Chemometrics and Intelligent Laboratory Systems 156 (2016) 157–165

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

Nonlinear diffusion filtering for peak-preserving smoothing of a spectrum signal

Yuanlu Li [a,b,*], Yaqing Ding [a], Tiao Li [a,b]

[a] B-DAT, School of Information and Control, Nanjing University of Information Science & Technology, Nanjing, China, 210044
[b] Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology, Nanjing University of Information Science & Technology, Nanjing, China, 210044

ARTICLE INFO

Article history:
Received 4 February 2016
Received in revised form 7 June 2016
Accepted 9 June 2016
Available online 11 June 2016

Keywords:
Spectra
Nonlinear diffusion
Peak-preserving smoothing
Regularization method
Wavelet method
Savitzky-Golay method

ABSTRACT

How to reduce the noise while preserving the peak is a challenging task in analytical techniques. In this paper, the nonlinear diffusion was proposed as a general method to accomplish peak-preserving smoothing. The implement of the nonlinear diffusion is simple. Taking the noisy signal as the initial condition of a nonlinear diffusion equation, the solution is a smoothed signal, and signal becomes increasingly smooth as iteration number increases. Details of the nonlinear diffusion filtering and its implementation were given clearly. Some simulated signals and an NMR spectrum has been used to verify the proposed method and compare the performance of other methods such as regularization method, Savitzky-Golay method and wavelet method. Results indicated that the nonlinear diffusion is an excellent smoothing method, it can reduce the noise while preserve the peak shape.

1. Introduction

In most of analytical methods, the position of a peak is important in the spectrum. In chromatography, e.g. the peak position is characteristic of the compound, and in spectrophotometry, a peak-shift may indicate the interaction of two compounds or even a chemical modification of the absorbing species. In many cases, the spectra are noisy [1–4]. Therefore smoothing is often desired to help with further analysis. For example, if derivatives of the signal are needed, the signal must be especially smooth otherwise the noise will be severely amplified in the derivative signal [1,2].

At present, there are many methods that can be used to reduce noise in a spectrum [1,3,5–11]. Moving average method is one of the simplest one. The Savitzky–Golay method is one of improvements of sliding average method [9], where a low-order polynomial is used to fit the data within a moving window rather than just taking their average. Usually, signal becomes increasingly smooth as the window size increases. On the other way, too broad of a window will reduce the effect of the resolution enhancement and distort the derivative spectra. The best parameters for the Savitzky–Golay method are selected usually by a trial-and-error method [12]. Regularization method, as an improved alternative to the Savitzky–Golay smoother, is a good smoothing method [1,11]. Similar method includes using a spline to smooth the spectra [8].

Other smoothing methods include wavelet method [6,7,13–15], in which the spectrum is decomposed by the discrete wavelet transform (DWT) and the high frequency components under threshold value are discarded and then the inverse wavelet transform used to reconstruct the spectrum.

In this paper, we use the nonlinear diffusion filtering to smooth the spectra. In fact, the nonlinear diffusion filtering is not new, and has been widely used in image processing. The nonlinear isotropic diffusion equation was firstly used to image denoising and edge detection by Perona and Malik [16], and it has been considered as an edge-preserving denoising method. Anisotropic diffusion can be looked as an energy-dissipating process that finds the minimum of the energy functional. If the energy functional is the total variation norm of an image, then the model is called as the well-known total variation model [17,18]. Although these methods can be able to achieve a good trade-off between noise reduction and edge preservation, the resulting image will easily lead to piecewise constant. To overcome the blocky effect, while preserving edges, many other nonlinear filtering methods based on the partial differential equation have been suggested in the literature [18–24].

It is surprising that the nonlinear diffusion filtering is generally unknown and little used among spectra smoothing, perhaps, in part, because the details for implementation are glossed over in the literature. For this purpose, details of the nonlinear diffusion filtering and its implementation are given as clearly as possible in this paper. The advantage of the nonlinear diffusion filtering will be highlighted by comparison with other smoothing methods.

