Compound Faults Diagnosis Method Based on Adaptive GST-NMF for Rolling Bearing

Hongwei Luo, Liuyang Song, Mengyang Wang, Huaqing Wang*
School of Mechanical & Electrical Engineering
Beijing University of Chemical Technology
Chao Yang District, Beijing, 100029, P. R. China
hqwang@mail.buct.edu.cn

Lingli Cui
Beijing Engineering Research Center of Precision Measurement Technology and Instruments
Beijing University of Technology
Chao Yang District, Beijing, 100124, P. R. China
cuilingli@bjut.edu.cn

***Abstract*—In order to solve the difficulty of features extraction of compound faults in underdetermined state, this research proposes an approach to extract signal features by combining adaptive generalized S transform (GST) and non-negative matrix factorization algorithm (NMF). The adaptive function (AF) is introduced to optimize GST. The optimized GST is used to process monitored signals to get the time-frequency features matrix. The NMF is improved by Itakura-Saito (IS) divergence. And the dimensionality of the signal time-frequency matrix is reduced by it. After iterative updating, several low-dimensional matrices are obtained. The time-domain waveforms of low-dimensional matrices are reconstructed, and the envelope spectrum analysis is performed to realize compound faults diagnosis. The simulation test and the actual bearing compound fault signals experiment prove that this method can effectively extract compound fault features in underdetermined state and realize bearing compound faults diagnosis.**

Keywords—GST; NMF; features extraction; compound faults diagnosis

I. INTRODUCTION

In the past, Fourier transform (FT), time-frequency analysis [1], variational mode decomposition, wavelet analysis [2], envelope analysis [3], empirical mode decomposition [4, 5], and other techniques have been used to analyze signals widely. Fault features extraction can be realized based on the above techniques combined with statistical methods [6], sparse component analysis, independent component analysis [7, 8], principal component analysis [9], singular value decomposition, autoregressive algorithm and so on. At the same time, automatic diagnosis and effective classification of faults can be realized by using neural network, support vector machine and other machine learning related algorithms [10-13]. The S transform (ST) [14] is proposed by Stockwell et al. in the 1990s. It combines the unique advantages of short-time Fourier transform (STFT) and wavelet transform. It achieves a variable signal decomposition scale for time-frequency analysis. For non-stationary signals, the method can make frequency resolution high when frequency is low, and can make time resolution better when frequency is high. Non-negative matrix factorization algorithm (NMF) [15] became a new mainstream method of blind source separation (BSS) in recent years. It overcomes the shortcomings of some traditional algorithms and is widely used in image processing, compression classification, computer vision, signal features extraction, etc. The theoretical support of NMF is strong, and the decomposition form is simple. The result of the decomposition is more physically meaningful due to the "non-negative" characteristic. In dealing with BSS, NMF requires fewer constraints, but faster convergence and higher decomposition efficiency. However, under the actual working conditions of the rotating machine, the collected bearing vibration signals are complex, the multi-source information is coupled, the SNR is low, and the characteristic components are weak. So when the traditional NMF algorithm is used for processing actual vibration signal, data redundancy will occur, and the features extraction is not ideal.

In summary, this research proposed an mathematical model of signal processing based on improved NMF and ST. The simulation and the bearing compound faults experiment verified the effectiveness of this proposed algorithm.

II. S TRANSFORM ALGORITHM THEORY

A. S Transform

For a collected vibration signal sample $x(t)$, the ST of it is defined as $S(\tau,f)=\int_{-\infty}^{+\infty}x(t)w(t-\tau)e^{-j2\pi ft}\mathrm{d}t$. In this equation, $w(t-\tau)$ is Gaussian window. $w(t)=e^{-t^2/2\sigma^2}/\sigma\sqrt{2\pi}$ and $\sigma=1/|f|$. It is obvious that $\int_{-\infty}^{+\infty}\frac{|f|}{\sqrt{2\pi}}e^{-\tau^2f^2/2}\mathrm{d}\tau=1$. In summary, the classical form of the ST is:

$$S(\tau,f)=\int_{-\infty}^{+\infty}x(t)\frac{|f|}{\sqrt{2\pi}}e^{-f^2(t-\tau)^2/2}e^{-j2\pi ft}\mathrm{d}t \tag{1}$$

