

An Efficient Profile Detection Method for Fiber Spectrum Images with Low SNR Based on Wigner Bispectrum

Jia Zhu^A, Zhangqin Zhu^A, and Zhongfu Ye^{A,B}

^A Institute of Statistical Signal Processing, Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei 230027, China

^B Corresponding author. Email: yezf@ustc.edu.cn

Received 2011 March 18, accepted 2011 April 6

Abstract: A novel profile detection method is proposed for astronomical fiber spectrum data with low signal-to-noise ratio. This approach can be applied to the pretreatment for 2-D astronomical spectrum data before the extraction of spectra. The Wigner bispectrum, a classical higher-order spectrum analysis method, is introduced and applied to deal with the spectrum signal in this article. After analyzing the Wigner higher-order spectra distribution of the target profile signal, the combination of the Wigner bispectrum algorithm and the fast Fourier transform algorithm is used to weaken the effect of the noise to obtain more accurate information. Both the reconstruction method of the Wigner bispectrum and inverse fast Fourier transform are used to acquire the detection signal. At the end of this paper, experiments with both simulated and observed data based on the Large Sky Area Multi-Object Fiber Spectroscopy Telescope project are presented to demonstrate the effectiveness of the proposed method.

Keywords: line: profiles — methods: data analysis — techniques: spectroscopic

1 Introduction

It is well known that 2-D astronomical fiber spectrum images are a kind of special signal, which is the optical signal through the telescope and other instruments projected onto the charge-coupled device (CCD). It is often affected by many factors, including instrument aberration, atmospheric turbulence, and detector noise, during the observation process. Hence, it makes the astronomical spectrum data have the following characteristics.

There are the spatial orientation and wavelength orientation, which are orthogonal in the fiber spectrum image. In the spatial orientation, the spectrum energy of each fiber diffuses to its adjacent pixels as a form of point-spread function (PSF) to form the respective profiles. These profiles have similar distributions, which are usually considered to be Gaussian distributions (Rhoads 2000). In the wavelength orientation, smoothing processing is usually not allowed because it will destroy the integrity of the spectrum information. Therefore, we usually process the spectrum image in its spatial orientation.

The noise which can easily pollute the observed astronomical spectrum images is mainly considered as having a Poisson distribution in statistics (Pych 2004), which is related to the intensity of the signal. If the signal has a low signal-to-noise ratio (SNR), the spectrum profiles will be polluted so heavily that the centers of profiles deviate from the ideal locations, and the shapes of contours are not smooth and regular anymore, especially when the observed targets are dark celestial bodies. This

situation makes extracting the profile information by using traditional filtering algorithms difficult. Moreover, these adverse factors will hinder the extraction of spectra, which is an important step in the processing of 2-D astronomical fiber spectrum images.

There are two main kinds of spectra extraction methods. A direct and rapid method is the aperture extraction method (de Boer & Snijders 1981), which just accumulates the flux within a certain aperture around the center of the fiber profile. Another modified algorithm was proposed later in which each pixel has a different weight during the flux extraction (Horne 1986; Robertson 1986; Marsh 1989). The benefit of this kind of method is that it does not need profile detection, which makes the processing simple. But the drawback is that the noise is also extracted with the signal in this algorithm, which makes the result of spectrum extraction imprecise, especially when the signal has low SNR. The other kind of approaches are profile-fitting methods (Piskunov & Valenti 2002; Blondin et al. 2005), which fit the fiber profile by using the flux in the spatial orientation. These kind of methods can reach a better accuracy but depend on the quality of fiber profiles and the SNR of the data (Sanchez 2006).

Due to the effects mentioned above, the profile detection becomes necessary to get more accurate profile information for low SNR signals. To meet this challenge, a detection method based on the fast Fourier transform (FFT) algorithm (Cui et al. 2008) was proposed, in which the fiber profile signal in the spatial orientation is transformed to the

frequency domain by the FFT algorithm, and then 0.2% of the energy of the high-frequency component is filtered out as noise. Its effectiveness is limited and not all the components of noise are filtered out. Zhu et al. (2010) described an algorithm in which the stationary wavelet transform was used to detect fiber profiles. The method shows good performance in detection.

