ASTRONOMICAL PLATES SPECTRA EXTRACTION OBJECTIVES AND POSSIBLE SOLUTIONS IMPLEMENTED ON DIGITIZED FIRST BYURAKAN SURVEY (DFBS) IMAGES

Aram Knyazyan, Areg Mickaelian, Hrachya Astsatryan

Abstract: The process of spectra extraction into catalogs from astronomical images, its difficulties and usage on the Digitized First Byurakan Survey (DFBS) plates are presented. The DFBS is the largest and the first systematic objective prism survey of the extragalactic sky. The large amount of photometric data is useful for variability studies and revealing new variables in the observed fields. New high proper motion stars can be discovered by a comparison of many observations of different observatories having large separation in years. The difficulty of DFBS images extraction is that extraction tools and programs are not adapted for such kind of plates. Astronomical images extraction process with usage of the Source Extractor (SE) tool is presented in this paper. The specificity of DFBS plates is that objects are presented in low-dispersion spectral form. It does not allow extraction tools to detect the objects exact coordinates and there is need of coordinates' correction. Apart this, it is required to configure SExtractor for current type of the plates so, that the output results be as close to real as possible. The extraction of DFBS plates will allow the creation of astronomical catalogs' database, which can be cross-correlated with known catalogs for investigation of the changes on sky during the years.

Keywords: IVOA, ArVO, DFBS, Plates extraction, SExtractor, VizieR, Astronomical catalogs

ACM Classification Keywords: I.4.1 Imaging geometry, Scanning I.4.3 Geometric correction, H.2.8 Data Mining, Scientific databases.

Introduction

New generation of astronomical instruments produce terabytes of images and catalogs, which fundamentally change the way astronomy. In this context Moore's law is driving astronomy even further. Now Armenian Astronomers are facing different kind of problems, such as modeling and simulation of physical observations, data reduction, sharing and analysis, etc. It has been clear for many years that a national approach, to building such a data infrastructure is insufficient. The datasets and services which astronomers wish to use are spread all around the world. The good solution of data management, analysis, distribution and interoperability are virtual observatories (VO). About seventeen VO projects, including Armenian VO (ArVO), are now funded through national and international programs, and all projects work together under the International Virtual Observatory Alliance (IVOA) to share expertise and develop common standards and infrastructures for astronomical data (images, spectra, catalogs, literature, etc.) exchange and interoperability.

ArVO is a project of the Byurakan Astrophysical Observatory (BAO) (in collaboration with the Institute for Informatics and Automation Problems of NAS RA), which being developed since 2005. ArVO is aimed at construction of a modern system for data archiving, extraction, acquisition, reduction, use and publication. The

ArVO was created to utilize the Digitized First Byurakan Survey (DFBS) as an appropriate spectroscopic database.

The DFBS is the largest and the first systematic objective prism survey of the extragalactic sky. It covers 17,000 deg^2 in the Northern sky together with a high galactic latitudes region in the Southern sky. The FBS has been carried out by B.E. Markarian, V.A. Lipovetski and J.A. Stepanian in 1965-1980 with the Byurakan Observatory 102/132/213 cm (40"/52"/84") Schmidt telescope using 1.5 deg. prism. Each FBS plate contains low-dispersion spectra of some 15,000-20,000 objects; the whole survey consists of about 20,000,000 objects. A number of different types of interesting objects may be distinguished in the DFBS [DFBS; Mickaelian, 2007].

The same fields of the sky are observed by different observatories of the world, in different time periods. The large amount of photometric data is useful for variability studies and revealing new variables in the observed fields. New high proper motion stars can also be discovered by a comparison of many observations of different observatories having large separation in years. Hence, investigation of those observations differences is very useful. This investigation is being done by astronomical plate catalogs cross-correlation between each other. The VizieR Catalogue Service [VizieR] is an astronomical catalog service provided by the Strasbourg astronomical Data Center (CDS) [CDS], which collects and distributes astronomical data catalogs, related to observations of stars and galaxies, other galactic and extragalactic objects, solar system bodies and atomic data. It includes hundreds of catalogues from the all world and provides web access to those catalogs.

The use of high performance computing resources is crucial to fulfill the needs of astronomers that use numerical simulations for their research activity. Recent years many astronomical communities use the Grid infrastrcutuctures, as "cyber-infrastructures" that support Astronomers in any aspect of their research activity, from data discovery and query to computation, from data storage to sharing resources and files. As an example, EuroVO covers interconnection between operational data and service grid, and supports interoperability between EuroVO and European Grid Initiative. In Armenia this problem has been successfully faced in the framework of the International Projects funded by the International Science and Technology Center.

The main purpose of this article is the presentation of DFBS astronomical plates cataloging objectives and its possible solutions based on investigation and testing results.

DFBS astronomical plates extraction into catalogs

Source Extractor (SE) [SE] is a program that builds a catalogue of objects from an astronomical image (photographic plates as well as CCDs). It is used for the automated detection and photometry of sources in files of FITS format. [Bertin, 1996] SE is particularly oriented towards reduction of large scale galaxy-survey data, but it also performs well on moderately crowded star fields. It is chosen for DFBS images extraction to catalogs.

