

An Active Flexure Compensation Method for LAMOST spectrograph based on BP-Neural Network

Shuo Huang, Hua Zou, Tian-Jiao Liu, Shi-Yu Zhao and Hang Zhang

College of Science, Hohai University, Nanjing 211100, China; zouhua96@163.com

Received 2019 July 9; accepted 2019 September 2

Abstract The Large sky Area Multi-Object fiber Spectroscopic Telescope (LAMOST) is a Chinese national scientific research facility operated by National Astronomical Observatories, Chinese Academy of Sciences (NAOC). The LAMOST survey for the Milky Way Galaxy and extra-galactic objects has been carried out for several years. The accuracies in measuring radial velocity are expected to be 5 km s^{-1} for the low resolution spectroscopic survey ($R = 1800$), and 1 km s^{-1} for the medium resolution mode. The stability of spectrograph is the main factor affecting the accuracies in measuring radial velocity, so an Active Flexure Compensation Method (AFCM) based on Back Propagation Neural Network (BPNN) is proposed in this paper. It utilizes a deep BP (4-layer, 5-layer etc.) model of thermal-induced flexure to periodically predict and apply flexure corrections by commanding the corresponding tilt and tip motions to the camera. The spectrograph camera system is adjusted so that the positions of these spots match those in a reference image. The simulated calibration of this compensation method analytically illustrates its performance on LAMOST spectrograph.

Key words: instrumentation: spectrographs — methods: data analysis — techniques: imaging spectroscopy — telescopes

1 INTRODUCTION

The Large sky Area Multi-Object fiber Spectroscopic Telescope (LAMOST) is a Chinese national scientific research facility that is operated by National Astronomical Observatories, Chinese Academy of Sciences (NAOC), which is a special meridian reflecting Schmidt telescope with a clear aperture of about 4-meter, a focal length of 20-meter and a field of view of 5 degree (Cui 2006). Its optical system consists of a reflecting Schmidt Mirror (named Ma), a spherical primary mirror (named Mb) and a focal plane. Both the Ma and the Mb are segmented. The Ma consists of 24 hexagonal plane sub-mirrors, and is capable to synchronously change specific aspherical figure accordingly to give the excellent optical quality. The Mb consists of 37 hexagonal spherical sub-mirrors. Each of the Ma and Mb is 1.1 meters in the diagonal dimension. 4000 fibers are mounted on the focal plane. Parallel controllable optical fiber positioning technique is adopted to solve the problem of simultaneously and accurately locating 4000 astronomical objects (Cui et al. 2012).

A total of 16 spectrographs are fed to the LAMOST, and each spectrograph is fed by 250 fibers with 320 micron

in diameter (corresponding 3.3 arcsec). The spectrographs design employs a dual-beamed and bench-mounted, with large-beamed, fast Schmidt cameras and Volume Phase Holographic (VPH) transmission gratings. Each spectrograph has two detectors of 4096 by 4136 pixels, one covering 370 – 590 nm and the other 570 – 900 nm, and the blue-red split is done with a dichroic coating on a beam splitter, the blue side reflected and the red side transmitted. The resolution of spectrograph is 1800 (with $2.2''$ slit), and the resolution can reach up to 7500 in the medium resolution mode.

The LAMOST survey has two major components: the LAMOST ExtraGalactic Survey (LEGAS) and the LAMOST Experiment for Galactic Understanding and Exploration (LEGUE). Details of the LAMOST spectral survey can be found in Zhao et al. (2012). The science plan for the LEGUE portion of the survey can be found in Deng et al. (2012). The LAMOST pilot survey started in October 2011, and ended in June 2012. The data have been released to the public in August 2012, named PDR. The regular survey started in September 2012. The DR6 was released in March 2019, and the DR6 included 11.25 million spectra, which included 9.37 million high-quality

spectra with $\text{SNR} > 10$. Besides, a catalogue which has provided stellar parameters of 6.36 million stars was also released in this data set.

The $R = 1800$ (2/3 slit width) spectra will be similar in quality to the longer exposure SEGUE spectra, with radial velocities and metallicities determined to about 5 km s^{-1} and 0.2 dex, respectively. This requirement results in the specification of image movement of spectrograph less than $0.6 \mu\text{m/h}$ (0.1 pixel/2h). The accuracies in measuring radial velocity and $[\text{Fe}/\text{H}]$ are expected to be 1 km s^{-1} and 0.1 dex respectively for the LAMOST medium resolution ($R = 7500$) spectroscopic survey.