In this paper, we propose a novel profile detection method based on the Wigner bispectrum analysis (Fonollosa & Nikias 1991) to process the fiber spectrum profile signals in the spatial orientation. We can analyze the distribution of noise in the Wigner bispectrum of the profile signals. A group of reference data (flat signals) with similar PSF and profile centers for each observed target can be applied to design filters with adaptive cut-off frequencies. The reconstruction method of the Wigner bispectrum is used to reconstruct the detection signal. We can acquire more accurate signal by the FFT filtering method. The experiment data are simulated based on the parameters of the Large Sky Area Multi-Object Fiber Spectroscopy Telescope (LAMOST, China) and the observed LAMOST data are also used.

In Section 2, the introduction and analysis of the Wigner bispectrum are provided and the profile detection algorithm is described. Experiments and results are given in Section 3. The summary is presented in the last section of this article.

2 Profile Detection Algorithm for Spectrum Image

2.1 Wigner Distribution and Wigner Bispectrum

The third-order Wigner distribution called the Wigner bispectrum was proposed by Gerr (1988). This mixed time–frequency–frequency representation extends the standard Wigner distribution to higher-order statistics domains.

The Wigner distribution is the most commonly used time–frequency analysis tool. This distribution was proposed in quantum mechanics by Wigner in 1932 and first applied in the signal processing field by Ville in 1948. It is widely used because of its ability to concurrently describe the signal in terms of its intensity in both the time and frequency domains (Chernogor et al. 2006).

The Wigner time–frequency distribution function of a signal $x(t)$ is described as follows

$$W_{2x}(t, \omega) = \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right)x^*\left(t - \frac{\tau}{2}\right)e^{-j\omega\tau} d\tau, \quad (1)$$

where ω is the parameter of frequency domain, $x^*(t)$ is the complex conjugate of the input signal and τ is the lag parameter of the time domain.

Analogous to equation (1), the third-order Wigner distribution, or Wigner bispectrum, of a real-valued signal $x(t)$ is defined as follows

$$\begin{aligned} W_{3x}(t, \omega_1, \omega_2) = & \int \int x(t + a(\tau_1, \tau_2)) \\ & \cdot x(t + b(\tau_1, \tau_2))x(t + c(\tau_1, \tau_2)) \\ & \cdot \exp\{-j(\omega_1 + \omega_2)\tau_1 - j\omega_2\tau_2\}d\tau_1d\tau_2, \end{aligned} \quad (2)$$

where the lag functions are given by

$$\begin{aligned} a(\tau_1, \tau_2) &= -\frac{2}{3}\tau_1 - \frac{1}{3}\tau_2 \\ b(\tau_1, \tau_2) &= \frac{1}{3}\tau_1 - \frac{1}{3}\tau_2 \\ c(\tau_1, \tau_2) &= \frac{1}{3}\tau_1 + \frac{2}{3}\tau_2. \end{aligned} \quad (3)$$

The lag functions (3) should satisfy the constraints as follows

$$\begin{aligned} a(\tau_1, \tau_2) + b(\tau_1, \tau_2) + c(\tau_1, \tau_2) &= 0 \\ b(\tau_1, \tau_2) - a(\tau_1, \tau_2) &= \tau_1 \\ c(\tau_1, \tau_2) - b(\tau_1, \tau_2) &= \tau_2. \end{aligned} \quad (4)$$

The bispectrum can extract more detailed spectral information about a signal, while the Wigner distribution has the ability of obtaining greater time resolution than that provided by conventional spectrograms. The Wigner bispectrum is useful for extracting time-varying phase information or phase coupling between frequency components, as well as for detection and classification of deterministic signals in stochastic noise.

