CSIRO PUBLISHING

Publications of the Astronomical Society of Australia, 2010, **27**, 290–295 www.publish.csiro.au/journals/pasa

Exponential Polynomial Fitting for Fibre Spectrum CCD Profiles

Zhangqin Zhu[A], *Jia Zhu*[A], *Hanqin Qin*[A], *Chong Wang*[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 2009 October 13, accepted 2010 March 3

Abstract: A fibre spectrum profile fitting method based on the least-squares method is presented in this article. For each spectrum of one fibre in spatial orientation, two exponential functions are employed to approximate the profile. Experiments are performed with both simulated profiles and observed profiles to demonstrate the effectiveness of the algorithm. Specially, the proposed method has a better performance for profiles that are asymmetric or composed of multi-Gaussian functions.

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

1 Introduction

Spectrum extraction, a key part of spectrum processing, extracts the flux of a spectrum at each pixel along the spatial orientation. The accuracy of spectrum extraction plays an important role in 1-D spectrum processing. How to measure flux more accurately is an ongoing focus of research in this field.

The aperture extraction method (de Boer & Snijders 1981) is a simple approach that measures the flux within a certain aperture around the centre of the spectrum profile. This method is limited because it is easily affected by noise and cross-talk from flux of an adjacent profile. Some improved aperture extraction methods (Horne 1986; Robertson 1986; Marsh 1989) were developed by adding different weights to different pixels in order to weaken the effect of noise. Some other modified approaches were proposed by using an instrumental profile database (Becker 2001; Roth et al. 2005) to avoid the cross-talk of adjacent fibre energies. These methods require an accurate determination of the profiles (Sanchez 2006).

Another approach is the profile fitting method (Piskunov & Valenti 2002; Blondin et al. 2005), which has better accuracy than the aperture extraction method, but depends on the signal-to-noise ratio (SNR) of the data (Sanchez 2006). This method is widely used to extract spectra for fibre images, such as the Sloan Digital Sky Survey (SDSS). The profile fitting algorithm is the most important part of the method. Some functions (usually a Gaussian function) are used to fit the profile of fibre flux. An improved method named Gaussian suppression combines the aperture method and Gaussian fitting method (Sanchez 2006), but this does not work well if the profiles are non-Gaussian. The radial basic function method (Qin et al. 2009) employed two Gaussian functions to fit profiles for the Large Sky Area Multi-Object Fibre Spectroscopy Telescope (LAMOST). This method works well with mutli-Gaussian profiles but not for asymmetric profiles.

In this paper, we propose another profile fitting function — exponential polynomial. This fitting method uses two different exponential functions to fit each profile, and the parameters of the two functions are adaptively obtained. First, we obtain the approximate centre of each spectrum profile, and the profile data are separated along the centre into two parts. Second, the two parts of the data are used to fit the left exponential function and the right exponential function. The criterion for function fitting is based on the least-squares method. We use the parameters of the two exponential functions, which are obtained by function fitting, to reconstruct the new profile and count the flux of the profile.

The model of the spectrum profile is described in Section 2. In Section 3, the exponential polynomial fitting method is introduced in detail. Experiments and results are given in Section 4. The summary is presented in the last section of this article.

2 Profile Model

The parameters employed in this paper are described in Table 1.

The spectrum profile is usually considered to be a Gaussian function. In spatial orientation, the spectrum profile is mathematically expressed as

$$f(x) = \frac{A}{\sqrt{2\pi}\sigma} e^{-\frac{(x-x_{\mathrm{cen}})^2}{2\sigma^2}}, \tag{1}$$

where x_{cen} denotes the centre of the spectrum profile, σ denotes the Gaussian sigma, and A denotes the flux of the profile.

 10.1071/AS09075 1323-3580/10/03290

The profile fitting method usually used a single Gaussian function to fit the spectrum profile and resulted in an approximation of the actual profile. As the precision required of spectrum extraction is increasing, however, the actual profile of fibre spectrum flux on CCD is getting more attention, and researchers are focussing on more exact fitting of the spectrum profile.

