Chemometrics and Intelligent Laboratory Systems 156 (2016) 54–61

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemolab

Recovery of excitation-emission fluorescence spectroscopy based on structural similarity

Shaohui Yu [a,*], Xue Xiao [b]

[a] School of Mathematics and Statistics, Hefei Normal University, Hefei 230061, China
[b] Key Laboratory of Environmental Optics & Technology, Anhui Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Hefei 230031, China

ARTICLE INFO

Article history:
Received 22 January 2016
Received in revised form 13 May 2016
Accepted 17 May 2016
Available online 20 May 2016

Keywords:
Wavelet transform
Empirical mode decomposition
EEM fluorescence spectroscopy
Structural similarity

ABSTRACT

Three-dimensional excitation-emission fluorescence spectroscopy technique is very important for the detection of trace organic matters. However, the weak fluorescence intensity is often submerged by the strong background, which is mostly Gaussian white noise. Based on the multi-scale wavelet transform (WT) and empirical mode decomposition (EMD), the WT-EMD method is proposed to extract the weak fluorescence spectroscopy from the strong noise. Firstly, excitation-emission matrix (EEM) fluorescence spectroscopy is decomposed into a set of intrinsic mode functions (IMFs) with different frequencies by EMD, and then the high-order IMFs are denoised by WT. By selecting the wavelet function and the input signal to noise ratio, four different evaluation indices of the WT method and the WT-EMD method are discussed. Furthermore, a robust evaluation index based on structural similarity is proposed. Numerical experiments show that the WT-EMD method can recover the weak EEM fluorescence spectroscopy from the strong noise, even when the input signal to noise ratio is relatively low. In other words, this result means the WT-EMD method based on structural similarity is not only less affected by WT, but also eliminates the model mixing of EMD.

1. Introduction

In recent years, spectral technology has been applied to more and more fields because of its high sensitivity, good selectivity and so on. One of the important fields is the detection of trace organic matter in water [1,2]. Especially, EEM fluorescence spectroscopy, as a kind of fingerprint, has a unique advantage in the detection of trace organic matter because of its large data volumes and abundant information. EEM fluorescence spectroscopy includes not only the excitation spectroscopy but also the emission spectroscopy. Moreover, EEM fluorescence spectroscopy with the second-order correction algorithms of chemometrics [3–6] greatly improves the accuracy of qualitative and quantitative analysis, which is especially suitable for the detection of trace organics. Unfortunately, it is inevitable to generate some noise during the data acquisition, which is adverse to the detection of trace organic matter. For certain signal to noise ratio, some second order correction algorithms can complete qualitative and quantitative analysis as the noise is treated as the unknown interference [3,5]. In addition, the noise components can also be extracted by the Zernike moments to complete the quantitative analysis [7]. Because the low order Zernike moment involves the low frequency information, and the higher order Zernike moment gets the high frequency information, so the quantitative analysis model based on the progressive regression can be established by removing the noise corresponding to the Zernike moment. Oscar et al. also discusses the classification method [8], which is less affected by noise. This method combines the latent Dirichlet allocation and the parallel factor method to avoid the noise. Another part of the work is the direct denoising methods, such as the Monte Carlo [9], the morphological method [10] and the wavelet transform [11]. Undoubtedly, these methods well deal with EEM fluorescence spectroscopy under a certain signal to noise ratio. But under lower signal to noise ratio, the denoising of EEM fluorescence spectroscopy will become into the recovery of weak fluorescence spectroscopy which is similar but different to the denoising. Although the two-dimensional wavelet transform has become an advanced denoising method [12,13], its parameter selection often complicates the denoising process [14]. In this paper, a new method is presented for the recovery of the noisy EEM fluorescence spectroscopy, which is based on the two-dimensional wavelet transform and empirical mode decomposition. And a robust evaluation index based on structural similarity is subsequently proposed, which solves the inconsistency of the denoising evaluation index. The rest of the paper is organized as follows. The basic theory of wavelet transform (WT) is introduced in the second part. The third part mainly focuses on the decomposition steps of the EEM fluorescence spectroscopy by empirical mode decomposition (EMD). In the fourth part, we propose the WT-

* Corresponding author.
E-mail address: yushaohui2005@163.com (S. Yu).

http://dx.doi.org/10.1016/j.chemolab.2016.05.011
0169-7439/

EMD method. Numerical experiments will be discussed in the fifth part and the structural similarity is also proposed in this part. Finally, the summary to the full text is done.

2. Two-dimensional wavelet transforms (2D-WT)

After years of development, the wavelet transform not only makes great progress in theory, but also has attracted more and more attention on applications, especially on the infrared spectroscopy [15] and hyperspectral image [16]. EEM fluorescence spectroscopy $X(i,j) \in R^{I \times J}$ (I,J denote the number of excitation wavelength and emission wavelength, respectively) can be expressed as:

$$X = S + N \tag{1}$$

where S represents the pure spectra and N is Gauss white noise. The wavelet denoising is to eliminate N. For EEM fluorescence spectroscopy $X(i,j)$, the two-dimensional wavelet transform with two-dimensional wavelet basis $\varphi(i,j)$ is defined as

$$W_f(a,b_1,b_2) = \frac{1}{a}\iint X(i,j)\varphi\left(\frac{i-b_1}{a},\frac{j-b_2}{a}\right)didj \quad (a\neq 0) \tag{2}$$

where a is the scaling factor and b_1,b_2 are the translation factors. The wavelet denoising is usually completed by dealing with the wavelet coefficients at some scales, and then the wavelet coefficients are reconstructed.