For computer programming, the discrete version of the Wigner bispectrum is also designed. For a discrete time signal $x(m)$, the discrete time and frequency Wigner bispectrum is described as follows

$$\begin{aligned} W_{3x}(m, n_1, n_2) = & \frac{1}{N} \exp\left\{j\frac{2\pi m}{3N}(n_1 + n_2)\right\} \\ & \cdot \sum_{k_1=0}^{N-1} \sum_{k_2=0}^{N-1} x(m - k_1 - k_2)x(k_1)x(k_2) \\ & \cdot \exp\left\{-j\frac{2\pi}{N}(k_1n_1 + k_2n_2)\right\}, \end{aligned} \quad (5)$$

where N is the length of the input signal $x(m)$.

The spectrum profile signals with noise and without noise shown in Figure 1 can be transformed into new 3-D signals via the Wigner bispectrum function. The Wigner bispectrum of the spectrum profile signal has three components, one time domain and two frequency

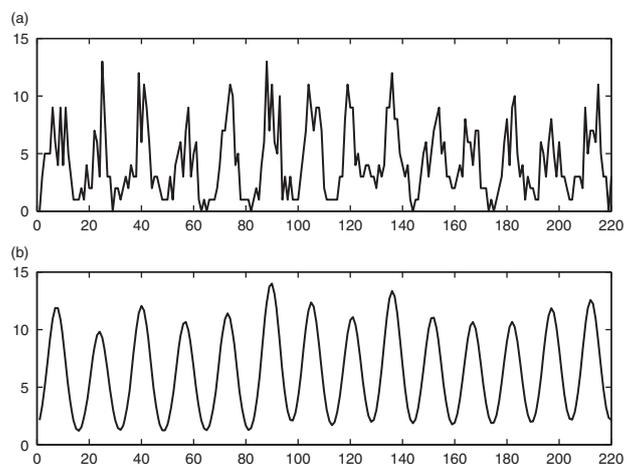


Figure 1 The spectrum profile signals. (a) The profile signal with noise; (b) The profile signal without noise.

domains. Some slices of the Wigner bispectrum (time component $m = 10$ and 50) are illustrated in Figure 2. More details of the Wigner bispectrum are shown in Figure 3.

We can see that the signal is mainly distributed in the corners while most of the noise converges in the middle area. The frequency spectra of signal and noise overlap in a certain frequency range.

2.2 Adaptive Filtering

To separate the noise from the profile signal, we design a filter to weaken the spectrum of noise, especially in the overlap part in two frequency domains as follows:

$$H(n_1, n_2) = \begin{cases} a, & \omega_1 < n_1, n_2 < N - \omega_1 \\ b, & \omega_2 < n_1, n_2 < \omega_1 \text{ or} \\ & N - \omega_1 < n_1, n_2 < N - \omega_2 \\ 1, & \text{other} \end{cases} \quad (6)$$

where N denotes the length of the signal, ω_1, ω_2 are the cut-off frequencies of the filter and $0 < \omega_2 < \omega_1 < (N + 1)/2$. a, b are the coefficients of the filter and $0 \leq a \leq b \leq 1$.