Fibre grating has some width so it is not an idealized line; therefore, each CCD pixel has some area rather than being a point. Hence, the spectrum profiles are more complex than a single Gaussian function. Moreover, the profiles are probably asymmetric. The profile can be described by

$$f(x) = \sum_i \frac{A_i}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x-x_{\text{cen}}-\triangle x_i)^2}{2\sigma_i^2}} + \sum_j \frac{A_j}{\sqrt{2\pi}\sigma_j} e^{-\frac{(x-x_{\text{cen}}+\triangle x_j)^2}{2\sigma_j^2}}. \quad (2)$$

In Figure 1, three different profiles are given.

Table 1. Parameters exposition

Parameters	Exposition
D	The image after initial processing
C	The original image
F	A balance factor
$f(x)$	The profile
x_{cen}	The position of profile centre
$\sigma(\sigma_i, \sigma_j)$	The sigma of Gaussian function
$A(A_i, A_j)$	The flux of Gaussian function
$\triangle x_i, \triangle y_i$	Excursion of centre, positive value
x_i	The ith point of profile
y_i	Flux of the ith point of profile
M	The order of polynomial
N	The number of control points
K	The point number of a profile
a_k	The coefficients of polynomial
$p(x_i)$	The ith fitting value
n_i	The noise of the ith point

3 Exponential Polynomial Fitting Method

The initial image processing includes the measurement errors (Horne 1986)

$$D = \frac{C - B}{F}. \quad (3)$$

F, the balance factor, is allowing direct calculation of Poisson variance estimates, which include the estimating of the readout-noise and the CCD gain. In this paper, all input data are assumed after the initial processing.

3.1 Least-Squares Method

The theory of the least-squares method (Deng 1997) is to minimize the sum of difference squares between the fitted

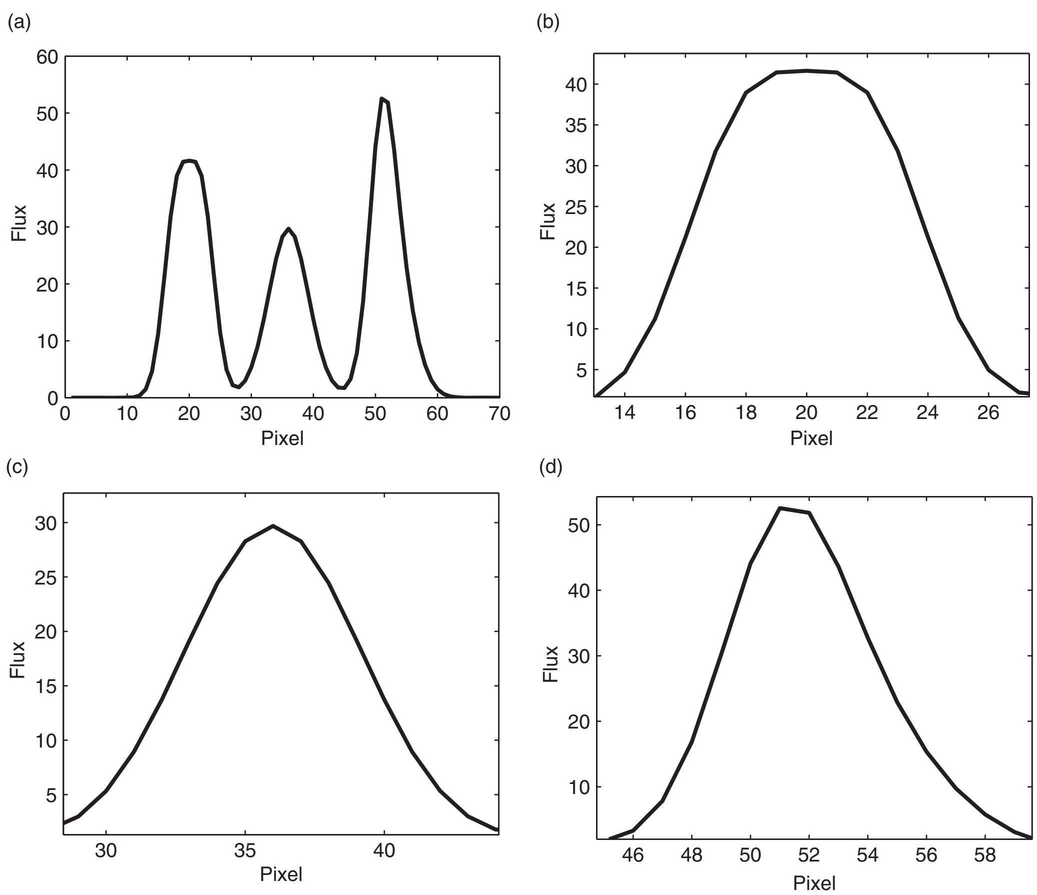

Figure 1 (a) Three types of profiles. (b) A profile with a flat top. (c) A profile with slight cross-talk. (d) An asymmetric profile.

