Extraction of Baseline Based on Second-Generation Wavelet Transform

Jiancai Wang(✉)

Office of Academic Affairs, Hei Longjiang University of Science and Technology, Harbin 150022, China
154539860@qq.com

Abstract. In the analysis of signals processing, due to the various kinds of interference in the transformation and sampling of the analytical instruments, the baseline of the signals is presented in the upper and lower drift. The upper and lower baseline could affect the accuracy of quantitative calculation, analysis and evaluation. In the study, the principle of second-generation wavelet is discussed and introduced to extract the baseline. The features of signals are analyzed and the quantitative accuracy of components has been significantly improved by the baseline extraction. The second-generation wavelet method successfully realizes the split of baseline from the signal peak with high efficiency and is easy to be implemented.

Keywords: Second-generation wavelet transform · Signals processing · Baseline extraction

1 Introduction

In the process of collecting signals, due to there are the various kinds of interference in the transformation and sampling of the analytical instruments, the baseline of the signals is presented in the upper and lower drift. The baseline shift can change the shape of the signal peak, affecting peak height and peak area calculation, which has a bad effect on the components calculation, quantitative analysis and evaluation [1–3].

There are various kinds of methods for baseline shift correction, such as Digital filtering, Baseline fitting, Adaptive filtering and Wavelet transform [4] etc. Baseline fitting is by means of mathematics, a function model (usually n order polynomial) is established to describe the baseline and then to correct, but it's difficult to deal with in technology and real-time processing is not high. In addition, the curve fitting method is also used to correct baseline shift, but the TLC baseline spectrum is a slowly varying frequency signal, when the signal is weak, the difference was very difficult to extract.

Recently, the method of wavelet transform is widely applied in traditional Chinese medicine fingerprint and high performance liquid chromatography fingerprint to correct baseline shift. In 1995, Sweldens proposed a new wavelet construction method, which does not depend on Fourier transform, the lifting scheme (Scheme Lifting), which is called the second generation wavelet transform (SGWT) [5]. In the transformation process, through the design of predictor and updater, the wavelet function is obtained with

© ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2020
Published by Springer Nature Switzerland AG 2020. All Rights Reserved
X. Jiang and P. Li (Eds.): GreeNets 2020, LNICST 333, pp. 156–161, 2020.
https://doi.org/10.1007/978-3-030-62483-5_17

expected properties [6]. Using the lifting wavelet to carry out baseline shift correction, signal denoising, it has the characteristics of simple structure, fast operation, saving storage space and realizing the integer wavelet transform.

In the study, the D4 wavelet is used in the process of exacting baseline, which will make a correction effect based on the discussion of the theory of second generation wavelet transform, with the advantages of simple coefficients and a few floating-point calculations, the method has a good inhibiting effect on baseline shift.

2 Methods

There are three steps to realize the second generation wavelet transform (SGWT) constituted by the lifting process: Split, Predict and Update. The reconstruction process is the inverse [7]. The specific implementation is shown in Fig. 1.

