# Accurate Instantaneous Frequency Estimation with Iterated Hilbert Transform and Its Application

SHUREN QIN, YI QIN, YONGFANG MAO Mechanical Engineering Department Chongqing University Chongqing, 400030 CHINA

*Abstract:* - Iterated Hilbert transform (IHT) is a new method for multicomponent demodulation. The principle of IHT is introduced, and some of its properties are researched. Although the amplitudes of the demodulated components obtained by IHT are accurate, there are limitations in the direct estimation of instantaneous frequencies via the extracted phases, so a smoothed instantaneous frequency estimation (SIFE) method based on difference operator and zero-phase digital low-pass filtering is proposed. The simulation results show that the proposed approach has higher performance than the adaptive segmentation algorithm and Hilbert-Huang transform. Finally, IHT and SIFE are applied to fault diagnosis of a rolling bearing, and the results show that the multicomponent fault vibration signal can be demodulated correctly and the weak feature of fault signals can be extracted efficiently with IHT and SIFE.

*Key-Words:* - Multicomponent demodulation, Hilbert transform, Amplitude envelopes, Filtering, Instantaneous frequency estimation, Fault diagnosis.

# **1** Introduction

When there's fault on the rotary component in the mechanical system, periodical impulse force will occur, which makes the vibration signals present the feature of modulation. The amplitude envelopes and phase signals of those modulated signals contain rich fault information. Additionally, most machinery fault vibration signals are multicomponent AM-FM signals, therefore, multicomponent demodulation is an effective way to extract the fault characteristics and diagnose the fault type. Currently, the typical multicomponent demodulation methods include: (i) Multiband energy separation algorithm (MESA) [1]; Periodic algebraic separation (ii) energy demodulation (PASED) [2]; (iii) Hilbert-Huang transform (HHT) [3,4]; (iv) Iterated Hilbert transform (IHT) [5], which is a new method for multicomponent analysis proposed by Gianfelici etc., hence called Gianfelici transform in this paper. Compared with the former three methods, Gianfelici transform is of higher demodulation accuracy, and of lower computational complexity.

By the research of Gianfelici transform, the authors find that although exact amplitude envelope of each component can be obtained with Gianfelici transform, however, to directly compute the instantaneous frequencies (IFs) from the extracted phase signals preserves some limitations. And the limitations are theoretically analyzed taking a generic two-component AM-FM signal as an example. Then a smoothed instantaneous frequency estimation (SIFE) method is proposed, which computes the derivative of the phase signal obtained by Gianfelici transform, and then carries out zero-phase digital low-pass filtering. Simulation shows that this method can exactly estimate the IF of corresponding component. Finally, Gianfelici transform and SIFE is applied to fault diagnosis for a rolling bearing, whose results indicate that the method is able to achieve multicomponent demodulation accurately and extract the weak fault characteristics in the vibration signal.

# 2 Gianfelici Transform

### 2.1 Hilbert Transform

For an arbitrary signal x(t), its Hilbert transform is defined as [6]

$$H[x(t)] = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} \mathrm{d}\tau \tag{1}$$

And the analytical signal z(t) of x(t) can be defined as

$$z(t) = x(t) + jH[x(t)] = A(t)e^{j\phi(t)}$$
(2)

In this case, the envelope A(t) is given by

$$A(t) = |z(t)| = \sqrt{x^{2}(t) + H^{2}[x(t)]}$$
(3)

Proceedings of the 7th WSEAS International Conference on SIGNAL PROCESSING, ROBOTICS and AUTOMATION (ISPRA '08) University of Cambridge, UK, February 20-22, 2008

and the phase function  $\phi(t)$ 

$$\phi(t) = \arg[z(t)] \tag{4}$$

The instantaneous frequency can be further computed by

$$f(t) = \frac{1}{2\pi}\omega(t) = \frac{1}{2\pi} \times \frac{d\phi(t)}{dt}$$
(5)

### 2.2 Multicomponent AM-FM Decomposition

Given an arbitrary signal x(t), and taking advantage of (2) it can be represented as

