ABSTRACT

The DFBS is the digitized version of the famous Markarian Survey. It is the largest low-dispersion spectroscopic survey of the sky covering 17,000 square degrees at galactic latitudes |b|>15. DFBS provides images and extracted spectra for all objects present in the FBS plates. Programs were developed to compute astrometric solution, extract spectra, and apply wavelength and photometric calibration for objects. DFBS database and catalog has been assembled containing data for nearly 20,000,000 objects. A classification scheme for the DFBS spectra is being developed. The Armenian Virtual Observatory is based on the DFBS database and other large-area surveys and catalogue data.

Keywords: surveys, Markarian galaxies, low-dispersion spectra, databases, Virtual Observatories

1 INTRODUCTION

The First Byurakan Survey (FBS) was the first systematic objective prism survey of the extragalactic sky. It was conducted by B.E. Markarian, V.A. Lipovetski and J.A. Stepanian in 1965-1980 at the Byurakan Astrophysical Observatory (BAO) with the 1m Schmidt telescope and 1.5° prism. Until now, the FBS is the largest area spectral survey, covering more than 17,000 deg² of all the northern sky and part of the southern sky at high Galactic latitudes (|b|>15°), with a total of ~ 40,000,000 spectra for ~ 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). Studies of Markarian galaxies early in the survey led to the spectral classification of Seyfert Galaxies (Khachikian & Weedman 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; Abrahamian et al. 1999; Mickaelian 2008); a survey for late-type stars (Gigoyan et al. 2003), 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; 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 The DFBS Project

The Digitized First Byurakan Survey (DFBS; Mickaelian et al. 2007) is the digitized version of the Markarian survey, the First Byurakan Survey. This project is relevant for preserving a very large database that was a milestone in the history of astronomy and to offer to the scientific community a new tool for investigating the properties of the nearby universe. In the following subsections, we describe the steps to construct the DFBS and give also guidelines for web access to the DFBS. The DFBS was carried out in 2001-2007 and consisted of the following tasks:

1. Scanning of the plates
2. Archiving on HDDs and DVDs
3. Astrometric plate solution
4. Extraction of images and spectra
5. Wavelength calibration
6. Density and flux calibration
7. Multiband (UBVR and O/E) photometry
8. Determining template spectra and numerical classification
9. DFBS catalogue and database
10. Web page and user interface

The participating teams are: Byurakan Astrophysical Observatory (BAO, Armenia), Dipartimento di Fisica, ‘La Sapienza’ Università di Roma (Italy), Cornell University (Ithaca, NY, USA), MIGG Informatica & Ricerca (Italy), Hamburger Sternwarte (HS, Germany), and the Institute for Informatics and Automation Problems (IIAP, Yerevan, Armenia).

2.1.1 Scanning

Digitizing of the FBS plates started in 2002. After several test scans on a set of plates, all of the FBS plates were digitized with an Epson Expression 1680 Pro scanner at the Byurakan observatory. DFBS was the first digitization 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 ~1.7mm, 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 (Mickaelian et al. 2005) were digitized and stored on DVDs. An example of a small field in DFBS is given in Figure 1.

2.1.2 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 2.1.4), but it is mildly sensitive to the brightness and spectral type (colour) of the object. The accuracy of definition for this cutoff is ~0.2". The software for the astrometric solution was written by H.J. Hagen by adapting the dedicated software for the Hamburg Quasar Survey (HQS, 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"/pixel. The best accuracy achieved was 0.87" rms, and the typical accuracy, ~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 ~5 min, and the procedure was applied for all scanned 1874 DFBS plates.
2.1.3 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 (Figure 2). 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): 

\[ I = \frac{(V-B)}{(T-B)} \]

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≤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 barycenter (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.
2.1.4 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, He, and other lines. We obtained only an approximate wavelength calibration scale because the dispersion is strongly non-linear, from about 22 Å/pixel at the blue edge of the spectrum to about 60 Å/pixel at the red edge, with mean dispersion 32.7 Å/pixel and dispersion of about 28.5 Å/pixel at Hα. Photometric information from the DFBS spectra can be obtained by integrating the spectra blueward and redward of the green sensitivity gap typical of I1F 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 \(4000 \times 4000\) 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 \(\pm 0.5^m\) in the range 13\(^m\)-16\(^m\) (Figure 3).
2.2 The DFBS Database

The DFBS database (version 3.0) presently is stored on a dedicated PC at the ‘La Sapienza’ University in Rome and can be accessed through a web interface (see subsection 2.3). A mirror site is available in Yerevan at the IIAP. The database includes all the digitized plates and the automatically extracted spectra. Only for F-type plates (IIaF, IIAF, IIF and 103aF) does the automatic procedure give reliable results; for such plates also approximate B and R magnitudes have been computed and are available. Visual software named bSpec was created in order to get automated extraction and classification of the spectral data in a DFBS plate, based on the experience gained with the IRAF tasks discussed above. This software performs all the operations necessary to build the DFBS database and was developed by one of the authors (GC) and the MIGG Informatica & Ricerca team in the frame of a collaboration project with the 'La Sapienza' University group; it was coded under Linux using the Borland Kylix compiler.