The stability of spectrograph is the main factor affecting the accuracies in measuring radial velocity. The larger instruments are more possible to suffer from gravitational flexure and thermal-induced flexure, and the flexure result in image motion at the detector which reduces spectral and spatial resolution, degrades line profiles and radial velocities. This image motion requires observers to regularly obtain calibration arc spectra at representative telescope positions during observing, thereby wasting significant amounts of telescope time. Flexure might be reduced to acceptable levels using a careful-tuned passive mechanical design, and some reasonable constraints on mass, space, and cost may preclude such a solution, and some form of active flexure compensation may prove more practical.

Many 4 – 10 meters telescopes instruments adopt various AFCMs to meet the high stability prerequisite in the past 15 years. The DEIMOS (Kibrick et al. 2004) on Keck II adjusts focal plane to compensate for CCD image (less than 0.2 pixels in spatial direction and 0.7 pixels in spectral direction). The MOSFIRE (McLean et al. 2012) on Keck I is equipped with piezo transducers to provide tip/tilt control for flexure compensation (at the 0.1 pixel level). The ESI (Kibrick et al. 2003) on Keck II adjusts collimator mirror via three driven actuators (about 0.25 pixels).

For LAMOST spectrographs, to dent thermal-induced flexure and meet the technical requirement of less than 1 km s^{-1} radial velocity measurement accuracy, the LAMOST medium resolution spectrograph (MRS) is calibrated more frequently. It is less conducive to carry on observations and survey tasks continuously.

This paper proposes an AFCM based on BPNN for LAMOST spectrograph. The stability performance of LAMOST-MRS and low-resolution spectrograph (LRS) is presented. The auto-focusing camera system (ACS), the compensation driving device, is also described in detail. We explain the core principle of this compensation algorithm and finally simulate the compensation work on LAMOST spectrograph to test the compensation method.

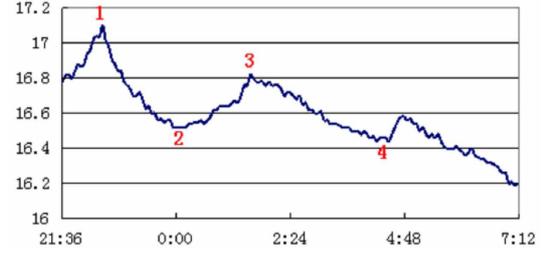


Fig. 1 The test temperature variation curve of spectrograph room, X-axis is time and Y-axis is temperature.

2 ANALYSIS OF LAMOST SPECTROGRAPH STABILITY PERFORMANCE

The 16 low and medium resolution spectrographs (MLRS) of LAMOST are mounted in the spectrograph room. During observation time, it is unnecessary for MLRS to rotate when the telescope tracking objects and environmental temperature in the spectrograph room is the main factor affecting the stability of the spectrograph. The changes of environmental temperature cause linear or non-linear mechanical flexure, resulting in image shift of the CCD. These thermal-induced behaviors degrade measurement accuracy of radial velocity and other astronomical parameters (Hou et al. 2012).

Figure 1 is the temperature variation curve detected via a thermal probes in the spectrograph room. Figure 2 shows the image shift of the spectrograph blue channel CCD and red channel CCD which reveal that image shift is highly correlated with environmental temperature in the spectrograph room. The spectral image changes farther along the spatial direction than along the dispersion direction. Besides, the image quality, also affected by environmental temperature during testing, bears 0.1 mm defocusing when the temperature changes more than 5 degrees. Therefore, monitoring the image quality of the spectrograph and compensating for the spectrograph in time are crucial steps.

The temperature variation in the LAMOST spectrograph room is less than 0.5 degree over one observation night and the image is of about 0.3 to 0.5 pixels displacement. This is the direct behavior result from mechanical flexure. This thermal-induced flexure is very small, but with the increasing demand of spectroscopic survey, the current spectrograph stability can no longer satisfy high precision observation requirement.