It is easy to get that $A(f)=\int_{-\infty}^{+\infty}S(\tau,f)\mathrm{d}\tau$ is the FT of $x(t)$ is, so the inverse ST is defined as

This research is supported by National Nature Science Foundation of China (Grant No.51675035 and No.51805022).

978-1-7281-0860-5/19/$31.00 ©2019 IEEE

$$x(t)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}S(\tau,f)e^{j2\pi ft}\mathrm{d}\tau\mathrm{d}f \tag{2}$$

It can be seen that the essence of the ST is to replace the window function $g(t)$ of the STFT a Gaussian window. σ is a deviation, which is related to f. So the window width has good adjustability, and it can expands and contracts with frequency. It can produce better frequency resolution in a lower frequency environment, and can achieve more accurate time-scale positioning in a higher frequency environment. Although the ST has considerable advantages, it still has its shortcomings. The shortcoming is that the standard deviation σ is constant in any frequency segment, which directly affects the adaptability of the ST algorithm. Moreover, in some applications, time-frequency analysis using ST has the disadvantage of energy dispersion and non-concentration. In response to this, the proposed more adaptive generalized ST improves the inadequacies of the standard ST: $\sigma=1/|f|^{p}$, that is, an adjustment factor is introduced, which functions to achieve controllable standard deviation σ, obtain better window width, and improve time-frequency energy concentration.

Therefore, it can be concluded that after performing GST on signal $x(t)$, the following equation is obtained:

$$S_x^p(\tau,f)=\int_{-\infty}^{+\infty}x(t)\frac{|f|^p}{\sqrt{2\pi}}e^{-f^{2p}(t-\tau)^2/2}e^{-j2\pi ft}\mathrm{d}t \tag{3}$$

In this equation, when $p=1$, the equation is the standard ST, when $p<1$, the window will be widened, and when $p>1$, the window will be narrowed.

B. *Adaptive Generalized S Transform Algorithm*

According to Djurovi's research [16], if $p<0$, the window width will be too large when the frequency is increased, if $p>1$, the window width will be too small, and if $0<p\le 1$, a better processing result can be obtained and the energy distribution of time-frequency diagram will be more concentrated. However, the results of Djurovi's research do not clearly give the criteria for determining the adjustment factor p. Therefore, this paper introduces the density factor ε as the evaluation function of p:

$$\varepsilon=1-N_{[E_p(d,k)=0]}/d\cdot k \tag{4}$$

where d and k are the rows number and columns number of the complex matrix obtained by GST, respectively. $d\cdot k$ is the elements number of the complex matrix. $p=p_1,p_2,...,p_n$. $E_p(d,k)$ represents one element in the complex matrix, d is the row and k is the column, when p is set to a value of $(p_1,p_2,...,p_n)$. After transforming each vector element into scalar element, a threshold will be set according to experience, and the element smaller than the threshold will be set to zero. At this point, the number of zero in all the elements of matrix is $N_{[E_p(d,k)=0]}$. Therefore, the definition of the density factor ε is the ratio of the non-zero elements number and total elements number in the matrix, that is, the degree of the difference between the elements in the matrix, which is helpful for judging the time-frequency concentration.

After initializing a series of p values, a corresponding series of ε values can be obtained according to (10). The smaller the ε value is, the stronger the time-frequency concentration is, and the corresponding p value is better. In theory, the smaller the threshold set by experience, the better the correct selection of p. The more appropriate the threshold is set, the more obvious the comparison between the data will be, and the faster the judgment process will be.

