Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 41 (2016) 717 – 722

www.elsevier.com/locate/procedia

48th CIRP Conference on MANUFACTURING SYSTEMS - CIRP CMS 2015

Improved signal characterization via Empirical Mode Decomposition to enhance in-line quality monitoring

Andrea Carroccia[a], Marco Grasso[a,b,*], Matteo Maggioni[a], Bianca M. Colosimo[a]

[a]*Dipartimento di Meccanica, Politecnico di Milano, Via La Masa 1, 20156 Milan, Italy*
[b]*Laboratorio MUSP, Via Tirotti 9, 29122 Piacenza, Italy*

* Corresponding author. Tel.: (+39) 0523-623190; fax: (+39) 0523-645268. *E-mail address:* marco.grasso@musp.it

Abstract

The machine tool industry is facing the need to increase the sensorization of production systems to meet evolving market demands. This leads to the increasing interest for in-process monitoring tools that allow a fast detection of faults and unnatural process behaviours during the process itself. Nevertheless, the analysis of sensor signals implies several challenges. One major challenge consists of the complexity of signal patterns, which often exhibit a multiscale content, i.e., a superimposition of both stationary and non-stationary fluctuations on different time-frequency levels. Among time-frequency techniques, Empirical Mode Decomposition (EMD) is a powerful method to decompose any signal into its embedded oscillatory modes in a fully data-driven way, without any ex-ante basis selection. Because of this, it might be used effectively for automated monitoring and diagnosis of manufacturing processes. Unfortunately, it usually yields an over-decomposition, with single oscillation modes that can be split into more than one scale (this effect is also known as "mode mixing"). The literature lacks effective strategies to automatically synthetize the decomposition into a minimal number of physically relevant and interpretable components. This paper proposes a novel approach to achieve a synthetic decomposition of complex signals through the EMD procedure. A new criterion is proposed to group together multiple components associated to a common time-frequency pattern, aimed at summarizing the information content into a minimal number of modes, which may be easier to interpret. A real case study in waterjet cutting is presented, to demonstrate the benefits and the critical issues of the proposed approach.

Peer-review under responsibility of the scientific committee of 48th CIRP Conference on MANUFACTURING SYSTEMS - CIRP CMS 2015

Keywords: Monitoring; Quality Assurance; Waterjet Machining

1. Introduction

Continuous advances of sensor technology and real-time computation capabilities provide increasing opportunities for the development of smart factories characterized by data-rich environments. In-process analysis and monitoring of discrete manufacturing processes represents a key issue to determine how a process is performing during the process itself, by means of in-process analysis of sensor signals. One goal consists of achieving a fast detection of faults and unnatural process behaviours without waiting for post-process inspections on the manufactured part. Another goal consists of exploiting as much as possible the information acquired during the production of the current part, without the need for a training phase that involve the collection of data for a number of analogous parts in the same lot. Nevertheless, the analysis of in-process signals implies several challenges, due to the complexity of signal patterns, the superimposition of stationary and non-stationary behaviours, and the presence of features belonging to different time-frequency scales. Those kinds of signals are referred to as "multiscale", and they are quite common in industrial applications. Different authors [1, 2] showed that a proper multi-resolution characterization of multiscale signal patterns could enhance the profile monitoring performances, because different faults may have different effects on distinct scales. An example of multiscale signal acquired during a cutting process is shown in Fig. 1: it is the water pressure signal acquired during a waterjet cutting process by using a single-acting ultra high pressure (UHP) pump with three parallel plungers. The signal, acquired at 2 kHz and then synchronously resampled, refers to a complete pumping cycle that includes the three consecutive acting strokes (for details about the plant and process, the

Peer-review under responsibility of the scientific committee of 48th CIRP Conference on MANUFACTURING SYSTEMS - CIRP CMS 2015
doi:10.1016/j.procir.2015.12.094

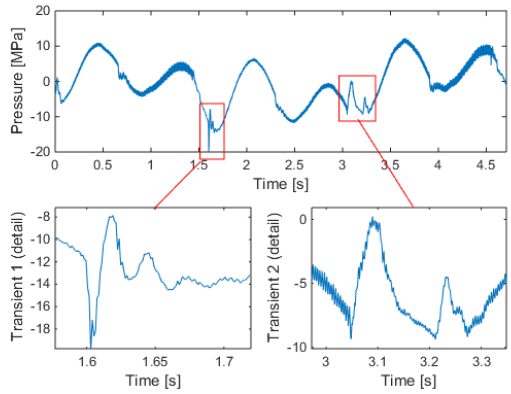

Fig. 1. Example of pressure profile in waterjet cutting (top panel) and details of high frequency transients (bottom panels)

interested reader may refer to [3, 4]). The pressure signal exhibits a superimposition of a noise component, some high-frequency transients that are localized in time (Fig. 1 – bottom panels), and a low frequency oscillation.

