Journal of Mechanical Science and Technology 22 (2008) 2151~2162

Journal of Mechanical Science and Technology

www.springerlink.com/content/1738-494x DOI 10.1007/s12206-008-0714-0

Fault diagnosis of partial rub and looseness in rotating machinery using Hilbert-Huang transform[†]

Seung-Mock Lee and Yeon-Sun Choi*

School of Mechanical Engineering, Sungkyunkwan University, 300 Chunchun-dong, Jangan-gu, Suwon, Kyunggi-do, 440-746, Korea

(Manuscript Received January 14, 2008; Revised July 17, 2008; Accepted July 17, 2008)

Abstract

Partial rub and looseness are common faults in rotating machinery because of the clearance between the rotor and the stator. These problems cause malfunctions in rotating machinery and create strange vibrations coming from impact and friction. However, non-linear and non-stationary signals due to impact and friction are difficult to identify. Therefore, exact time and frequency information is needed for identifying these signals. For this purpose, a newly developed time-frequency analysis method, HHT (Hilbert-Huang Transform), is applied to the signals of partial rub and looseness from the experiment using RK-4 rotor kit. Conventional signal processing methods such as FFT, STFT and CWT were compared to verify the effectiveness of fault diagnosis using HHT. The results showed that the impact signals were generated regularly when partial rub occurred, but the intermittent impact and friction signals were generated irregularly when looseness occurred. The time and frequency information was represented exactly by using HHT in both cases, which makes clear fault diagnosis between partial rub and looseness.

Keywords: Fault diagnosis; HHT (Hilbert-Huang Transform); HT (Hilbert transform); EMD (Empirical Mode Decomposition); Rotating machinery; Partial rub; Looseness

1. Introduction

The performance of rotating machinery such as motors, compressors, pumps and gas turbines can be raised when the clearance between the rotor and the stator is made as small as possible. However, the rotor can come in contact with the stator during whirling due to imbalance or misalignment of the rotor.

A protrusion at the stator leads to a partial rub during the whirling motion. The partial rub characterizes intermittent contact which persists, becomes more severe, and leads to higher vibration levels. A partial rub can be one of the most damaging malfunctions in rotating machinery. Looseness is also common in rotating machinery. The looseness caused by manufacturing error, thermal expansion or wear exists between the shaft and the bearing or the pedestal in rotating machinery. The effects of clearance due to a partial rub or looseness must be understood to reduce the noise and the vibration, and maintain the performance of the rotating machinery. Therefore, an early diagnosis of the partial rub and the looseness would be beneficial for the safe operation of rotating machinery. However, the partial rub and the looseness cannot be distinguished easily by measured signals.

Rubbing has been studied extensively. Black [1] treated the rubbing phenomenon that occurs when rotating elements contact non-rotating elements in a jet engine. Choi [2] demonstrated the super-harmonics and sub-harmonics in the partial rub with an experimental set-up. Peng [3] discriminated the occurrence of a partial rub by Fourier spectrum and time-frequency analysis. Peng [4] also diagnosed a partial rub by using wavelet transform. However,

[†] This paper was recommended for publication in revised form by Associate Editor Eung-Soo Shin

^{*}Corresponding author. Tel.: +82 31 290 7440, Fax.: +82 31 295 1937 E-mail address: vschoi@yurim.skku.ac.kr

[©] KSME & Springer 2008

looseness has not been studied as extensively as the partial rub and there are only few researches on it. Choi [5] diagnosed the looseness of rotating machinery by using the correlation dimension.

When the partial rub and looseness occur, output signals become non-linear and non-stationary. Therefore, to diagnose an occurrence of the partial rub and the looseness, the time and the frequency information must be detected exactly to diagnose the faults. The conventional frequency analysis method, Fourier spectrum, does not contain time information about a signal. Therefore, it is not proper to use to detect the faults. As time-frequency analysis methods, STFT (short time Fourier-transform) and Wavelet are widely used. However, these methods exactly determine the characteristics of the partial rub and the looseness.

HHT (Hilbert Huang transform) is a newly developed time-frequency analysis method for the nonlinear and non-stationary. [7] In this paper, to identify the clearance effect in rotating machinery, HHT is applied properly with some modifications for the original procedure of HHT.