The calculation process of the adaptive GST is as follows:

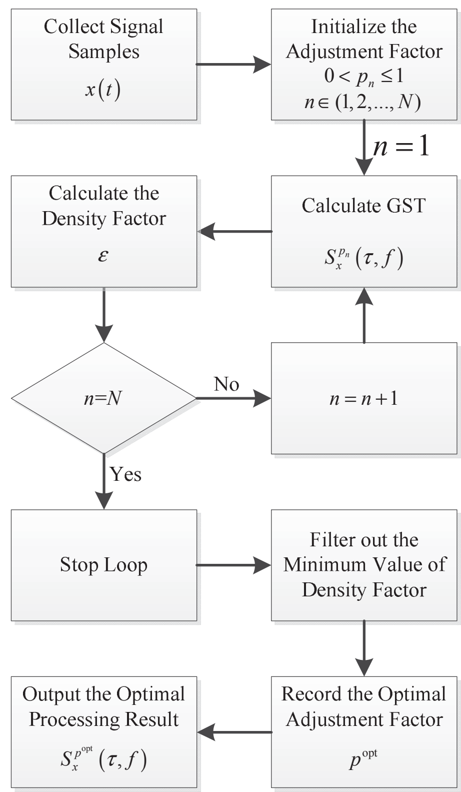

Figure 1. Adaptive GST algorithm flow.

III. NON-NEGATIVE MATRIX FACTORIZATION ALGORITHM

A. *Traditional NMF Algorithm*

Lee formally proposed the NMF algorithm in Nature [17]. The implementation process of the algorithm is mainly based on finding the appropriate matrix $W\in R_+^{m\times r}$ and $H=(h_1,...,h_n)\in R_+^{r\times n}$ for any matrix $V=(v_1,...,v_n)\in R_+^{m\times n}$ to satisfy the following equation: $V_{m\times n}=W_{m\times r}H_{r\times n}$. The three matrices are non-negative. m is sample dimension. r is the rank. n is samples number. The determination of r must usually satisfy $(m+n)r<mn$. And because of $m>>r$, the dimensions of W and H are both smaller than the dimension of V, that is, the compression of the original sample dimension is achieved.

B. Optimized NMF Based on Itakura-Saito Divergence

For the NMF, the selection of the cost function is an important research direction. When using NMF algorithm to separate features from multi-source mixed signal, the merits of loss function selection are related to the strength of features, the quality of data compression, the number of redundant components, and so on. Using Itakura-Saito (IS) divergence as the NMF cost function can effectively enhance fault features, enhance data dimensional compression, reduce redundant components of features extraction. It facilitate features extraction, signals reconstruction and faults identification. The expression of IS divergence is $q_{IS}(a,b)=\sum_i(\frac{a_i}{b_i}-\ln\frac{a_i}{b_i}-1)$, therefore, it can be concluded that the optimized cost function of the NMF algorithm based on IS divergence is: $S_H(W)=\frac{1}{N}\sum_{n=1}^{N}q_{IS}(v_n,Wh_n)$.

The optimized NMF algorithm [18] flow is as follows:

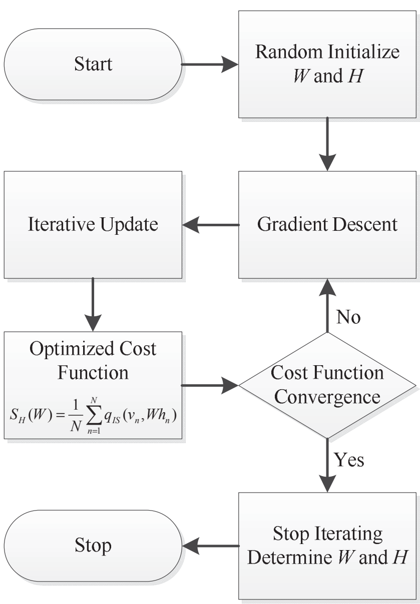

Figure 2. Optimized NMF algorithm flow.

