - 2k)S_j^i(n)
\end{align*}
$$

(1)

In Eq. (1),

$$
J = \log_2 n; \quad j = 0,1,\ldots,J; \quad n = 1,2,\ldots,2^{-j}; \quad i = 1,2,\ldots,2^j.
$$

The binary wavelet packet reconstruction algorithm is:

$$
S_j^{i-1}(n) = 2\sum_k S_{j+1}^i(n)h_k(n - 2k) + \sum_k S_{j+1}^i(n)g_k(n - 2k)
$$

(2)

In Eq. (2),

$$
J = \log_2 n; \quad j = J - 1,\ldots,0; \quad n = 1,2,\ldots,2^{j-1}; \quad i = 2^j,2^{j-1},\ldots,1.
$$

Fig. (2) is the three layers wavelet packet decomposition tree. The node $(0, 0)$ is a source node which is the signal used to decompose, $h_0$ and $g_0$ are shown in Fig. (1). As shown in Fig. (2), we will get eight node signals from $(0, 0)$. 

![Wavelet Tree](image-url)
to (3, 7) if we decompose the original signal to third layer. We also call the node signal as single sub-band signal and put the node (3, 7) as node 7 for convenience. Next, we need to reconstruct the specific single sub-band signal and call the result as single sub-band reconstructed signal.

2.2. The Improved Wavelet Packet Algorithm

"Wavelet packet decomposition and reconstruction" are implemented by convolving the signal with the filter. When convolving one-dimensional discrete signal with the wavelet filter, if the length of the boundary valid data is less than the wavelet filter length, we cannot perform the convolving operation effectively. So, in engineering, we usually expand the original data to meet the requirements of convolution.

The traditional wavelet packet decomposition algorithm uses symmetric extension, that is, the extended data are symmetrical with the continuation end data. Provided \( x(m) \), \( m = 0, 1, \ldots, N - 1 \) is the one-dimensional discrete signal, the length of signal and filter are \( N \) and \( L \) respectively. The signal \( x(m) \) will be \( x(L-2), \ldots, x(0), \ldots, x(N-1) \) after extension. For this signal, because it has been convoluted, its front-end and back-end data of the low frequency coefficients will have a greater transition and the low frequency coefficient is not relatively consistent with the original signal, the error of the coefficient is relatively large. In order to overcome this defect, the authors in literature [13] compared the decomposition effect of the several common continuation modes and reached on the conclusion that the wavelet decomposition coefficients with the means of linear and parabolic continuation are relatively consistent with the original signal. In this paper, the "wavelet packet decomposition" will adopt parabola continuation method.

In the traditional wavelet packet reconstruction algorithm, since the wavelet filter is non-ideal, the up-sample and down-sample will lead to the spectrum expansion and compression, which will make the reconstructed sub-band signal contain unwanted frequency components, all of these will make the feature vectors extracted from the single sub-band reconstructed signal inaccurate. This article uses FFT and IFFT transformation to process the high and low frequency coefficients, so the single sub-band reconstructed signal will not contain the unwanted frequency components [14].

To sum up, this paper proposes a new wavelet packet algorithm to extract the feature vectors. The proposed new algorithm can be illustrated as in Fig. (3) and the abbreviation RCST stands for reconstruction.

Fig. (3) also describes the four main steps of the new algorithm:

**Step 1. Signal extension.**

The high and low frequency coefficients to be decomposed are extended in the way of parabolic extension. Assuming the signal is \( x(a), x(a+1), x(a+2) \), we extend both the left and right ends of the signal, with the extended values expressed as shown in Eq. (3). The extended signal is represented by \( E \) in Fig. (3).

\[
\begin{align*}
\{ x(a-1) &= 3x(a)-3x(a+1)+x(a+2) \\
\{ x(a+3) &= 3x(a+2)-3x(a+1)+x(a)
\end{align*}
\]