This requirement of the radial velocities accuracy of LAMOST MRS results in less image movement, thereby demanding for higher spectrograph stability performance. Schmidt optical design is adopted by camera system, and CCD and its refrigeration system are mounted on the focal plane of the camera. However, the CCD controller and CCD power are two main heat sources when operating. They are installed near the focal plane of the camera sys-

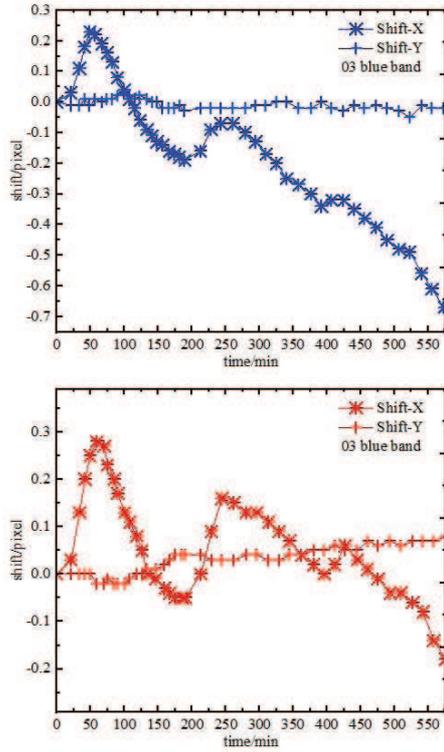


Fig. 2 The image shift of the spectrograph blue channel CCD (*top panel*) and red channel CCD (*bottom panel*) of LAMOST No.03 spectrograph, + represents the dispersion direction, * represents the spatial direction. Horizontal coordinates are time and vertical coordinates are pixel.

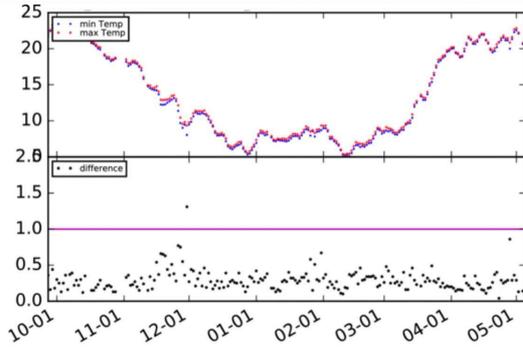


Fig. 3 The daily changes of the environmental temperature in the spectrograph room. The above panel shows the maximum temperature and minimum temperature every testing day. The bottom panel shows the difference between initial and final temperature in the spectrograph room every observation night.

tem and the temperature there is difficult to control (Hou et al. 2012). To diminish thermal-induced flexure and meet the technical requirement of 1 km s^{-1} measurement accuracy, LAMOST-MRS calibrates more frequently, thereby wasting significant amounts of telescope time during spectroscopic survey.

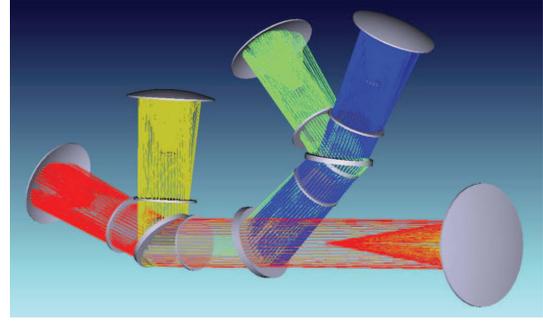


Fig. 4 The LAMOST LMRS optical layout. The *red area* represents the LRS red channel, the *blue area* represents the LRS blue channel, the *green area* represents the MRS blue channel and the *yellow area* represents the MRS red channel.

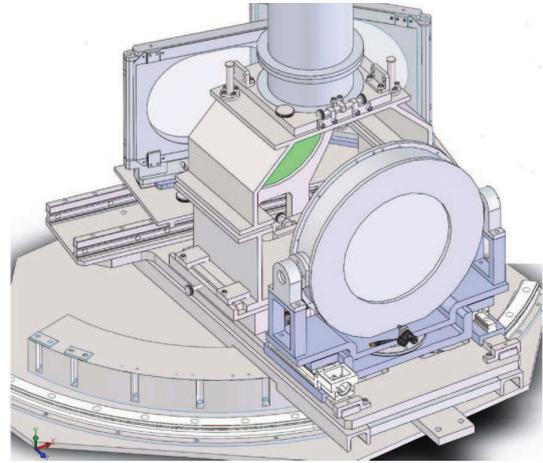


Fig. 5 The diagram of LAMOST ACS structure.

3 THE COMPONENT AND ADJUSTMENTS OF LMRS ACS

The LAMOST spectrographs are of double-beam full Schmidt design by using VPH gratings. Spectrum of light beam from 4000 fibers of 320 micron diameter (arranged along a 144 mm height slit from each spectrograph) is imaged onto 32 pieces of 4096×4096 EEV CCD chips with 12 nm square pixel size. The full width half maximum (FWHM) of monochromatic images is about 7 – 9 pixels. Dispersion of the red band is about 0.084 nm/pixel, and the dispersion of the blue band is 0.057 nm/pixel (Cui et al. 2012). The center distance between two adjacent optical fibers is 0.578 mm in the slit height direction. The slit is an arc with a bending radius of 691 ± 1 mm. The focal ratio of collimator is 4, and the focal length is 800 mm. The collimator employs Schmidt system and its spherical aberration and out-of-axis aberration are corrected by two Schmidt correcting mirrors.