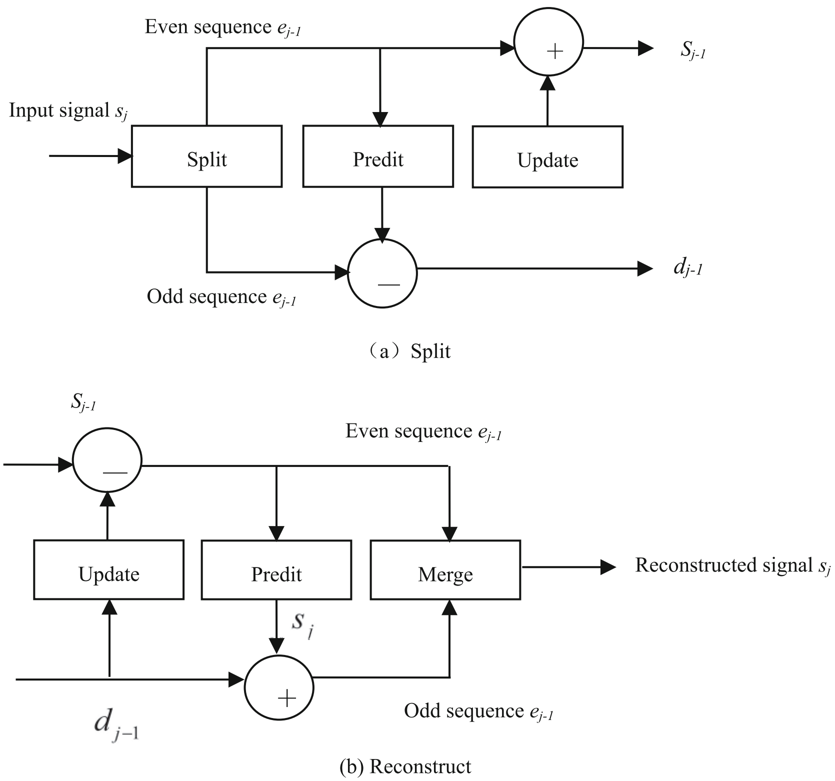

Fig. 1. Split and reconstruction process of SGWT

2.1 Split

The original signal is divided into two subsets, which is usually divided into even sequences and odd sequences. The length of each subset is half of the atom sets and does not intersect each other. That is:

$$split(s_j) = (e_{j-1}, o_{j-1}) \tag{1}$$

In the formula, $e_{j-1} = \{e_{j-1,k} = s_{j,\,2k}\}$, $o_{j-1} = \{o_{j-1,\,k} = s_{j,\,2k+1}\}$

2.2 Predict

According to the correlation between the sequences, generally, the odd sequence is predicted by the even sequence. In the prediction, there is a deviation dj − 1 between the actual value Oj − 1 and the predicted value p (ej − 1), the difference is called the detail coefficient or the wavelet coefficient, which reflects the degree of approximation. The amplitude change of wavelet coefficient is inversely proportional to the correlation of the data. The prediction process as follows:

$$d_{j-1} = o_{j-1} - P(e_{j-1}) \tag{2}$$

In fact, though it is impossible from the subset e_{j-1} to predict the subset o_{j-1} accurately, p (e_{j-1}) may be close to o_{j-1}. In this way, d_{j-1} contains less information than the original. Repeat the process of split and predict, after n steps, the original signal sets can be represented by $\{e_n, o_n, \ldots\ldots, e_1, o_1\}$.

2.3 Update

After the procedure of split, some of the overall characteristics (such as the mean) of the subsets may not be consistent with the original data. At this time, the procedure of update is needed to maintain the overall characteristics of the original data. The process as follows:

$$s_{j-1} = e_{j-1} + U(d_{j-1}) \tag{3}$$

P and U take different functions, can be constructed by different wavelet transforms.

Sweldens has been proved to make Integers be set to an integer wavelet transform based on the lifting algorithm [8]. An integer set is obtained by SGWT and it also proves that any traditional wavelet can be converted into the corresponding second generation integer wavelet. Select the appropriate algorithms of prediction and lifting for discrete data samples to be processed, the required frequency band is obtained [9]. For the signal data, DB wavelet as the prediction and lifting algorithm is chosen in the study, the second generation of discrete wavelet construction scheme is used to extract the baseline from the original signal.

3 Results and Discussion

3.1 Description of the Algorithm

There are many commonly used wavelet functions such as Haar wavelet, Shannon wavelet and Daubechies wavelet. Due to the frequency spectrum of TLC baseline signal is low, the trend is relatively flat, and the detection effect of Daubechies wavelet to non-stationary signals is better, D4 wavelet function is used to split and reconstruct TLC signal to realization the extraction of the baseline.