**Step 2. Wavelet packet decomposition.**

The new signal obtained from the previous step convolves with \( h_0 \) and \( g_0 \) respectively. With the results of two filtered signals, we use the HF-cut-IF (corresponding to the \( h_0 \)) and LF-cut-IF (corresponding to the \( g_0 \)) operators to process the signals respectively before down-sampling, with the results to be down-sampled to get the single sub-band signal. The conduct of operators HF-cut-IF and LF-cut-IF is as follows: firstly, we make an FFT transform to the two convolving resultant signals. Secondly, we set the amplitude of the high frequency (HF) or low frequency (LF) components to be zero (HF-cut, LF-cut) according to the decomposition filter (high pass or low pass decomposition filter). Finally, an IFFT transform is performed on the components-discarded data. The HF-cut-IF and LF-cut-IF operators process each layer of the signal, which can be expressed as Eqs. (4) and (5).
Fig. (4). Acoustic emission detection schematic diagram.

\[
X(k) = \sum_{n=0}^{N_j-1} x(n)W^{kn}, \quad 0 \leq k \leq \frac{N}{4} \text{ or } \frac{3N}{4} \leq k \leq N_j
\]

\[
X(k) = 0, \quad \text{others}
\]

\[
x(n) = \sum_{k=N}^{N_j-1} X(k)W^{-kn}, n = 0, 1, \ldots, N_j - 1
\]

\[
X(k) = \sum_{n=0}^{N_j-1} x(n)W^{-kn}, \quad \frac{N}{4} \leq k \leq \frac{3N}{4}
\]

\[
X(k) = 0, \quad \text{others}
\]

\[
x(n) = \sum_{k=N}^{N_j-1} X(k)W^{-kn}, n = 0, 1, \ldots, N_j - 1
\]

In Eqs. (4) and (5), \( x(n) \) is the wavelet packet coefficient in \( 2^j \) scale, \( N_j \) is the length of the data in \( 2^j \) scale, \( W = e^{-\frac{j2\pi k}{N}} \), \( k = 0, 1, \ldots, N_j - 1; n = 0, 1, \ldots, N_j - 1 \).

**Step 3. Reconstruction of single sub-band signal.**

The single sub-band signal obtained from Step 2 is firstly up-sampled and then convolved with \( h_1 \) (corresponding to the \( h_0 \)) or \( g_1 \) (corresponding to the \( g_0 \)), with the results to be processed by the operator HF-cut-IF (corresponding to the \( h_1 \)) or LF-cut-IF (corresponding to the \( g_1 \)). Finally, we can get the single sub-band reconstructed signal.

**Step 4. Extracting Eigenvector.**

According to the signal spectrum in normal and leakage conditions and the frequency range of single sub-band reconstructed signal, we select the single sub-band reconstructed signal that has obvious change in normal and leakage conditions and then extract the corresponding characteristic parameters from it to form the feature vectors.

3. EXPERIMENT AND RESULT ANALYSIS

3.1. Experimental Facility

We performed a number of experiments to extract the feature vectors from the measured acoustic emission signal, and the experimental results were analyzed. The “schematic diagram” of the experiment is shown in Fig. (4), we use a 50 m long pipeline with the inner diameter to be 0.03 m. An air compressor is fixed on the front of the pipeline. The valve, pressure pan and flow meter are positioned successively. Pipeline leakage would occur after opening the needle valve in the pipeline. The acoustic emission sensors are installed at both ends of the leak hole, with the distance between the left sensor 1 and leak point being 30 m and the distance between the right sensor 2 and leak point 10 m. Sensor 1, which is at the gas inlet, has the sensitivity of 10 mv/Pa. The sensed data are inputted into the computer after being filtered, pre-amplified and acquired by acquisition card. The diameter of the leakage hole is 0.004 m. When the experiment starts, we turn off the needle valve in the pipeline and fill the pipeline with the 0.5 MPa high pressure gas via the compressor, the data acquisition card starts to work at 200 kHz sampling frequency. After a period of time, we quickly open the needle valve to let the leakage occur and the acquisition card continues to collect the leakage signal data. We repeat the process many times and analyze the collected data.

3.2. Feature Parameter

The pipeline leakage signal is extremely different from the normal signal in a certain frequency band [15]. We can extract the feature vectors from the single sub-band reconstructed signal to detect the leakage state. The feature vectors usually contain time-domain and frequency-domain parameters. The time-domain parameters can be peak value, arrival time [16], rise time [17], time-domain energy (ET) [18], the mean (A), variance (V), skewness (k₄) and kurtosis (k₅).
The frequency-domain parameters can be frequency spectrum distribution [19], frequency-domain energy (EF), peak value (PF) and so on. Although these parameters can represent the characteristics of the signal in some aspects, not all the parameters extracted from the single sub-band reconstructed signal have differences in normal and leakage conditions. In this paper, we choose the obvious changed parameters through experiments analysis to form feature vectors and detect the pipeline leakage by comparing the feature vectors in the normal and leakage conditions. The main parameters are introduced as follows.