The whole design scheme of the LAMOST ACS is shown in Figure 5. The prototype of ACS was upgraded from the spectrograph camera system. It consists of two

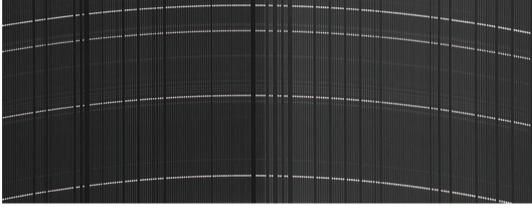


Fig. 6 A part of the LAMOST MRS Th-Ar arc lamp on a 4K×4K CCD. The *white spots* are emission lines, and the light is dispersed along the vertical direction.

symmetrical support structure rails to adjust the camera mirror. One rail provides power and the other is driven. The parallelism of the two rails is better than 3 microns. The AM2224 motor manufactured by Swiss Maxon Motor Company is adopted as the control motor with a torque of 0.1 Nm and a reduction ratio of 989:1. This performance meets the stability requirements of camera system. The ACS of the camera system can provide pitch adjustment, tilt adjustment and focusing adjustment for camera mirror to correct its focus plane.

4 SPECTROSCOPIC DATA PROCESSING AND ANALYSIS

Figure 6 shows an example of the raw arc image obtained in a LAMOST MRS exposure. Each fiber projects a spectrum onto the CCD, which has 4160×4136 array of 12 μm pixels. Abscissa X illustrates the spatial direction, and Y indicates the spectral direction.

4.1 Emission Lines Positioning

For one arc image, there are thousands of emission line spots and each spot should be distinguished, positioned and analyzed. The center coordinates of all spots are calculated by:

$$X_k = \frac{\sum_{i=1}^{13} i \cdot v_{ij}}{\sum_i \sum_j v_{ij}}, \quad (1)$$

$$Y_k = \frac{\sum_{j=1}^{13} j \cdot v_{ij}}{\sum_{i=1}^{13} \sum_{j=1}^{13} v_{ij}}, \quad (2)$$

where v_{ij} is the observation flux of pixel position (i, j) . X_k is the center coordinate of spot k in the spatial direction, and Y_k is in the dispersion direction. They all vary with time because of image shift.

When there are more than one image frames, set the earlier image as reference image. Each spot in image moves less than 1 pixel during one observation night even when spectrograph room temperature is not under control (Hou et al. 2012). Because of tiny displacement of spots, the coordinates of positioned spots can be utilized for spot matching between images and spot shift vectors calculation.

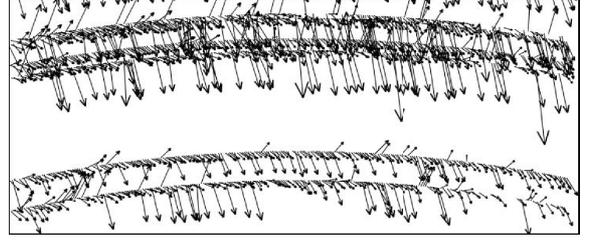


Fig. 7 A part of the uncompensated spot shift vectors (shift within 140 min).

4.2 Calculation, Simulation and Analysis of Image Shift

Shift vector has both direction and magnitude, which can describe displacement precisely. It begins at the original position of the relative spot and ends at the final position after a period of experimental time. In terms of the spot positioning principle, coordinates of all the emission lines in different images at different times are determined, according to which the shift vectors image is simulated in order to analysis the characteristics of data collection from an intuitive interface. The spot shift vectors are determined by:

$$\mathbf{S}_{kmn} = [X_k(t_m) - X_k(t_n) \quad Y_k(t_m) - Y_k(t_n)]. \quad (3)$$

where t_n is starting time, t_m is ending time and \mathbf{S}_{kmn} is shift vector of spot k . In the arc spectrum exposure, when the temperature is not in control, the maximum image displacement is about 0.1 pixels in 2.5 hours much less than 4000 pixels. The measured image shift along the dispersion direction is about 10 times shorter than that along the spatial direction. Therefore, we magnify image shift 200 times along dispersion direction and shift Y 50 times along spatial direction in the simulation for the convenience of analysis, as shown in Figure 7. In terms of that every spot shift vector is of even magnitude and similar direction approximately, every part of one CCD image shift is assumed to be even in this paper. Fiber slit motion and other inhomogeneous motion can be ignored.

