Comparison between the efficiency of L.M.D and E.M.D algorithms for early detection of gear defects

THAMEUR KIDAR\textsuperscript{1,2}, MARC THOMAS\textsuperscript{1,a}, RAYNALD GUILBAULT\textsuperscript{1} and MOHAMED EL BADAOUI\textsuperscript{2}

\textsuperscript{1} Department of Mechanical Engineering, \textsuperscript{a}Corresponding author: marc.thomas@etsmtl.ca
\textsuperscript{2} University of Lyon, University of Saint Etienne, LASPI EA-3059, 20 Av de Paris, 42334 Roanne Cedex, France

Received 27 November 2012, Accepted 19 December 2012

Abstract – In recent years, the adaptive decomposition methods have become the center of interest of researchers in many fields and especially in the vibration diagnosis of rotating machines. This paper compares the sensitivity of defect detection of two adaptive methods: local mean decomposition (L.M.D) and empirical mode decomposition (E.M.D). The efficiency of L.M.D and E.M.D methods for detecting defects is investigated for two cases, one from numerical signals and the other from signals recorded during a fatigue test on gear bench. The time descriptors Kurtosis, Thalaf and Thikat are applied on the signal and on its Hilbert transform. The results reveal that both techniques seem to be suitable and have good efficiency for the fault detection. From experimental signals, the comparative results reveal that both methods allow for monitoring wear, but that L.M.D is more sensitive to detect rapid changes of degradation than E.M.D method for the considered case. Consequently these features have potential to become powerful tools for the monitoring of rotating machinery.

Key words: Gear defect / local mean decomposition / empirical mode decomposition / defect descriptor / Hilbert transform

1 Introduction

In gear maintenance, it is well known that the gear defaults generate vibrations and harmonics of the gear mesh frequency, often accompanied by modulation frequencies of the rotation frequency. Consequently, gear vibration measurement is usually used for monitoring defect occurrence. In industrial applications, scalar time descriptors (Peak, R.M.S, Kurtosis, crest factor, etc.) remain the most used tools to detect a fault, but they don’t allow for an early detection of degradation. Furthermore some of these time descriptors proved to be unreliable when the defect size increases and becomes important. Consequently, new descriptors such as Talaf and Thikat [1] or shock filter [2] have been developed to overcome these drawbacks. The analysis in the frequency domain allows for an identification of the defect source (worn, bent shaft, corrosion, cracks, etc.) and spectral, cepstral analysis or time-frequency techniques have been developed [3,4]. As random fluctuations can occur in a periodic phenomenon particularly due to friction, cyclostationarity analysis [5] is also a good approach that allows for detecting and localizing a defect in the angular domain. However, early detection of the defect occurrence remains a major problem. In this study, we propose an approach by Huang-Hilbert transform and a comparison between the efficiency of empirical mode decomposition (E.M.D) and local mean decomposition (L.M.D) [6,7] to solve the problem by using time descriptors. Empirical mode decomposition is an adaptive method which was proposed by Huang et al. [8,9] for the analysis of non-stationary and nonlinear signals. This technique allows for decomposing any complex signal into a collection of intrinsic mode functions (I.M.Fs). E.M.D has attracted the attention of researchers and has been widely used in the fields of monitoring and fault detections of machinery [10–12]. Another new adaptive analysis technique called local mean decomposition (L.M.D) was initially developed by Smith [13]. L.M.D decomposes signals into a set of product functions (P.Fs). Since L.M.D and E.M.D are both adaptive techniques, they do not involve the concept of time and frequency-resolutions but instead instantaneous frequency. L.M.D is ostensibly similar to E.M.D; it is worthwhile to compare these two methods as well as their effectiveness in the applications of machinery health diagnosis. As will be revealed in this work, L.M.D differs from E.M.D from few aspects: local mean, decomposed components, instantaneous frequency (I.F). This paper
is organized as follows: in Section 2.1 we introduce the time descriptors TALAF and THIKAT. A brief introduction to the L.M.D and E.M.D algorithms is provided in Sections 2.2 and 2.3. A comparative study is made between L.M.D and E.M.D in Section 3. Vibration signals measured from bench gear with fatigue test are used to compare the sensitivity of L.M.D and E.M.D in extracting meaningful information in Section 4. Conclusions are given in Section 5.

