

## THE SPECTROSCOPIC DATABASE OF THE DIGITIZED FIRST BYURAKAN SURVEY

A.M. Mickaelian<sup>1</sup>, R. Nesci<sup>2</sup>, G. Cirimele<sup>3</sup>, L.A. Sargsyan<sup>1</sup>, L. D'Amante<sup>3</sup>, P. Germano<sup>3</sup>, E. Massaro<sup>2</sup>, C. Rossi<sup>2</sup>, S. Gaudenzi<sup>2</sup>, D. Weedman<sup>4</sup>, G.A. Mikayelyan<sup>5</sup>, L.K. Erastova<sup>1</sup>, K.S. Gigoyan<sup>1</sup>, L.R. Hovhannisyan<sup>1</sup>, P.K. Sinamyan<sup>1</sup>, J. Houck<sup>4</sup>, and D. Barry<sup>4</sup>

<sup>1</sup>Byurakan Astrophysical Observatory (BAO), Armenia

<sup>2</sup>Università di Roma "La Sapienza", Italy

<sup>3</sup>MIGG s.r.l., Italy

<sup>4</sup>Cornell University, Ithaca, USA

<sup>5</sup>Yerevan State University (YSU), Armenia

### ABSTRACT

The First Byurakan Survey (FBS), also known as Markarian Survey, is the largest low dispersion spectroscopic survey of the sky covering 17,000 square degrees at Galactic latitudes  $|b| > 15$ . The FBS led to the discovery of 1500 UV-excess (Markarian) galaxies and it was also used to search for UV-excess stellar objects (QSOs, WDs, CVs, etc.), late-type stars, and for the identification of IRAS sources. The Digitised First Byurakan Survey (DFBS) provides the astronomical community with a digitised version of the FBS images and the extracted spectra for the objects present in the plates. Nearly 2000 plates have been scanned and stored and programs were developed to compute the astrometric solution, extract the spectra, and apply wavelength and photometric calibration to the objects. The DFBS database and catalogue has been assembled containing data for 20,000,000 objects. A classification scheme for the DFBS spectra is being developed. A work on making the DFBS available through the Virtual Observatory standards and access to spectroscopic data has been started. From the point of view of VO, the DFBS is a new database needing both image and spectra access tools and an interchange between these two standards. Algorithms, tools, and facilities needed for efficient use of the DFBS are discussed, in particular the spectral visualisation and analysis tools. New scientific projects as well as existing surveys will benefit by the digitised images and the ready-to-use extracted spectra which will allow an efficient computer-based analysis of the dataset.

Key words: Astronomical data bases: miscellaneous; Surveys; Virtual Observatory.

### 1. INTRODUCTION. THE FBS AND PREVIOUS RESULTS

The First Byurakan Survey (FBS) was the first systematic objective prism survey of the extragalactic sky. It was conducted by B. Markarian, V. Lipovetski and J. Stepanian in 1965-1980 at the Byurakan Astrophysical Observatory with the 1m Schmidt telescope and 1.5° prism. Until now, the FBS is the largest area spectral survey, covering 17,000 deg<sup>2</sup> of all the northern sky and part of the southern sky at high Galactic latitudes, with a total of about 20,000,000 objects in the entire survey (Markarian et al. , 1989).

The original aim was the search for galaxies with UV excess (Mazzarella & Balzano , 1986), (Markarian et al. , 1989), (Markarian et al. , 1997) - catalogue No. VII/172 at CDS. Studies of the Markarian galaxies early in the survey led to the spectral classification of Seyfert Galaxies (Weedman & Khachikian , 1971), and to the first definition of starburst galaxies (Weedman , 1977).

The huge amount of spectral information contained in the plates allowed the development of several other projects based on the FBS, the most important being the discovery and investigation of blue stellar objects (Abrahamian & Mickaelian , 1996) (CDS catalogue No. II/223); a survey for late-type stars (Gigoyan et al. , 2002) and the optical identifications of sources from the IRAS catalogue (IRAS , 1988). The sample of stellar objects is available at CDS (Byurakan-IRAS Stars - BIS, catalogue No. III/237, (Mickaelian & Gigoyan , 2006), and a similar catalogue for IRAS galaxies (Byurakan-IRAS Galaxies - BIG, (Mickaelian & Sargsyan , 2004)) will soon appear at CDS.

All of these results were obtained by eye inspection of the plates at the Byurakan Observatory using a lens. The number and classes of new objects discovered within the

FBS made clear the need for open access to this information by the entire astronomical community.

