Extraction & Classification of Stellar Spectra from CTIO Plates

By: Lajja Pancholy

Project Supervisors: Dr. Peter Freeman, Thurburn Barker

Introduction

Stellar classification is the process of identifying a star based on its spectral characteristics. Prior to the advent of charge-coupled devices (CCDs), spectra were recorded by dispersing light onto a photographic plate. Stellar spectra contain valuable information about stars, including their temperature, metallicity, velocity, etc. Given these characteristics, we can coarsely classify each spectrum into seven spectral classes: O, B, A, F, G, K and M. Our aim is to automate the process of spectral extraction and identification.

Example of stellar spectrum (class A0I)

Data Pre-Processing & Extraction

Data

- In this project, we will demonstrate our algorithm using data from the Cerro-Tololo Inter-American Observatory (CTIO). Below are data from Plate 6099 of CTIO, as provided by the Pisgah Astronomical Research Institute (PARI).
- The plate data are jpeg images, that are then rendered in grayscale. Pixel values range from 0 (black) to 255 (white)

Image Convolution

- As can be seen, a typical plate contains upwards of 100 potentially overlapping spectra. We want to extract the spectra associated with the brightest stars.
- To determine the location of these spectra, we performed an image convolution on the plate. We defined the kernel to be the size of an average spectrum (21 x 399) and used a box linear filter. This allows us to more clearly see the locations of the spectra associated with brighter stars.

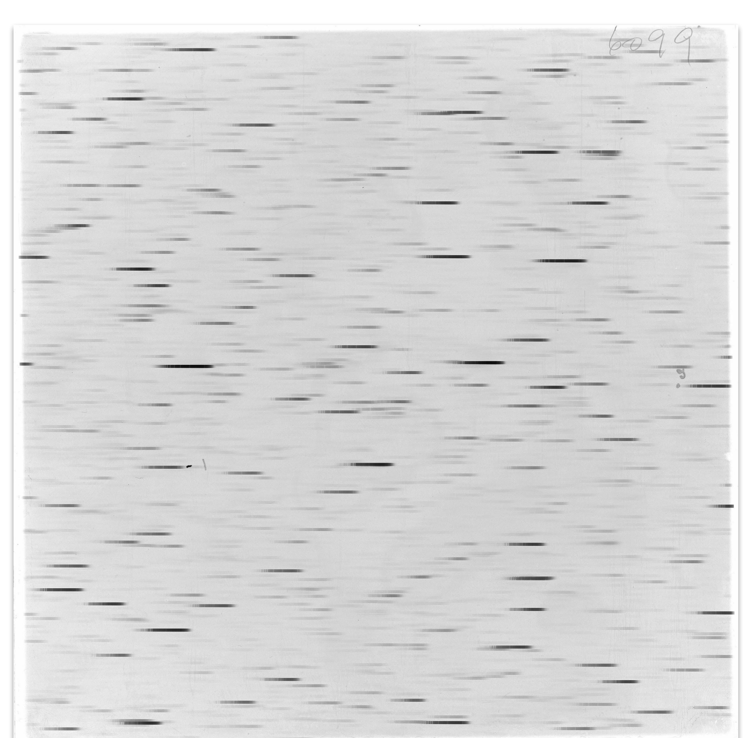
Identifying Locations

- To identify the location of the spectra in the convolved image, we first identify which pixel values fall below a particular threshold.
- We compare the pixel to surrounding pixels to ensure only one spectrum is identified per source (i.e. no overlapping bounding boxes).

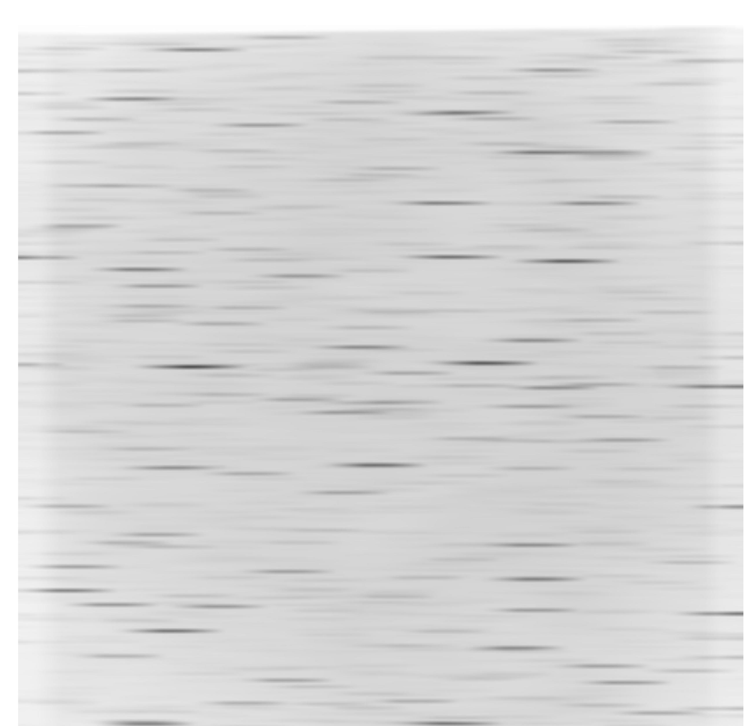
Extraction

- 98 out of the 21,081,830 pixels of the plate were identified as grayscale minima lying below our defined threshold value of 190. We center the bounding boxes (21 x 399) at the minima and extract the spectra from the image.