In the reduction procedure, the catalogue driven approach has been used for the finding algorithm: an object list in each plate area was downloaded from the USNO-A2 catalogue, with an upper magnitude bound in B≤17m. Starting from the USNO coordinates, each spectrum was re-centered with a combination of two parameters a) peak position, b) barycenter. For each image, a mean spectrum position angle was computed through the distribution of the whole set of objects and adopted as the spectrum direction for all the spectra of the plate. To avoid noisy objects, a limit in brightness and ellipticity was imposed to the spectra to be used in the computation of the mean direction. The local background around each object was estimated using the median value of the pixel distribution in two strips (left and right side) parallel to the spectrum direction. The software performs automatically the transformation from DN to intensity (in arbitrary units) for each plate, finds the “red head” of each spectrum and extracts the spectrum subtracting the local background; the abscissa of each extracted spectrum is set to have pixel 20 at the red head.

A preliminary magnitude calibration for each plate is made using the objects with USNO magnitudes between 12m and 16m in the central part (1600×1600 pixels) of the plate: several hundred stars are generally available for this purpose. Instrumental B and R magnitudes are evaluated integrating the spectrum between pixels 20-40 (R) and 55-90 (B). A linear fit of these magnitudes against their USNO-A2 magnitudes provides a calibration curve; objects deviating more than two σ are excluded from the relation and a second iteration is made to get the final calibration. This linear fit is then used to compute the DFBS magnitudes of all the objects in the plate. Objects brighter than 13m are generally overexposed in the DFBS so their luminosity is strongly underestimated and unreliable; objects fainter than about 16.0m have a large uncertainty due to the low Signal to Noise ratio.
The calibration plots for both B and R magnitudes can be displayed from the box showing any extracted spectrum.

For each object, the database contains the following information: USNOA2 identification, RA and Dec, position of the red head in plate pixels, B and R magnitudes in the USNO-A2 and from the DFBS, the local background value, a quality flag, the spectrum length (defined as the distance between the red edge and the first pixel below the sky level blueward), and the extracted spectrum (142 pixels long including 20 pixels of sky before the red edge). The present version of *bSpec* (1.78) is not able to extract separately two spectra if they are too nearby in RA (a few pixels); in such cases only one spectrum appears in the database, associated with the brighter object, and is not reliable. If two spectra have nearly the same RA and less than 2.5′ distant in Dec they partially overlap; in these cases the spectra are extracted but are not physically very meaningful. Human judgement is necessary to derive any information from such cases.

The automatic spectral classification is currently under test and will be included in a future version of the database. Presently only 4 flags are used: OK; the spectrum has a good S/N and is not overlapped. F; the spectrum is faint but good NL; the spectrum overlaps another one. The overlap test is performed using the spectrum length as distance criterion (in declination) from another spectrum. U; unclassified. The classification is sometimes faulty; it can be used for statistical purposes but eye inspection is recommended for studies on a given object.

### 2.3 The DFBS web page and user interface

A web page and user interface have been created to allow access to the DFBS database for the astronomical community. It can be queried at [http://byurakan.phys.uniroma1.it/](http://byurakan.phys.uniroma1.it/) or [http://arvo.sci.am/ARVO/DFBS/](http://arvo.sci.am/ARVO/DFBS/). The user interface provides access to general information on the FBS and DFBS, an easy comparison of a spectroscopic DFBS plate with the corresponding direct plates from the DSS1 and DSS2, the access to the database and to the digitized plates. It does not allow the user to run the *bSpec* software on the remote machine. Users may download portions of the plates in FITS format and perform their own analysis. Each spectrum is identified by its coordinates in the USNO-A2 catalogue. Given that often more than one plate contains the object, the plate number is added at the beginning of each identifier. The full identification is therefore, for example, FBS1053–DFBS J213143.70-124622.2. Spectra can be downloaded as a single ASCII file containing all the selected spectra or as separate files for each spectrum (a tar or zip archive is actually downloaded). In case of separate files the full identification of each spectrum is used as filename with .spec extension. The user interface (the DFBS portal) presently allows the following operations to the guest user:

i) **Sky coverage**: in a RA, Dec rectangular sky map shows the position on the sky of each plate. Basic data about each plate are available by clicking on the corresponding sky position.

ii) **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. Basic data about each plate are available by clicking on the plate number.

iii) **Explore**: it is the general purpose window. 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 guest computer. If more than one plate covers the requested sky area these plates are listed in a combo box and can also be displayed.

iv) **Get Image**: allows users to select a portion of a plate (presently up to 1024×1024 pixels, i.e. about 26.5′×26.5′) 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.

v) **Get Spectra**: allows downloading all the spectra in the database within a given distance from a selected central position (cone search); 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. The user can choose to download only "OK class" spectra, spectra just from the plate with the center nearest to the requested sky position, or all the plates covering the selected area. 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. If less than 50 objects are selected, also the 2D images can be downloaded selecting a check box. The table containing the identifier and the main data for each spectrum (RA, Dec, B, R, Plate number, Class, length) can be also downloaded as an ASCII file. The webshot of this option is given in Figure 4.