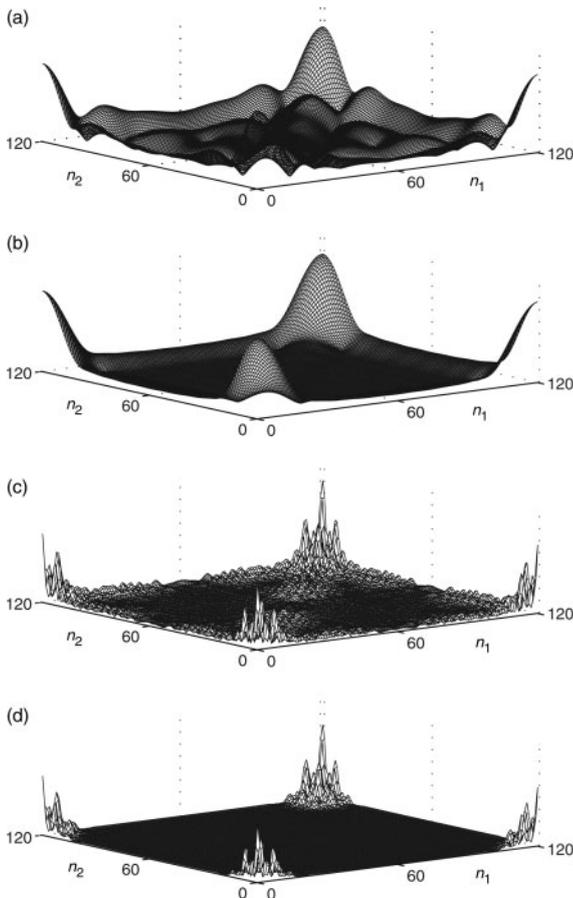


Figure 2 Magnitude of the Wigner bispectrum. (a) Wigner bispectrum of the profile signal with noise ($m = 10$); (b) Wigner bispectrum of the profile signal without noise ($m = 10$); (c) Wigner bispectrum of the profile signal with noise ($m = 50$); (d) Wigner bispectrum of the profile signal without noise ($m = 50$).

The Wigner bispectrum has three components, including one time domain and two frequency domains, and there are a lot of slices of the Wigner bispectrum when the time parameter m takes different values. For each slice, the cut-off frequencies and coefficients of the filter may be different.

The reference signal (flat signal) is applied to determine the optimal cut-off frequencies and coefficients because it has similar PSF and profile centers for the input profile signal. The Pearson correlation coefficient (PCC; Benesty et al. 2008) is adopted here by satisfying the condition: the detection signal after filtering and reconstruction has the best correlation with the reference signal. The PCC between detection signal and the corresponding reference signal is defined by

$$R_{\hat{x}x} = \frac{N \sum_{i=1}^N \hat{x}(i)x(i) - \sum_{i=1}^N \hat{x}(i) \sum_{i=1}^N x(i)}{\sqrt{\text{Var}(\hat{x}) \cdot \text{Var}(x)}}, \quad (7)$$

where

$$\text{Var}(\hat{x}) = N \sum_{i=1}^N \hat{x}^2(i) - \left[\sum_{i=1}^N \hat{x}(i) \right]^2, \quad (8)$$

where N denotes the length of the signal, \hat{x} denotes the detection signal and x denotes the reference signal.

2.3 Signal Reconstruction of Wigner Bispectrum

The reconstruction algorithm of the Wigner bispectrum is used to acquire the detection signal of the fiber profile signal. Considering the Wigner bispectrum distribution function defined by equation (5), the signal reconstruction method is given as follows:

An inverse discrete Fourier transform (IDFT) is applied to deal with both sides of equation (5). Therefore, equation (5) is changed as follows

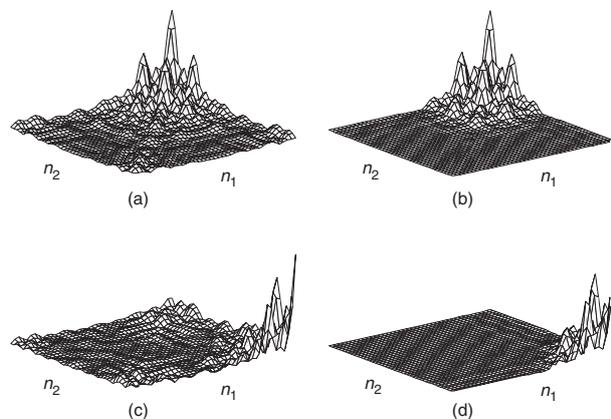


Figure 3 Details of the Wigner bispectrum ($m = 50$). (a) and (c) are parts of the Wigner bispectrum of the profile signal with noise; (b) and (d) are parts of the Wigner bispectrum of the profile signal without noise.

$$\begin{aligned}
 & N \sum_{n_1=1}^N \sum_{n_2=1}^N W_{3x}(m, n_1, n_2) \exp\left\{j\frac{2\pi}{N}(k_1n_1 + k_2n_2)\right\} \\
 & \cdot \exp\left\{-j\frac{2\pi m}{3N}(n_1 + n_2)\right\} \\
 & = x(m - k_1 - k_2)x(k_1)x(k_2)
 \end{aligned} \tag{9}$$