 $x(t) = \operatorname{Re}[z(t)] = a_0(t) \cos[\alpha_0(t)]$ (6)

where Re[·] represents the real part of a complex number, amplitude is  $a_0(t) = |z(t)|$ , and phase is  $\alpha_0(t) = \arg[z(t)]$ . In order to obtain the multicomponent model, we could iteratively operate Hilbert transform to the amplitude envelope. Assume the amplitude envelope after *j* iterations is  $a_j(t)$ . It's easy to get that  $a_j(t)$  will exhibit oscillating behavior and be always non-negative if the signal contains a mixture of sinusoids. Therefore,  $a_j(t)$  and alternating component  $\tilde{a}_j(t)$  by means of a certain filtering algorithm, that is

$$a_{j}(t) = \overline{a}_{j}(t) + \widetilde{a}_{j}(t)$$
(7)

In order to assure that the result of the next Hilbert transform is reasonable,  $\tilde{a}_j(t)$  should be a zero-mean oscillating signal, and it can be obtained by applying high-pass filtering to  $a_j(t)$ . Then the trend  $\bar{a}_j(t)$  can be easily computed by virtue of (7).

Formally, starting with j = 0 as the first step of this iterative algorithm, from (6) and (7) we obtain

$$x(t) = a_0(t) \cos[\phi_0^1(t)] = \left[\overline{a}_0(t) + \widetilde{a}_0(t)\right] \cos[\phi_0^1(t)]$$
(8)

where  $\phi_0^1(t) = \alpha_0(t)$ . Denote the analytical signal  $z_{i+1}(t)$  of the alternating component  $\tilde{\alpha}_i(t)$  as

$$z_{j+1}(t) = \widetilde{a}_j(t) + iH[\widetilde{a}_j(t)]$$
(9)

And denote

$$a_{j+1}(t) = \left| z_{j+1}(t) \right| \tag{10}$$

$$\alpha_{j+1}(t) = \arg[z_{j+1}(t)]$$
(11)

Thus,  $\tilde{a}_{i}(t)$  can be rewritten as

$$\tilde{a}_{j}(t) = a_{j+1}(t)\cos[\alpha_{j+1}(t)]$$
 (12)

When j = 0, there is

$$\widetilde{a}_0(t) = a_1(t)\cos[\alpha_1(t)]$$
(13)

Substitute the above equation into (8), yields

$$x(t) = \overline{a}_0(t) \cos[\phi_0^1(t)] +$$
(14)

$$a_1(t)\{\cos[\alpha_1(t)]\cos[\phi_0^1(t)]\}$$

And with Werner trigonometry formula, we obtain

$$x(t) = \overline{a}_0(t)\cos[\phi_0^1(t)] + \frac{a_1(t)}{2}\sum_{i=1}^{2^*}\cos[\phi_1^i(t)]$$
(15)

where

$$\phi_1^1(t) = \alpha_1(t) - \phi_0^1(t) \tag{16}$$

$$\phi_1^2(t) = \alpha_1(t) + \phi_0^1(t) \tag{17}$$

After the first iteration has been done, continue to deal with  $a_1(t)$  in the above way to finish the second iteration. In the same way, after N iterations and  $a_N(t)$  is filtered, we can obtain [5]

$$x(t) = \sum_{j=0}^{N} \frac{\overline{a}_{j}(t)}{2^{j}} \sum_{i=1}^{2^{j}} \cos[\phi_{j}^{i}(t)] + \frac{\widetilde{a}_{N}(t)}{2^{N}} \sum_{i=1}^{2^{N}} \cos[\phi_{N}^{i}(t)]$$
(18)

where  $\phi_i^i(t)$  can be iteratively calculated as

$$\phi_{j}^{2l-1}(t) = \alpha_{j}(t) - \phi_{j-1}^{l}(t)$$
(19)

$$\phi_{j}^{2l}(t) = \alpha_{j}(t) + \phi_{j-1}^{l}(t)$$
(20)

Where  $l = 1, 2, \dots, 2^{j-1}$  and  $j = 1, 2, \dots, N$  Assume that

$$r_{N}(t) = \frac{\widetilde{a}_{N}(t)}{2^{N}} \sum_{i=1}^{2^{N}} \cos[\phi_{N}^{i}(t)]$$
(21)

and define the norm as

$$\left\|F(\omega)\right\|^{2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left|F(\omega)\right|^{2} \mathrm{d}t \qquad (22)$$