* Corresponding author at: B-DAT, School of Information and Control, Nanjing University of Information Science & Technology, Nanjing, 210044, China.
E-mail address: lyl_nuist@nuist.edu.cn (Y. Li).

http://dx.doi.org/10.1016/j.chemolab.2016.06.007
0169-7439/

2. Linear diffusion filtering

2.1. Homogeneous linear diffusion

Witkin [25] found that the solution of a heat diffusion equation with a signal as initial data was equivalent to the convolution of the signal with Gaussians function at each scale. This process can be described as the solution of the homogeneous linear diffusion equation:

$$\frac{\partial U(x,t)}{\partial t} = \frac{\partial^2 U(x,t)}{\partial x^2}, \quad U(x,0) = f(x). \tag{1}$$

The solution of the Eq. (1) is given by the convolution integral

$$U(x,t) = G_t * f, \tag{2}$$

where G_t denotes the Gaussian function with standard deviation σ:

$$G_\sigma(x) = C\sigma^{-\frac{1}{2}} \exp\left(\frac{-x^2}{4\sigma}\right) \tag{3}$$

The original signal corresponds to the scale $t = 0$ and larger values of t, correspond to a more smoothed signal. But the Gaussian smoothing has a main disadvantage for signal smoothing. It does not only reduce noise, but also blurs the peaks. Since Gaussian smoothing cannot take into account any a priori information on the structures which are worth being preserved.

2.2. Inhomogeneous linear diffusion

In order to reduce the problem of homogeneous linear diffusion filtering, one of the simplest models for including a priori information is inhomogeneous linear diffusion filtering [26]. To preserve signal peaks in a better way than Gaussian smoothing, we hope smoothing is weak at peaks. So we can take $|f''(x)|$ as a fuzzy peak detector: locations with large $|f''(x)|$ have a higher probability to be a peak. Therefore, one can make the diffusion is low for larger value of $|f''(x)|$, for example by setting

$$g\left(\left|f''(x)\right|\right) = \frac{1}{1 + \left|f''(x)\right|^2/\lambda^2}. \tag{4}$$

However, the diffusion equation remains linear:

$$\frac{\partial U(x,t)}{\partial t} = div\left(g\left(\left|f''(x)\right|\right)\frac{\partial U(x,t)}{\partial x}\right) \tag{5}$$

From the Ref. [27], we can know the inhomogeneous linear diffusion filtering can improve peaks preserving. However, for larger t the filtered signal reveals some spurious details which reflect the differential structure of the initial signal.

3. Nonlinear diffusion filtering

3.1. The Perona-Malik model

In order to reduce spurious details of inhomogeneous linear diffusion filtering, Perona and Malik [16] proposed the nonlinear diffusion equation by adapting the diffusion function g to the gradient of the actual image $U(x,t)$ instead of the original image $f(x)$. Their main creative is to introduce the gradient of the actual image $U(x,t)$ as the feedback into the diffusion process. The process can be formulated as follows:

$$\frac{\partial U(x,t)}{\partial t} = div\left(g(x,t)\frac{\partial U(x,t)}{\partial x}\right), \tag{6}$$

with initial and boundary conditions:

$$U(x,0) = f(x), 0 \le x \le L \tag{7}$$

$$U(0,t) = U(L,t) = 0, 0 < t < T \tag{8}$$

where $g(x,t)$ is the diffusion function. In general, $g(x,t)$ is a smooth decreasing function with $g(0,t) = 1$, $g(x,t) \ge 0$ and $g(x,t)$ tending to zero at infinity. The diffusion strength is controlled by $g(x,t)$.

Two different diffusion functions have been proposed [16]:

$$g_1(x,t) = exp\left[-\left(\frac{|\nabla U(x,t)|}{\lambda}\right)^2\right] \tag{9}$$

$$g_2(x,t) = \frac{1}{1 + \left(\frac{|\nabla U(x,t)|}{\lambda}\right)^2} \tag{10}$$

If the gradient $|\nabla U(x,t)|$ is large, $g(x,t)$ is small. This means the diffusion is weak. Conversely, if the gradient $|\nabla U(x,t)|$ is small, $g(x,t)$ is large, the diffusion is strong. The parameter λ is chosen by the noise level and the edge strength. A proper choice of the diffusion function can preserve the edges and even enhance them while being numerically stable [16].

In this paper, we take use of the spectrum to be smoothed as the reference signal of the diffusion function. So the diffusion functions (9) and (10) will be correspondingly replaced by

$$g_3(x,t) = exp\left[-\left(\frac{|u(x,t)|}{\lambda}\right)^2\right] \tag{11}$$