The Research of Interactive Emission-line Stars Recognition for LAMOST Low Quality Spectra

Lipeng Bi[1,a], Jingchang Pan[1,b*], Cong Liu[1,c] and Guozhou Ge[1.d]

[1]School of Mechanical Electronic and Information Engineering, Shandong University, Weihai, China

[a]bilipeng0418@126.com, [b]pjc@sdu.edu.cn, [c]onionsheep@gmail.com, [d]geguozhou@139.com

*Corresponding author

Keywords: LAMOST; Emission-line stars; Continuum normalization; Spectral lines feature detection

Abstract. Massive sky survey data are produced by the LAMOST large-scale survey project, which provides opportunities to search for some unusual and rare astronomical spectra. The low quality spectra still accounted for about half of total LAMOST observed data at present. For these low quality spectra, research of unusual and rare spectra recognition using computer technology must be combined with artificial check for more efficient recognition. Therefore, this paper focused on the research of interactive emission-line stars recognition. First, we use min-max normalization and continuum normalization for experimental spectra, so that there are only spectral lines features in the spectra. Then, we do feature detection on some particular spectral lines, and recognize emission-line stars through the results of spectral lines feature detection.

Introduction

Since ancient times, astronomical research has never stopped, and has reached its pinnacle in the support of modern science. At present, there are many large aperture with wide field of view survey telescopes in the world like the Sloan Digital Sky Survey (SDSS) Telescope of America[1], the Anglo-Australian Telescope of Australia[2] and the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST) of China[3,4], etc. These large-scale survey projects obtained numerous astronomical data, which provides the basis for the recognition and discovery of unusual and rare astronomical objects such as emission-line stars.

On December 30, 2014, the data set of LAMOST data release two (DR2)[5,6], including spectra of the pilot survey and spectra of the past two years of the regular spectroscopic survey, is published to domestic data users and foreign partners. The DR2 releases massive spectra data, which includes a lot of unusual and rare astronomical spectra. These unusual and rare astronomical spectra include, but are not limited to these types: emission-line stars[7], binary stars[8], stars with unusual continuum, carbon stars [9], white dwarf stars[10], and so on.

For LAMOST low quality spectra, we design an interactive platform to complete the visual identification of emission-line stars. In this paper, we unify dimension for experimental spectra through min-max normalization at first. After that, we remove the continuous spectrum of spectra through continuum normalization to extrude spectral lines features. Finally, based on the characteristics of that most of emission-line stars have hydrogen (H) and helium (He) emission lines, we recognize emission-line stars from LAMOST DR2 data through spectral lines feature detection.

The characteristic of emission-line stars

Most star spectra show the obvious characteristics of atomic absorption lines or molecular absorption band, only some very special stellar spectra exhibit emission-line features. These stellar spectra with emission-line characteristic generally characterize some special types of objects, such

861

as cataclysmic variable (CV), planetary nebulae, etc. The discovery and identification of these emission-line stars is one of the important tasks of large-scale LAMOST digital sky survey.

The emission-line stars has obvious emission-line features. We can notarize whether the spectrum has obvious emission lines through the method of spectral lines detection, to determine whether the spectrum is emission-line stars. The emission-line stars are more representative of cataclysmic variable stars and planetary nebulae, compared with other celestial bodies, which have very strong spectral characterization of accretion process of hydrogen (H) and helium (He) emission lines, so we can identify emission-line stars with hydrogen (H) and helium (He) spectral lines. Thereinto, the most significant and easily identifiable spectral lines are the hydrogen Ballmer lines, He I and He II spectral lines. Table 1 describes center wavelength of these spectral lines and selected bandpass range.

Table 1 Spectral lines information of emission-line stars

Serial numbers	Spectral lines	Center wavelength / Å	Bandpass range / Å
1	H α	6564.61	6562.61-6566.61
2	He I	5876.00	5874.00-5878.00
3	H β	4862.68	4860.68-4864.68
4	He II	4686.00	4684.00-4688.00
5	H γ	4341.68	4339.68-4343.68
6	H δ	4102.89	4100.89-4104.89

Continuum normalization

The observed LAMOST spectra contained spectral lines, continuous spectrum and noise. In order to facilitate the spectral lines detection, we just need to keep spectral lines information in spectra, so the continuous spectrum and noise removal is preparations for spectral lines detection. For LAMOST spectra, the noises like the skylight residuals, the device noise and cosmic rays are difficult to eliminate, which have almost no effect on spectral emission-line feature detection, so the spectral normalization process is the process to eliminate continuous spectrum.