Then, let *k* be the filter energy loss, define as

$$k = \frac{\left\|A_N(\omega)H(\omega)\right\|^2}{\left\|A_N(\omega)\right\|^2}$$
(23)

where  $\widetilde{A}_N(\omega)$  and  $A_N(\omega)$  are respectively the Fourier transform of  $\widetilde{a}_N(t)$  and  $a_N(t)$ ,  $H(\omega)$  is the transfer function of the filter which is applied for filtering  $a_N(t)$ . And it can be proved that  $r_N(t) \to 0$  when  $N \to \infty$  and k < 1/2 [5]. Therefore, (18) can be rewritten as

$$x(t) \approx \sum_{j=0}^{N} \frac{\overline{a}_{j}(t)}{2^{j}} \sum_{i=1}^{2^{j}} \cos[\phi_{j}^{i}(t)] = \sum_{k=1}^{2^{N+1}-1} \overline{A}_{k}(t) \cos[\Phi_{k}(t)]$$
(24)

where

$$\Phi_k(t) = \phi_j^i(t) \quad k = 2^j - 1 + i$$
 (25)

$$\overline{A}_{k}(t) = \frac{\overline{a}_{j}(t)}{2^{j}} \quad 2^{j} - 1 < k \le 2^{j+1} - 1$$
 (26)

Equation (24) is an asymptotically exact decomposition of the signal in terms of amplitude and phase envelopes. It is worth noting that  $\overline{a}_j(t)$  is the amplitude envelope of one component in the original signal. Please refer to [5] for the discussion of convergence of (24).

### 2.3 Further Research

The high-pass filter used in the iterative procedure can be designed with FDATool in MATLAB, and k < 1/2 must be satisfied. For a certain kind of signal such as mechanical vibration signals, speech signals, earthquake signals, radar signals, experiments can be done to take the suitable parameters for the filter and in the iteration, just adopt the same filter. Additionally, zero-phase digital filtering is the only filtering way in the above procedure of Gianfelici transform [7].

Since k < 1/2, the decomposition is to continuous extract the component with great energy, i.e. the components are ranged according to the energy from great to small. The number of components to be extracted is determined by the number of iterations. Equation (24) represents N + 1components are extracted, and the envelope of *j*th component is  $\overline{a}_j(t)$ , while the instantaneous frequency is given by one or more phase signals among  $\{\phi_j^i(t), i = 1, 2, \dots 2^j\}$ , which will be shown in section 3 with examples.

# **3** Accurate Instantaneous Frequency Estimation

### 3.1 Research on Limitations of Direct Computation of IF

The instantaneous frequency of each component are computed according to the associated phase signals in (24). A self-adaptive segmentation algorithm to compute the instantaneous frequency is used in [5]. It's a method for calculating the derivative of the unwrapped phase by linear regression which can effectively reduce the influence of noise upon IF estimation, but is mainly suitable for phase signal preserving piecewise linearity. For some signals whose frequency modulation parts are usually sinusoidal signals, e.g. mechanical vibration signals, if using self-adaptive segmentation algorithm to estimate IF, the accuracy will decline. Besides, in the model shown in (24), each phase signal is the result of a combination of several components' phase, hence to directly compute the IF from the derivative of the phase signal has certain limitations. We will show you the limitations by analyzing a two-component AM-FM signal's phase characteristics after the Hilbert transform.

A generic two-component AM-FM signal can be denoted as

$$x(t) = x_1(t) + x_2(t) = a_1(t) \cos[\phi_1(t)] + a_2(t) \cos[\phi_2(t)]$$
(27)

where for  $\forall t$ , there's  $a_1(t) > a_2(t)$ , so the first step of Gianfelici transform is to demodulate the component  $x_1(t)$ . Assume that Bedrosian theorem is satisfied by both  $x_1(t)$  and  $x_2(t)$  [8], then the Hilbert transform of x(t) is given by

$$H[x(t)] = H[x_1(t)] + H[x_2(t)] = a_1(t)\sin[\phi_1(t)] + a_2(t)\sin[\phi_2(t)]$$
(28)

And the corresponding analytical signal is

$$z(t) = x(t) + jH[x(t)]$$
(29)

According to the characteristics of Gianfelici transform, the phase signal used to estimate the instantaneous frequency of  $x_1(t)$  is

$$\phi(t) = \arg(z(t)) = \arctan\frac{H[x(t)]}{x(t)} + C \qquad (30)$$

where C is a constant corresponding to the sign of H[x(t)] and x(t).