The low frequency pressure ripples are due to the plunger kinematics, whereas the high frequency transients correspond to transitions between the end of a plunger active stroke and the beginning of the next one. When a piston reaches the top dead centre, a valve commutation occurs, with a consequent oil flow-rate reduction in the next cylinder, resulting in a dynamic pressure discontinuity. Minor transients are present too, in correspondence of points in time where each plunger completes its suction step. They are caused by a flow-rate modification when the piston reaches the bottom dead centre. Refer to [3, 4] for further details.

Different faults of the most critical components (e.g., components of the UHP pump or the cutting head) may differently affect the scale-dependent features of the signal. Thus, a proper decomposition of those features is expected to provide multiple benefits: (i) a better characterization of the natural pattern variability, (ii) enhanced signal monitoring performances provided by a faster detection of small shifts affecting a single scale or a sub-set of scales, and (iii) an enhanced diagnostic capability, thanks to a better interpretability of fault effects.

The mainstream literature focuses on wavelet analysis for such a kind of signals [1-2; 5-6]. Nevertheless, the performances of wavelet-based methods rely on several problem-dependent and subjective choices, including the thresholding technique, the selected wavelet basis, the decomposition settings, the wavelet coefficient selection, etc. A more flexible alternative is represented by the Empirical Mode Decomposition (EMD) methodology [7], thanks to its data-adaptive nature. It is a data-driven and adaptive method that allows decomposing any signal into a number of Intrinsic Mode Functions (IMFs), which represent its natural oscillatory modes. It requires neither any integral transform nor the definition of any basis function. The IMFs are determined by the signal itself and they work as basis functions. Moreover, the EMD may yield a finer time-frequency resolution, and hence a better multiscale variability characterization [8]. However, the EMD may yield an over-decomposition of the signal, and the extracted IMFs may be affected by the so called "*mode mixing*" issue [7]. An open issue in the reference literature consists of finding a way to automatically achieve a synthetic decomposition in terms of a minimal number of relevant modes. This paper briefly reviews a methodology to convert the original IMF decomposition into Combined Mode Functions (CMFs) and it proposes a novel way to automatically determine the optimal decomposition into a minimal number of CMFs. The proposed method consists of combining together adjacent IMFs characterized by high similarity degree. An increasing number of CMFs is iteratively evaluated and the algorithm is stopped when an optimality criterion is met.

This paper extends the previous study of Grasso *et al*. [3], by comparing different indexes for the similarity estimation between IMFs and by introducing a novel algorithm for the automatic selection of the CMFs and its stopping criterion.

The proposed method is demonstrated by means of a real case study, i.e., the multiscale pressure signals acquired during a waterjet cutting operation.

Section 2 reviews the EMD and CMF methodologies; Section 3 presents the proposed approach; Section 4 discusses the performances of the method in a waterjet cutting case study; Section 5 concludes the paper.

Nomenclature

$\boldsymbol{Y}_j$	j[th] profile pattern, *j=1,2,…*
p	number of datapoints in each profile
$\boldsymbol{h}_{i,j}$	i[th] difference between $\boldsymbol{Y}_j$ and the envelope mean
$\boldsymbol{m}_{i,j}$	i[th] envelope mean for the signal $\boldsymbol{Y}_j$
$\boldsymbol{c}_{1,j}$	i[th] IMF of the signal $\boldsymbol{Y}_j$
$\boldsymbol{r}_{n_j,j}$	residue of the EMD decomposition of $\boldsymbol{Y}_j$
n_j	number of IMFs for the signal $\boldsymbol{Y}_j$
$\boldsymbol{c}_{s_k,j}$	k[th] CMF of the signal $\boldsymbol{Y}_j$
k	number of CMFs
$E_{j,i}$	energy of the i[th] IMF of the signal $\boldsymbol{Y}_j$
$C_{j,i}$	cross-corr. between $\boldsymbol{Y}_j$ and its i[th] IMF
$R_{j,i}$	peak-to-peak range of the i[th] IMF of the signal $\boldsymbol{Y}_j$
$S_{j,i}$	skewness of the i[th] IMF of the signal $\boldsymbol{Y}_j$
$K_{j,i}$	kurtosis of the i[th] IMF of the signal $\boldsymbol{Y}_j$
$\boldsymbol{V}_{j,i}$	multivariate vector of indexes for the i[th] IMF of $\boldsymbol{Y}_j$
$D_{l,j}$	dissimilarity between two adjacent IMFs
$SSW_{k,j}$	sum of squares within the CMF