When $k_1 = k_2 = 1$, equation (9) can be simplified by

$$\begin{aligned}
 x(m - 2) &= \frac{N}{x^2(1)} \sum_{n_1=1}^N \sum_{n_2=1}^N W_{3x}(m, n_1, n_2) \\
 & \cdot \exp\left\{j\frac{2\pi(3 - m)}{3N}(n_1 + n_2)\right\}
 \end{aligned} \tag{10}$$

which shows that the detection profile signal $x(m)$ can be reconstructed by equation (10).

2.4 Summary of Our Algorithm

For the observed spectrum image, each 1-D signal in the spatial orientation is processed as follows:

- (i) The original observed signal is extended to a 3-D signal via the Wigner bispectrum function. We can obtain several slices of the Wigner bispectrum by changing the value of the time parameter m .
- (ii) A filter (described in equation (6)) is applied to weaken the noise spectrum. The optimal cut-off frequencies and coefficients are determined by the adaptive method. For different slices, the cut-off frequencies and coefficients of the filters may not be the same.
- (iii) The reconstruction algorithm of the Wigner bispectrum is used to acquire the processed signal.
- (iv) The FFT algorithm is used to transform the processed signal acquired in step (iii) to the frequency domain.
- (v) An ideal low-pass filter with a certain cut-off frequency, such as

$$H(\omega) = \begin{cases} 0, & \omega > \omega_c \\ 1, & \omega \leq \omega_c \end{cases} \tag{11}$$

where ω_c is the cut-off frequency, is used to further remove residual noise from the signal. The adaptive method to estimate the cut-off frequency is applied.

- (vi) The inverse fast Fourier transform (IFFT) algorithm is used to restore the detection signal by dealing with the filtered result.

3 Experiment and Results

The Pearson correlation coefficient has the ability to reflect the similarity and smoothness between two signals. Therefore, it is adopted to evaluate the experimental results in this section. The variance between the detected signal and the original signal without noise is also applied for evaluation, which is defined by

$$V_{\hat{x}x} = \frac{1}{N} \sum_{i=1}^N \left(\text{err}_i - \frac{1}{N} \sum_{i=1}^N \text{err}_i \right)^2, \tag{12}$$

$$\text{err}_i = \hat{x}(i) - x(i), \tag{13}$$

where N denotes the length of the signal, \hat{x} denotes the detection signal and x denotes the reference signal.

Owing to the fact that the original signals without noise are unknown for the observed data, the simulated signals are adopted for the experiments. The simulated profile signal is based on the parameters of LAMOST. An experiment for the observed LAMOST data is also given to demonstrate the effectiveness of our method.

According to the detection method based on the FFT algorithm (Cui et al. 2008), only 0.2% of the energy of the high-frequency component is filtered out as noise. We adopt an improved FFT method as the contrast experiment by combining the adaptive filtering in which more high-frequency components can be filtered out and performance can be improved.

3.1 Experiments Based on Simulated LAMOST Data

In this subsection, the model of the simulation data with a Gaussian PSF is based on the LAMOST project, which is one of the national major scientific projects of China. There are 4000 optical fibers in LAMOST, which are medially distributed to 16 multi-object fiber spectrographs. The full width at half-maximum (FWHM) is approximately 8 pixels. The spacing of two adjacent fiber centers is in a range of 14 to 16 pixels. The noise of the simulated signals is mainly Poisson distributed with a low SNR. Ten groups of profile signals and reference signals are employed randomly with the length of 121 pixels.

For the simulated data, the profile signals without noise are easily obtained for determining the optimal cut-off frequencies ω_1 , ω_2 , ω_c and coefficients a , b . We can get the approximate values of cut-off frequencies and coefficients by analyzing the Wigner bispectrum and Fourier spectrum of the profile signals without noise. For the detection signals, several Pearson correlation coefficients between the processed signals and the signals without noise can be obtained by changing the cut-off frequencies and coefficients around the approximate values. The choice of the threshold is to find the maximum value point of the Pearson correlation coefficient.

