31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Midpoint-Based Empirical Decomposition for Nonlinear Trend Estimation

Qingbo He[1], Robert X. Gao[1*], and Patty Freedson[2]

Abstract—**This paper presents a new method for nonlinear trend estimation of non-stationary signals, by which the trend can be self-adaptively decomposed through calculating the midpoint-based local means. In this method, the so-called midpoints are proposed to construct the local mean of a signal instead of two envelopes in the classical empirical mode decomposition (EMD) algorithm, thus resulting in the midpoint-based empirical decomposition. Furthermore, a negentropy-based statistical method is presented to justify decomposition of the trend. Simulation results indicate that the new algorithm improves the performance of signal decomposition and trend estimation in comparison with the classical EMD algorithm. The proposed method also shows the value in self-adaptively estimating the nonlinear respiratory component from non-invasively measured ventilation signals.**

I. INTRODUCTION

TREND estimation is significant for analyzing, estimating and predicting the slow-varying trend of physical systems, including communication, climate, electrical power, economy, and biomedical applications [1]–[6]. Linear trend analysis is the commonly used model [1], [3]. Nonlinear trend is more general and useful in many applications [2], [5], [6]. This paper examines the use of nonlinear trend estimation from measured non-stationary signals, specifically from the respiratory signals measured by the piezoelectric sensor belts.

The nonlinear trend is often corrupted or even buried by noise. Effective estimation is thus very challenging. There are traditional methods to estimate such a trend. Since the slow-varying trend is mainly concentrated in the lower frequency range, low-pass filtering (LPF) can be used [5]. However, it needs a priori information of a signal's frequency characteristics in order to choose the cutoff frequency properly. Wavelet-based methods have shown to be effective in nonlinear trend estimation [7]. However the effectiveness depends on the proper choice of a basis wavelet function, because the choice of the wavelet filters determines the trend model. A data-driven technique can overcome these shortcomings, and empirical mode decomposition (EMD) [8] has shown the merit of producing basis functions from the signal itself. EMD can self-adaptively decompose the non-stationery signals, and thus has attracted increasing attention to trend estimation [6]. The EMD process is essentially a filtering operation, through which the signal is represented by the sum of a series of frequency bands [9]. The EMD can be thus applied to construct the low-pass filtering for estimating the slow-varying trend.

In the classical EMD algorithm, decomposition of a oscillation mode mainly depends on calculation of the upper and lower envelopes based on the maxima and minima of the signal. The mean envelope is then empirically determined as the local mean of the analyzed signal. However, the cubic spline fitting, a core step to construct the two envelopes, has both overshoot and undershoot problems [8]. These problems can be alleviated by applying higher-order spline fitting, which however increases the computational load. This paper proposes to represent the local mean by the midpoints of successive extrema instead of the envelope mean. The midpoints have the rational meaning of 'signal mean' and can thus be seen as the discrete points of the local mean of the signal. In this way, only one spline fitting is required to form the local mean rather than two in the classical algorithm. At the same time the number of interpolation points is increased to two times in comparison to the classical EMD method. Therefore, the midpoint-based decomposition is simpler and the interpolation effect can also be improved.

Furthermore, this paper also presents a new statistical method to automatically select one middle residue signal in the decomposition as the useful trend of the raw data. Through this technique, trend extraction and signal decomposition are integrated into one systematic frame. The performance of the proposed method is validated by simulation and the practical respiratory signal trend estimation.

II. METHOD

A. Signal Acquisition

The respiratory signals are obtained from two elastic piezoelectric sensor belts encircling the rib cage and abdomen. The experimental system is illustrated in Fig. 1. Healthy subjects were recruited from the University of Massachusetts Amherst for testing. The subjects performed a continuous test including 7-minute of standing and four, 7-minute treadmill exercise conditions: slow walking (2.4 km/h), fast walking (4.8 km/h), jogging (7.2 km/h) and running (9.6 km/h). A 2-minute rest separated all treadmill conditions. As shown in Fig. 1(b), the measured respiratory signals contain tissue artifact noise due to physical activity. The respiratory component should be extracted as the nonlinear trend by removing the tissue artifact noise for further ventilation estimation [10].

Manuscript received April 7, 2009. This work was supported by a grant from the National Institute of Health (NIH) under award UO1-CA130783.