SE builds catalogs in ASCII format and its package works in a series of steps. It determines the background and whether pixels belong to background or objects. Then it splits up the area that is not background into separate objects and determines the properties of each object, writing them to a catalog. All the pixels above a certain threshold are taken to belong to an object. If there is a saddle point in the intensity distribution (there are two peaks in the light distribution distinct enough), the object is split in different entries in the catalogs. Photometry is done on these by dividing up the intensity of the shared pixels. There is an option "clean" the catalog in order to eliminate artifacts caused by bright objects. Afterward, there is a list of objects with a series of parameters measured (ellipticity, size etc.). These are classified into stars (point-like objects) and galaxies (extended objects, everything non-star) by a neural network.

SE gets the positional information from the FITS header but most part of parameters must be specified in its input configuration file, which is an ASCII file (plain text) with the name of the parameters and the value on separate lines. Hence, it is very essential to choose the parameters such, that the output results be as close to real as possible. The threshold parameters indicate the level from which SE should start treating pixels as if they were part of objects, determining parameters from them. The main threshold parameter is counted by this formula:

$$T = k * \sigma \quad (1)$$

where σ is the standard deviation of the image von, k is a coefficient which is given in input configuration. The bigger is k, the fewer objects will be detected.

Before the detection of pixels above the threshold, there is an option of applying a filter. Filter essentially smoothes the image. There are some advantages to applying a filter before detection. It may help detect faint, extended objects. However, it may not be so helpful if the data is very crowded. There are four types of filter (each one has its subtype) – Gaussian (the best for faint object detection), Mexican hat (useful in very crowded star fields, or in the vicinity of a nebula), Tophat (useful to detect extended, low-surface brightness objects, with a very low threshold) and "block" function of various sizes (is a small "block" function for re-binned images), all normalized.

As DFBS plates objects magnitudes are <18 (the objects are faint), the best filters for its extraction are Gaussian filters. It is also obtained by tests results.

De-blending is the part of SE where a decision is made whether or not a group of adjacent pixels above threshold is a single object or not. For example if there is a little island of adjacent pixels above the threshold it is not clear if it is an object or maybe several really close next to each other. This is set by the de-blend threshold parameter. This option is more essential especially for DFBS plates, as the objects are presented in spectral forms, their von is not homogeneous and the same object can be detected twice.

These are the most important parameters of SE configuration file. During the simulations and multiple tests the best configuration of SE for DFBS plates is found. But the main objective of DFBS plates' extraction by SE is detected objects positions precision. This is illustrated on fig. 1.

SE sets detected object position by creating an ellipse around it and then found the center by programming calculations. The same is doing for DFBS spectrums and hence the coordinate is set somewhere on the center of spectrum both by x and y coordinates. Position by x is detected correctly, as spectrum is stretched only by y, by x it is the same as for usual images. As can be seen on fig.2, the real position by y is on the center of drawn circle, but SE finds it much lower. SE has options to find also detected object x and y minimum and maximum coordinates, which is very helpful. y(real) coordinate can be calculated with the formula below:

$$y(real)=y(max)-[x(max)-x(min)]/2 \quad (2)$$

Where x(max) and x(min) are the detected object maximum and minimum coordinates by x, y(max) is maximum coordinate by y.

Thus, detected objects y coordinate will be reset by running a script on SE output ASCII catalog, which will calculate y(real) coordinate. But the coordinates can be calculated only by pixels using this method, and then it has to be converted to astronomical coordinates (RA and DEC) to be able to cross-correliate it with other catalogs.