\[
ET = \left[ \int |S_i(t)|^2 dt \right] = \sum_{k=1}^{N} |S_{i,k}^j|^2
\]

Where \(S_{i,k}^j\) is the \(k\)-th coefficient of the node \(i\) single sub-band reconstructed signal in the \(j\)-th layer. \(N\) is the length of the single sub-band reconstructed signal.

\[
A = \frac{1}{N} \sum_{k=1}^{N} S_{i,k}^j
\]

\(V\) is the variance and the expression is shown in Eq. (8).

\[
V = \frac{1}{N} \sum_{k=1}^{N} (S_{i,k}^j - A)^2
\]

\(k_3\) stands for the value of probability density distribution curve in the average and the expression [20] is shown by Eq. 9:

\[
k_3 = \sqrt{\frac{1}{6N} \sum_{k=1}^{N} \left( \frac{S_{i,k}^j - A}{\sqrt{V}} \right)^3}
\]

\(k_4\) is the measurement of deflection direction and degree of the statistics data distribution and the expression [20] is:

\[
k_4 = \sqrt{\frac{1}{24N} \sum_{k=1}^{N} \left( \frac{S_{i,k}^j - A}{\sqrt{V}} \right)^4 - 3}
\]

\(EF\) is the single sub-band reconstructed signal energy in frequency-domain and the expression is:

\[
EF = \int |F_{j}^i(f)|^2 df = \sum_{k=1}^{N} |F_{j,k}^i|^2
\]

Where \(F_{j,k}^i\) is the frequency spectrum value of the single sub-band reconstructed signal.

### 3.3. Reconstruction of the Single Sub-band Signal

We analyzed the measured data from sensor 2, with the spectra of the measured signal in normal and leakage conditions shown in Fig. (5) and found that the spectrum distribution of acoustic emission signal is very wide, regardless of whether a leak occurs or not. The value of the spectrum at low frequencies near 60 Hz is big, which is caused by the hose connected with pressure pan vibration, but it does not affect the detection and extraction of leakage signal, because the spacing between the low frequency and leakage characteristic signal spectrum is vast. In addition, we can find the difference between the spectrum of normal and leakage signal is mainly concentrated in the vicinity of 64 kHz. The frequency band nearly 64 kHz is important for us to conduct pipeline detecting, locating and feature vectors extracting.

We use the traditional and improved wavelet packet algorithms to decompose the leak signal and reconstruct the single sub-band signal, and then we compare the performance of the two algorithms. We choose the db10 wavelet to decompose the leak signal to third layer, with the frequency interval of the sub-band signal to be 12.5 kHz. Just as shown in Fig. (2), we get eight single sub-band signals from node 0 to node 7 and then we need to reconstruct the single sub-band signal. Taking the node 7 for example, the flow chart of reconstruction of the single sub-band signal is as shown in Fig. (6) and the \(S_{7}^3\) in the input end is the single sub-band signal of node 7, the output end is the single sub-band reconstructed signal. The frequency ranges of each single sub-band reconstructed signal are 0–12.5 kHz (node 0), 12.5–25 kHz (node 1), 25–37.5 kHz (node 2), 37.5–50 kHz (node 3), 50–62.5 kHz (node 4), 62.5–75 kHz (node 5), 75–87.5 kHz (node 6), 87.5–100 kHz (node 7) if the sampling frequency is 200 kHz. In Fig. (5), we find the spectrum difference between normal and leakage signals mainly occurs in the 62.5–75kHz frequency band, so we reconstruct the single sub-band signal of this frequency band. The reconstructed signal spectra are as shown in Fig. (7).

As shown in Fig. (7a), the single sub-band signal reconstructed by the traditional wavelet packet algorithm contains many redundant frequency components, with some frequency components having relatively large amplitudes; the traditional algorithm brings relatively large disturbance components around 60 kHz and 35kHz. The improved wavelet packet algorithm as shown in Fig. (7b), however, can accurately reconstruct the single sub-band signal without the redundant frequency components.