2. Hilbert Huang transform

HHT is composed of HT (Hilbert transform) and EMD (empirical mode decomposition). To obtain a meaningful parameter through HT, the signal must be a single oscillation. Vibration signals are generally composed of numerous single oscillations. To analyze the general vibration signals by using HT, the signal should be divided into several single signals by using EMD. Then each divided signal becomes a single oscillation, called IMF (intrinsic mode function). The HHT method analyzes each IMF by using HT.

2.1 Hilbert transform

HT (Hilbert transform) can verify just one instantaneous frequency for a mono-component signal. The instantaneous frequency is an inherent characteristic of the signal. The Hilbert transform y(t) of a signal x(t) is defined as:

$$y(t) = \frac{P}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau$$
(1)

Where x(t) and y(t) are formed into complex

numbers, then analytical signal z (t) is constructed from these numbers.

$$z(t) = x(t) + iy(t)$$
⁽²⁾

Where the analytical function z(t) can be represented as an exponent form, and then z(t) is expressed as follows,

$$z(t) = a(t)e^{i\theta t}$$
(3)
Where $a(t) = \sqrt{x^2(t) + y^2(t)}$
 $\theta(t) = \tan^{-1}\left(\frac{y(t)}{x(t)}\right)$

a(t) is the instantaneous amplitude that indicates the energy of the signal x(t) in the time domain and $\theta(t)$ is the instantaneous phase. Then, the instantaneous frequency can be defined by differentiating $\theta(t)$. If the signal is mono-component, the instantaneous frequency becomes a meaningful parameter.

2.2 Empirical mode decomposition

EMD (empirical mode decomposition) converts a multi-component signal into several monocomponent signals, which are called IMF (intrinsic mode function). IMF requires two necessary conditions: (1) all the maxima of the signal should be larger than 0, and all the minima should be smaller than 0, and (2) the envelope of the maxima and the minima must be symmetric. Then, IMF can be analyzed by using HT.

The procedures of the EMD method are as follows. First, find all the maxima of the signal and envelop these maxima. Second, repeat this procedure on the minima of the signal. Then, the upper and the lower boundary of the signal can be made. Third, calculate the mean value between the upper and lower boundary. Fourth, subtract the mean value from the original signal. Then, the residue becomes another signal after the first shifting.

The EMD procedure is shown in Fig. 1(a)-(c). The result of the first part of the procedure is shown in Fig. 1(c); however, the mean values of the signal have peaks near the 22th point. This means that the signal does not satisfy the second condition of symmetry. In this case the result of HHT is not clear. Therefore, to obtain a more exact IMF, the procedure needs to be repeated as Eq. (4).





Fig. 1. Illustration of the sifting processes.

$$c_{11} = x(t) - m_{10}$$

$$c_{12} = c_{11} - m_{11}$$

$$\vdots$$

$$c_{1k} = c_{1(k-1)} - m_{1(k-1)}$$
(4)

Where the m_{lk} is the mean value between the maxima and the minima of c_{lk} . These procedures are repeated under two criteria. The first criterion is that all the maxima are greater than 0, and all the minima are less than 0. This is an absolutely necessary condition because IMF must be mono-component. However, the symmetry of the signal, which is the necessary condition of IMF, can distort the physical sense of both the amplitude and frequency modulations; therefore, there is a limit to the size of the standard deviation, SD, which is computed from the two consecutive sifting results by Eq. (5). Generally, the SD value is 0.2-0.3 [7].

$$SD = \sum_{t=0}^{T} \left[\frac{\left| c_{1(k-1)}(t) - c_{1k}(t) \right|^2}{c_{1(k-1)}^2} \right]$$
(5)

After the *k*th shifting, if c_{1k} satisfies these conditions, then that is the first component c_1 of the signal. It can be also called *IMF*₁. The result after five shiftings is shown in Fig. 1(d).