There are mainly three steps in the denoising of EEM fluorescence spectroscopy by two-dimensional wavelet transform.

Algorithm 1.

Step1 Decomposition of EEM fluorescence spectroscopy. Select a wavelet function $\varphi(i,j)$ and a proper decomposition level M, and then calculate the multilayer decomposition of the EEM fluorescence spectroscopy.

Step2 Threshold processing of high frequency coefficients. For each layer from 1 to M, select a threshold value and do threshold processing on the high frequency coefficients.

Step3 Reconstruction of EEM fluorescence spectroscopy. Reconstruct EEM fluorescence spectroscopy by the low frequency coefficients of the last layer and the other coefficients of each layer after the threshold processing.

During the denoising process, there are two important factors. One is how to select the wavelet function and the other is the threshold value of each layer.

Any function satisfying the wavelet conditions can be regarded as the wavelet function, and therefore there are many wavelet functions, such as Haar wavelet, Daubechies wavelet, SymletsA wavelet, Biorthogonal wavelet, coiflet wavelet, Morlet wavelet, and Meyer wavelet. For all their complexity, there are many different properties such as the length of the compact support, the length of the filter, the symmetry and the vanishing moments that makes it difficult to choose. So most of the time, they depend on experience.

Secondly, the threshold value is also a key problem for the denoising results. At present, there are two widely methods, hard threshold and soft threshold [17,18].

Hard threshold

$$\tilde{x} = \eta(x) = \begin{cases} x, & |x| \geq \tau \\ 0 & |x| < \tau \end{cases} \tag{3}$$

Soft threshold

$$\tilde{x} = \eta(x) = \begin{cases} \mathrm{sgn}(x)(|x| - \tau), & |x| \geq \tau \\ 0, & |x| < \tau \end{cases} \tag{4}$$

Each of the two thresholds has their advantages and disadvantages. Hard threshold can retain the edge and other local features, but there will be a pseudo Gibbs effect and other distortions. While the denoising results of the soft threshold are relatively smooth, the edge will appear distorted. In view of the importance of fluorescence peak in the EEM fluorescence spectroscopy, the soft threshold processing is adopted in all the following numerical experiments.

3. Two-dimensional empirical mode decomposition (2D-EMD)

Two-dimensional empirical mode decomposition has been applied to the denoising [19]. EEM fluorescence spectroscopy can also be decomposed by the following steps.

Algorithm 2.

Step1 Initialize the noisy EEM fluorescence spectroscopy, $H_0 = X$,$R_0 = X$.

Step2 Extract the i-th imf(intrinsic mode function) component.

2.1 Find out the maximum point set Max and the minimum point set Min ofH_{i-1}. Complete the surface fitting, and then calculate the envelope surface and the mean value $Mean$ of the envelope surface.

2.2 Estimate whether the value $H_i = H_{i-1} - Mean$ meets the sifting criteria. If not, go to (2.1).

Step3 Set $imf_i = H_i$,$R_i = R_{i-1} - imf_i$. If the maximum points of R_i is more than 2, then reset $H_{i-1} = R_i$ and go to step 2.

Finally, EEM fluorescence spectroscopy is decomposed into a set of IMFs by 2D-EMD.

$$X = \sum_{i=1}^{I} imf_i + R_I \tag{5}$$

Although external points are found out by the morphological method and the neural network of the radial basis function accomplishes the curve surface fitting [20–22], the denoising process is not effective because of the mode mixing problem [23]. So the mode mixing attracts more and more attentions. Ensemble empirical mode decomposition method (EEMD) [24,25] eliminates the mode mixing by adding the noise and averaging the noise. While the noise-assisted multivariate empirical mode decomposition (N-A-MEMD) [26] and the masking signal method [27] solve the mode mixing problem by adding the noise and the masking signal. In short, the pretreatment of original fluorescence spectroscopy is currently the primary method to avoid the mode mixing.

4. WT-EMD method

Undoubtedly, the wavelet transform has certain advantage in the denoising, but the parameter selection often complicates its application. Conversely, the self adaptability of EMD avoids too many parameters, though the modal mixing affects the denoising results. Therefore, a good method should make full use of WT and EMD. Firstly, EEM fluorescence spectroscopy is decomposed into IMF components with different frequency by EMD. And then the high-frequency IMF components are denoised by WT. It seems very helpful to denoise and recover weak fluorescence spectroscopy.

Recovery of EEM fluorescence spectroscopy by the WT-EMD method is carried out as followed.