Compute the deritives of both the two ends of (30) and we get

$$\dot{\phi}(t) = \frac{\dot{H}[x(t)]x(t) - H[x(t)]\dot{x}(t)}{|z(t)|^2}$$
(31)

From (27) and (28), we can obtain

$$\dot{H}[x(t)]x(t) - H[x(t)]\dot{x}(t) = a_1^2(t)\dot{\phi}_1(t) + a_2^2(t)\dot{\phi}_2(t) +$$
(32)

$$\begin{aligned} [\dot{a}_{1}(t)a_{2}(t) + a_{1}(t)\dot{a}_{2}(t)]\sin[\phi_{2}(t) - \phi_{1}(t)] + \\ a_{1}(t)a_{2}(t)[\dot{\phi}_{1}(t) + \dot{\phi}_{2}(t)]\cos[\phi_{2}(t) - \phi_{1}(t)] \\ |z(t)|^{2} = a_{1}^{2}(t) + a_{2}^{2}(t) + \\ 2a_{1}(t)a_{2}(t)\cos[\phi_{2}(t) - \phi_{1}(t)] \end{aligned}$$
(33)

Then assume

$$\dot{\phi}_2(t) = \dot{\phi}_1(t) + \varepsilon(t)$$
(34)  
And substitute (32), (33), (34) into (31), yields

substitute (32), (33), (34) into (31), yields  $\dot{\phi}(t) = \dot{\phi}_1(t) + E(t)$  (35)

where

$$E(t) = [a_1(t)a_2(t)]' \sin[\phi_2(t) - \phi_1(t)] + \varepsilon(t)[a_2^2(t) + a_1(t)a_2(t)] \cos[\phi_2(t) - \phi_1(t)]$$
(36)

From (35) and (36) we can see that it will arouse error to estimate the IF of  $x_1(t)$  by differentiating  $\phi(t)$  and the error E(t) preserves high frequency characteristics. The alternating component  $\widetilde{a}_{i}(t)$ can also be seen as a multicomponent signal, and the phase obtained by Hilbert transform also have the similar errors. Furthermore, from (19) and (20), it can be seen that the demodulation error by the former step will spread into the next decomposition. In the same way, these limitations will be extended to generic multicomponent signals, i.e. the IFs computed with the phaze signal derived in (24) has certain high frequency error.

#### 3.2 Smoothed Instantaneous Frequency Estimation

In order to reduce the error of instantaneous frequency estimation by directly computing the derivative of the phase, a suitable low-pass filter can be used to filter the derivative of the phase signal.

segmentation Although the self-adaptive algorithm can effectively diminish the influence of noise, it involves a lot of computation, hence not suitable for engineering application. Therefore, the difference operator is applied to compute the derivative of the phase signal. Assume the discrete signal corresponding to the phase signal  $\Phi_{k}(t)$  in (24) to be  $\Phi_k(n)$ , then the derivative of  $\Phi_k(t)$  can be approximated by backward difference, i.e.  $\omega_{\iota}$ 

$$(n) = \Phi_k(n) - \Phi_k(n-1)$$
 (37)

where  $\omega_{k}(n)$  is the discrete instantaneous angular frequency corresponding to  $\Phi_k(n)$ . If  $\omega_k(n)$  is negative, its absolute value  $|\omega_k(n)|$  is taken. Then carry out zero-phase digital low-pass filtering to  $|\omega_k(n)|$  to eliminate the high frequency error. And the relatively exact discrete instantaneous angular frequency  $\overline{\omega}_k(n)$  is obtained. The method is named smoothed instantaneous frequency estimation(SIFE), and its flowchart is depicted in Fig. 1.



Fig. 1 Flowchart of SIFE.

The low-pass filter used in SIFE is designed in the same way as the high-pass filter mentioned in the former section. Usually, the low-pass filter and the high-pass filter are complementary.