The authors are with [1]Electromechanical Systems Laboratory, University of Connecticut, Storrs, CT 06269, USA, and [2]Exercise Physiology Laboratory, University of Massachusetts Amherst, Amherst, MA 01003, USA (correspondence: 860-486-7110, rgao@engr.uconn.edu).

978-1-4244-3296-7/09/$25.00 ©2009 IEEE

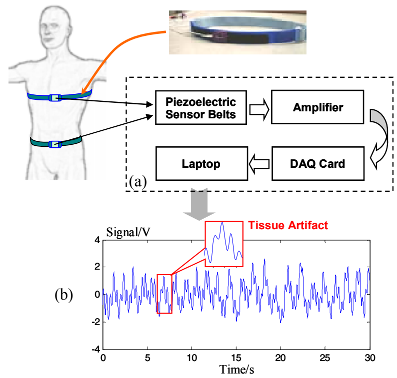

Fig. 1. Respiratory signal measurement: (a) data acquisition system; (b) sample of a measured signal.

B. Midpoint-Based Decomposition

For the classical EMD algorithm, the intrinsic mode functions (IMFs) is decomposed by using the envelopes to calculate the local mean. The overshoots and undershoots are the common phenomenon in the envelope calculation [8]. Considering the uncertainty of construction of two envelopes in the classical algorithm, it is ideal to determine the discrete point series on the local mean line, and then interpolate these points to obtain the local mean curve. Since the local mean is not a priori knowledge, the interpolation points are difficult to calculate. Inflection points of the signal, defined by the zeros of the second derivative $d^2x(t)/dt^2 = 0$, have been used to interpolate the local mean, but they were considered to deviate from the optimal interpolation points [11]. In our study, the midpoints of successive extrema are proposed for local mean estimation.

Mathematically, to the successive extrema of a signal at the time t_i and t_{i+1} (i=0,1,…,p) with the zeros of the first derivative $dx(t)/dt = 0$, their midpoint is defined as:

$$M_i^t = \frac{t_i + t_{i+1}}{2}$$
$$M_i^x = \frac{x_i + x_{i+1}}{2} \tag{1}$$

The midpoints (M_i^t, M_i^x) represent the geometrical mean points of the signal. As seen in Fig. 2, in the triangle connected by three successive extrema, the distance of each extreme to the line connecting two midpoints is the same. Hence, the midpoints can be seen as the discrete points of the signal local mean and used to construct the local mean.

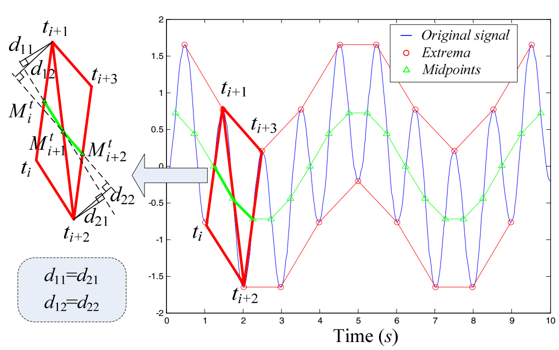

Fig. 2. Demonstration of the midpoint theory.

While the specific values of the maxima and minima of the signal being analyzed may change when noise is present, the nature of the midpoint calculation remains the same. The midpoint algorithm will not be disproportionally affected as compared to the traditional envelope-based method, as long as both methods are used to analyze the same signal, with or without noise. To the calculated p+1 midpoints, the piecewise cubic spline can be constructed in each interval [M^t_i, M^t_{i+1},], i = 0, 1, …, p–1 by

$$s_i(t) = a_i t^3 + b_i t^2 + c_i t + d_i \tag{2}$$

to formulate the local mean with the following boundary conditions:

$$\begin{cases} s_i(M_i^t) = M_i^x \\ s_i(M_{i+1}^t) = M_{i+1}^x \\ s'_{i-1}(M_i^t) = s'_i(M_i^t) & i = 1,2,\ldots,p-1 \\ s''_{i-1}(M_i^t) = s''_i(M_i^t) & i = 1,2,\ldots,p-1 \end{cases} \tag{3}$$

Since the total number of the maxima and minima is p+2, the midpoints have the amount p, which is nearly two times of that of either the maxima or the minima. As a result, compared to the envelope-based method, the midpoint-based method for local mean calculation has two advantages: simplicity as there is less interpolation operation and reliability of interpolation, since the intensity of interpolation points increases.