2 Signal processing tools

2.1 Time descriptors

In literature, many descriptors (Peak, R.M.S, crest factor, Kurtosis, etc.) have been used for fault detection. The Ku and the C.F are two particularly well adapted indicators for defect detection at the early stage of degradation. However, after a certain level of degradation, the use of some of these descriptors becomes unreliable for monitoring the defect evolution. New descriptors Talaf and Thikat have therefore been developed to improve the monitoring [1]. Talaf combines the traditional descriptors Kurtosis and R.M.S (Eq. (1))

\[
TALAF = \log \left[ \text{kurtosis} + \frac{RMS}{RMS_0} \right]
\]

where \(RMS_0\) is the root mean square value defined under healthy conditions.

If \(RMS_0\) is not known, the method considers any initial value obtained at the beginning of the monitoring or deduced as an acceptable level from the ISO 10816 for healthy machines. Thikat incorporates information collected from several descriptors (Kurtosis, R.M.S, C.F, Peak) into a single unit of information (Eq. (2))

\[
TIKHAT = \log \left[ (\text{kurtosis})^{cf} + \left( \frac{RMS}{RMS_0} \right)^{\text{Peak}} \right]
\]

2.2 Local mean decomposition (L.M.D)

L.M.D was originally developed to decompose modulated signals into a small set of product functions (P.Fs), each of which being the product of an amplitude envelope signal and a frequency modulated (F.M) signal. Furthermore, the time–frequency distribution of the original signal can be obtained. For any signal \(x(t)\), it can be decomposed as follows [13]:

1. Determine all local extrema \(n_i\) of the original signal \(x(t)\), and then the mean value \(m_i\) of two successive extrema \(n_i\) and \(n_{i+1}\) can be calculated by:

\[
m_i = \frac{n_i + n_{i+1}}{2}
\]

2. A corresponding envelope estimate \(a_i\) is given by:

\[
a_i = \left| \frac{n_i - n_{i+1}}{2} \right|
\]

The local mean and local magnitude at the endings cannot be computed by using equations (3) and (4). The smooth continuous local mean function \(\tilde{m}_i(t)\) and the amplitude function \(\tilde{a}_i(t)\) are obtained by the moving average method.

3. Let \(h_{11}(t) = x(t) - \tilde{m}_{11}(t)\) and \(s_{11}(t) = h_{11}(t)/\tilde{a}_{11}(t)\). If \(s_{11}(t)\) is a purely normalized F.M, then go to step (4). Otherwise, steps (1) and (2) will be repeated, until \(s_{1r_1}(t) = h_{1r_1}(t)/\tilde{a}_{1r_1}(t)\) is the flat F.M signal which fluctuates between –1 and 1, where \(r_1\) denotes the number of iterations for computing the 1st P.F, and the corresponding instantaneous amplitude (I.A) is

\[
a_1(t) = \tilde{a}_{11}(t) \tilde{a}_{12}(t) \ldots \tilde{a}_{1r_1}(t)
\]

The instantaneous phase (I.P) \(\varphi_1(t)\) and the LF \(f_1(t)\) can be obtained by:

\[
\varphi_1(t) = \arccos (s_{1r_1}(t))
\]

\[
f_1(t) = \frac{f_s}{2\pi} dt
\]

where \(f_s\) is the sample rate.

4. With the envelope function and the final frequency demodulated signal, the first P.F is written as

\[
P F_1(t) = a_1(t) s_{1r_1}(t)
\]

5. Then \(u_1(t) = x(t) - P F_1(t)\) is treated as the smoothed version of the original data and the procedure is repeated from (1) to (4), until \(u_{i-1}(t) - P F_i(t)\) is a monotonic function or no more than five oscillations. The final result can be obtained as

\[
x(t) = \sum_{i=1}^{P} P F_i(t) + u_p(t).
\]

2.3 Empirical mode decomposition (E.M.D)

The E.M.D method decomposes the time signal into a finite set of oscillatory functions called the intrinsic mode functions (I.M.F). An I.M.F is a function respecting the following conditions:

1. The number of extrema and the number of zero crossings are either equal or differ at most by 1.
2. The value of the moving average between the superior envelope (defined by local maxima) and the inferior envelope (defined by local minima) is zero.