We define the average shift of all emission lines in one image as its image shift, which is calculated by:

$$\text{shift}X_{mn} = \frac{\sum_{k=0}^l (X_k(t_m) - X_k(t_n))}{l}, \quad (4)$$

$$\text{shift}Y_{mn} = \frac{\sum_{k=0}^l (Y_k(t_m) - Y_k(t_n))}{l}, \quad (5)$$

where l is the total number of emission lines in one image frame, $\text{shift}X_{mn}$ is the image shift along the spatial direction and $\text{shift}Y_{mn}$ is the image shift along the dispersion direction. During the observation, the temperature of spectrograph room is also recorded every 15 min via

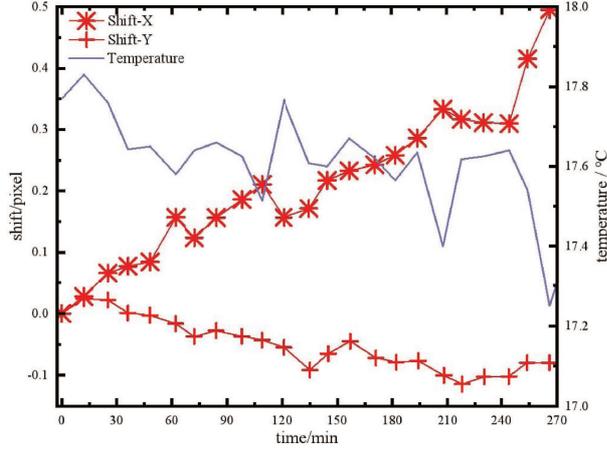


Fig. 8 The calculated image shift along the dispersion direction (+) and along the spatial direction (*) in 270 min.

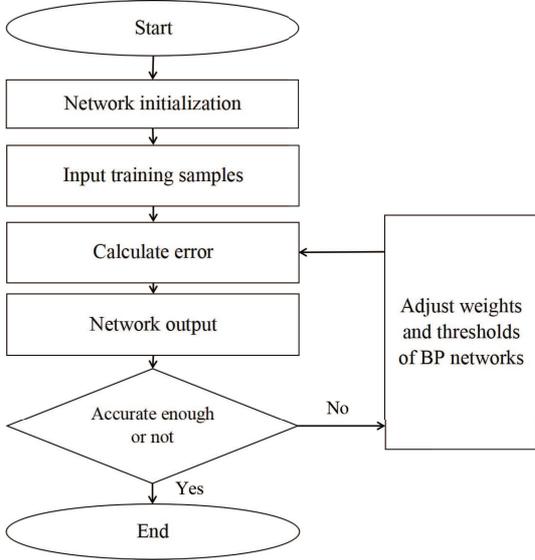


Fig. 9 Flow chart of BPNN training.

thermo-detector mounted nearby. Figure 8 gives the variation curve of image shift and spectrograph room temperature within 270 min where the left axis is the magnitude of image shift (pixel) and the right axis is temperature ($^{\circ}\text{C}$). It reveals that the extreme points of temperature, shift-X and shift-Y occur almost simultaneously. Besides, when temperature increases or decreases, the shift-X and shift-Y changes correspondingly.

The correlation coefficients between temperature and shift-X and between temperature and shift-Y are calculated by:

$$R_x = \text{cov}(ST, S_x) / \sqrt{\text{cov}(ST, ST) \text{cov}(S_x, S_x)}, \quad (6)$$

$$R_y = \text{cov}(ST, S_y) / \sqrt{\text{cov}(ST, ST) \text{cov}(S_y, S_y)}, \quad (7)$$

where $R_x, R_y \in [-1, 1]$ is correlation coefficient, ST is temperature data sets, S_x is shift-X data sets and S_y is

shift-Y data sets. In this way, R_x is calculated as -0.8369 and R_y is calculated as 0.7145 . They prove the strong linear correlation between temperature and image shift, which means the mapping relation modeling is able to be found and is available for image shift predicting and spectrograph compensation.