In this subsection, the thresholds of adaptive filters are different for different groups of profile signals. ω_1 is in a range of 20 to 24, ω_2 is about 14 to 17 and ω_c is about 20. The coefficients a , b of adaptive filters are about 0 to 0.2 and 0.7 to 0.8.

Two groups of the detection results are illustrated in Figure 4, and the data of contrast experiments are shown in Table 1, Figure 5, and Figure 6. In the table and figures, R_1 , V_1 represent the correlation coefficient and the variance between the observed signal and the original signal without noise, respectively. R_2 , V_2 represent the correlation coefficient and the variance between the

signal detected by our method and the original signal without noise, and R_3 , V_3 denote the results by the FFT method.

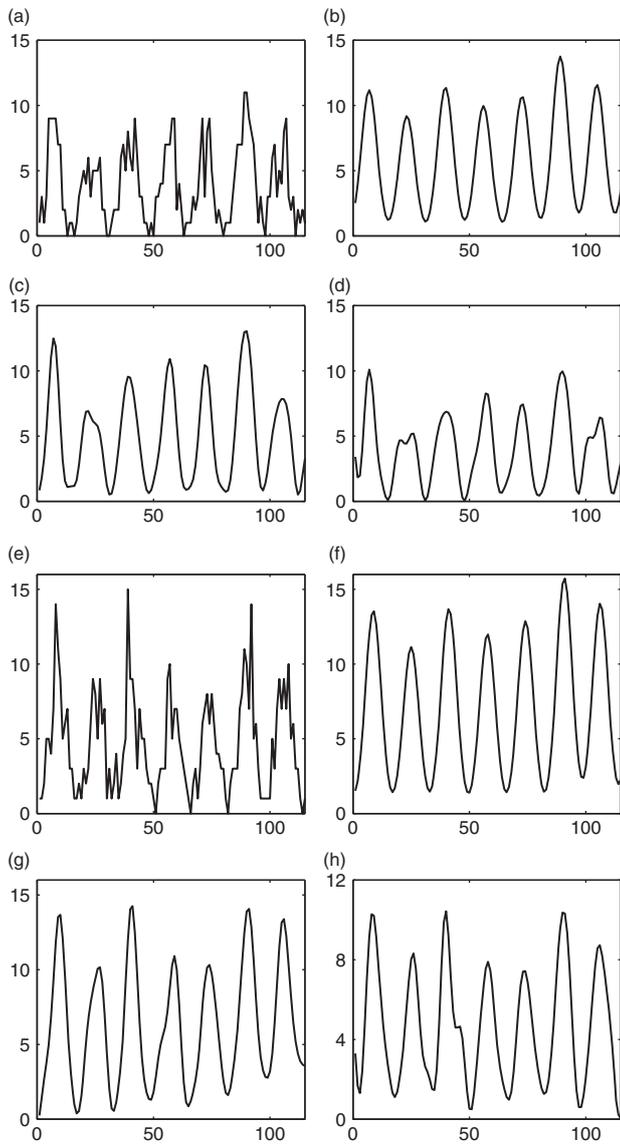


Figure 4 The detection signals based on simulated LAMOST data. (a) and (e) are the spectrum profile signals with noise; (b) and (f) are the spectrum profile signals without noise; (c) and (g) are the detection signals based on our method; (d) and (h) are the detection signals based on the FFT method.

It can be seen from the results that our method can detect the fiber profile signal more accurately than the FFT method. The detection profile based on our method has a better smoothness and similarity with the original signal without noise.