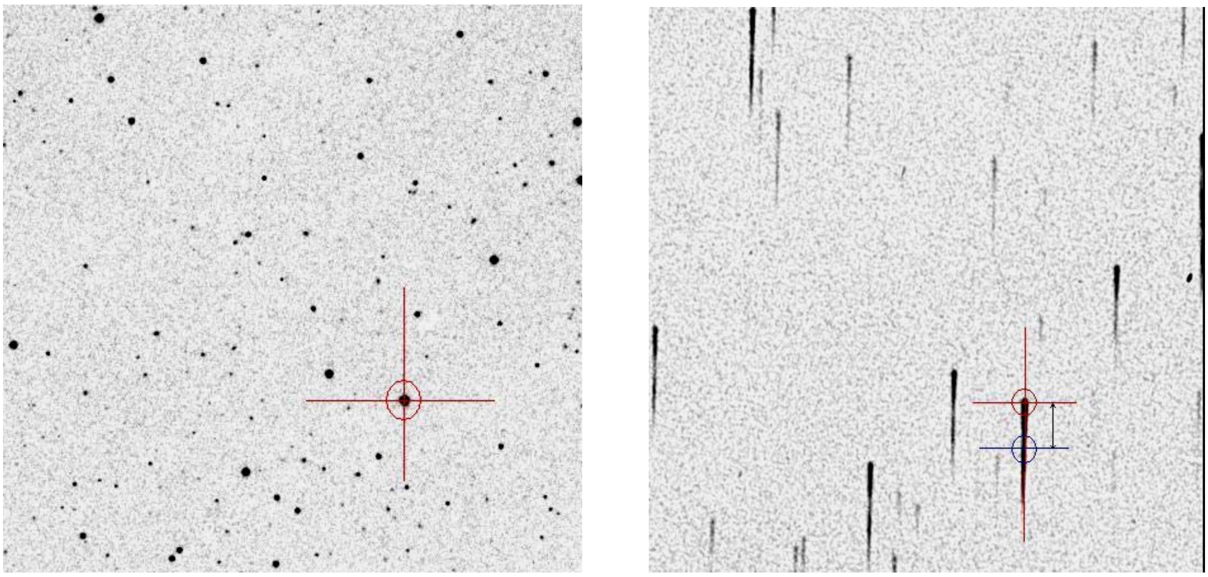

Fig. 1: Illustration of detected object position on the same sky field (left – DSS2 survey image, right - DFBS image)

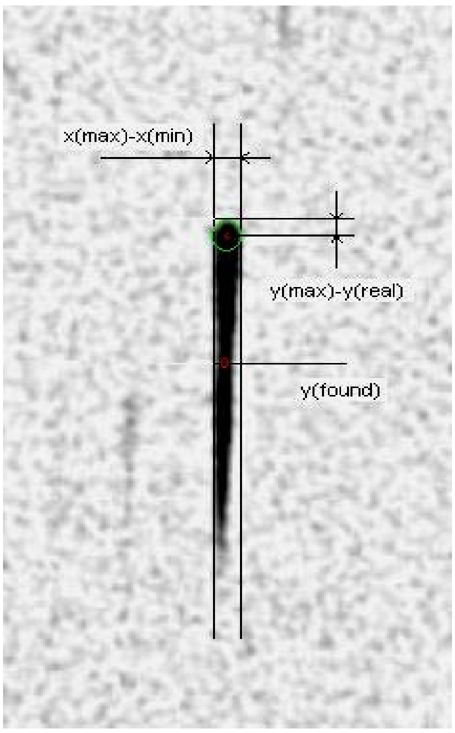

Fig.2: Illustration of DFBS detected spectrum coordinate correction

Conclusion

Astronomical catalogs research is very essential for discovery of new variables and high proper motion stars by a comparison of many observations of different observatories having large separation in years. In the future the aim is the extraction of DFBS all plates to catalogs, its integration in ArVO Portal and sharing with other virtual observatories. The cross-correlation of DFBS catalogs with known calalogs will allow to investigate the changes on sky during the years, as well as develop a classification scheme to estimate the nature of each object.

Acknowledgement

The paper is published with financial support by the project ITHEA XXI of the Institute of Information Theories and Applications FOI ITHEA (www.ithea.org) and the Association of Developers and Users of Intelligent Systems ADUIS Ukraine (www.aduis.com.ua).

Bibliography

[IVOA] International Virtual Observatories Alliance, http://www.ivoa.net

[ArVO] Armenian Virtual Observatory, http://arvo.sci.am

[BAO] Byurakan Astrophysical Observatory, http://www.bao.am

[DFBS] Digitized First Byurakan Survey, http://byurakan.phys.uniroma1.it

[Mickaelian, 2007] Mickaelian, A. M.; Nesci, R.; Rossi, C.; Weedman, D.; Cirimele, G.; Sargsyan, L. A.; Erastova, L. K.; Gigoyan, K. S.; Mikayelyan, G. A.; Massaro, E.; Gaudenzi, S.; Houck, J.; Barry, D.; D'Amante, L.; Germano, P., The Digitized First Byurakan Survey – DFBS, Journal Astronomy and Astrophysics, Vol 464, pp. 1177-1180, 2007

[VizieR] The VizieR Catalogue Service, http:// vizier.u-strasbg.fr

[CDS] Strasbourg astronomical Data Center, http://astro.u-strasbg.fr

[SExtractor] SExtractor software, http://www.astromatic.net/software/sextractor

[Bertin, 1996] Bertin, E., Arnouts, S., SExtractor: Software for source extraction, Astronomy & Astrophysics Supplement 317, 393, 1996

Authors' Information

Aram Knyazyan - Institute for Informatics and Automation Problems, National Academy of Sciences of the Republic of Armenia, 1, P. Sevak str., Yerevan 0014, Armenia; e-mail: aram.knyazyan@ipia.sci.am

Areg Mickaelian - Byurakan Astrophysical Observatory, National Academy of Sciences of the Republic of Armenia, Byurakan 0213, Aragatzotn province, Armenia; e-mail: aregmick@aras.am

Hrachya Astsatryan - Institute for Informatics and Automation Problems, National Academy of Sciences of the Republic of Armenia, 1, P. Sevak str., Yerevan 0014, Armenia; e-mail: hrach@sci.am