5 ACTIVE COMPENSATION METHOD PRINCIPLE AND PERFORMANCE

The flexure corrections of CCD image (defocus, pitch and tilt) can be calculated by only two image frames. In term of the flexure corrections, we utilize ACS: tilt adjustment, pitch adjustment and focusing adjustment, to adjust camera lens so that the position of every spot matches those position in a reference image. We have demonstrated the strong linear correlation between temperature and image shift (shift-X and shift-Y). Since that the longer no compensation measure is taken, the worse image quality becomes, shifting time is also incorporated into the model as an important factor in this paper. Therefore, temperature and shifting time have the main influence on image shift.

A neural network can help find the function best maps a set of inputs (flexure-induced image shift, temperature in spectrograph environment and shift time) to their correct output via modifying weights. Besides, LAMOST has accumulated amounts of observational data from history survey programs, offering advantages for machine learning. This paper adopts Deep Neural Network, i.e., 4-layer BP to simulate and predict image shift. We also list the prediction accuracy of 5-layer BP, 6-layer BP, etc. at the end of Section 5.2.

5.1 BP Neural Network Training Process

The simplest BP network is a three-layer network including input layer, hidden layer and output layer. We may add more hidden layers to fit more complex and non-linear functions while difficulties (e.g., overfitting) are more possible to be confronted as well. Fortunately, there are some algorithms (e.g., dropout method, regularization) to overcome these issues.

The output vector of the network is $S_x = (\Delta x_{11}, \Delta x_{12}, \dots, \Delta x_{mn}, \dots, \Delta x_{qq})^T$ and $S_y = (\Delta y_{11}, \Delta y_{12}, \dots, \Delta y_{mn}, \dots, \Delta y_{qq})^T$; the input vector in input layer is $St = (t_1 - t_1, t_1 - t_2, \dots, t_m - t_n, \dots, t_{qq})$ and $ST = (T_1 - T_1, T_1 - T_2, \dots, T_m - T_n, \dots, T_{qq})$; where q is the sample size of image shift data. We adopt two hidden layers and the number of nodes are 3 and 4, i.e., 2–3–4–2 network, i.e., there are two nodes in input layer, three nodes in first hidden layer, four nodes in second hidden layer and two nodes in output layer.

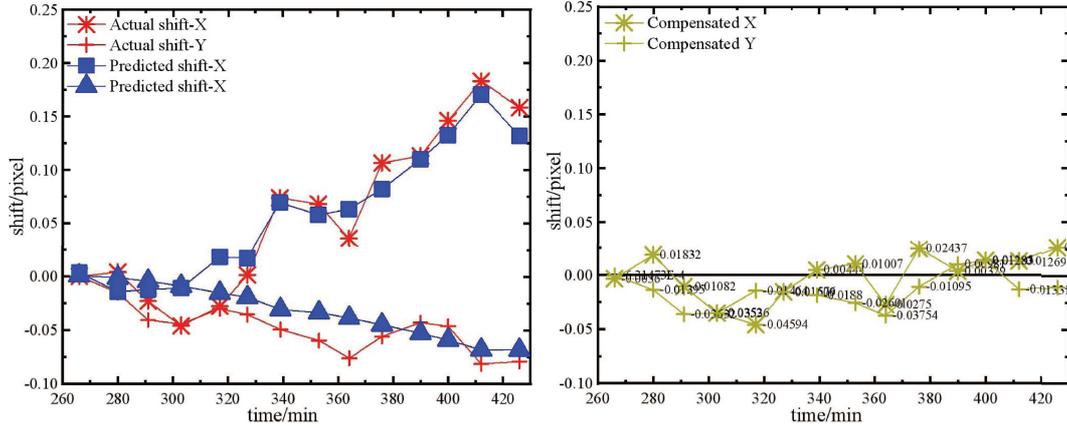


Fig. 10 The average prediction accuracy performance (*left panel*) and compensation residual testing result (*right panel*).

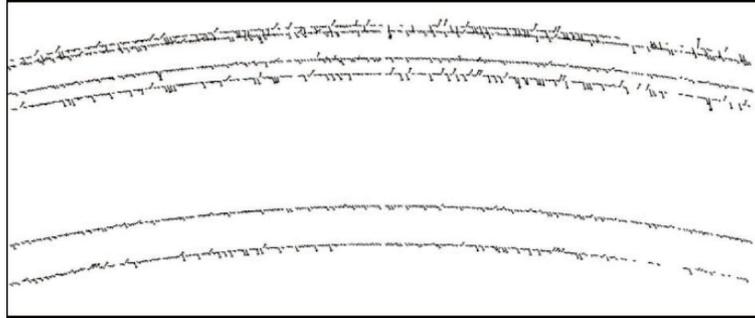


Fig. 11 A part of the compensated spot shift vectors.