3.2 Experiments Based on Observed LAMOST Data

In order to prove that our method still performs well in engineering practice, the observed data of the LAMOST

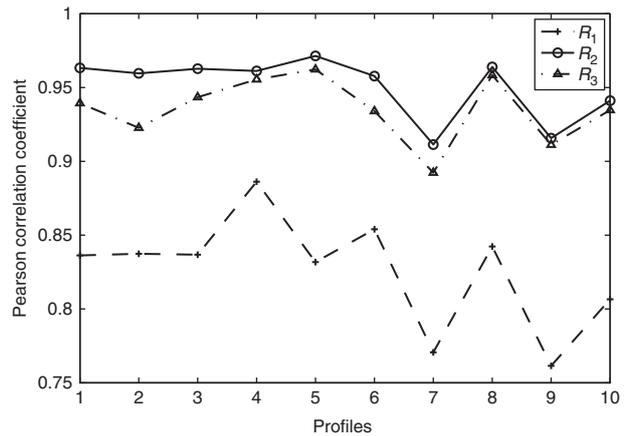


Figure 5 The Pearson correlation coefficients of contrast experiments based on simulated LAMOST data.

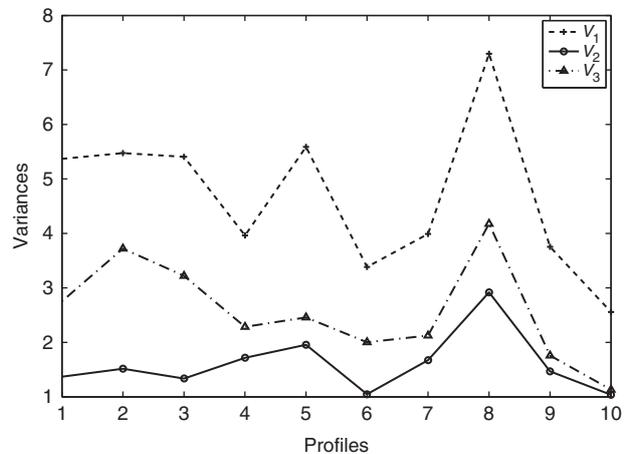


Figure 6 The variances of contrast experiments based on simulated LAMOST data.

Table 1. The Pearson correlation coefficients and the variances of contrast experiments based on simulated LAMOST data

Profile	No. 1	No. 2	No. 3	No. 4	No. 5	No. 6	No. 7	No. 8	No. 9	No. 10
R_1	0.8363	0.8374	0.8368	0.8862	0.8318	0.8540	0.7706	0.8423	0.7615	0.8066
R_2	0.9633	0.9596	0.9627	0.9612	0.9714	0.9578	0.9114	0.9639	0.9157	0.9411
R_3	0.9394	0.9227	0.9434	0.9555	0.9623	0.9340	0.8923	0.9584	0.9112	0.9347
V_1	5.3681	5.4745	5.4054	3.9623	5.5873	3.3844	3.9907	7.2962	3.7540	2.5570
V_2	1.3712	1.5166	1.3367	1.7165	1.9568	1.0493	1.6731	2.9193	1.4686	1.0379
V_3	2.7558	3.7221	3.2216	2.2863	2.4576	2.0031	2.1257	4.1764	1.7588	1.1330

R_1 , V_1 : observed signal; R_2 , V_2 : our method; R_3 , V_3 : FFT method.