IV. BEARING COMPOUND FAULTS DIAGNOSIS PROCESS BASED ON ADAPTIVE GST-NMF

The bearing compound faults diagnosis can be realized by combining the proposed adaptive GST with the NMF algorithm optimized by the IS divergence [19].

The proposed method of faults diagnosis based on adaptive GST-NMF is shown as follows:

1) The bearing vibration signal is monitored and collected, and the time domain signal samples are obtained, and the adaptive GST algorithm is performed to get the time-frequency features matrix.

2) The energy value of the features matrix is taken, and the optimized NMF algorithm is used for dimensionality reduction to obtain low-dimensional matrix factors W and H.

3) The base matrix W and the coefficient matrix H are reconstructed in a space of lower dimension, and the time-frequency information is returned to the time domain by using inverse generalized S transform (IGST) to obtain the reconstructed waveform of the feature components.

4) Envelope spectrum analysis is performed on the reconstructed waveform to realize faults diagnosis finally.

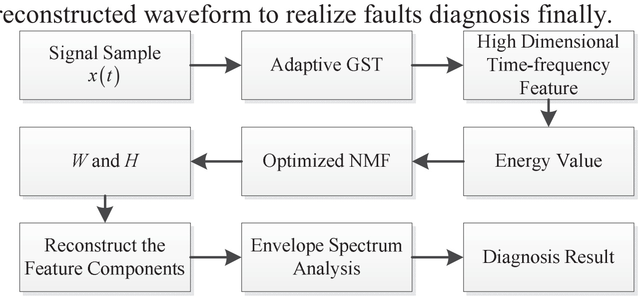

Figure 3. Bearing compound faults diagnosis flow.

V. ADAPTIVE GST SIMULATION ANALYSIS

According to the flow shown in Fig. 3, the performance of the algorithm is tested and verified by using the simulated signal $x(t)$. Six feature signal components with a sampling frequency of 1000Hz are randomly generated, and the time segment of 1s is taken to form the matrix B. At the same time, the mixing matrix A is randomly generated:

$$x_1(t)=\sin\left(5^t\cdot 30\pi t\right) \qquad x_2(t)=\sin\left(9^t\cdot 70\pi t\right)$$

$$x_3(t)=\sin\left(600\pi t\right) \qquad x_4(t)=\sin\left(500\pi t\right)$$

$$x_5(t)=\sin\left(8\pi t^4+60\pi t\right) \qquad x_6(t)=\sin\left(6\pi t^4+20\pi t\right)$$

The simulated signal $x(t)$ is a mixture of matrix A and matrix B:

$$x(t)=A\cdot B \tag{5}$$

$$A=[0.957,0.485,0.800,0.142,0.422,0.916] \tag{6}$$

$$B=[x_1(t),x_2(t),x_3(t),x_4(t),x_5(t),x_6(t)]^{\mathrm{T}} \tag{7}$$

The simulation experimental results are as follows:

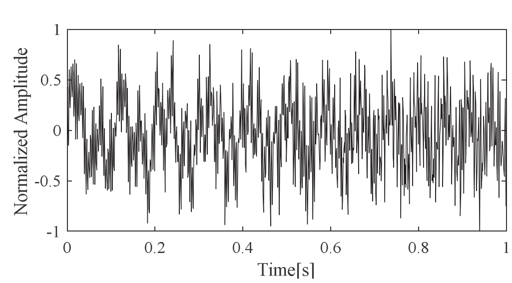

Figure 4. Simulated signal $x(t)$ waveform.

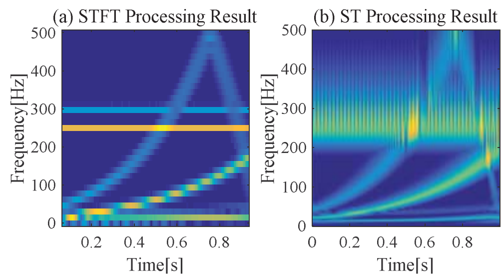

Figure 5. STFT time-frequency features diagram and standard ST time-frequency features diagram.