https://doi.org/10.1071/AS09075 Published online by Cambridge University Press

values and original values. The method can be described as follows: for one group of known data (x_i, y_i) where $(i = 1, \ldots, N)$, the data are fitted in the form of some function to get a new group of data $p(x_i)$ where $(i = 1, \ldots, N)$, and the sum of difference squares, err, between y_i and $p(x_i)$ is minimised. That is,

$$\min \text{err} = \sum_{i=1}^{N} [y_i - p(x_i)]^2. \tag{4}$$

Polynomial fitting is a common method for profile fitting based on the least-squares method. In the polynomial fitting method, the function $p(x_i)$ has a certain format:

$$p(x_i) = \sum_{k=0}^{M} a_k \times x_i^k. \tag{5}$$

In the polynomial fitting method, the fitting error can be described by

$$\text{err}_p = \sum_{i=1}^{N} \left(y_i - \sum_{k=0}^{M} a_k \times x_i^k \right)^2. \tag{6}$$

The problem of polynomial fitting based on the least-squares method is solving for the coefficients, a_k, under extremal conditions. Then,

$$\frac{\partial \text{err}_p}{\partial a_k} = 2 \sum_{i=1}^{N} \left(y_i - \sum_{k=0}^{M} a_k x_i^k \right) \times x_i^k = 0 \tag{7}$$

The equations can also be displayed in matrix form:

$$\begin{pmatrix} N & \sum_{i=1}^{N} x_i & \cdots & \sum_{i=1}^{N} x_i^M \\ \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 & \cdots & \sum_{i=1}^{N} x_i^{M+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{N} x_i^M & \sum_{i=1}^{N} x_i^{M+1} & \cdots & \sum_{i=1}^{N} x_i^{2M} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_M \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} (y_i \times x_i) \\ \vdots \\ \sum_{i=1}^{N} (y_i \times x_i^M) \end{pmatrix}. \tag{8}$$

The coefficients are estimated by solving these equations.

3.2 Exponential Polynomial Fitting

Fibre profiles are always non-Gaussian and asymmetric, so a Gaussian function will not provide the optimal fit to the profiles. The profile fitted by an exponential function can avoid the problems of a Gaussian function, which has a fixed format profile. The exponential polynomial function can be described by

$$e^{a_0 + a_1 x_i + a_2 x_i^2} = y_i \quad (i = 1, \ldots, N). \tag{9}$$

To fit the profile using an exponential polynomial function based on the least-squares method, the function must be transformed to a linear function by taking the logarithm:

$$a_0 + a_1 x_i + a_2 x_i^2 = \ln y_i \quad (i = 1, \ldots, N). \tag{10}$$

For common profiles, third-order polynomials provide the optimal approach. If the profile has a very flat top, third-order polynomials may not approach it very well. In this case, fourth-order polynomials can be used. Fourth-order polynomials can be applied to more complex profiles:

$$a_0 + a_1 x_i + a_2 x_i^2 + a_3 x_i^3 = \ln y_i \quad (i = 1, \ldots, N) \tag{11}$$

However, the higher the order of polynomial the easier the approached profile is affected by the control points.

The algorithm for exponential polynomial fitting for fibre spectrum CCD profiles is described in detail below.

1. For each profile, the profile centre can be obtained by the gravity method. The centre of the profile should be accurate to within a pixel. If the half-peak width of the point spread function is narrow, as was the case for the SDSS, the centre must be accurate to a width smaller than a pixel to ensure the precision of the fitting result. If the half-peak width is wider, like LAMOST, the accuracy of the centre can be lessened slightly.
2. For each profile, the profile points are divided into two parts along the centre point, creating a left part and a right part. The number of pixel-points in the two parts depends on the SNR and the cross-talk of the profile.
3. For each part, the exponential polynomial fitting method is used to get the coefficients.
4. For each profile, the two groups of coefficients are used to get two exponential functions. The two functions are then combined to complete the profile fitting.