The wavelet function of Daubechies wavelet is:

$$\phi(t) = \sqrt{2}\sum_{n=0}^{2N-1} g_n \phi(2t - n) \tag{4}$$

The scaling function of Daubechies wavelet is:

$$f(t) = \sqrt{2}\sum_{n=0}^{2N-1} h_n f(2t - n) \tag{5}$$

Thereinto,the coefficient of High-pass filter is g_n, the coefficient of Low-pass filter is h_n. The z transoform of D4 wavelet is $h(z) = h_0 + h_1 z^{-1} + h_2 z^{-2} + h_3 z^{-3}$,$g(z) = -h_3 z^{-2} + h_2 z^1 - h_1 + h_0 z^{-1}$. Thereinto, the coefficients of Low-pass filter are $h_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}$, $h_1 = \frac{3+\sqrt{3}}{4\sqrt{2}}$, $h_2 = \frac{3-\sqrt{3}}{4\sqrt{2}}$, $h_3 = \frac{1-\sqrt{3}}{4\sqrt{2}}$. The improved Laurent polynomial Euclidean algorithm is used to factorize the multi-phase matrix p (z) of D4 wavelet filter.

$$p(z) = \begin{pmatrix} h_0 + h_2 z^{-1} & -h_3 z^1 - h_1 \\ h_1 + h_3 z^{-1} & h_2 z^1 + h_0 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & -\sqrt{3} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \frac{\sqrt{3}-2}{4} z^{-1} & 1 \end{pmatrix}\begin{pmatrix} 1 & \frac{2}{\sqrt{3}} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & \frac{1}{4}z - \frac{1}{\sqrt{3}} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \frac{1+\sqrt{3}}{2\sqrt{2}} & 0 \\ 0 & \frac{2\sqrt{2}}{1+\sqrt{3}} \end{pmatrix}$$

Therefore, the procedures of baseline extraction are described as follows:

(1) The input signal is split into even and odd. The original signal x with length N is split into even sequence s_l^0 and odd sequence d_l^0 respectively. $s_l^0 = x_{2l}$, $d_l^0 = x_{2l+1}$,l = 0,1,…,N/2-1.

(2) The steps of predict and update are carried out alternately. The predictive value of the P dual signal of the filter is used as the odd signal. There are four steps to enhance and dual lifting obtained:

$s_l^1 = s_l^0 + \sqrt{3} d_l^0$, $d_l^1 = d_l^0 - \frac{\sqrt{3}-2}{4} s_{l-1}^1$, $s_l^2 = s_l^1 - \frac{2}{\sqrt{3}} d_l^1$, $d_l^2 = d_l^1 - \frac{\sqrt{3}}{2} s_l^2$, $s_l^3 = s_l^2 - \frac{1}{4} d_{l+1}^2 + \frac{1}{\sqrt{3}} d_l^2$.

(3) Finally, scale transformation is carried out. Scale transform is used to get $s_l = \frac{\sqrt{3}+1}{2\sqrt{2}} s_l^3$, $d_l = \frac{2\sqrt{2}}{\sqrt{3}+1} d_l^3$, here s and d are the low frequency and the

high frequency of the wavelet decomposition respectively. Among them, $s = \{s_0, s_1, \ldots, s_{N/2-1}\}$, $d = \{d_0, d_1, \ldots, d_{N/2-1}\}$, $l = 0, 1, \ldots, N/2-1$.

(4) The reconstruction process of the second generation wavelet decomposition is basically the inverse process. First performed scaling, and then update and predict steps, finally carry out the parity inversion to reconstruct the signal.

3.2 Extraction of the Baseline

Before extraction of the baseline, the signal denosing is carried out first. The separation of the high-frequency signals and the low-frequency signals is achieved by Wavelet transform. After separating the high frequency and low frequency signals, the high frequency part of the corresponding peak is zero, and the reconstructed data contains low frequency data so that the baseline can be obtained from the reconstructed data. The method is applied in the TLC. The sample of TLC is obtained by Thin layer chromatography scanner, and then light density detection. The original signal and the result after baseline extraction are shown in Fig. 2.