The obtained midpoint-mean curve is then taken as $m(t)$ for performing the empirical decomposition. The difference between the signal and the local mean, $x(t)$–$m(t)$, is designated as a new time series $h(t)$. The $h(t)$ represents the first IMF of the signal $x(t)$ if it satisfies two constraints as follows: the number of extrema and zero-crossings differs at most by one and the local mean is zero at any point [8]. Otherwise the above iteration process is repeated by taking $h(t)$ as the signal itself until it becomes an IMF. Through the iteration process, the IMFs $c_1(t)$, …, $c_n(t)$, and the residue signal $r_n(t)$, are achieved by consecutively subtracting the local mean from the signal $x(t)$. As a result, the signal $x(t)$ can be expressed as:

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) \tag{4}$$

C. Trend Decomposition

To address the problem of slow-varying trend mixed with higher-frequency noise, the slow-varying trend can be decomposed as a residue signal by the self-adaptive midpoint-based method. While the decomposed trend is non-Gaussian, the noise, which is composed of high-frequency IMFs, can be considered as having a nearly Gaussian distribution. This inherent difference can be explored to separate the trend from the noise. In this study, a statistical approach has been developed to automatically decompose the slow-varying trend. The statistics applied here is the negentropy, which is a measure for non-Gaussianity, defined as follows:

$$Ng(y) = H(y_{Gauss}) - H(y) \tag{5}$$

where y_{Gauss} is a Gaussian random variable with the same variance as random variable y, and $H(\bullet)$ is an entropy function. The negentropy is zero for Gaussian variable and always non-negative.

Specifically, negentropy is used to measure the non-Gaussianity of the residue signal $r_i(t)$ and the sum of several IMFs $HFC_i(t)$ in the ith order decomposition. Then the negentropy ratio of $HFC_i(t)$ and $r_i(t)$ is proposed as follows:

$$NR_i = \frac{Ng\left[r_i(t)\right]}{Ng\left[HFC_i(t)\right]} \tag{6}$$

where

$$HFC_i(t) = \sum_{j=1}^{i} c_j(t) \tag{7}$$

A big NR_i value corresponds to a higher possibility for the existence of the slow-varying trend in $r_i(t)$. In each order of decomposition, the NR_i ($i = 1, 2, \ldots, n$) is calculated. The NR_i curve will gradually drop and there will be an abrupt change on the curve when the trend is not included in the residue signal. The abrupt change just corresponds to decomposition of the slow-varying trend, and can be identified by the ratio of the NR_i in two decomposition processes:

$$RNR_i = \frac{NR_{i-1}}{NR_i} \tag{8}$$

Let $NR_0 = 0$, then $RNR_1 = 0$. The point with an abrupt increasing RNR_i value is regarded as the singular point to decompose the slow-varying trend $r_i(t)$.

Setting the singular point as the break condition in the decomposition algorithm, the trend extraction can be integrated into the systematic frame of the midpoint-based decomposition. As a result, the signal $x(t)$ can be decomposed to be:

$$x(t) = HFC_i(t) + r_i(t) \tag{9}$$

where $r_i(t)$ is the finally estimated slow-varying trend.

III. Simulation

To quantitatively investigate the effectiveness of the proposed midpoint-based decomposition method and the trend decomposition method, simulations were conducted as follows.

A. Signal Construction

A test signal is constructed as formulated below:

$$X(t) = \sum_{i=1}^{N} A_i \times \sin[2\pi f_i t + \theta_i] \tag{10}$$

where A_i, f_i, and θ_i are the amplitude, frequency and initial phase of applied N frequency components. The relevant parameters of the test signal for validating the new decomposition algorithm are listed in Table I.

TABLE I
PARAMETERS OF SIMULATED SIGNAL

i	1	2	3	4	5	6	7	8
A_i	0.1	0.2	0.3	0.3	0.3	0.3	0.4	0.5
f_i (Hz)	100	50	20	10	5	2	1	0.2
θ_i (π)	-0.5	0.5	0	0.6	1	-0.5	0	0.5

To investigate the effectiveness of the proposed trend decomposition method, a test nonlinear trend signal is first needed as a reference base. In this study, such a test signal was formulated based on a real respiratory signal measured in the standing activity. Through the inverse Fourier transform, 11 major frequency components identified in the respiratory signal spectrum were formulated by Eq. (10) to construct the test slow-varying trend. Together, the 11 frequency components represented over 92% of the energy content of the original signal. The constructed trend signal is considered to be nonlinear and not mono-component.