Extracted stellar spectrum



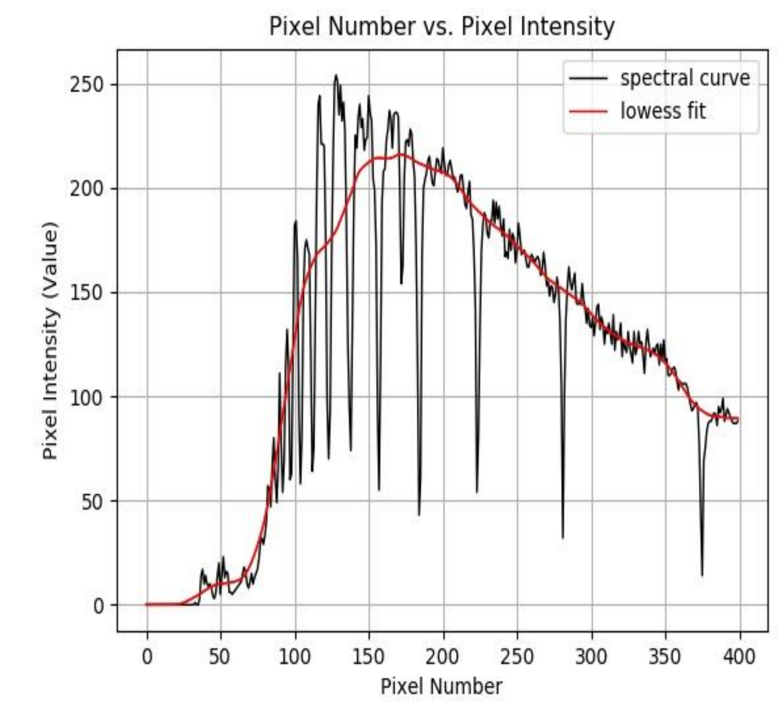
Orgirinal Plate CTIO-6099



CTIO-6099 after image convolution

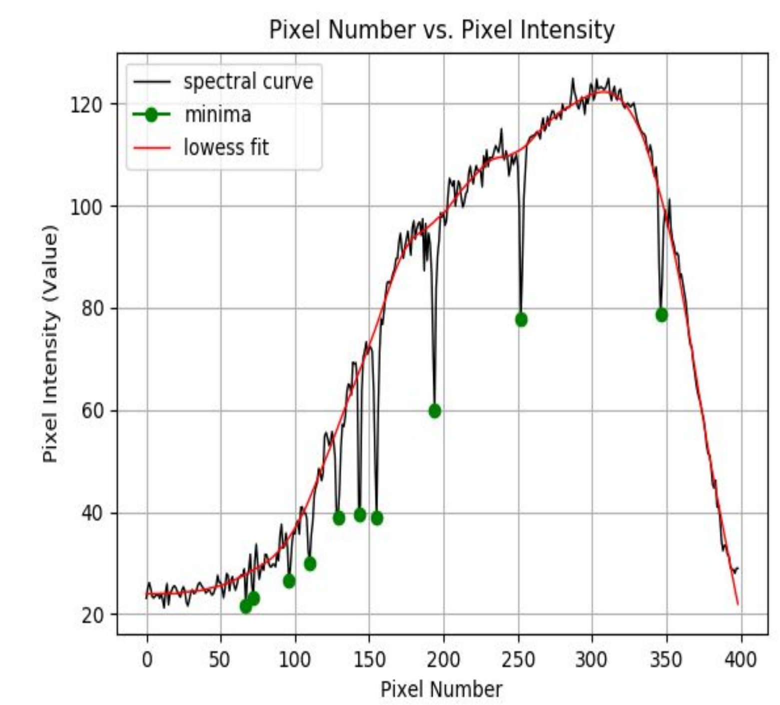
Spectral Identification Methodology

- **Wavelength Calibration**
  - For identification, we will utilize spectra of standard stars, which have known labels.
  - Spectra typically exhibit absorption lines (i.e. dips) associated with, e.g., electron transitions in hydrogen atoms. In properly calibrated spectra, such as the standard stars, the locations of these dips are known.
  - Proper wavelength calibration of an extracted spectrum is necessary to identify it. We can compare the locations of the absorption lines on a standard star to the absorption lines on an extracted spectrum to identify how much an extracted spectrum needs to be shifted to be properly calibrated.



Graph of a standard star spectrum ((class A0I)

- **Identifying Absorption Lines**
  - We can create a plot of pixel value (intensity) vs. pixel number by computing the column-wise mean of the image pixel values. Absorption lines are located at the minima in the plot.
  - To identify the minima, we fit a locally weighted scatterplot smoothing (LOWESS) curve to the data. We find points that fall below the LOWESS curve by a certain threshold. We identify these points as absorption lines. If we infer which line is which, then we can properly calibrate wavelength.



Graph of an extracted spectrum

- **Shifting**
  - We can compare the spacing between observed absorption lines to those between known lines in order to shift the spectra, so as to be properly calibrated.

Standard Star O65V

Calibrated Spectrum

Results & Next Steps

- Now that we have properly calibrated spectra, we can compare them with the standard stars spectra.
- The standard star that most closely matches an extracted spectrum, based on a statistical measure such as mean squared error (mse), will be determined as the identification for the extracted spectrum.
- Once the methodology works for CTIO-6099, we would like to test this on other astronomical photographic plates. Our hope is that this process can be refined such that it works on many different types of plates and can make make stellar spectral data more available to scientists.