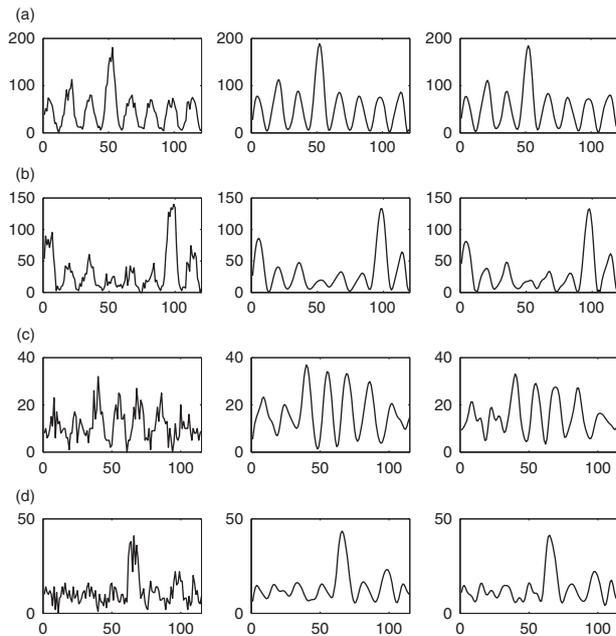


Figure 7 The detection signals based on observed LAMOST data. (a) ro-04r-20100121 (MJD=79514244), Domain (2500,700:820); (b) ro-13b-20100121 (MJD=79514169), Domain (3600,710:830); (c) ro-13b-20100121 (MJD=79514169), Domain (3400,1300:1420) and (d) ro-13b-20100121 (MJD=79514169), Domain (2240,3695:3815). In (a), (b), (c) and (d), from left to right, the three subfigures are the spectrum profile signal, the detection signal based on our method, and the detection signal based on the FFT method, respectively.

project are used. The observed 2-D spectrum images are 4136×4160 pixels in size. The FWHM of the spectra is about 8 pixels.

For the observed data, the original signal without noise is unknown. The flat signals are used to substitute for the original signal without noise. By analyzing the Wigner bispectrum and Fourier spectrum of the flat signals, the approximate values of cut-off frequencies and coefficients can be acquired. The optimal threshold of adaptive filter is determined when the Pearson correlation coefficient gets the largest value. In this subsection, ω_1 is in a range of 18 to 20, ω_2 is about 13 to 15 and ω_c is about 21. The coefficients a , b of adaptive filters are about 0 to 0.15 and 0.5 to 0.7.

The applied data are selected from the LAMOST test database in which most of the observed targets are dark

celestial bodies. The effect of processing the observed data by our method is displayed in Figure 7. For the signal of dark celestial bodies which has low SNR, the detection profile signal based on our method has a better smoothness than the FFT method.

4 Conclusion

An efficient detection method is proposed based on the Wigner bispectrum for the astronomical fiber spectrum images with low SNR in this paper. The profile signals in the spatial orientation are extended to 3-D signals via the Wigner bispectrum function. Adaptive filters are designed by using the corresponding reference signal to filter out parts of noise. Combining the reconstruction method of the Wigner bispectrum and the FFT algorithm, the target profile signals are restored. Experiments based on the LAMOST project are presented to demonstrate the effectiveness of the proposed method. From the experimental results, we can see that the proposed method obtains a better result.

Acknowledgments

The authors acknowledge support from the National Natural Science Foundation of China (NSFC 11078016).

References

- Benesty, J., Chen, J. D. & Huang, Y. T., 2008, *IEEE Trans. Audio Speech Lang. Process.*, 16(4), 757
- Blondin, S., Walsh, J. R., Leibundgut, B. & Sainton, G., 2005, *A&A*, 431, 757
- Chernogor, L. F., Lazorenko, O. V. & Vishnivezky, O. V., 2006, in *Proc. Int. Conf. Ultrawideband Ultrashort Impulse Signals* 3, 297
- Cui, B., Ye, Z. F. & Bai, Z. R., 2008, *AcASn*, 49(3), 327
- de Boer, K. S. & Snijders, M. A. J., 1981, *IUENN*, 14, 154
- Fonollosa, J. R. & Nikias, C. L., 1991, in *ICASSP*, 5, 3085
- Gerr, N. L., 1988, *Proc. IEEE*, 76(3), 290
- Horne, K., 1986, *PASP*, 609, 617
- Marsh, T. R., 1989, *PASP*, 98, 609
- Piskunov, N. E. & Valenti, J. A., 2002, *A&A*, 385, 1095
- Pych, W., 2004, *PASP*, 116, 148
- Rhoads, J. E., 2000, *PASP*, 112, 703
- Robertson, J. G., 1986, *PASP*, 1220, 1231
- Sanchez, S. F., 2006, *AN*, 327, 850
- Ville, J., 1948, *Cables Transm.*, 2A, 61
- Wigner, E. P., 1932, *PhRv*, 40, 749
- Zhu, Z. Q., Zhu, J. & Ye, Z. F., 2010, *Image Signal Process.*, 3, 4118