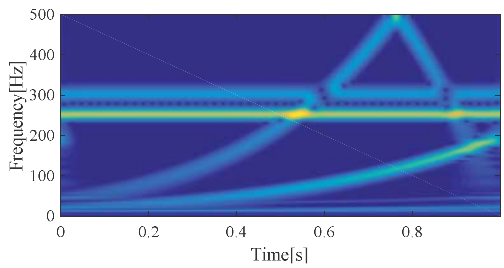

Figure 6. Adaptive GST time-frequency features diagram.

Fig. 4 is the waveform of the simulated mixed signal. Fig. 5 is the result obtained by STFT and ST. As can be seen from Fig. 5(a), the result obtained after STFT processing is relatively low resolution in both. In Fig. 5(b), the standard ST can make the resolution of time-frequency higher, but there is data redundancy in signal processing. These two situations are not conducive to the realization of BSS, and it is not conducive to accurate faults diagnosis in the later stage.

Fig. 6 is the result of processing with the proposed adaptive GST algorithm. At this point, the theoretical optimal p value is 0.8, and the corresponding theoretical minimum density factor is 51.27%. Comparing Fig. 6 with Fig. 5, the conclusion could be got that the time-frequency resolution of the proposed adaptive GST algorithm is relatively high, and the processing result is more excellent, the data redundancy is greatly reduced. It facilitates more accurate faults diagnosis in the later stage.

Four groups of standard GST comparative experimental groups were set up so as to further prove the excellent performance of this proposed adaptive GST. The experimental parameters and results are as follows:

TABLE I. COMPARATIVE EXPERIMENTS PARAMETERS INFORMATION

	Test Group of Adaptive GST	**Four Comparative Groups of Standard GST**			
p	$p^{opt}=0.8$	0.1	0.4	1.1	1.4
ε	$\varepsilon_{min}=51.27\%$	99.28 %	69.33 %	53.64 %	67.09 %
Result	Fig. 6	Fig. 7(a)	Fig. 7(b)	Fig. 7(c)	Fig. 7(d)

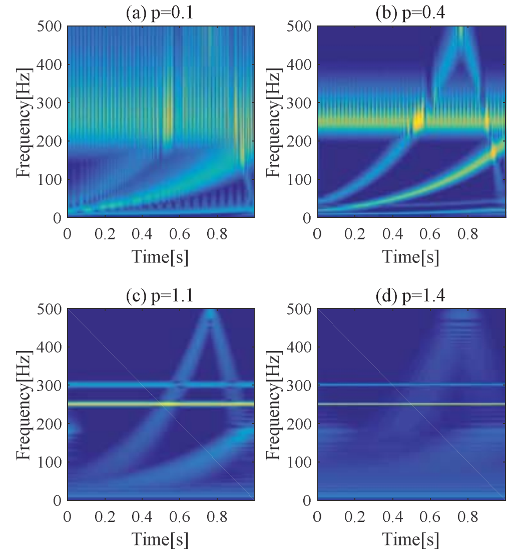

Figure 7. The time-frequency diagram of comparative experimental result.

According to Table I, in the four sets of standard GST comparative experiments, the ε-values calculated according to the artificially preset p-values are larger than the ε-values automatically obtained by the adaptive GST algorithm, and the validity of the minimum ε-value calculated by the proposed adaptive GST algorithm is verified. Moreover, comparing Fig. 7 with Fig. 6, it can be concluded that the performance of the standard GST is not as good as the adaptive GST proposed in this research, both in terms of data redundancy and time-frequency resolution.

VI. EXPERIMENTAL VERIFICATION

To prove the advantage of this proposed method again, the faults diagnosis experiment was carried out using the NTN N204 cylindrical roller bearing vibration signal. The bearing structure parameters are as follows:

TABLE II. EXPERIMENTAL BEARING RELATED PARAMETERS

Number of Bearing Rollers z	**Contact Angle** α	**Diameter of Inner** D	**Diameter of Outer** d
10	0 rad	20 mm	47 mm

Figure 8. The bearing used in the experiment and the bearing vibration test bench.