The number of control points can be selected based on the SNR and the cross-talk of the spectrum profiles. For a second-order polynomial, at least five points are needed (the centre point of each profile, two left and two right points that are adjacent to the centre point), and at least seven points are necessary for a third-order polynomial. More control points can be used to keep the method robust, using more information from the profile, if these points are not affected by cross-talk and heavy noise.

If middle points that should be adopted as control points are negative due to noise or other reasons, the point should not be used as a control point. If many points of one profile are negative after reducing the bias, one constant value can be added to the profile, then after the fitting, the constant value can be subtracted from the extracted flux.

4 Experiments and Discussion

Experiments are performed by applying the exponential polynomial fitting method to different forms of fibre spectrum profiles, and the results are compared to the Gaussian function fitting method (which is used by SDSS; in the experiments, different sigmas of Gaussian functions are given to replace iteration and other processing). The effectiveness is demonstrated by the ratio of the absolute error to the original values.

In this section, dual Gaussian functions are adopted to simulate different kinds of profile, which will be used in following experiments. Through modifying the distance between the two centres and the parameters of the two functions, different species of profile — which are required for discussion in this section — can be obtained. For each kind of profile, Poisson noise and Gaussian noise are added into the profiles with different SNR. The error ratio can be described by

$$\mathrm{err} = \frac{\sum_{i=1}^{K} |y_i - p(x_i)|}{\sum_{i=1}^{K} y_i}. \tag{12}$$

The SNR is defined as

$$\mathrm{SNR} = \frac{\sum_{i=1}^{K} y_i}{\sum_{i=1}^{K} |n_i|}. \tag{13}$$

For the simulated profiles, there are nine control points.

4.1 Multi-Gaussian Function Profile

Due to the width of the fibre grating, the profile might be composed of several Gaussian functions. For this kind of profile, a Gaussian function will not obtain optimal approximation profiles, whereas the exponential polynomial fitting method works well. The simulated profile consists of two equal Gaussian functions, $\sigma = 2$, with a distance between the two centres of 2 pixels. This profile has a flat top approximately like the multi-Gaussian function profile. Figure 2 gives the fitting results of the Gaussian method and exponential polynomial method for this simulated profile.

Figure 2 shows that for profiles composed of multi-Gaussian functions, the single Gaussian fitting is not a good approach, while the exponential polynomial fitting method works well if the data has a suitable SNR. Table 2 gives the errors for both methods for various SNR.

Table 2 shows that SNR plays an important role in the fitting results. The exponential polynomial fitting method is more effective than the Gaussian fitting method only for suitable SNR. Because the parameters of exponential polynomials are very flexible and completely depend on the control points, this method is more substantially affected by noise than the Gaussian fitting method.

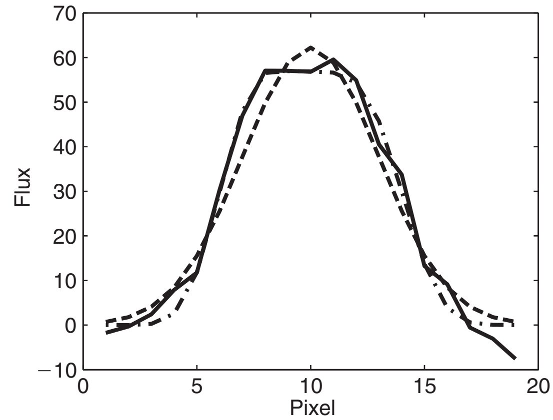

Figure 2 Fitting profiles to a multi-Gaussian profile. The solid line is the profile with noise, the dash-dot line is the fitting profile using the exponential polynomial fitting method, and the dashed line is the fitting profile using a single Gaussian function.

Table 2. Error comparison for a multi-Gaussian function profile

SNR	EPT[a] (err)	Gaussian fitting (err)		
		3.2[b]	3.0[b]	3.4[b]
16.719	0.050	0.103	0.101	0.103
10.538	0.059	0.108	0.101	0.108
8.640	0.069	0.105	0.102	0.109
6.621	0.129	0.110	0.101	0.108

[a]Exponential polynomial fitting.
[b]Sigma of Gaussian function.