2. The EMD and CMF methodology

2.1. The EMD algorithm

Let $\boldsymbol{Y}_j = [Y_{j,1}, Y_{j,2}, ..., Y_{j,p}]^T$ be a waveform, hereafter denoted by the term "*profile*", that characterizes a repeating pattern, such that $j = 1, 2, ...$ is the index of the repeating profile, and p is the number of datapoints in each profile.

Then, the IMFs that capture intrinsic oscillation modes can be extracted by means of the "*sifting*" process, which consists of the following steps [7]:

1) All the local minima and maxima of the profile $\boldsymbol{Y}_j = [Y_{j,1}, Y_{j,2}, ..., Y_{j,p}]^T$, $j = 1, 2, ...$ are identified and they are interpolated respectively by an upper and a lower envelope expressed on a cubic spline basis (notice that, without loss of generality, an equal sample size, p, is assumed for all the acquired profiles);
2) the mean of the two envelopes is computed and designated as $\boldsymbol{m}_{1,j}$; then, the difference between the signal $\boldsymbol{Y}_j$ and $\boldsymbol{m}_{1,j}$ is computed and designated as $\boldsymbol{h}_{1,j}$:

$$\boldsymbol{h}_{1,j} = \boldsymbol{Y}_j - \boldsymbol{m}_{1,j}, \quad j = 1, 2, ... \tag{1}$$

If $\boldsymbol{h}_{1,j}$ is an IMF, i.e., if $\boldsymbol{h}_{1,j}$ satisfies the following conditions:
   a) in the entire dataset, the number of extremes and the number of zero crossings must be either equal or different at most by one;
   b) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero;

then, $\boldsymbol{h}_{1,j}$ is taken as the first IMF of the signal and designated as $\boldsymbol{c}_{1,j}$. If $\boldsymbol{h}_{1,j}$ is not an IMF, $\boldsymbol{h}_{1,j}$ replaces the original signal and the above steps are repeated until an IMF is obtained.
3) The first IMF, $\boldsymbol{c}_{1,j}$, is separated from the signal $\boldsymbol{Y}_j$ by:

$$\boldsymbol{r}_{1,j} = \boldsymbol{Y}_j - \boldsymbol{c}_{1,j}, \quad j = 1, 2, ... \tag{2}$$

The residue $\boldsymbol{r}_{1,j}$ is treated as the original signal and the above steps are repeated, leading to the extraction of the following IMFs $\boldsymbol{c}_{2,j}, ..., \boldsymbol{c}_{n_j,j}$ such that:

$$\begin{aligned} \boldsymbol{r}_{1,j} - \boldsymbol{c}_{2,j} &= \boldsymbol{r}_{2,j} \\ \boldsymbol{r}_{n_j-1 j} - \boldsymbol{c}_{n_j,j} &= \boldsymbol{r}_{n_j,j} \end{aligned} \quad j = 1, 2, ... \tag{3}$$

At the end of the process, the signal is decomposed into n_j intrinsic modes and a residue $\boldsymbol{r}_{n_j,j}$:

$$\boldsymbol{Y}_j = \sum_{i=1}^{n_j} \boldsymbol{c}_{i,j} + \boldsymbol{r}_{n_j,j}, \quad j = 1, 2, ... \tag{4}$$

The residue is a signal such that no further decomposition is possible. In this study, the Amplitude Ratio criterion proposed by Rilling *et al*. [9] is advocated. The EMD algorithm usually converges rapidly in few iterative passes, producing a nearly orthogonal adaptive basis. As an example, the decomposition of one waterjet pressure profile is shown in Fig. 2. In this case, $n_j = 8$.