### 3.4. The Extraction of Feature Vector and Result Analysis

In this paper, we reconstruct the single sub-band signal based on the traditional and improved algorithms and extract the feature vectors from the single sub-band reconstructed signals. We find that, for the single sub-band reconstructed signals, there are big differences in \(ET, EF, PF, k_3\) and variance \(V\) between normal and leakage conditions. So, we choose the five characteristic parameters to form the feature vectors.

In Fig. (5), we can see the difference is mainly concentrated in the vicinity of 64 kHz in normal and leakage conditions, and the frequency range of single sub-band recon-
Study on Extracting Pipeline Leak Eigenvector Based on Wavelet Packet

The Open Automation and Control Systems Journal, 2014, Volume 6

Fig. (5). Sensor measured signal spectrum.

(a) Sensor 2 without leakage

(b) Sensor 2 leakage

Fig. (6). Flow chart of single sub-band signal reconstruction.
Fig. (7). Spectrum of reconstruction signal.

(a) Sensor 2 without leakage

(b) Sensor 2 leakage

Fig. (8). contd…

(a) The time-domain energy
Fig. (8), contd...

(b) The frequency-domain energy

(c) Kurtosis

(d) Frequency-domain peak
structured signal on node 7 is 62.5 ~ 75 kHz, so the difference between normal and leakage signals is contained in the reconstructed signal of node 7, meanwhile, we can find that there is no difference in the reconstructed signal of node 3 (25 ~ 37.5 kHz) and node 6 (50 ~ 62.5 kHz) between normal and leakage signals.

Fig. (8a, b, c, d) and (e) are ET, EF, k₃, PF and V respectively of the single sub-band reconstructed signal in node 3, 6 and 7 in normal and leakage conditions, with the solid line representing the parameters extracted by the traditional algorithm in normal condition, the dashed representing the parameters extracted by the traditional algorithm in leakage condition, the line with plus sign representing the parameters extracted by the improved algorithm in normal condition, and the line with triangle sign representing the parameters extracted by the improved algorithm in leakage condition. According to the previous analysis, we can find that the characteristics parameters of the single sub-band reconstructed signal on node 3 and 6 do not change in normal and leakage conditions. However, the characteristics parameters of the reconstructed signal on node 3 and 6 extracted by the traditional algorithm have big difference under normal and leakage conditions. The unreasonable phenomenon is caused by the mixing effects of the traditional algorithm in reconstructing single sub-band signal. On the contrary, the characteristic parameters extracted by our new algorithm in the two nodes do not change under the two conditions. In addition, the feature parameters extracted by traditional and improved algorithms have differences in amplitude, which are also caused by mixing effects. The single sub-band reconstructed signal of node 7 has extreme difference between normal and leakage conditions, so the characteristic parameters should also have larger difference under the two conditions. As shown in Fig. (8), the characteristic parameters extracted by the new algorithm under the two conditions have significantly large difference. According to the previous analysis, we can get the conclusion that the improved algorithm can overcome the mixing phenomenon in reconstructing signal, exactly extract the feature vectors and improve the accuracy of leak detection, which provides a reliable basis for pipeline leakage detection.

CONCLUSION

The mixing phenomenon will occur when we use the traditional algorithm to reconstruct the single sub-band signal, which will make the extracted feature vectors not the accurate ones. In this article, the authors proposed an improved wavelet packet algorithm to reconstruct single sub-band signal and extract the feature vectors to overcome the disadvantage of traditional wavelet packet algorithm. The proposed algorithm adopts the parabola continuation method and uses FFT and IFFT to process the convolved signal. We performed a number of experiments to extract feature vectors based on the proposed algorithm and the results were compared with that of the traditional algorithm. The eigenvector is formed by five time-frequency parameters chosen through analysis. The experimental results show that the improved algorithm can overcome the mixing effects brought by the up-sample and down-sample process in the traditional “wavelet packet decomposition and reconstruction” and accurately extract the feature vectors. The study of this article provides good foundation for leak detection and positioning.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

ACKNOWLEDGEMENTS

The project is supported by the Shanghai Science and Technology Committee (Grant No. 11510502400).

REFERENCES

Study on Extracting Pipeline Leak Eigenvector Based on Wavelet Packet


The Open Automation and Control Systems Journal, 2014, Volume 6


Received: December 26, 2013 Revised: April 25, 2014 Accepted: May 14, 2014

© Liu et al.; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.