# 4 Simulation and Comparison

Consider the two-component AM-FM signal  

$$x(t) = x_1(t) + x_2(t) =$$
  
 $[1 + 0.5 \cos(2\pi 10t)]\sin[2\pi 200t + 2.5 \cos(2\pi 10t)] +$   
 $0.3 \sin(2\pi 800t)$ 
(38)

where  $t \in [0,1]$ ,  $x_1(t)$  is the AM-FM signal, and  $x_2(t)$  is the sinusoidal signal. The time domain waveform of x(t) is shown in Fig. 2.



After Gianfelici transform, we obtain the amplitude envelope  $\overline{a}_0(t)$ ,  $\overline{a}_1(t)$  respectively of  $x_1(t), x_2(t)$ , which is shown in Fig. 3 and phase signal  $\phi_0^1(t)$ ,  $\phi_1^1(t)$ ,  $\phi_1^2(t)$ . Then using SIFE, we can compute the instantaneous frequencies respectively corresponding to  $\phi_0^1(t)$ ,  $\phi_1^1(t)$  and  $\phi_1^2(t)$ . According to the spectral characteristics of x(t), we can judge that  $\phi_1^1(t)$  is the false phase signal. Therefore, the instantaneous frequencies should be calculated by  $\phi_0^1(t)$  and  $\phi_1^2(t)$ . The result is shown in Fig. 4. The instantaneous frequencies of the two components extracted by self-adaptive segmentation algorithm are shown in Fig. 5. From Fig. 4 and Fig. 5, we can see that SIFE is apparently better than self-adaptive segmentation algorithm. Although SIFE is of great performance, however, zero-phase digital filtering will arouse fluctuation at the ends of the instantaneous frequency curve. Therefore, in occasions which require high precision, expansion at the two ends of time series should be done before zero-phase digital filtering.

In order to better appreciate the validity of Gianfelici transform and SIFE, a comparison with the performance of HHT is carried out. For the same signal, the amplitude envelopes and instantaneous frequencies of the two components obtained by HHT are respectively shown in Fig. 6 and 7. As can be seen from Fig. 3, 4, 6 and 7, the curves obtained by Gianfelici transform and SIFE is smoother and of higher precision. In additon, since there's fast algorithm for Hilbert transform, the calculation speed of Gianfelici transform is faster.



Fig. 3 Amplitude envelopes of the two components in x(t) obtained by Gianfelici transform (solid—the amplitude envelope estimation of  $x_1(t)$ , dashed—the amplitude envelope estimation of  $x_2(t)$ ).



Fig. 4 Instantaneous frequencies of the two components in x(t)obtained by SIFE (solid—he IF estimation of  $x_1(t)$ , dashed—the IF estimation of  $x_2(t)$ ).



Fig. 5 Instantaneous frequencies of the two components in x(t) obtained by self-adaptive segmentation algorithm (solid—the IF estimation of  $x_1(t)$ , dashed—the IF estimation of  $x_2(t)$ ).



Fig. 6 Amplitude envelopes of the two components in x(t) obtained by HHT (solid—the amplitude envelope estimation of  $x_2(t)$ , dashed—the amplitude envelope estimation of  $x_1(t)$ ).



Fig. 7 Instantaneous frequencies of the two components in x(t)obtained by HHT (solid—the IF estimation of  $x_2(t)$ , dashed—the IF estimation of  $x_1(t)$ ).

In the above example, only one phase signal is used to compute each component's instantaneous frequency. However, if some components' frequency ranges overlap with each other, then more than one phase signals should be taken into account to compute one component's instantaneous frequency. Please refer to [5] to see the example.

# 5 Application to Fault Diagnosis

When there's weak damage to the rolling bearing, the signal containing the fault characteristics is usually drowned by the background signals relevant to the rotary speed of the rotor and other noises. In this case, it's hard to extract the fault characteristics by envelope analysis, and multicomponent demodulation should be employed to do fault diagnosis.

Fig. 8 shows the vibration acceleration signal of a SKF6203-type rolling bearing with out-race fault. The sampling frequency is 12 kHz, the rotary frequency ( $f_r$ ) is 29.95 Hz, and by calculation the out-race fault characteristic frequency ( $f_{oc}$ ) is 107.4 Hz. The spectrum of the signal is shown in Fig. 9, from which we can see that there're several frequency families (monocomponent signals), and hence it's hard to judge whether there's fault on the bearing and further analysis is needed.