## 2. THE DIGITISED FIRST BYURAKAN SURVEY (DFBS)

### 2.1. Scanning

Digitising of the FBS plates started in 2002. After several test scans on a set of plates, all of the FBS plates were digitised with an Epson Expression 1680 Pro scanner at the Byurakan observatory. DFBS was the first digitisation project conducted by scanner and its results allowed a number of other similar projects. The scanning resolution was 1600 dpi so the pixel size is about 16 microns. The typical length of an FBS spectrum is  $\sim 1.7$ mm, which gives 107 pixels along the wavelength direction.

An “ad hoc” program SCANFITS written by Stefano Mottola (DLR - Institute of Planetary Research, Berlin) allows the resulting image to be written directly in FITS format with corresponding information about the plate in the header. All 1874 plates in 1139 fields were digitised and stored on DVDs, as well as they are accessible through our web page at <http://astrot1.phys.uniroma1.it/byurakan/index.php>.

### 2.2. Astrometric solution

The red cutoff of the FBS spectra, defined as the point where the intensity is half of the peak value, is relatively sharp and can be used as a reference point for position (as well as for wavelength calibration; see subsection 2.4), but it is mildly sensitive to the brightness and spectral type (colour) of the object. The accuracy of definition for this cutoff is  $\sim 0.2''$ . The software for the astrometric solution was written by H.J. Hagen by adapting the dedicated software for the Hamburg Quasar Survey (Hagen et al. , 1995). The Second Guide Star Catalogue (GSC2.2 , 2001) was used as the reference positional input. Beginning with the plate centre and the brightest stars in the field, an iterative procedure converges to the astrometric solution. Up to 800 stars of progressively decreasing brightness were used in each field, and up to 7 approximations were carried out to achieve the best solution. The plate scale is  $1.54''/\text{pixel}$ . The best accuracy achieved was  $0.87''$  rms, and the typical accuracy,  $\sim 1''$  rms or 0.6 pixel, is sufficient for confident object identification (a spectrum is typically 5-7 pixels wide). The total time needed for the astrometric solution of each plate was  $\sim 5$  minutes, and the procedure was applied for all 1874 DFBS plates.

### 2.3. Density calibration and spectral extraction

Dedicated software named *bSpec* was developed to perform automatic extraction and classification of the spectral data in a DFBS image. This software was developed by the MIGG s.r.l. team as a collaborative project with the “La Sapienza” University group; it was coded under Linux using the Borland Kylix compiler.

The FBS plates do not have photometric calibration, and it is not possible to build a characteristic curve for each plate. Therefore, we converted the original Data Numbers ( $DN$ ) output of the scanning to relative intensity according to the formula (De Vaucouleurs , 1968):  $I = (V - B)/(T - B)$ , where  $I$  is the intensity (in arbitrary units),  $V$  is the average  $DN$  value for the unexposed plate,  $B$  is the average for the darkest pixels, and  $T$  is the data number for a given pixel. In the immediate future, we will begin developing an absolute calibration using the energy distribution of known stars within the plates.

The extraction of individual spectra from the two dimensional images was performed using several steps. To define the object position and spectral contours, we downloaded a list of objects with  $B \leq 17$  from the USNO catalogue (Monet et al. , 2003). Starting from the catalogue coordinates, each spectrum was recentered using a combination of the two parameters peak position (the pixel of maximum density in the spectrum) and barycentre (the centre of the two-dimensional fit to the spectrum). For each image, a mean spectrum direction angle was computed using the distribution of the entire set of catalogue objects and then adopted as the spectrum direction for all spectra on the plate.

All spectra of the objects from the catalogue were then extracted automatically, corrected for the adjacent sky. The abscissa of each spectrum was set to have pixel 20 at the ‘red head’, where the sensitivity of the plates sharply drops to zero. A disadvantage of this approach is that we lose variable or moving objects present in the plates but fainter than the adopted threshold in the USNO catalogue (or absent in the catalogue).

Therefore, we have tested a second procedure for extraction of the spectra based on the software SExtractor (Bertin & Arnouts , 1996). We found this method to be good for finding all objects, although defects and artifacts are taken as objects and some faint objects are missed. We concluded that this procedure can be used for relatively low-density fields and brighter objects (e. g., for overlapped images, always the brightest spectra is taken). In the future, it may be applied to produce a database somewhat modified from the present one.

### 2.4. Wavelength calibration and broad band photometry

To determine the wavelength calibration, we used some well-exposed spectra of planetary nebulae, white dwarfs,

sub-dwarfs, cataclysmic variables, and QSOs, which have broad Balmer, helium, and other lines. We obtained only an approximate wavelength calibration scale because the dispersion is strongly non-linear, from about 22Å/pixel at the blue edge of the spectrum to about 60Å/pixel at the red edge, with mean dispersion 32.7Å/pixel and dispersion of about 28.5Å/pixel at H $\gamma$ .