To get another component, remove the first component c_1 from the original signal, and then repeat procedure of Eq. (4). All EMD procedures are stopped when the number of the maxima or the minima of the residue signal is smaller than 2. If the residue signal cannot satisfy this number, the residue signal is assigned as the final residue signal. Then, n components and the final residue signal are generated and can be represented as Eq. (6).

$$x(t) = \sum_{i=1}^{n} c_i + r_n$$
(6)

Then, each IMF can be analyzed by using the Hilbert transform since the IMF is nearly monocomponent. And the instantaneous amplitude and frequency can be calculated from Eq. (3). The signal can be represented in the following form:



Fig. 2. Decomposed signal using EMD method.

$$x(t) = \sum_{i=1}^{n} a_i(t) \exp\left(i \int \omega_i(t) dt\right)$$
(7)

Eq. (7) can be represented as a three-dimensional map with instantaneous amplitude, frequency and time. For example, the signal of the rotor, which has a partial rub, can be decomposed as shown in Fig. 2 by using EMD. The signal consists of three components and a residue.

2.3 Examples of HHT

To demonstrate the advantage of HHT, two signals were analyzed by using HHT and compared with CWT (continuous wavelet transform).

$$sig_{1}(t) = 2\sin(6\pi t) + \sin(72\pi t), \ 0 \le t \le 2 \ (s)$$

$$sig_{2}(t) = \begin{cases} \sin(6\pi t), & 0 \le t \le 1 \ (s) \\ 2\sin(24\pi (t-1)^{2}), & 1 \le t \le 2 \ (s) \end{cases}$$
(8)

 Sig_1 consisted of 3 Hz and 36 Hz linearly, and the 3 Hz component had a larger magnitude than that of the 36 Hz component. The analyzed results of sig_1 by using CWT and HHT are shown in Fig. 3. The result of CWT shows that more energy is concentrated at 3 Hz than 36 Hz; however, the energy is spread widely, so it is difficult to analyze exactly. On



Fig. 3. Comparison for sig1.

the other hand, the result of HHT gave sharp and constant resolutions in both the high and low frequency regions. The 3 Hz having the larger energy component was also identified exactly.

 Sig_2 consisted of only 3 Hz until 1 sec., and the frequency shifted with the time function after 1 sec. The results are shown in Fig. 4. The result of CWT is also spread in the relatively lower frequency region, and at about 1 sec, which was the boundary of shifting frequency, was not clear. However, in the result of HHT, the shift of the frequency was shown clearly at 1 sec representing frequency information clearly. As stated above, HHT gave sharper and higher resolution than CWT in both the time and frequency regions. Therefore, complex signals were analyzed exactly and clearly by using HHT.

2.4 Improvement of HHT

Even though HHT could represent complex signals with high resolution, the EMD procedure had some problems.



Fig. 4. Comparison for sig2.

First, if there were no values at the end of the procedure enveloping the maxima and the minima, the spline line would become divergent as shown in Fig. 5(a). Then, the wrong frequency information might result at the ends of the signal; this is called the end effects. In this research, the maxima and the minima were calculated as numbers n and m, respectively, at the EMD procedures, and the end values were set at both sides, and then the maxima and the minima could be represented as Eq. (9),

$$maxima = [\max_0, \max_1, \max_2, \cdots, \max_{n+1}]$$
$$minima = [\min_0, \min_1, \min_2, \cdots, \min_{m+1}]$$
(9)

Where the real values ranged from 1 to n and m, respectively, and four values were added at the ends to prevent the end effects. These values were assigned as the mean value of the near two values as Eq. (10).

$$\max_{0} = \frac{\max_{1} + \max_{2}}{2}, \ \max_{n+1} = \frac{\max_{n-1} + \max_{n}}{2}$$

$$\min_{0} = \frac{\min_{1} + \min_{2}}{2}, \ \min_{n+1} = \frac{\min_{n-1} + \min_{n}}{2}$$
(10)



Fig. 5. End effects.

As a result of this assignment, the end effects could be removed, as shown in Fig. 5(b). This simple method can reduce the calculation time and prevent frequency distortion at the ends.

Second, HHT needs enough data since the instantaneous frequency was defined by differentiating the instantaneous phase. Huang [7] stated that there should be more than 5 data at least to define the instantaneous frequency. In this research, the frequency information was verified empirically. The reference signals of 30 Hz, 100 Hz and 1100 Hz were prepared, and the signals were measured with the different sampling numbers, which are shown in Fig. 6. As a result, the concerned frequency band was determined by the sampling number, and the frequency of interest was sampled 12.5 times. The sampling number of 12.5 times was more than double that in the case of FFT. Therefore, HHT takes a long time to calculate for a signal of relatively high frequency and long time span since the data size becomes larger. Nevertheless, it



Fig. 6. Frequency definition of HHT.