2.2. Combined Mode Functions

A critical issue regarding the EMD methodology is represented by the possible appearance of the so-called *mode mixing* effect [7]. It causes an intrinsic mode to be split into two or more adjacent IMFs, or disparate scales to be superimposed into a single IMF, which deprives single IMFs of physical meaning. Different approaches have been proposed to cope with such an issue. Wu and Huang [10] proposed a method called Ensemble EMD (EEMD) whose main limitation is the computational cost, as it requires the computation of a sufficient number of ensemble trials. A more efficient variant of the EEMD was proposed by Zhang *et al*. [11], but the computational cost is still considerably higher than the basic

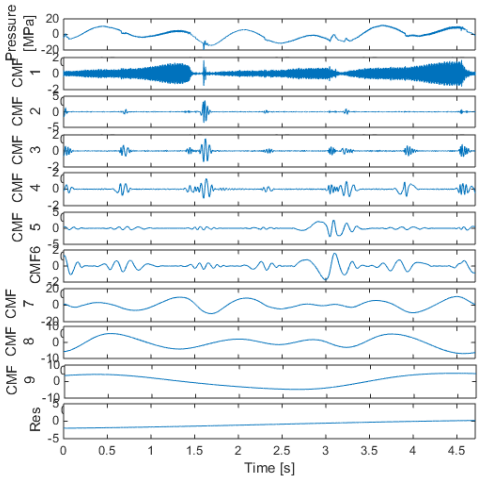

Fig. 2. EMD of a pressure profile in waterjet cutting

EMD, and this makes EEMD-based methods poorly attractive as far as in-process applications are concerned.

A more interesting and effective approach from an in-process implementation perspective is based on the Combined Mode Function (CMF) approach [12]. It consists of combining neighboring IMFs $\boldsymbol{c}_{i,j}, \boldsymbol{c}_{i+1,j}, ..., \boldsymbol{c}_{1+q_k,j}$ to obtain a new CMF, $\boldsymbol{c}_{s_k,j}$, as follows:

$$\boldsymbol{c}_{s_k,j} = \boldsymbol{c}_{i,j} + \boldsymbol{c}_{i+1,j} + ... + \boldsymbol{c}_{i+q_k,j}, \quad j = 1, 2, ...; i \in [1, n_j] \tag{5}$$

where $q_k + 1$ is the number of IMFs combined into the k^{th} CMF, being $0 \le q_k \le n_j - 1$ and $k \le n_j$. Such a combination of subsets of IMFs can be interpreted as a new adaptive filter bank, which is based on the intrinsic time scales of the signal, resulting in an accuracy increase of the EMD [12]. Grasso *et al*. [3] showed that by combining adjacent IMFs into CMFs two main benefits may be achieved: (i) a more synthetic representation of multiscale signal features and (ii) a mitigation (or even the avoidance) of the mode mixing effect.

Nevertheless, there is no clear procedure to automatically choose the number k of CMFs and to select the best CMF decomposition. There is a wide literature devoted to the automatic selection of subsets of IMFs for different goals that include (i) signal de-noising (when only low frequency modes are retained), (ii) signal de-trending (when the profile is reconstructed by removing the last IMFs), or (iii) the isolation of physically relevant IMFs (band-pass filtering). The mainstream approaches rely on synthetic indexes (e.g., average value [13], energy [14], [15] correlation between the IMF and the original signal, [16] peak frequency, and others [17]). However, the literature lacks an automated strategy to determine the best (minimal) decomposition.

3. The proposed approach for CMF selection

The proposed approach for the generation of the CMFs inherits the index-based techniques proposed in the literature, but it leads to a fully automated procedure. It consists of the following steps:

1) Generation of an iteratively increasing number of CMFs for the current profile, $\boldsymbol{Y}_j$, from $k = 1$ (i.e., all IMFs

combined into a single CMF) up to $k = n_j$ (i.e., CMFs that coincide with the IMFs), by using a dissimilarity-based criteria;

2) A metric consisting of the sum of squares of within-CMF differences is computed to determine to what extent IMFs combined into the same CMFs are similar to each other;
3) The optimal CMF configuration is chosen such that the sum of squares decrease by passing from *k-1* to *k* CMFs is maximized: this means that the addition of the k^{th} CMF enhances the decomposition, whereas a configuration with a larger number of CMFs do not provide relevant improvements.