Fig. 8 Vibration signal of a rolling bearing with out-race fault.



Fig. 9 Amplitude spectrum of the signal shown in Fig. 8.

Gianfelici transform and SIFE are employed to demodulate the first three components and their instantaneous frequencies  $f_1(t)$ ,  $f_2(t)$  and  $f_3(t)$ are respectively shown in Fig. 10(a)-(c), while the corresponding envelopes are denoted with  $a_1(t)$ ,  $a_2(t)$  and  $a_3(t)$ . Then spectrum analysis is carried out to  $a_1(t)$ ,  $a_2(t)$  and  $a_3(t)$ , and the amplitude spectrums are shown in Fig. 11(a)-(c) respectively. From Fig. 10 it is easy to note that the center frequencies of the 3 components are respectively about 1000 Hz, 2100 Hz and 2700 Hz. Combined with Fig. 9, we can see that the method introduced in this paper correctly extracted the information of three frequency families in the original signal. And according to Fig. 11 we learn that the first two components are AM-FM signals corresponding to the rotary frequency, and the third component is the AM-FM signal containing the fault information. In Fig. 11(c), there's obvious spectral line at the out-race fault characteristics frequency  $f_{\rm or}$ .



Fig. 10. Instantaneous frequencies of first three components in the fault signal shown in Fig. 8: (a) IF  $f_1(t)$  of first component, (b) IF  $f_2(t)$  of second component, (c) IF  $f_3(t)$  of third component.



Fig. 11. Amplitude spectrums of amplitude envelopes  $a_1(t)$ ,  $a_2(t)$ and  $a_3(t)$ : (a) Amplitude spectrum of  $a_1(t)$ , (a) Amplitude spectrum of  $a_1(t)$ , (c) Amplitude spectrum of  $a_1(t)$ .

## 6 Conclusion

Gianfelici transform (IHT) is a new signal analysis method for multicomponet signals. It has higher performance and doesn't need complex filter optimizations, hence has great application potential. The main efforts of the paper are as follows:

1) Introduces the principle of Gianfelici transform, states how to design the filters, and points out the decomposition is a process of continuously extracting the component with great energy.

2) Shows the limitations of directly computing the instantaneous frequency of each component by theoretical analysis. Hence, introduces a smoothed instantaneous frequency estimation method and validate its effectiveness via simulation.

3) Applies Gianfelici transform and SIFE into early fault diagnosis for bearing and obtains good effect, thus providing mechanical fault diagnosis with a new method.

References:

- [1] A.C. Bovik, P. Maragos, T.F. Quatieri, AM–FM energy detection and separation in noise using multiband energy operators, *IEEE Transacitons* on Signal Processing, Vol.41, No.12, 1993, pp. 3245-3265.
- [2] B. Santhanam, P. Maragos, Multicomponent AM–FM demodulation via periodicity-based algebraic separation and energy-based demodulation, *IEEE Transacitons on Communications*, Vol.48, No.3, 2000, pp. 473-490.
- [3] N.E. Huang, Z. Shen, S.R. Long, M.L.C. Wu, H.H. Shih, Q.N. Zheng, N.C. Yen, C.C. Tung, H.H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis, *Proc. R. Soc. Lond. A*, Vol.454, 1998, pp. 903-995.
- [4] S.R. Qin, Y.M. Zhong, A new envelope algorithm of Hilbert–Huang Transform, *Mechanical System and Signal Processing*, Vol.20, No.8, 2006, pp.1941-1952.
- [5] F. Gianfelici, G. Biagetti, P. Crippa, C. Turchetti, Multicomponent AM-FM representations: an asymptotically exact approach, *IEEE Transactions on Audio, Speech, and Language Processing*, Vol.20, No.3, 2007, pp. 823-837.
- [6] S. L. Hahn, *Hilbert transforms in signal processing*, Artech House, 1996.
- [7] A. Antoniou, *Digital filters, analysis, design, and applications*, 2nd ed, McGrawHill, 1993.
- [8] B. Boashash, Estimating and interpreting the instantaneous frequency of a signal——Part 1: Fundamentals, *Proceedings of the IEEE*, Vol.80, No.4, 1992, pp. 520-538.