To conduct the simulation of trend estimation, white noise was generally considered to be the noise source to corrupt the slow-varying trend. The noisy signal is expressed as

$$S(t) = X(t) + e(t) \tag{11}$$

where $e(t)$ is the noise component defined by the signal-to-noise ratio (SNR):

$$SNR = 10\log_{10}\frac{\|X(t)\|_2}{\|e(t)\|_2} \tag{12}$$

where $\|X(t)\|_2 = \sum X(t)^2$, $\|e(t)\|_2 = \sum e(t)^2$.

B. Assessment Criteria

The decomposition error of the predefined signals is evaluated by the *Root Mean Square Error* (RMSE) as

$$RMSE = \sqrt{\frac{1}{L}\sum_{i=1}^{L}\left(x_i - \hat{x}_i\right)^2} \tag{13}$$

where x_i is the predefined value, $\hat{x}_i$ is the estimated value by the decomposition methods, and L is the signal length.

To assess the performance of the proposed method, the classical envelope-based EMD method is applied in the simulation study as a comparison.

C. Results

The constructed test signal with eight frequency components is decomposed by the proposed midpoint-based decomposition algorithm and the classical envelope-based EMD algorithm, respectively. As shown in Table II, the midpoint-based method revealed a smaller RMSE value than the envelope-based method by over 50% which indicates the midpoint-based method can obtain more accurate decomposition results.

TABLE II
DECOMPOSITION ACCURACY BY RMSE

Frequency (Hz)	100	50	20	10	5	2	1	0.2
Envelope (10^{-2})	3.54	3.63	6.06	8.83	7.88	19.17	16.48	23.82
Midpoint (10^{-2})	1.13	1.61	2.02	2.78	3.85	4.98	6.55	5.99
Improvement (%)	68.1	55.6	66.7	68.5	51.1	74.0	60.3	74.9

To verify the performance of trend decomposition, more SNRs are considered to formulate the noisy signal. As seen in Fig. 3, the results show that the midpoint-based algorithm is more effective for smaller RMSE than the classical envelope-based algorithm, and the improvement of the proposed method is more obvious for lower SNR cases. The analysis indicates the midpoint-based decomposition algorithm is valuable for nonlinear trend estimation from the noisy data.

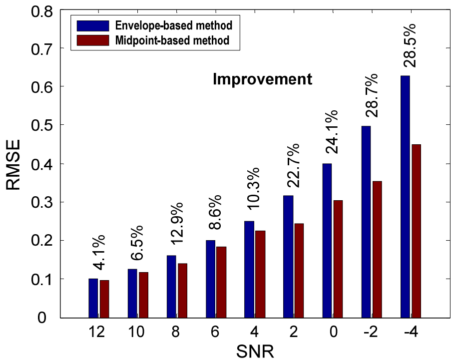

Fig. 3. Trend estimation performance comparison.

IV. Experimental Results

The proposed midpoint-based trend decomposition method is used to remove the tissue artifact noise from the measured respiratory signal. Considering the jogging activity, the measured respiratory signal was corrupted by severe tissue artifact noise riding on the respiratory component as small waves as shown in Fig. 4. Using the proposed method, at the fourth order decomposition, the *RNR* value has a clear abrupt increase with 200 times over the previous one. Therefore, the respiratory component of the signal was self-adaptively estimated at the third order decomposition. As seen in Fig. 4, the estimated respiratory component smoothly represents the global tendency in the local mean sense of the raw signal. The classical envelope-based EMD method was also compared. Two statistics, kurtosis and negentropy, of the removed tissue artifact noise were used as a quantitative measure. The negentropy is defined by Eq. (5) and the kurtosis is defined by

$$Kurt = \frac{E[HFC_i(t)-\mu]^4}{\sigma^4} - 3 \tag{14}$$

where μ is the mean and σ is the standard derivation of $HFC_i(t)$. They characterize the nature of the noise which represents the Gaussian property. Smaller statistical value corresponds to better effect. As seen from Table III, the kurtosis of classical method is 8.9 times that of the proposed method, and the negentropy ratio is 3.9. Therefore, the proposed method is more valuable for self-adaptively estimating the effective nonlinear trend of the measured respiratory signal.