The different steps of the proposed methods are explained and discussed in the following sub-sections.

3.1. Computation of the dissimilarity between IMFs

An estimate of similarity/dissimilarity between adjacent IMFs can be based on synthetic indexes like the ones mentioned in the previous section. The following indexes are considered: the IMF energy, $E_{j,i}$, the cross-correlation coefficient between each IMF and the original signal, $C_{j,i}$, the IMF peak-to-peak range, $R_{j,i}$, the IMF skewness, $S_{j,i}$, and the IMF kurtosis, $K_{j,i}$, for $i = 1,...,n_j$ and $j = 1,2,...$.

Two options are compared in this paper: (i) a univariate approach, where only one of those indexes is considered, and (ii) a multivariate approach, where a vector $\boldsymbol{V}_{j,i} = [E_{j,i}, C_{j,i}, \sigma_{j,i}, S_{j,i}, K_{j,i}]^T$ is considered.

Let $I_{j,i}$ be any of the indexes mentioned above. With regard to the univariate approach, the dissimilarity between two adjacent IMFs, $\boldsymbol{c}_{i,j}$ and $\boldsymbol{c}_{i+1,j}$, is computed as $D_{l,j} = \left\| I_{j,i+1} - I_{j,i} \right\|$ for $l = 1,...,n_j - 1$. With regard to the multivariate approach, instead, the dissimilarity is computed as $D_{l,j} = \left\| \boldsymbol{V}_{j,i+1} - \boldsymbol{V}_{j,i} \right\|$.

It is evident that the choice of the index plays a critical role. The next section will show the impact of this choice on the overall performances of the method.

3.2. Iterative generation of CMFs

The iterative generation of CMFs works as follows:

1) Compute the distance $D_{l,j}$ between each pair of adjacent IMFs for the current profile $\boldsymbol{Y}_j$. As an example, the $D_{l,j}$ values for the 9 IMFs shown in Fig. 2 are depicted in Fig. 3, where $D_{l,j} = \left\| E_{j,i+1} - E_{j,i} \right\|$;
2) Initialize a counter *q=1*;
3) Set *k=q*:
   a. if *q=1*, then a single CMF is generated, which is the sum of all the IMFs;
   b. if *q=2*, then two CMFs are generated, such that $\boldsymbol{c}_{s_1,j} = \sum_{i=1}^{T_q} \boldsymbol{c}_i$ and $\boldsymbol{c}_{s_2,j} = \sum_{i=T_q+1}^{n_j} \boldsymbol{c}_i$, where $T_q = \arg\max_l (D_{l,j})$. In the example shown in Fig. 3, two CMFs would be generated by summing up the IMFs from *i=1* to *i=7* and the ones from *i=8* to *i=9*,

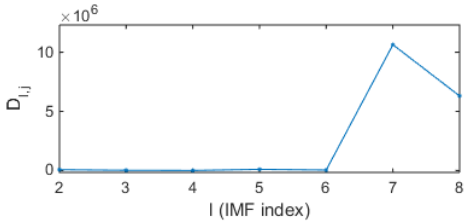

Fig. 3. Values of the distance $D_{l,j}$ as a function of *l* for the pressure profile in Fig. 2 (based on energy index)

because the maximum dissimilarity occurs at the 7th IMFs;

   c. if *q>2*, then the CMF that includes the $(q-1)^{th}$ largest difference $D_{l,j}$ is divided into two distinct CMFs analogously to the previous point, where now T_q is the argument of the $(q-1)^{th}$ largest difference.
4) Set *q=q+1* and repeat the step 3) until $k = n_j$.

The result is a collection of CMF configurations: each configuration consists of a signal decomposition into *k* CMFs, from *k=1* to $k = n_j$. The last step of the proposed procedure consists of selecting the best configuration among them.