The learning process of BP network is shown in Figure 9. Data is transmitted from the input layer to the middle (hidden) layer and then to the output layer to output the result. If the output value and the expectation are not equal, the output error will be transmitted from back to front, i.e., from the output layer to the hidden layer and then to the input layer. The weights in each layer are modified in the process of back propagation. The process of Figure 9 will be repeated until the network output accuracy is better than the set value.

5.2 Experiments and Analysis

To compute the prediction accuracy of this compensation algorithm, the predicted image shift is compared to actual image shift. We need test many times and utilize average prediction results to evaluate the appropriateness of this compensation method. By inputting temperature and shifting time data into trained networks, predicted image shift data are obtained. With 50 different trained networks found and temperature and shifting time data input, 50 different predicted image shift results come out. We then simulate the compensation residual (image shift after compensating) for CCD image correspondingly using the predicted image shift data (the mechanical hysteresis induced by ACS is negligible). The average prediction accuracy

performance and compensation testing result are shown in Figure 10.

The thermal-induced flexure prediction model provides a good fit to the observed data during test. The worst-case residuals are: 0.045 pixels along spatial direction; 0.040 pixels along dispersion direction. The average compensation residuals are 0.02167 pixels in the spatial direction and 0.02129 pixels in the dispersion direction within 140 min, which meet the spectrograph stability requirement of less than 0.1 pixels every 2 hours. The spots shift vectors after compensating are shown in Figure 11. By comparing these to the corresponding measurements obtained with no compensation procedures (see Fig. 7), it is clear that the compensation method has significantly reduced flexure-induced image shift during the test.

It is remarkable that the image shift along spatial direction is about 10 times longer than along the dispersion direction while the accuracy of shift-X compensation is approximately equal to shift-Y. This may be caused by the limit of measurement accuracy of LAMOST and inaccurate spot positioning algorithm.

We also utilized deeper neural network, i.e., more hidden layers to solve our problems. The average compensation residuals were: 0.0225 pixels along spatial direction and 0.0245 pixels along dispersion direction in 2-3-4-3-