An intrinsic mode is the embedded time scale in the signal. It is defined as the time between two successive extrema, and is not necessarily a harmonic function. In fact, it may include non-stationary amplitudes and modulated frequencies. The signal is decomposed into I.M.F as follows:

– First, identify all the local extrema, and connect all the local maxima by a cubic spline to form the superior envelope.
Then, the mean \( \text{moy}_1(t) \) is subtracted from the signal \( x(t) \), leading to the first component \( d_1(t) \), i.e:

\[
d_1(t) = x(t) - \text{moy}_1(t)
\]

Ideally, if \( d_1(t) \) is an I.M.F, then \( d_1(t) \) is the first component of \( x(t) \).

- If \( d_1(t) \) is not an I.M.F, \( d_1(t) \) is treated as the original signal and repeat the previous steps.

\[
d_{1k}(t) = d_{1(k-1)}(t) - \text{moy}_{1k}(t)
\]

where \( \text{moy}_{11}(t) \) is the mean of the upper and lower envelopes of \( d_1(t) \).

- After repeating the procedure up to \( k \) times, \( d_{1k}(t) \) becomes an I.M.F such as,

\[
d_{1k}(t) = d_{1(k-1)}(t) - \text{moy}_{1k}(t)
\]

then, it is designated like:

\[
C_1(t) = d_{1k}(t)
\]

the first component (I.M.F) of the original signal. \( C_1(t) \) represents the shortest period of signal or the highest frequency band.

- The computed difference between \( C_1(t) \) and \( x(t) \) is:

\[
r_1(t) = x(t) - C_1(t)
\]

where \( r_1(t) \) may be considered as the original signal.

- The above process is repeated \( n \) times, the \( n \)-I.M.Fs of signal \( x(t) \) is then:

\[
\begin{align*}
  r_2(t) &= r_1(t) - C_2(t) \\
  &\vdots \\
  r_n(t) &= r_{(n-1)}(t) - C_n(t)
\end{align*}
\]

- The process is stopped when \( r_n(t) \) becomes a monotonic function for which it is not possible to extract other I.M.Fs.

- The signal \( x(t) \) can then be written as:

\[
x(t) = \sum_{k=1}^{n} C_k(t) + r_n(t)
\]

where \( r_n(t) \) is the mean trend of \( x(t) \).

The I.M.Fs \( \{C_1(t), C_2(t), \ldots, C_n(t)\} \) represent different frequency bands ranging from the high frequencies to low frequencies. Frequency components contained in each I.M.F relate changes within the signal itself.

### 3 Numerical analysis

In order to simulate the vibratory signals of a gearbox, a signal generator [3] has been used. This generator is based on a multiplication of signals. It can therefore produce signals similar to gear meshes. This signal generator is aimed to simulate modulations around the harmonics of the gear mesh frequency which manifest with the presence of defects. The meshing is modulated in amplitude. Repetitive shocks \( f(t) \) have also been added to simulate damages (Fig. 1):

\[
S(t) = \left( \sum_{n=-\infty}^{+\infty} s_e(t - n\tau_e) \right) \left( 1 + \sum_{m=-\infty}^{+\infty} s_{r1}(t - m\tau_{r1}) + \sum_{p=-\infty}^{+\infty} s_{r2}(t - p\tau_{r2}) \right) + f(t)
\]

where \( \tau_e, \tau_{r1} \) and \( \tau_{r2} \) represent the meshing period and the rotational periods of wheel 1 and 2, respectively, and where \( s_e(t), s_{r1}(t) \) and \( s_{r2}(t) \) represent the meshing signal with its modulation, respectively.

Figure 1 shows the gear modulated signal and the repetitive damped shocks.

The decomposition of the generated signal into I.M.Fs and P.Fs is shown in (Fig. 2). There is not much difference between the components of each technique but the L.M.D is very sensitive to sampling frequencies because the growth of this latter increases the number of zero crossings which decreases the local magnitude and may diverge because the term \((s_{ij}(t) = \frac{\partial s_{ij}(t)}{\partial t})\). In (Fig. 3), it appears for example, the divergence of P.F2, P.F3, P.F4 and P.F5 to 10^{10} when increasing the sampling frequency.

In this part, gear signals with variable amplitude shocks are simulated by a signal generator and processed by decomposition of each signal through the two methods (E.M.D and L.M.D). Then, Hilbert transform and the time descriptors (Kurtosis, Talaf and Thikat) are applied to each component. (Fig. 4) shows that all these descriptors are not enough sensitive to the defect with the first I.M.F.