3.3. Selection of the optimum number of CMFs

The proposed approach for the selection of the optimum number of CMFs is inspired by the cluster validation criteria used in unsupervised learning [18]. Analogously to clustering problems, we want to find a CMF configuration such that both the similarity within each CMF and the dissimilarity between distinct CMFs are large. A suitable index for such a purpose is the sum of squares within the CMF (SSW), computed as follows:

$$SSW_{k,j} = \sum_{q=1}^{k} \sum_{i \in c_{q,j}} \left\| \theta_{j,i,q} - \bar{\theta}_{j,q} \right\| \tag{6}$$

where $\theta_{j,i,q}$ is the statistic used to estimate dissimilarities among IMFs, i.e., $\theta_{j,i,q} = I_{j,i,q}$ in the univariate case, or $\theta_{j,i,q} = \boldsymbol{V}_{j,i,q}$ in the multivariate case, for $q = 1,...,k$ and $j = 1,2,...$.

As an example, Fig. 4 (top) shows the values of $SSW_{k,j}$ for the signal in Fig. 2 (in this case the univariate approach based on the energy index is applied). Generally speaking, $SSW_{k,j}$ is a monotone decreasing function. Assume that the signal includes *k* known modes (i.e., *k* separable time-frequency scales): then, the separation of original IMFs into *k* CMFs will produce a great reduction of $SSW_{k,j}$, but if one goes on by decomposing the signal into more than *k* CMFs, the further reduction of the index will be negligible. Because of this, the proposed criterion for the automatic selection of *k* consists of choosing *k* such that the successive differences of $SSW_{k,j}$ are maximized, i.e. $k_j = \arg\max_k \left\| SSW_{k,j} - SSW_{k-1,j} \right\|$, where k_j is the optimal choice of *k* for the j^{th} profile. Fig. 4 (bottom) shows the successive differences of the $SSW_{k,j}$ index: in this case, the

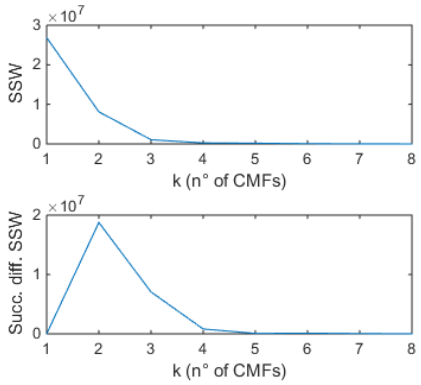

Fig. 4. SSW (top panel) and its successive differences as a function of the number *k* of CMFs (for the pressure profile in Fig. 2) – energy based approach

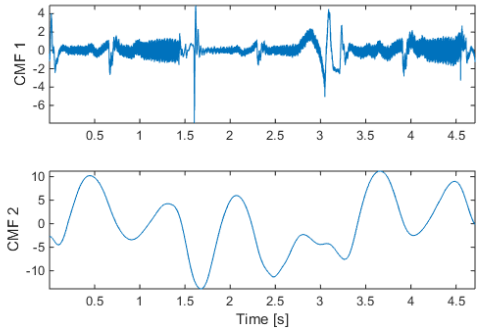

Fig. 5. Resulting CMF decomposition of the pressure profile in Fig. 2 when the univariate approach based on the energy index is used

maximum occurs at *k=2*, and hence the two CMFs shown in Fig. 5 are generated.

Fig. 5 shows that, when $\theta_{j,i,q} = E_{j,i,q}$ is used in Eq. 6, two CMFs are generated: the first one captures the noise term and the high frequency transients that are localized in time, whereas the second CMF captures the low frequency ripples. In the next section we discuss how different choices of the synthetic indexes influence the CMF decomposition.

4. A real case study

The waterjet cutting process introduced in Section 1 is used as a real case study to discuss the performances of the proposed method in a real industrial application. The pressure profiles were collected by installing a water pressure transducer on the high pressure water duct. Signal data were collected on a 45 kW UHP with a nominal working point characterized by a water pressure of 350 MPa, a water flow rate of 5 l/min and a 0.25 mm orifice. The signal was acquired at 2 kHz during repeated cutting on a laminate, and it was segmented into profiles such that one profile corresponds to a complete pumping cycle [3, 4]. In this study, 50 profiles under in-control conditions are considered, and each profile is decomposed via the proposed approach. The goal is to evaluate the resulting decomposition and to what extent such a decomposition is stable (in terms of number and nature of the CMFs) for all the in-control profiles.