Fig. 7. Measured points for impact test.

was faster than CWT, which was defined with a convolution integral.

3. Experiment of the rotor with clearance

3.1 Impact test of the rotor

An RK-4 Rotor kit, manufactured by Bently Nevada Co., was used in the experiment as shown in Fig 7. The impact test for the shaft of the rotor was performed to find the natural frequency of the shaft. An accelerometer was attached at the end of the shaft, and the shaft was impacted with an impact hammer at the center of the shaft since the partial rub was occurring at the center. The signals measured by the accelerometers were stored in a computer through A/D conversion, and analyzed with MATLAB. The frequency response of the shaft is shown in Fig. 8. The highest peak is observed at 151.3 Hz, which is the dominant natural frequency affecting the acceleration responses at the top of the bushing.



Fig. 8. Result of impact test.



Fig. 9. Schematic diagram of the rotor system.



Fig. 10. Vibration responses with the variation of clearances and speeds.

3.2 Experiment of partial rub

The contact between the rotor and the stator was accomplished by making a protrusion of a brass screw bolt at the top of the stator, as shown in Fig. 9. If the runout of the rotor exceeds the given clearance between the protrusion and the shaft during the whirling motion of the rotor, the shaft comes into contact with the protrusion. The running speed of the rotor was measured by a photo sensor with a spur gear having 60 teeth. The accelerometer for measuring the vibration was attached at the top of the bushing. Two cases of the partial rub were investigated: one was a



Fig. 11. Different sizes of the bushing.



Fig. 12. Vibration Tesponse with the variation of looseness and speeds.

weak partial rub and the other was a severe partial rub. For the weak one, the clearance was set to 3 mm and for the severe one, 2 mm. The vibration response with the variation of the rotating speed for each clearance is shown in Fig. 10. The peaks moved to the higher speed region when the partial rub became more severe. The vibration peak for the weak partial rub occurred at 1700 rpm. However, as the clearance decreased to 2 mm, the peak moved to 1790 rpm. If a partial rub occurred, the impact between the rotor and the stator occurred. Therefore, to check the impact, the duration of data analysis was set to 0.2 sec, which corresponds to the interval of six revolutions.

3.3 Experiment of looseness

To characterize the vibration signal upon the looseness, an experiment was performed for six different bushings. The diameters of the bushings ranged from 10 mm to 10.5 mm at the interval of 0.1 mm, as shown in Fig. 11. A 10 mm diameter bushing means almost zero looseness. The clearance increased from 0.1 mm to 0.5 mm as the diameter of the bushing

increased. The position of the accelerometer was the same as it was in the partial rub experiment. The responses of the vibration with respect to the rotation speed are depicted in Fig. 12. These responses were similar, regardless of the change of looseness size. The peak vibration occurred near 1580 rpm, which is believed to be the critical speed of the rotor system. However, there were several small peaks due to the looseness. To characterize the looseness depending on lubrication, two conditions of lubrication and dryfriction were investigated in this study. The looseness generated complex signals which contained many frequencies due to impact and friction, which were different from those of the partial rub that showed a regular impact signal. Therefore, the duration of data analysis was set to be relatively long to verify the characteristic of the looseness.

4. Conventional signal processing

The Fourier spectrum is widely used for frequency

analysis, and STFT (short time Fourier transform) and CWT (continuous wavelet transform) are used as time-frequency analysis. Signal analysis by using HHT was compared with those methods in this study.

The results of the Fourier spectrum are shown in Fig. 13(a). The rotation frequency was 28. 3 Hz when the rotor rotated at the speed of 1700 rpm. The multiple components of the rotating frequency occurred due to the nonlinearity from impact. However, it showed limitations in recognizing the characteristics of the signal since there was no time information. Fig. 13(b) shows the result of the signal analysis of the rotor system by using STFT. To analyze the signal, the interval was divided by ten since six impacts occurred during 2 sec during the partial rub. The six impact components were generated in the partial rub. However, the resolution was not enough to confirm the time and frequency information. However, it provided constant resolution since it used the same window for the analysis. The signal analysis using CWT is shown in Fig. 13(c). There are many types of



Fig. 13. Comparison of HHT with other conventional methods.