The spectral normalization method we used is divided into two steps: min-max normalization and continuum normalization.

Min-max normalization

In order to unify dimension, to prevent the generation of negative flux, we normalize the flux of experimental spectra, normalized to [0, 1]. The flux of spectra is denoted as $f(n)$, and every flux point is denoted as f_i. Finding the maximum f_{max} and the minimum f_{min} in all spectral flux points, then processing every flux point f_i to make it become $(f_i - f_{min}) / (f_{max} - f_{min})$ after normalization. After that, the maximum of spectral flux values is 1, the minimum of spectral flux values is 0, and the spectral flux is recorded as $f^*(n)$ after min-max normalization.

Continuum normalization

The purpose of continuum normalization is to remove the continuous spectrum. We use the method of least square polynomial fit[11] to extract the continuous spectrum characteristics. The polynomial represents the fitting continuous spectrum, which is denoted as $f_c(n)$. According to this polynomial, the corresponding wavelength of the continuous spectral flux can be calculated, each item of continuous spectral flux is denoted as f_{ci}. Thus, each item of spectral flux after continuous spectrum removal is $f_{ui} = f_i^* / f_{ci}$. Finally the normalized spectral flux is denoted as $f_u(n)$, and $f_u(n) = f^*(n) / f_c(n)$.

We use min-max normalization to unify dimension, and use continuum normalization to remove continuous spectrum. So far, the experimental spectra contain only spectral lines and noise, which is very beneficial to the next spectral lines feature detection showing obvious characteristics of spectral lines. As shown in Fig.1, after the spectral normalization, the spectrum named *spec-56812-VB231N08V1_sp13-025.fits* shows obvious emission-line characteristics, which is judged to apparent emission-line star through observation on the interactive platform. But for some spectra of unobvious emission lines, spectral lines feature detection is an effective manner to verify and recognize emission-line stellar spectra.
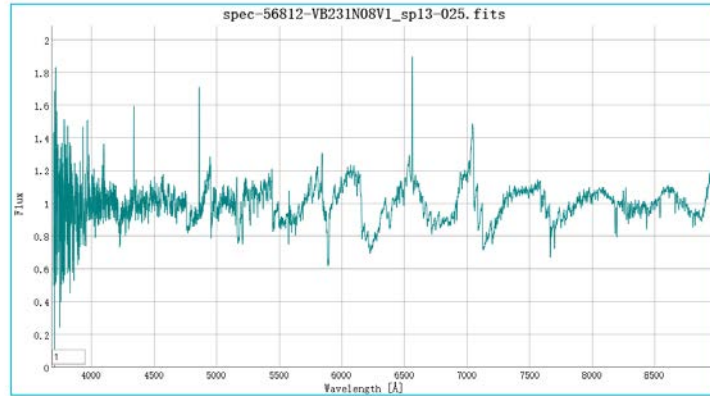

Fig.1. Apparent emission-line stellar spectrum

Spectral lines feature detection

There are six hydrogen and helium special spectral lines in table 1. We are able to determine that the spectral lines have obvious emission-line features, as long as we determine that the six spectral lines flux values in bandpass range increase significantly than flux values near a certain range (the bandpass of blue end and the bandpass of red end).

For a spectrum after spectral normalization, we define the bandpass range of blue end and the bandpass range of red end: if the center wavelength of a spectral line is λ, we set the bandpass range of blue end and the bandpass range of red end are respectively [λ-20, λ-10] and [λ+10, λ+20]. For example, the center wavelength of Hα line is 6564.61(Å), so its bandpass range of blue end is [6544.61, 6554.61] and its bandpass range of red end is [6574.61, 6584.61].

For any of the spectral lines, the normalized flux values set corresponding to the wavelengths in the bandpass range of blue end and the bandpass range of red end is denoted as F. Its mean value is μ and standard deviation is σ. These two values reflect the average situation and the difference of the flux on the bandpass range of blue end and the bandpass range of red end. Because all flow values are not negative after spectral normalization and we want to search for emission lines, we can use $\mu + \omega * \sigma$ to represent the characteristics of spectral flux, where ω is a variable parameter which can be used to adjust our judge selection criteria. The parameter ω is set to 3 in the interactive platform.