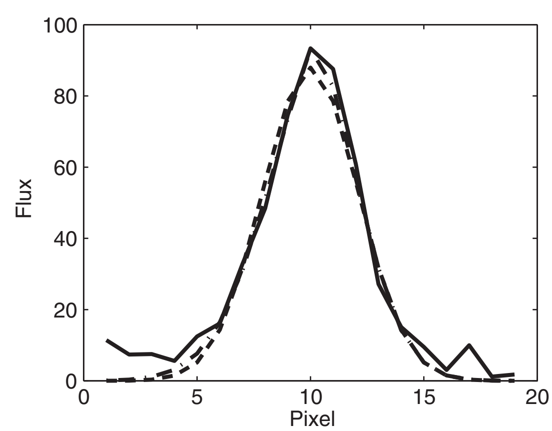

Figure 3 Fitting profiles for an asymmetric profile. The solid line is the profile with noise, the dash-dot line is the fitting profile using the exponential polynomial fitting method, and the dashed line is the fitting profile using a single Gaussian function.

4.2 Asymmetric Profile

Profiles are usually asymmetric due to the effect of spectrograph optics. In this section, experiments are performed on some asymmetric profiles to demonstrate the effectiveness of the exponential polynomial fitting method.

The simulated profiles are also composed of two Gaussian functions. The distance between the two centres is 1 pixel, and the two Gaussian functions have the same flux but different sigmas: $\sigma = 1.5$ and $\sigma = 2.5$. The experimental results are given in Figure 3 and Table 3.

As seen in Figure 3, for the asymmetric profiles, the single Gaussian function is not a good fit even using different centres and sigmas, whereas the exponential polynomial fitting method works well. Table 3 gives the errors. For asymmetric profiles, the exponential polynomial fitting

https://doi.org/10.1071/AS09075 Published online by Cambridge University Press

Table 3. Error comparison for an asymmetric profile

SNR	EPT[a] (err)	Gaussian fitting (err)		
		2.2[b]	2.1[b]	2.0[b]
16.719	0.094	0.140	0.143	0.158
10.538	0.066	0.143	0.139	0.147
7.982	0.058	0.140	0.134	0.151
5.142	0.062	0.139	0.135	0.153

[a]Exponential polynomial fitting.
[b]Sigma of Gaussian function.

method achieves better results than the Gaussian fitting method. When SNR reduces to a rather small value, the effectiveness evidently decreases. When the SNR in a suitable range, the exponential polynomial fitting results are robust.

4.3 Cross-talk

Experiments are carried out to evaluate the effect of reducing cross-talk. The distance between every two adjacent profile centres is 13 pixels. The middle profile is composed of two Gaussian functions with equal flux, $\sigma = 2.5$, and a distance between the two centres of 0.5 pixels. In fact, the middle profile is very close to a single Gaussian function, and the Gaussian fitting method performs adequately. To check the ability of the algorithm to counter cross-contamination between adjacent fibres, different degrees of cross-talk are added into the middle profile. The ratio of cross-talk flux, which is contributed to the middle profile by the two adjacent fibres, and the middle profile flux is used to denote the degree of cross-talk, d_c, given by

$$d_c = \frac{\text{flux}_{\text{cro}}}{\text{flux}_{\text{mid}}}, \quad (14)$$

where, flux_{cro} denotes the cross-talk flux, and flux_{mid} denotes the middle profile flux.

The experimental results are given in Figure 4 and Table 4.

The experimental results show that the two methods have similar results. The proposed method has a similar tolerance to the Gaussian fitting method for cross-talk.

4.4 Approximate Gaussian Profile

Two Gaussian functions are used to produce an approximate Gaussian profile. The centres of the two functions are very close, so the profiles are very similar to a single Gaussian distribution. The experiments demonstrate that the proposed method also works on this kind of profile and gives acceptable results.

The simulated profile is also composed of two Gaussian functions with the same flux, $\sigma = 3.0$ and a distance between the two centres of 0.5 pixels.

Figure 5 and Table 5 show that for profiles with Gaussian distribution, the Gaussian function fitting method performs better, as expected, and the proposed method also provides adequate results with sufficient SNR.