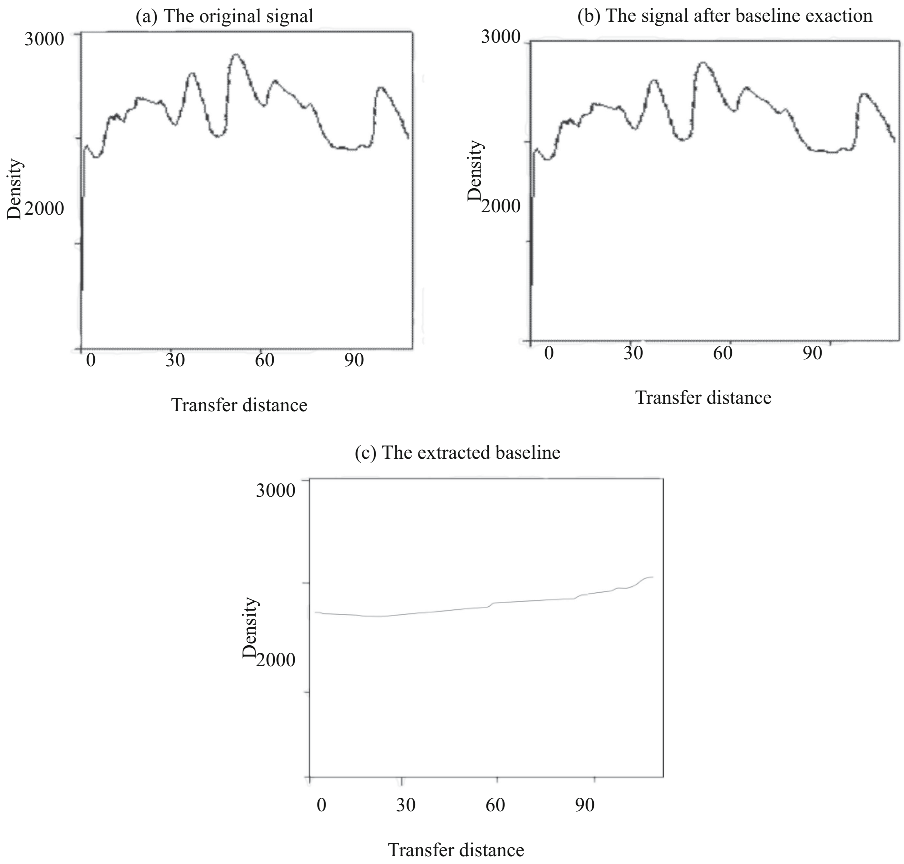

Fig. 2. The original signal and the result after baseline extraction

4 Conclusion

Experiments show that, with the basic principles of the second generation wavelet transform the second generation D4 wavelet is used to extract the baseline based on the lifting algorithm, which not only to speed up the processing speed, but also effectively suppress baseline shift. After extraction, the quantitative accuracy of crude oil components has been significantly improved. In the study, the method of baseline extraction on TLC signal has a certain practical significance and application values, but the general applicability and stability of this method is verified by a large number of real sample analysis.

Acknowledgements. The work was supported by Hei longjiang Fundamental Research Funds for the Local Universities in 2018 (No. 2018-KYYWF-1189).

References

1. Jiangtie, L.,Wei, L., Zhixia Z, Xianchun L. Application of thin layer chromatography fingerprint in the quality evaluation of Salvia miltiorrhiza. J. Xiamen Univ. (Nat. Sci.) **44**(6), 801–805 (2005)
2. Yun, Y., Xuefeng, Z.: Baseline spliting of traditional Chinese medicine fingerprint based on the second generation wavelet transform, J. South China University of Technol. (Nat. Sci. Ed.) **35**, 49–52 (2007)
3. Jianxin, C.: Study on the method of visible pattern analysis of fingerprint of traditional Chinese medicine based on the representation of multiple maps. Yanshan University, Hebei (2012)
4. Xiangkui, W.: ECG baseline drift depress algorithm based on multi-resolution analysis. Comput. Eng. Des. **29**(13), 3482–3484 (2008)
5. Sweldens, W.: The lifting scheme: a construction of second generation wavelets SIAM. J. Math. Aanal. **29**(2), 511–546 (1997)
6. Daubechies, Ingrid: Ten Lectures on Wavelets. National Defence Industry Press, Beijing (2004)
7. Xinjun, L.: An method of improved wavelet threshold denoising. Sci. Technol. Vis. **25**, 62–63 (2014)
8. Daubechinces, C.R., Seldens, W.: Wavelet transforms that ma Pinter-gers to intergers. Applied and computational hamonic analysis (1998)
9. Yongqiang, W., Haifeng, Y.: ECG signal denoising algorithm based on second generation wavelet transform. J. China Sci. Technol. Expo. **1**, 95–99 (2012)