For the center bandpass [λ-2, λ+2] of each spectral line, we get all normalized flux values corresponding to the wavelength range. Every item of obtained flux values f_{ui} compares with the standard flux $\mu + \omega * \sigma$, when $f_{ui} > \mu + \omega * \sigma$, its means that the flux value is higher than the standard flux, so emission-line high value counter N can add 1. Only when N>=2, we think this spectral line is emission line.

The result of spectral lines feature detection is that we can get a six-element array, the array element respectively represent high value counting results of Hδ, Hγ, HeII, Hβ, HeI, Hα six spectral lines. When the counting result is greater than or equal to 2, we believe that the spectrum in the vicinity of the spectral line exhibits emission-line characteristics. We can determine that the spectrum is emission-line stellar spectrum as long as one of the six spectral lines shows emission-line characteristics.

In order to better demonstrate the results of the spectral lines detection, we use a chart to display the counting result. Fig.2 is the result of spectral lines detection of the spectrum *spec-56812-VB231N08V1_sp04-222.fits*. We can know that the H δ , H γ , H β , H α are emission lines in this spectrum from the figure.

For unobvious spectral emission lines, interactive recognition can ensure the accuracy of the results of spectral lines detection. For example, Fig.3 is the result of spectral normalization of spectrum *spec-56812-VB231N08V2_sp05-083.fits*. It is difficult for us to directly identify whether this spectrum is the emission-line stellar spectrum, but the result of the spectral lines detection is reliable and evident in Fig.4.
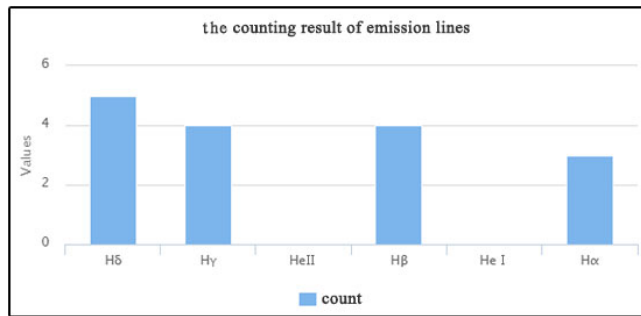

Fig.2. The result of spectral lines detection

In order to better observe, the six spectral lines are marked on normalized spectrum, users can zoom to further observe the result of spectral normalization, as shown in Fig.5. This interactive platform can not only be used to recognition of emission-line stellar spectra by spectral lines detection, but also is more intuitive and convincing for the visual verification of the known emission-line stars.
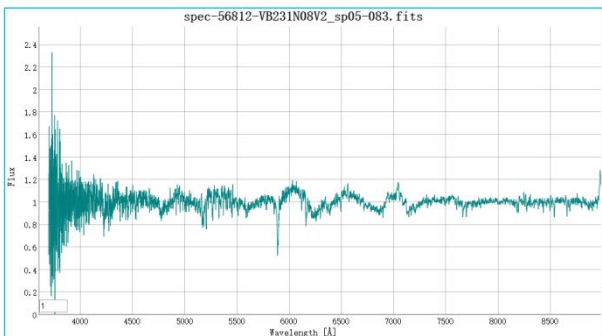