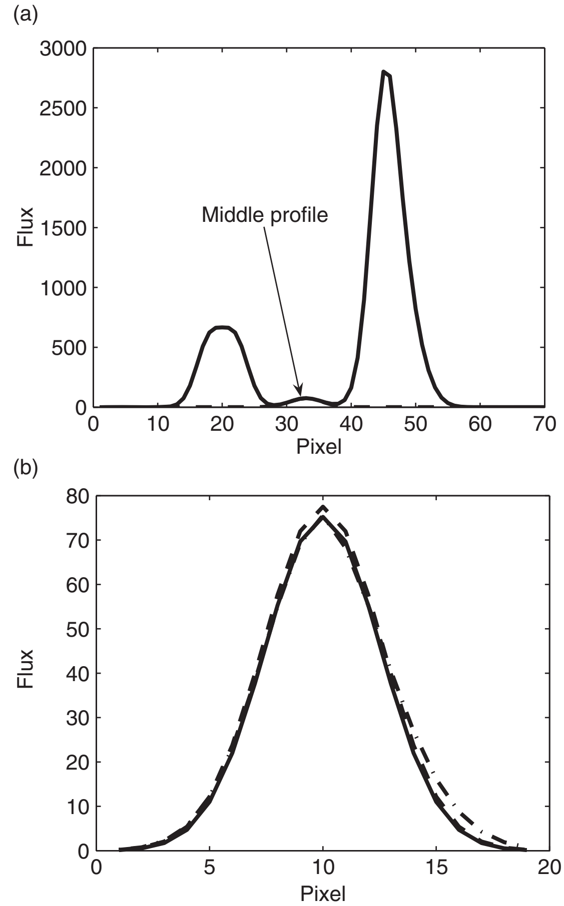

Figure 4 Fitting profiles with cross-talk. (a) The solid line contains three profiles with the middle one contaminated by the cross-talk of adjacent profiles, the dashed line is the original profile without cross-talk. (b) The solid line is the original profile without cross-talk in (a), the dash-dot line is the fitting profile using the exponential polynomial fitting method, and the dashed line is the fitting profile using a single Gaussian function.

Table 4. Error comparison for profiles with cross-talk

d_c	EPT[a] (err)	Gaussian fitting (err)		
		2.5[b]	2.6[b]	2.7[b]
1.130	0.014	0.052	0.017	0.021
3.674	0.047	0.051	0.040	0.046
7.065	0.098	0.086	0.086	0.088
8.761	0.128	0.114	0.114	0.115

[a]Exponential polynomial fitting.
[b]Sigma of Gaussian function.

4.5 Observed Profiles

In this section, the observed profiles from LAMOST and SDSS are used. The SDSS data number is sdR-r2-00031096 and the magnitude range is almost 14–18. Each CCD image is 2128 × 2069 pixels. The LAMOST data is a sky image. For LAMOST, each CCD image is 4096 × 4096 pixels, and the scopes of wavelengths are

https://doi.org/10.1071/AS09075 Published online by Cambridge University Press

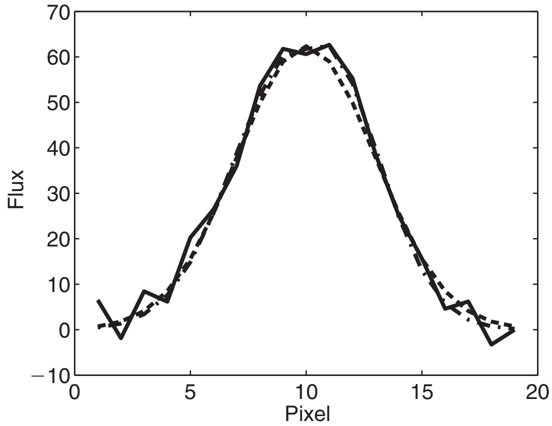

Figure 5 Fitting profiles for an approximate Gaussian profile. The solid line is the profile with noise, the dash-dot line is the fitting profile using the exponential polynomial fitting method, and the dashed line is the fitting profile using a single Gaussian function.