Fig. 14. Partial rub with clearance 3 mm.

mother wavelets. However, the Morlet wavelet is particularly attractive since the Morlet wavelet function is an exponentially sinusoidal signal, and a damped sinusoidal is the common response of many dynamical systems. Moreover, the Morlet wavelet has a single frequency. If a signal correlates highly with a scaled Morlet wavelet, then the frequency of the wavelet indicates the frequency of the signal to be analyzed. [4] In this study the Morlet wavelet was used for comparison. Although six impacts were discovered, the peaks were spread widely at both the time and the frequency region, and the scales, which were in the relatively lower frequency region, were not clear enough for recognition of the exact information. On the other hand, the result from HHT had clearer resolution than those obtained by any other method as shown in Fig. 13(d). The Gaussian filter [7] was used to smoothen the result, which can show the impacts exactly. Furthermore, the time and the frequency information were clear.





Fig. 15. Partial rub with clearance 2 mm.

5. Signal processing using HHT

5.1 Partial rub

The partial rub between the rotor and the stator was generated when the clearance was less than 3 mm in the experimental set-up. The experiment was performed for clearances of 3 mm and 2 mm, respectively. The partial rub became severe as the clearance decreased. Two cases, 1700 rpm and 1790 rpm, were chosen for the analysis. And Fig. 14 and Fig. 15 had the same scale for fault diagnosis, but Fig. 14(b) was enlarged since the amplitude was small.

The results for the clearance 3 mm are shown in Fig. 14. There was one impact per revolution at 1700 rpm; this impact is shown in Fig. 14(a). The rotating frequency and its third harmonic component were shown throughout the period and 151 Hz, which corresponds to the natural frequency of the shaft. As the protrusion contacted the shaft periodically, the shaft



Fig. 16. Frequency analysis for looseness.

was excited with its natural frequency. In Fig. 14(b), the contacts did not occur at 1790 rpm in the case of the weak partial rub, so the impacts were not found, and the rotating frequency and its second harmonic component appeared throughout the measurement. Although the *z*-axis scale was exaggerated more than that of 1700 rpm, the difference of the magnitude was almost negligible.

The results in the case of the clearance 2 mm are shown in Fig 15. This case represents a more severe partial rub than the case of the clearance 3 mm. There were similar components to the weak partial rub at the rotating frequency and its third harmonic component throughout the time and the impact components occurred regularly. However, two or three impacts were generated per revolution at clearance 2 mm, and these are shown in Fig. 15(a). The vibration level at 1790 rpm was almost the largest. There were more severe impacts, three or more as shown in Fig. 15(b), in this case than any other cases. The signal analysis for the partial rub using HHT showed better resolutions than any other conventional method.

5.2 Looseness

The Fourier spectrum at the dry-friction condition in the case of 1100 rpm is shown in Fig. 16. The frequencies between 120 Hz and 180 Hz are increased as the looseness becomes greater than 0.2 mm. The frequency components are related to the nonlinearity due to the looseness, which needs to be investigated.

5.2.1 Lubrication condition

The results of HHT analysis for the looseness cases at the lubrication condition are shown in Fig. 17. The



Fig. 17. HHT results at lubrication condition.

clearances ranged from 0.0 mm to 0.5 mm and the rotating speeds were the 1100 rpm and the relatively lower speed of 1580 rpm. The results of the lower speed are shown in Fig. 17(a). The frequency components, which were approximately between 120 Hz and 180 Hz, appeared weakly. In case of the 1580 rpm, there are more severe vibrations than in the case of 1100 rpm, as shown in Fig. 17(b). The HHT plots here are depicted as a top view for comparison. The rotating frequency, 26.3 Hz is shown as a strip throughout the measured interval. And the frequency components between 120 Hz and 180 Hz in the Fourier spectrum are believed to be the natural frequencies of the shaft and the RK-4 rotor kit itself. However, it is difficult to characterize the looseness size at the lubrication condition with this HHT analysis.