Photometric information from the DFBS spectra can be obtained by integrating the spectra blueward and redward of the green sensitivity gap typical of IIF plates, thereby deriving an instrumental “blue” and “red” magnitude; these spectral regions are very similar to the POSS O (4050Å) and E (6450Å) emulsion sensitivity ranges and can be reliably related, therefore, to the B and R magnitudes given in the USNO catalogue. A magnitude calibration for each plate was derived from the stars contained in an area of 4000x4000 pixels in the central part of the plate. Instrumental B and R magnitudes were evaluated by integrating the spectrum between pixels 20-40 (R) and 55-90 (B). A polynomial fit of these instrumental magnitudes compared to the USNO magnitudes provided a calibration curve which was then used to compute the DFBS magnitudes for all objects in the plates. Objects brighter than B=11 are generally overexposed in the DFBS so their magnitudes are not very accurate, nor are the faint objects having magnitudes near the plate limit. A comparison of data from two plates in the same field proved the internal consistency of the calibration process; for both B and R, the typical agreement is  $\sim 0.5^m$  in the range 13<sup>m</sup>-16<sup>m</sup>.

### 3. THE DFBS DATABASE AND THE WEB PAGE

All spectra extracted with *bSpec* were assembled in a database which contains for each object the one-dimensional spectrum plus additional information including an identification label in the coordinates hhmmss+ddmmss, the coordinates RA, DEC at equinox 2000, the position of the red head in plate in pixels, B and R magnitudes from the extracted spectra obtained as described above, the local background value, and a quality flag computed on the basis of the brightness of the object and the possible blending with other stars (before the spectral classification, it will serve as a rough one).

The DFBS database containing the extracted spectra is presently stored on a dedicated PC at the “La Sapienza” University in Rome and can be accessed through the web interface <http://astrot1.phys.uniroma1.it/byurakan/index.php>.

A web page and user interface allow access to the DFBS database for the astronomical community. Beside the Rome University, the DFBS web page is open in Armenia (<http://aras.am/dfbs.html>), and the Rome mirror site will be available their soon. The list of all FBS plates is also available in the Wide-Field Plate Database (WFPDB) at [http://draco.skyarchive.org/search\\_list/](http://draco.skyarchive.org/search_list/), as well

as at CDS (catalogue No. VI/116) at <http://cdsweb.u-strasbg.fr/viz-bin/Cat?VI/116> (Mickaelian et al. , 2005).

The DFBS database (v1.0) can be queried on line at the Rome web page. Several palettes give general information on the FBS and DFBS, the main characteristics of the observational material, and the main results obtained from the plates. In particular, the keys *sky – coverage* and *platelist* contain detailed information on individual plates, such as type of emulsion, observation date, observer, and accurate central coordinates from the plate centre computed after application of the astrometric solution. It is possible to retrieve a list of DFBS plates within given ranges in observing dates, sky area, emulsion types, and observers. The digitisation and reduction procedures are also described in detail, including the scanning, astrometric solution, extraction, density and wavelength calibration.

The user interface has a menu type structure. In the *explore* menu portion, the images are visualised and a spectroscopic DFBS plate can be compared with the corresponding direct plates from the Digital Sky Survey (DSS1 and DSS2). There is an easy access to the database and to portions of the digitised plates. The extracted spectra of individual objects can also be displayed and stored to be downloaded later. However, users cannot run the *bSpec* software on a remote machine to create a personal database, but they may download portions of the plates in FITS format and perform their own analysis of the images.

The user interface (the DFBS portal) presently allows the following operations to the guest user:

a) Sky coverage: in a RA, DEC rectangular interactive sky map shows the position on the sky of each plate; plates already processed are colour-coded. Basic data about each plate are available by clicking on the corresponding plate (<http://astrot1.phys.uniroma1.it/byurakan/index.php?page=cover>).

b) Plate list: the list of the plates is shown; it can be sorted by several parameters (e.g. plate numbers, RA or DEC, emulsion types, observers) and downloaded. Each plate number is also clickable, so that it is possible to check all data concerning any plate (<http://astrot1.phys.uniroma1.it/byurakan/index.php?page=platelist>).

c) Explore: it allows the display of a portion of plate around a given central RA, DEC position, comparison with the same portion of the DSS1 or DSS2 (blue, red, or IR), interactive selection of one or more spectra present in the database, their collection (saving in a list) and downloading (ASCII files) to the guests computer (<http://astrot1.phys.uniroma1.it/byurakan/explore.php>). This option is useful for studies of specific fields.