A defect having a width of 0.5mm and a depth of 0.15mm is machined on the outer ring and the roller by wire cutting technology. And the bearing rotation unbalance fault is set.

During the experiment, the signals of bearing vibration are collected by some acceleration sensors. The frequency of sampling is 100kHz. The time of a single sampling is 10s. And two sets of experiments were performed respectively by setting the motor speed to 1300r/min and 900r/min. According to Table II and theoretical fault features frequency calculation formula of bearing, they can be obtained as shown in Table III:

TABLE III. FAULT FEATURES FREQUENCY

Rotating Speed	Outer Ring	Rolling Element	Unbalance
900r/min	60Hz	74Hz	15Hz
1300r/min	86Hz	101Hz	22Hz

Firstly, the experiment with the motor speed of 1300r/min and 900r/min are performed. And 1s time slices of the two vibration signal are intercepted for processing.

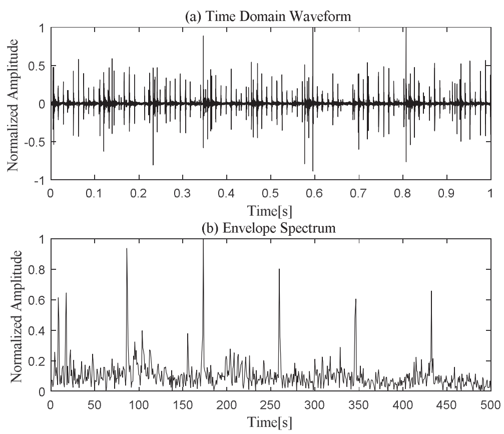

Figure 9. 1300r/min original signal waveform and envelope spectrum.

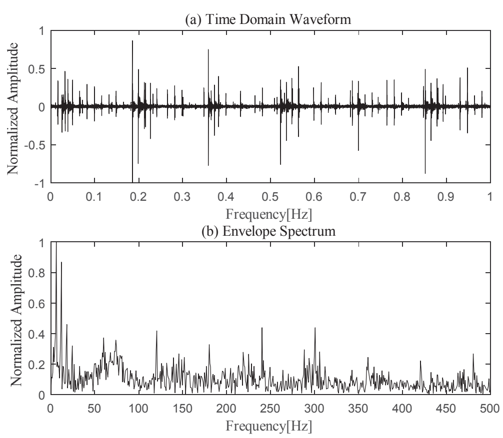

Figure 10. 900r/min original signal waveform and envelope spectrum.

From Fig.9 and Fig. 10, the signals contains impact components, indicating that the bearing has failed, but the features of the impact components are not obvious, and the faults diagnosis cannot be directly performed.

According to the proposed method, the adaptive GST is performed on the acquired two sets of vibration signals. A high-dimensional features matrix X of the signal is got. The energy matrix of X is used as the input of optimized NMF algorithm. After continuous iterative decomposition, the base matrix and the coefficient matrix are finally determined. The matrices W and H are reconstructed in the subspace, and the inverse GST is performed to put them into time domain so that the reconstructed signals could be obtained. The spectrum analysis of the reconstructed signals is as shown below:

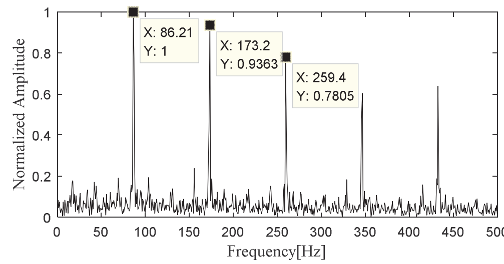

Figure 11. 1300r/min bearing outer ring fault feature spectrum.