On the other hand, all these descriptors are able to detect the fault from the normalized amplitude of 0.0135 except by using THIKAT with the Hilbert transform where the defect is early detected from a normalized magnitude of 0.009 at I.M.F5. It is obvious that some of these descriptors rise and fall with the evolution of shocks except Thikat with the Hilbert transform and Ku from the time signal where it rises gradually in I.M.F2-I.M.F5. In (Fig. 5), the descriptors of L.M.D levels in most cases tend to very high values from a default of 0.0135.

To detect the default at a very early stage of degradation (0.009), it seems that Thikat with the Hilbert transform is the best descriptor compared to others.

### 4 Experimental application

E.M.D and L.M.D methods were applied for monitoring and diagnosing the defects during a fatigue test on a
gear reducer by using the recorded vibration signals [3]. The signals delivered from the reducer operated 24 h over 24 h. The number of teeth is respectively 20 teeth on the first wheel and 21 teeth on the second wheel. The rotation speed of the input shaft of the reducer is 1000 rpm which gives rotation frequencies of \( f_{r1} = 16.67 \) Hz and \( f_{r2} = 15.87 \) Hz on the output shaft. The meshing frequency is \( f_e = 333.33 \) Hz. A measurement has been recorded every day during a period of 12 days. During experimentation, the testing gear was worn from a well operating state to a deteriorated gear state. For early detection of degradation, we have chosen the signals delivered from the 2nd to the 8th day. The acquisition is composed of 60000 samples and the sampling frequency is 20 kHz. The meshing signal is principally caused by shocks between teeth of wheels that compose the reducer.
meshing signal is modulated in amplitude and frequency by the signals emitted from the two wheels whose frequencies are respectively $f_{r1}$ and $f_{r2}$. Diagnosing a worn tooth is made by following the amplitude in frequency domain. For applying the E.M.D and L.M.D methods on the experimental signal, we have first re-sampled the signal at 5000 Hz in order to avoid the problem of divergence of data as in the case of L.M.D, since this latter is very sensitive to high frequencies and high noise which disrupts the local averaging. We have repeated the same procedure as described in the theoretical part on the experimental signals. We have monitored the evolution of the absolute error between each defect with the 2nd signal considered as healthy for each I.M.F and the absolute error between each day.

From Figure 6, it may be observed that the descriptors of I.M.F1 rise gradually in all figures with wear. This means that the default manifests in the high frequencies at the very early stage of degradation. In all other cases, it is difficult to put down clearly a diagnostic by using the E.M.D since the other I.M.Fs move up and down.

On the other hand, the descriptors of P.Fs of L.M.D method (Fig. 7) clearly react with the default evolution from the 4th day for the Kurtosis with the Hilbert transform while Thikat and Talaf applied to the envelope of P.F show a gradual evolution of the degradation from the
3th day due to the wear with a more sensitive degradation after the 6th day, due to a greater damage. After the 11th day, the gear train has been stopped and examined. A tooth break has been detected. However with this method (L.M.D), we only arrived to calculate the first 4 P.Fs due to the re-sampling of the recorded signal.

5 Conclusion

In this paper, we have compared the sensitivity of time descriptors (Ku, Thikat, Talaf) at different levels of decomposition by using two adaptive decomposition methods, i.e., L.M.D and E.M.D, and their Hilbert transform. The study has been conducted first from numerical signals and then on experimental signals during a fatigue test on a gear train. From the numerical simulations, both methods have shown the same efficiency for the early detection of degradation. However, fewer decomposed components can be obtained with L.M.D than with the E.M.D method, since L.M.D is very sensitive to the noise and sampling frequency, and can diverge. From the experimental data recorded during a fatigue test on the gear train during 12 days, the L.M.D method has shown more clear results than the E.M.D method by detecting a clear degradation due to the wear of teeth. All the three scalar descriptors that we have investigated have shown a significant increase after the 5th or 6th day due to degradation, but it seems that Talaf and Thikat are able to continuously monitor the increase of wear by using the Hilbert envelope with the L.M.D. A tooth break has been revealed after the 11th day. However, the efficiency of each method to detect degradation needs yet more investigation and experiments, but it seems that E.M.D and L.M.D methods present potential techniques for monitoring mechanical defects or wear and that L.M.D is more sensitive.
to rapid change of degradation. The expectation is that further research and applications of the E.M.D and L.M.D techniques will flourish in the near future.

Acknowledgements. The support of NSERC (Natural Sciences and Engineering Research Council of Canada), through Research Cooperative grants is gratefully acknowledged.

References