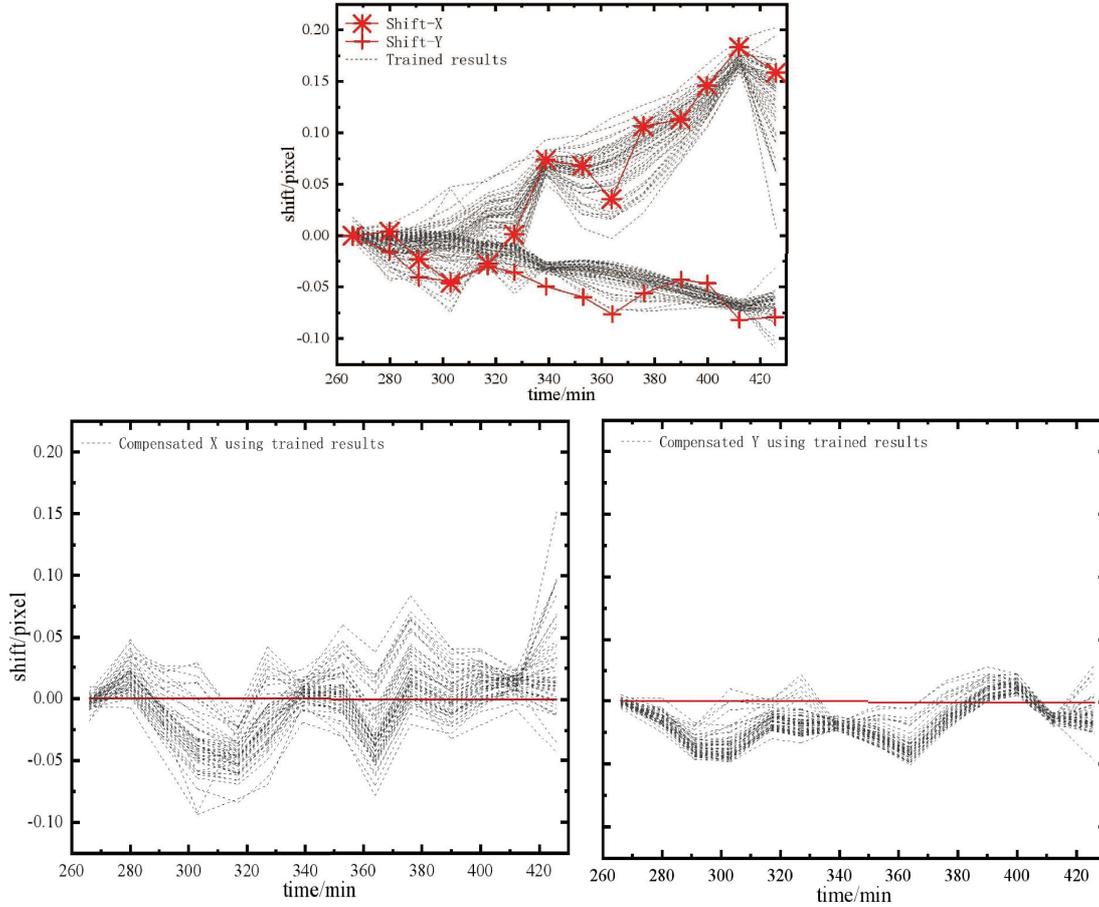


Fig. 12 Different prediction and compensation results. Every line is a different training result. The top panel is trained results, the left one of the bottom panel is compensated image shift along the spatial direction and the right one is compensated image shift along the dispersion direction.

2 BP network; 0.0224 pixels along spatial direction and 0.0251 pixels along dispersion direction in 2-3-4-3-3-2 BP network; 0.0238 pixels along spatial direction and 0.0255 pixels along dispersion direction in 2-3-4-3-3-3-2 BP network. They show similar accuracy to 4-layer BP but will cost more either on processing time or algorithmic complexity. When the relationship between variables is more non-linear and intricate, deeper neural network may work better.

5.3 Stability Performance of the Compensation Algorithm

The 50 different trained and compensation results are shown as the dotted curves in Figure 12. It is easy to tell the dense zone among shift-X curves and shift-Y curves. The distribution of image compensation residual-Y (compensated image shift-Y) is denser than residual-X. It may be due to that the image shift along spatial direction is about 10 times longer than along the dispersion direction.

The worst-case compensation residuals are: less than 0.09 pixels (only one datum is about 0.15 pixels) along

spatial direction and less than 0.05 pixels along dispersion direction. Generally, the proposal active compensation method works stably. Utilizing average results of a large amount of training results can avoid accidental error well. Besides, the prediction about image shift along the dispersion direction is more stable than the prediction along the spatial direction.

6 CONCLUSIONS

Through analyzing the raw arc images obtained with LAMOST spectrograph during survey, the image motion on the detector can be derived. We utilize Deep Neural Network (4-layer BP) to predict the image motion caused by thermal-induced flexure, and the image shift can be compensated based on the prediction. The proposed AFCM has been successfully tested on LAMOST spectrograph. Image shift due to thermal-induced flexure has been reduced by a factor of 4 to a few hundredths of a pixel for a typical exposure.

Acknowledgements The authors gratefully acknowledge the support of the National Natural Science Foundation of China (Grant No. 11503005), Jiangsu Students' Innovation and Entrepreneurship Training Program (201910294155Y and 201810294059X) and the National Undergraduate Training Program for Innovation and Entrepreneurship (201810294099).

References

- Cui, X. 2006, in Proc. SPIE, 6267, 626703
- Cui, X.-Q., Zhao, Y.-H., Chu, Y.-Q., et al. 2012, RAA (Research in Astronomy and Astrophysics), 12, 1197
- Deng, L.-C., Newberg, H. J., Liu, C., et al. 2012, RAA (Research in Astronomy and Astrophysics), 12, 735
- Evans, D.-C. 2008, Astronomy & Geophysics, 49, 4
- Hou, Y., Wang, L., Hu, Z., et al. 2012, in Proc. SPIE, 8446, Ground-based and Airborne Instrumentation for Astronomy IV, 844660
- Kibrick, R. I., Faber, S. M., Phillips, A. C., et al. 2003, in Proc. SPIE, 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes, eds. M. Iye, & A. F. M. Moorwood, 1385
- Kibrick, R. I., Allen, S. L., Clarke, D. A., et al. 2004, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 5492, eds. A. F. M. Moorwood, & M. Iye, 799
- McLean, I. S., Steidel, C. C., Epps, H. W., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 8446, 84460J
- Zhao, G., Zhao, Y.-H., Chu, Y.-Q., Jing, Y.-P., & Deng, L.-C. 2012, RAA (Research in Astronomy and Astrophysics), 12, 723