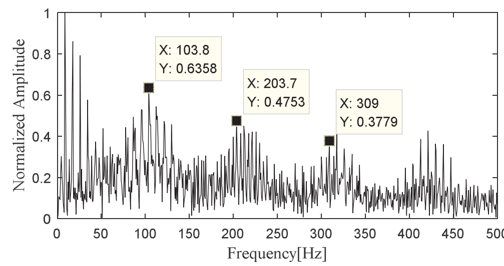

Figure 12. 1300r/min bearing rolling element fault feature spectrum.

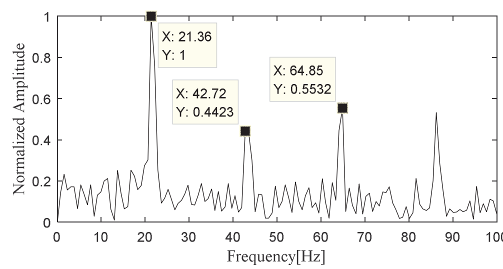

Figure 13. 1300r/min bearing unbalance fault feature spectrum.

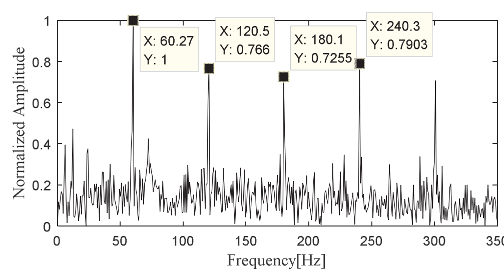

Figure 14. 900r/min bearing outer ring fault feature spectrum.

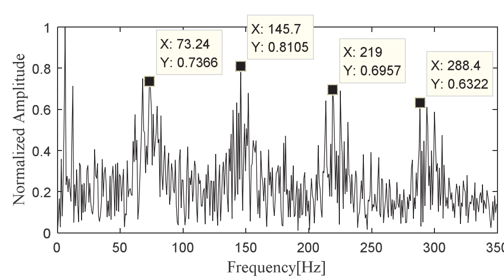

Figure 15. 900r/min bearing rolling element fault feature spectrum.

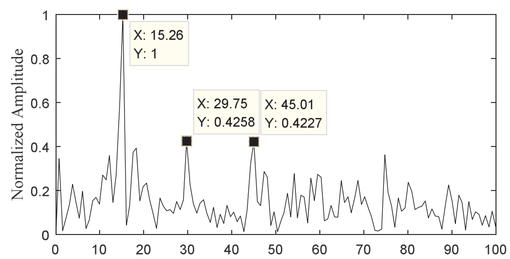

Figure 16. 900r/min bearing unbalance fault feature spectrum.

It can be clearly seen from the results of experiment that this method could separate the outer ring fault feature, rolling element fault feature and unbalance fault feature from the original signal better. The three feature frequencies obtained by the experiment are basically consistent with the theoretical calculation values shown in Table III, and the harmonic components are obvious. So the compound faults have been effectively diagnosed.

VII. Conclusions

In this research, a compound faults diagnosis approach based on the adaptive GST-NMF is proposed. The adaptive optimal generalized S transform is realized by introducing the density factor as the evaluation function of GST. At the same time, the Itakura-Saito divergence is introduced as the loss function to optimize and improve the NMF algorithm. The high-dimensional time-frequency features matrix obtained by adaptive GST is used as the input of the improved NMF. Iterative decomposition is performed under the control of the loss function, and finally the optimal matrix factors W and H are obtained. The matrix factors are reconstructed to obtain the reconstructed signal and the envelope spectrum analysis is performed. Finally, the separation of compound fault features is achieved.

In the algorithm performance simulation and experiment, the separation and extraction of compound fault features are achieved. Therefore, this proposed approach is of great significance of rotating machinery compound faults diagnosis, and it has a fairly broad application prospect. Once applied to actual engineering, it can generate high economic benefit.