5.2.2 Dry-friction condition

HHT analysis of the looseness in the dry-friction condition was performed at the same rotating speeds of the lubrication condition. At the lower speed of 1100 rpm, there was a rotating frequency component at the looseness 0.0 mm and 0.1 mm. However, when the looseness became larger than 0.1 mm, the frequencies between 120 Hz and 180 Hz were generated intermittently. As the clearance increased, the ampli-



Fig. 18. HHT results at dry-friction condition.

tude of the frequencies between 120 and 180 Hz became larger, as shown in Fig. 18(a). This was different from the lubrication conditions, since the rotor and the bushing contact severely and the effect of friction was enlarged. The HHT results at 1580 rpm are shown in Fig. 18(b). There were the rotating frequency 26.3 Hz and the frequencies approximately between 120 Hz and 180 Hz continuously. The whirling motion of the shaft became large at 1580 rpm since the shaft and the bushing contacted continuously and strongly.

When looseness occurred, the characteristic of friction was obvious. The rotating frequency existed clearly when the looseness was under 0.1 mm, and frequency band due to friction, which was between 120 Hz and 180 Hz, was weak. On the other hand, when the looseness was larger than 0.1 mm, the frequency band occurred strongly and non-periodically as the looseness increased.

6. Conclusion

Partial rub and looseness, the common faults of rotor systems with clearance, were diagnosed by a newly developed analysis method, HHT, and compared with results of the conventional signal processing methods of Fourier spectrum, STFT and CWT. The Fourier spectrum gave the exact frequency information but it showed limitations for non-linear and non-stationary signals. STFT was difficult to apply to non-linear and non-stationary signals because of the constant window function. Furthermore, it was not suitable for the data within a very short time. Signal processing using CWT could properly describe nonlinear and non-stationary signals but did not give clear resolution, and the information was spread at the low frequency region.

Although HHT needs a higher sampling number, it could analyze the several harmonic terms due to nonlinearity clearly and showed the clearest resolution regardless of the high and low frequency regions. As a result, some characteristics of the rotor system with clearance were investigated thoroughly.

In case of the partial rub due to a protrusion, impact signals were generated during revolution. The shaft contacted with the protrusion regularly per revolution, and then the shaft was excited at the natural frequencies. The impact occurred once per revolution at the weak partial rub. As the partial rub became severe, impact occurred two or more times per revolution.

In case of the looseness at the bushing, the characteristic of the looseness signal implied impact and friction. It was weak in lubrication condition, but it was severe in the dry-friction condition. This feature was represented when the looseness was larger than a specific clearance, and observed at the specific frequency band.

Consequently, HHT can diagnose the fault of rotor systems with clearance better than any conventional signal processing method. For the improvement of the HHT results, this study suggests that the mean values should be assigned at the ends to remove the end effects and the frequency of interest should be sampled more than 12.5 times.

References

- H. F. Black, Interaction of a Whirling Rotor with a Vibrating Stator Across a Clearance Annulus, J. Mech. Eng. 10 (1968) 1-12.
- [2] Y. S. Choi, Experimental Investigation of Partial Rotor Rub, KSME International Journal. 14 (11) (2000) 1250-1256.
- [3] Z. K. Peng, Detecting of the Rubbing-caused Impacts for Rotor-stator Fault diagnosis using reassigned Scalogram, Mechanical System and Signal Processing. 19 (2) (2003) 391-409.

2162

- [4] Z. Peng, Y. He, Q. Lu and F. Chu, Feature Extraction of the Rub-impact Rotor System by Means of Wavelet Analysis, Journal of Sound and Vibration. 59 (2003) 1000-1010.
- [5] Y. S. Choi and S. M. Park, Diagnosis on the Clearance of Rotating Machinery Using Correlation Dimension, Transactions of the Korean Society for Noise and Vibration Engineering. 15 (7) (2004) 134-139.
- [6] S. W. Hwang and S. K. Lee, An Analysis of the Signal of a Engine Valve Based on STFT, Higher

Order Time-Frequency Analysis and Wavelet Transform, The Korean Society of Automotive Engineers. 2 (2003) 732-737.

- [7] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung and H. H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proceedings of the Royal Society of London. 454 (1998) 903-995.
- [8] Using STAR MODAL version 6.2, Spectral Dynamics (2002).