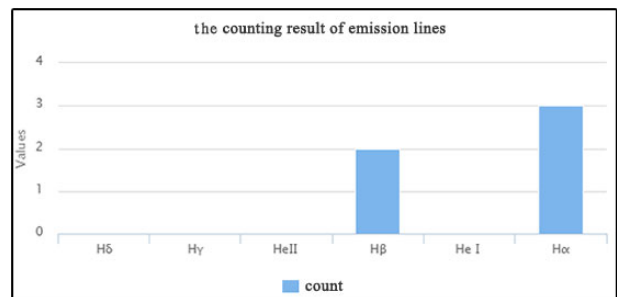Fig.3. Unobvious emission-line stellar spectrum
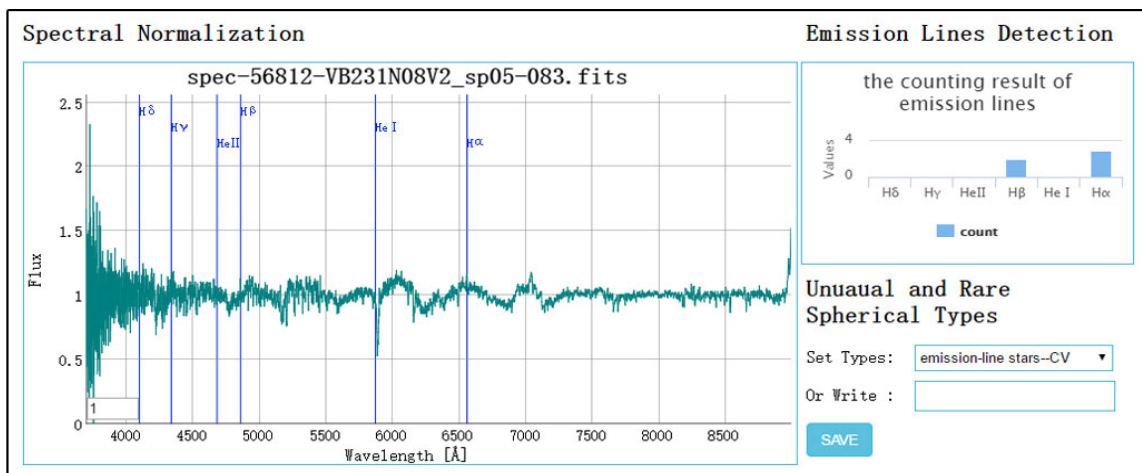

Fig.4. Unobvious emission-line stellar spectrum

Fig.5. The six spectral lines are marked on normalized spectrum

Summary

In view of the LAMOST low quality spectra, this paper design an interactive platform to research the emission-line stars recognition. We normalize experimental spectra through min-max normalization to unify dimension at first. Then we extract the continuous spectrum by polynomial fitting, remove the continuous spectrum by continuum normalization. Then we determine whether the spectral lines are emission lines by spectral lines detection for the six special hydrogen (H) and helium (He) spectral lines. The spectrum is emission-line stellar spectrum as long as one of the six spectral lines is emission line. The method of spectral normalization and spectral lines detection can commendably recognize emission-line stars.

Acknowledgement

References

[1] Donald G. York, et al. The Sloan Digital Sky Survey: Technical Summary[J]. The Astrophysical Journal, 2000,120: 1579-1587.

[2] S. C. Ellis, et al. Suppression of the near-infrared OH night-sky lines with fibre Bragg gratings - first results[J]. Monthly Notices of the Royal Astronomical Society, 2012, 425: 1682-1695.

[3] Xiang-Qun Cui, et al. The Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST)[J]. Research in Astronomy and Astrophysics, 2012,12: 1197-1242.

[4] Ali Luo, et al. Data release of the LAMOST pilot survey[J]. Research in Astronomy and Astrophysics, 2012,12: 1243-1246.

[5] Zhang B, Chen X Y, Liu C, et al. Member candidates of the star clusters from LAMOST DR2 data[J]. arXiv preprint arXiv:1506.04222, 2015.

[6] Zhao J K, Zhao G, Chen Y Q, et al. Halo stream candidates in the LAMOST DR2[J]. Research in Astronomy and Astrophysics, 2015, 15(8): 1378.

[7] Krabbe A, Genzel R, Drapatz S, et al. A cluster of He I emission-line stars in the Galactic center[J]. The Astrophysical Journal, 1991, 382: L19-L22.

[8] Chapman S, Pongracic H, Disney M, et al. The formation of binary and multiple star systems[J]. Nature, 1992, 359: 207-210.

[9] Jianmin Si, Ali Luo, Li Yinbi, and et al. Carbon stars found in LAMOST pilot survey using efficient manifold ranking. MNRAS, submitted.

[10] Zhang Y Y, Deng L C, Liu C, et al. DA white dwarfs observed in the LAMOST pilot survey[J]. The Astronomical Journal, 2013, 146(2): 34.

[11] Steinier J, Termonia Y, Deltour J. Smoothing and differentiation of data by simplified least square procedure[J]. Analytical Chemistry, 1972, 44(11): 1906-1909.