References

[1] W. J. Wang and P. D. McFadden, "Early detection of gear failure by vibration analysis—II. Interpretation of the time-frequency distribution using image processing techniques," Mechanical Systems and Signal Processing, vol. 7, no. 3, pp. 205-215, 1993.

[2] J. Lin and L. Qu, "Feature extraction based on morlet wavelet and its application for mechanical fault diagnosis," Journal of Sound & Vibration, vol. 234, no. 1, pp. 135-148, 2000.

[3] R. B. Randall, J. Antoni, and S. Chobsaard, "The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals," Mechanical Systems and Signal Processing, vol. 15, no. 5, pp. 945-962, 2001.

[4] Z. K. Peng, P. W. Tse, and F. L. Chu, "A comparison study of improved Hilbert-Huang transform and wavelet transform: Application to fault diagnosis for rolling bearing," Mechanical Systems and Signal Processing, vol. 19, no. 5, pp. 974-988, 2005.

[5] H. Oehlmann, D. Brie, M. Tomczak, and A. Richard, "A method for analysing gearbox faults using time-frequency representations," Mechanical Systems and Signal Processing, vol. 11, no. 4, pp. 529-545, 1997.

[6] Y. Lei, Z. He, Y. Zi, and X. Chen, "New clustering algorithm-based fault diagnosis using compensation distance evaluation technique," Mechanical Systems and Signal Processing, vol. 22, no. 2, pp. 419-435, 2008.

[7] C. Junsheng, Y. Dejie, and Y. Yu, "A fault diagnosis approach for roller bearings based on EMD method and AR model," Journal of Vibration Engineering, vol. 20, no. 2, pp. 350-362, 2004.

[8] A. Widodo, B. S. Yang, and T. Han, "Combination of independent component analysis and support vector machines for intelligent faults diagnosis of induction motors," Expert Systems with Applications, vol. 32, no. 2, pp. 299-312, 2007.

[9] W. Li, T. Shi, G. Liao, and S. Yang, "Feature extraction and classification of gear faults using principal component analysis," Journal of Quality in Maintenance Engineering, vol. 9, no. 2, pp. 132-143, 2003.

[10] B. Samanta and K. R. Al-Balushi, "Artificial neural network based fault diagnostics of rolling element bearings using time-domain features," Mechanical Systems and Signal Processing, vol. 17, no. 2, pp. 317-328, 2003.

[11] Y. Lei, Z. He, Y. Zi, and Q. Hu, "Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAS," Mechanical Systems and Signal Processing, vol. 21, pp. 2280-2294, 2007.

[12] V. Sugumaran and K. I. Ramachandran, "Automatic rule learning using decision tree for fuzzy classifier in fault diagnosis of roller bearing," Mechanical Systems and Signal Processing, 2007.

[13] B. Samanta and C. Nataraj, "Application of particle swarm optimization and proximal support vector machines for fault detection," Swarm Intelligence, 2009.

[14] R. G. Stockwell, L. Mansinha, and R. P. Lowe, "Localization of the complex spectrum: The S transform," IEEE Transactions on Signal Processing, vol. 44, no. 4, pp. 998-1001, 2002.

[15] P. O. Hoyer, "Non-negative Matrix Factorization with Sparseness Constraints," Journal of Machine Learning Research, vol. 5, no. 1, pp. 1457–1469, 2004.

[16] I. Djurović, E. Sejdić, and J. Jiang, "Frequency-based window width optimization for S-transform," AEU-International Journal of Electronics and Communications, vol. 62, no. 4, pp. 245-250, 2008.

[17] L. Stankovi, "A Measure of some time-frequency distributions concentration," Signal Processing, vol. 81, no. 3, pp. 621-631, 2001.

[18] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," Nature, vol. 401, pp. 788-791, 1999.

[19] H. W. Luo, M. Y. Wang, L. Y. Song and H. Q. Wang, "Feature Extraction Method Based on Improved NMF for Rolling Bearing Compound Fault," International Journal of Comprehensive Engineering, Part A, vol. 7, no. 1, pp. 1-10, 2018.