TABLE III
Statistics of Removed Tissue Artifact Noise

	Kurtosis	Negentropy
Envelope method	0.2417	2.71×10^{-5}
Midpoint method	0.0272	6.94×10^{-6}
Ratio (Envelope/Midpoint)	8.9	3.9

V. Conclusion

A new method has been proposed to estimate the nonlinear trend with an application to remove tissue artifact from the non-invasive measured respiratory signal. The proposed midpoint-based decomposition algorithm has shown to be simpler and more accurate in calculating the local mean of a signal for further iterative IMF decomposition. The presented negentropy-based method has shown to be effective in justifying decomposition of the trend. The advantage of the proposed midpoint-based trend decomposition method lies in its ability to self-adaptively, automatically and reliably estimate the nonlinear trend. Simulation performance using the formulated signals has confirmed the effectiveness of the proposed method for signal decomposition and nonlinear trend estimation from the contaminated signal outperforming the classical EMD method. Application to the respiratory component extraction has also shown the value of the proposed method for self-adaptive nonlinear trend estimation.

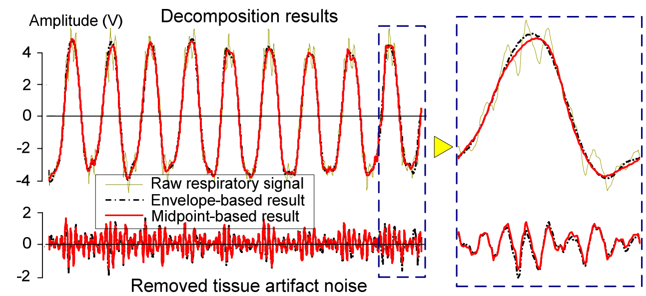

Fig. 4. Respiratory component decomposition by the proposed method and the performance comparison with the traditional EMD method.

Acknowledgment

The authors would like to thank Mr. Shaopeng Liu and Dr. Todd Hagobian for their assistance with the data collection.

References

[1] T. Trump, "Maximum likelihood trend estimation in exponential noise," *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 2087-95, 2001.

[2] J. L. Minetti, W. M. Vargas, A. G. Poblete, L. R. Acuna, and G. Casagrande, "Non-linear trends and low frequency oscillations in annual precipitation over Argentina and Chile, 1931–1999," *Atmosfera*, vol. 16, pp. 119-135, 2003.

[3] E. González, M. A. Jaramillo, and D. Carmona, "Monthly electric energy demand forecasting based on trend extraction," *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1946-1953, 2006.

[4] L. Peretto, R. Sasdelli, and R. Tinarelli, "A statistical model for estimating the trend of electrical quantities in power systems," *IEEE Trans. Instrum. Meas.*, vol. 52, no. 4, pp. 1143-1147, 2003.

[5] D.S.G. Pollock, "Trend estimation and de-trending via rational square-wave filters," *J. Econometrics*, vol. 99, no. 2, pp. 317-334, 2000.

[6] H. Liang, Q.-H. Lin, and, J. D. Z. Chen, "Application of the empirical mode decomposition to the analysis of esophageal manometric data in gastroesophageal reflux disease," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 10, pp. 1692-1701, Oct., 2005.

[7] R. Tona, C. Benqlilou, A. Espuna, and L. Puigjaner, "Dynamic data reconciliation based on wavelet trend analysis," *Ind. Eng. Chem. Res.*, vol. 44, no. 12, pp. 4323-4335, 2005.

[8] N. E. Huang, Z. Shen, and S. R. Long, "The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. London A*, vol. 454, pp. 903-995, 1998.

[9] P. Flandrin, G. Rilling, and P. Gonçalvés, "Empirical mode decomposition as a filter bank," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 112-114, 2004.

[10] K. Konno and J. Mead, "Measurement of the separate volume changes of rib cage and abdomen during breathing," *J. Appl. Physiol.*, vol. 22, no. 3, pp. 407-422, 1967.

[11] Y. Kopsinis and S. McLaughlin, "Investigation and performance enhancement of the empirical mode decomposition method based on a heuristic search optimization approach," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 1-13, 2008.