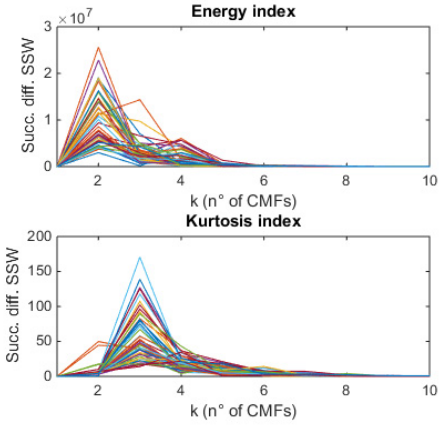

Fig. 6. Successive differences of the energy-based (top panel) and kurtosis-based (bottom panel) SSW index

With regard to the univariate approach for the IMF dissimilarity computation, Fig. 6 shows that different indexes may yield different decompositions. As an example, Fig. 6 shows the SSW successive differences based on the energy index, $E_{j,i}$, and on the kurtosis index, $K_{j,i}$, respectively, for 50 pressure profiles.

Apart from few exceptions, the use of index $E_{j,i}$ yields the selection of two CMFs, which correspond to the ones shown in Fig. 5, whereas the use of index $K_{j,i}$ yields the selection of three CMFs, shown in Fig. 7. The kurtosis index is more appropriate to determine differences between stationary fluctuations and transient ones, and hence it allows one separating the high frequency transients both from the noise term and the low frequency ripples. In terms of multiscale pattern interpretation, the CMF decomposition shown in Fig. 7 might be preferred, as it enhances the characterization of the local transients, although the low frequency ripples results less smooth than in Fig. 5.

Table 1. Frequency of selection of *k* CMFs for the different kinds of indexes

	Index	Frequency of selection of *k* CMFs (%)			
		k=2	*k=3*	*k=4*	*k>4*
Univar.	Energy, E	94	4	2	0
	Corr, C	96	2	2	0
	Range, R	94	4	2	0
	Kurt, K	6	88	6	0
	Skew, S	4	58	22	16
Multivariate		50	20	12	18

Table 1 shows the frequencies of the number *k* of CMFs generated for all the in-control profiles, either by using univariate indexes or the multivariate approach (i.e. the entire set of indexes included into a single vector).

It is worth to notice that, whenever *k=2*, the CMF decomposition resembles the one in Fig. 5, whereas, whenever *k=3*, it resembles the one in Fig. 7.

For the waterjet data, *k=2* and *k=3* are the two most likely choices, as far as the univariate approach is used. The main difference between these two CMF decompositions is the superimposition or separation of the noise term and the transient mode. The choice *k=2* is achieved by using either the

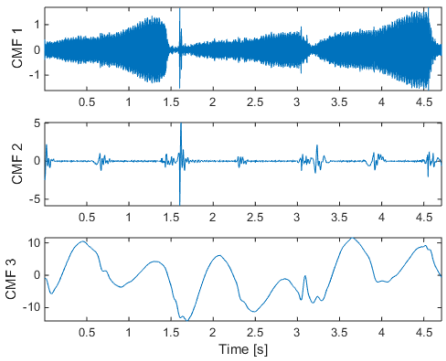

Fig. 7. Resulting CMF decomposition of the pressure profile in Fig. 2 when the univariate approach based on the kurtosis index is used

$E_{i,j}$, the $C_{i,j}$ or the $R_{i,j}$ index, and such a choice results to be stable as the EMD is applied to repeating profile patterns.

The choice $k=3$ is achieved by using either the $K_{i,j}$ or the $S_{i,j}$ index, but the stability is lower (especially regarding the $S_{i,j}$ index). Eventually, the multivariate approach results the less stable one, since the choice depends on which subset of indexes prevails on the other one, case by case.

The results summarized in Table 1 show that the choice of the index used to determine the dissimilarity between IMFs plays a crucial role. Some engineering knowledge about the process might be required to choose the proper index. Otherwise, more reliable and effective approaches may be investigated in future study, to avoid the use of one (or more) synthetic indexes and to focus on statistics that are more informative.

5. Conclusions

Multiscale signal decomposition is a key issue to analyze and monitor sensor signals whose salient features belong to different time-frequency scales. The EMD is a flexible technique for such a task, thanks to its data-driven and adaptive nature, but it does not guarantee a synthetic and easy to interpret decomposition of major signal modes. This study proposed an automated approach to find an optimal and synthetic decomposition in terms of a minimal number of CMFs that are expected to be easier to interpret than the original IMFs. The results showed that this method is suitable to deal with real data and complex patterns, but the choice of the indexes for determining the dissimilarity between IMFs plays a crucial role. Future studies will be aimed at investigating more effective solutions that do not rely on synthetic indexes, and make a better use of the overall information enclosed in each IMF.