d) Get image: it allows users to select a portion (presently up to 1024x1024 pixels, i.e. about 26.5'x26.5') of a plate in FITS format and all the spectra of this portion present in the database for downloading (spectra are ASCII files), as well as

downloading of the whole selected field in FITS format (<http://astrot1.phys.uniroma1.it/byurakan/getimage.php>). This option is useful for downloading portions of DFBS.

e) Get spectra: it allows downloading all the spectra in the database in a given field (within a given distance from a selected central position); the query may be either interactive, with the RA, DEC position, or made by uploading an ASCII file containing one or more RA, DEC positions (one per line). Objects may be selected by B, R or B-R values. This option displays also an interactive (“clickable”) table of the selected objects, which allows looking at each object individually (both 1D and 2D spectra) for a quick evaluation of the data (<http://astrot1.phys.uniroma1.it/byurakan/getspectra.php>).

#### 4. FUTURE DEVELOPMENTS

The electronically-accessible DFBS is the primary result of our project. Use of the digitised spectra will increase the efficiency of object selection for various scientific goals. For example, inspection of low dispersion spectra is often the most efficient way for initial identification of optical counterparts of sources discovered in X-ray, IR, and radio surveys.

The next step of our work will consist of the development of criteria for automatic spectral classification. Different methods are currently being tested and will be included in a future version of the database. One approach is based on developing template spectra for different types of objects using available catalogues and averaging the DFBS spectra for each type (e.g., QSOs, white dwarfs, cataclysmic variables, carbon stars, etc.). A second approach is based on a numerical classification scheme which utilises multivariate relations such as relations between magnitudes and widths of the spectra (for separation of stellar and diffuse objects), spectral energy distribution, ratio of the red/blue part of the spectra, length of the spectra, etc., as well as the broad absorption and emission lines (seen in DFBS for QSO/Sey, white dwarfs, many sub-dwarfs). Such classifications derive from criteria developed during the selection of blue stellar objects, red stars, and identification of IRAS sources.

In the Introduction, we briefly mentioned some of the previous surveys based on the FBS plates. The digitised images of the DFBS and the resulting template spectra will allow faster, more quantitative identification of targets. We are confident that these new tools will not only benefit previously initiated projects, but that new research ideas will also originate from the Digitised First Byurakan Survey.

#### ACKNOWLEDGEMENTS

We acknowledge grants from ANSEF in 2002 and 2007, USA NSF in 2003-2004, REGIONE CALABRIA and

U.E. (FESR) in 2005-2006, and CRDF grant ARP1-2849-YE-06 in 2007.

The University “La Sapienza” partially supported the project with special funds for International collaboration.

#### REFERENCES

- Abrahamian, H.V., and Mickaelian, A.M. 1996, Ap 39, 531
- Bertin, E., Arnouts, S. 1996, A&AS 117, 393
- De Vaucouleurs, G. 1968, Applied Optics, 7, 1513
- Gigoyan K.S., Abrahamian H.V., Azzopardi M., and Russeil D. 2002, Ap 45, 322
- GSC2.2 'The Guide Star Catalogue, Version 2.2.01 Space Telescope Science Institute (STScI) and Osservatorio Astronomico di Torino', 2001
- Hagen, H.-J., Groote D., Engels, D., Reimers, D. 1995, A&AS 111, 195
- IRAS 1988, The Point Source Catalogue. Joint IRAS Science Working Group, NASA, Washington, DC
- Markarian, B.E., Lipovetski, V.A., Stepanian, J.A., et al. 1989, Commun. Special Astrophys. Obs., 62, 5
- Markarian, B.E., Lipovetski, V.A., Stepanian, J.A. 1997, First Byurakan Survey (FBS), CDS, Strasbourg;
- Mazzarella, J.M., Balzano, V. 1986, ApJS 62, 751
- Mickaelian, A.M., Sargsyan, L.A. 2004, Ap 47, 251
- Mickaelian, A.M., Hagen, H.-J., Sargsyan, L.A., Mikayelyan, G.A. 2005, Digitised First Byurakan Survey plate database, CDS, Catalogue No. VI/116, Strasbourg.
- Mickaelian, A.M., Gigoyan, K.S. 2006, A&A 464, 1177
- Monet, D.G., Levine, S.E., Casian, B., et al. 2003, AJ 125, 984
- Weedman, D.W. 1977, Vistas in Astronomy 21, 55
- Weedman, D.W., Khachikian, E.Ye., 1971, Ap 7, 389