Table 5. Error comparison for an approximate Gaussian profile

SNR	EPT[a] (err)	Gaussian fitting (err)		
		3.2[b]	3.1[b]	3.0[b]
17.359	0.052	0.055	0.031	0.032
13.617	0.051	0.054	0.037	0.037
8.393	0.077	0.101	0.080	0.073
5.824	0.106	0.081	0.057	0.044

[a]Exponential polynomial fitting.
[b]Sigma of Gaussian function.

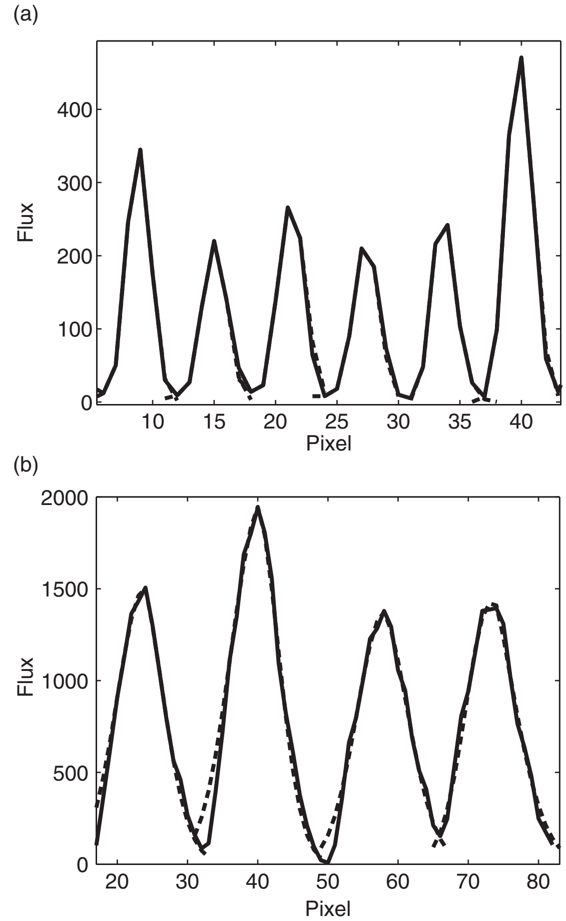

Figure 6 Fitting profiles for observed profiles. (a) SDSS profile. (b) LAMOST profile. The solid line is the observed profile, and the dashed line is the fitting profile by the exponential polynomial fitting method.

3700–5900 Å and 5700–9000 Å. For SDSS data, the number of control points is five, because the half-peak width of the contours is narrow, whereas the half-peak width of LAMOST profiles is relatively wide, and the number of the control points is nine.

We examine the exponential polynomial fitting method with the observed data. The results are compared to the original profiles in Figure 6, which shows that our proposed method is practical. Our method also has the flexibility to fit the profiles for different sigmas — from the SDSS data with a relatively small sigma to the LAMOST data with a relatively large sigma.

5 Conclusion

In this paper, an exponential polynomial fitting method for fibre spectrum CCD profiles based on the least-squares method is proposed. Two exponential functions are used to fit one profile in spatial orientation.

Due to the flexible parameters, the proposed method can fit complex profiles. The proposed method is particularly effective for profiles with flat tops or asymmetric contours. Although the method is more easily affected by noise than the Gaussian fitting method, the proposed method can still achieve good performance for suitable SNR (that is, SNR approximately equal to or greater than 8, based on experiments). Experiments demonstrate the effectiveness of the proposed algorithm on both multi-Gaussian and asymmetric profiles, and it obtains better results than the single Gaussian fitting method provided the data have suitable SNR. For approximate Gaussian profiles, the proposed method also works well and gets applicable results.

Acknowledgments

The authors acknowledge support from the Graduate Innovation Fund of USTC KD2008060.

References

Becker, T., 2001, PhD Thesis, Potsdam University
de Boer, K. S. & Snijders, M. A. J., 1981, IUE NASA Newsletter, 14, 154
Deng, T. B., 1997, ICICS, 1288
Horne, K., 1986, PASP, 609, 617
Qin, H. Q., Ye, Z. F. & Luo A.-L., 2009, PASP, 121, 408
Marsh, T. R., 1989, PASP, 98, 609
Robertson, J. G., 1986, PASP, 1220, 1231
Roth, M. M. et al., 2005, PASP, 117, 620
Sanchez, S. F., 2006, AN, 327, 850

https://doi.org/10.1071/AS09075 Published online by Cambridge University Press