References

[1] Jin J., Shi, J., Feature-Preserving Data Compression of Stamping Tonnage Information Using Wavelets, Technometrics, 1999, 41:4, 327 – 339

[2] Jin, J., Shi, J., Automatic Feature Extraction of Waveform Signals for In-process Diagnostic Performance Improvement, Journal of Intelligent Manufacturing, 2001, 12, 257 – 268

[3] Grasso, M., Pennacchi, P., Colosimo, B. M., Empirical mode decomposition of pressure signal for health condition monitoring in waterjet cutting, International Journal of Advanced Manufacturing Technology, 2014, 72:1-4, 347-364

[4] Grasso M., Goletti M., Annoni M., Colosimo B.M., A New Approach for Online Health Assessment of Abrasive Waterjet Cutting Systems, International Journal of Abrasive Technology, 2013, 6:2, 158-181

[5] Zhou S., Sun B., Shi J., An SPC Monitoring System for Cycle-Based Waveform Signals using Haar Transform, IEEE Transactions on Automation Science and Engineering, 2006, 3:1, 60-72

[6] Lada, E. K., Lu, J. C., Wilson, J. R., A wavelet-based procedure for process fault detection. Semiconductor Manufacturing, IEEE Transactions on, 2002, 15(1), 79-90

[7] Huang, N.E. Shen, Z. Long, S.R. Wu, M.L. Shih, H.H. et al., The Empirical Mode Decomposition and Hilbert Spectrum for Nonlinear And Nonstationary Time Series Analysis, Proc. R. Soc. London Ser. A, 1998, 454, 903–95

[8] Feng, Z. Liang, M. Chu, F., Recent Advances in Time-Frequency Analysis Methods for Machinery Fault Diagnosis: A Review with Application Examples, Mechanical Systems and Signal Processing, 2013, 38, 165-205

[9] Rilling, G. Flandrin, P. Goncalves, P., On empirical mode decomposition and its algorithms. IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03, 2003, Grado (Italy)

[10] Wu, Z. Huang, N.E., Ensemble Empirical Mode Decomposition: a Noise Assisted Data Analysis Method, Advances in Adaptive Data Analysis, 2009, 1:1, 1-41

[11] Zhang, J. Yan, R. Gao, R.X. Feng, Z., Performance Enhancement of Ensemble Empirical Mode Decomposition, Mechanical Systems and Signal Processing, 2010, 24, 2104-2123

[12] Gao, Q. Duan, C. Fan, H. Meng, Q., Rotating Machine Fault Diagnosis using Empirical Mode Decomposition, Mechanical Systems and Signal Processing, 2008, 22, 1072-1081

[13] Liang, H. Lin, Q-H. Chen, J.D.Z., Application of the Empirical Mode Decomposition to the Analysis of the Esophageal Manometric Data in Gastroesophageal Reflux Disease, IEEE Transactions on Biomedical Engineering, 2005, 52:10, 1692-1701, 2005

[14] Flandrin, P. Goncalves, P. Rilling, G., Detrending and Denoising with Empirical Mode Decomposition, XII European Signal Processing Conference (EUSIPCO), September 6-10, 2004, Vienna, Austria

[15] Ayenu-Prah, A. Attoh-Okine, N., A Criterion for Selecting Relevant Intrinsic Mode Functions in Empirical Mode Decomposition, Advances in Adaptive Data Analysis, 2010, 2:1, 1-24

[16] Bu, N. Ueno, N. Fukuda, O., Monitoring of Respiration and Heartbeat during Sleep using a Flexible Piezoelectric Film Sensor and Empirical Mode Decomposition, Proceedings of the 29th Annual International Conference of the IEEE EMBS, 2007, Lyon (France)

[17] Ricci, R. Pennacchi, P., Diagnostics of Gear Faults Based on EMD and Automatic Selection of Intrinsic Mode Functions, Mechanical Systems and Signal Processing, 2011, 25:3, 821-838

[18] Jain, A. K., Murty, M. N., Flynn, P. J., Data clustering: a review. ACM computing surveys (CSUR), 1999, 31(3), 264-323.