![DFBS webpage showing the list of requested spectra and an extraction of a given spectrum on the right.](image)

Figure 4. Operation “GETSPECTRA” at the DFBS webpage showing the list of requested spectra and an extraction of a given spectrum on the right.

### 3 THE ARMENIAN VIRTUAL OBSERVATORY

#### 3.1 AVOs and the Armenian VO project

The Astrophysical Virtual Observatories (AVOs) have been created in a number of countries using their available databases and current observing material as a collection of interoperating data archives and software tools to form a research environment in which complex research programs can be conducted. Among all these data, a large spectroscopic database for all objects will be especially useful. The Armenian Virtual Observatory (ArVO, [http://arvo.sci.am/ARVO/](http://arvo.sci.am/ARVO/)) has been created to utilize the DFBS as an appropriate spectroscopic database, as well as the Digitized Second Byurakan Survey (DSBS; Stepanian 2005) and other archive data. The DFBS database has to be used together with other available data on the Internet, which are accessed through AVOs. The ArVO project includes the creation of a database of DFBS spectra and its integration in AVOs, creation of a user interface with a full access to all DFBS data as well as all existing data from other databases. DFBS may be a major contribution to AVOs, giving full access to its data by the astronomical community, including the possibility of preliminary classification of any <18\(\text{m}\) object for further study, and the possibility of selecting objects of needed types from the DFBS. A quick optical identification of radio, IR or X-ray sources will be possible by plotting their positions on the DFBS plate and matching all available data.

VO standards include the Simple Image Access Protocol (SIAP) and Simple Spectra Access Protocol (SSAP). These allow homogeneous access to various image and spectral data in the world databases. To make the DFBS data available in VO, we have started preparing the DFBS images and spectra in VO standards. DFBS is promoted as an
The Armenian Virtual Observatory (ArVO) is a part of the International Virtual Observatories Alliance (IVOA, http://www.ivoa.net).

### 3.2 Astrophysical projects in ArVO

The main goal of the Armenian Virtual Observatory is to develop efficient methods for science projects based on the DFBS and other large astronomical databases, both Armenian and international. Thus ArVO project also includes several astrophysical applications, subprojects that will make use of the DFBS and the ArVO in general. Examples are creation of joint digitized low-dispersion spectroscopic database in the North; development of an automatic identification procedure for X-ray, IR, and radio sources; development of an automatic search procedure for modeled objects, etc.

Two groups of projects are especially productive: search for new interesting objects of definite types by low-dispersion template spectra, and optical identifications of new γ, X-ray, IR and radio sources. The first one is based on modeling of spectra for a number of types of objects: QSOs, Seyfert galaxies, white dwarfs, subdwarfs, cataclysmic variables, planetary nebulae, late-type stars (K-M, C, S), etc. Each kind of object appears in the DFBS with its typical SED and spectral lines (for objects having broad lines only, like white dwarfs and subdwarfs, quasars and Seyferts, etc.), however affected also by its brightness, so that each template works for definite range of magnitudes. The search criteria define how many objects will be found for further study, and may restrict these numbers leaving with the best candidates. Optical identifications have been proven to be rather efficient for IR sources from IRAS PSC and FSC. Tests have been carried out for X-ray and radio sources as well.

A special emphasis is put on search for bright QSOs missed by the Sloan Digital Sky Survey (SDSS, Adelman-McCarthy et al. 2008), rather important for making up their complete sample, studies of the properties of the Local Universe, a comparison of their X-ray, IR, and radio properties and making up their multiwavelength SEDs, as well as for a refinement of the AGN classification. A project for search for new bright QSOs using the DFBS has been started in the region with DEC>0° and |b|>20°. The Byurakan 2.6m telescope is being used for the spectral identification of the candidates. The first test resulted in 145 objects found, 81 being known QSOs/Sys, and 64 new candidates (including 23 NVSS and FIRST radio sources).

One of the popular VO software tools, SkyBoT (Sky Body Tracker), has already been used for search for asteroids in the DFBS (Thuillot et al. 2007). All known bright (<15°-16°) asteroids have been grouped into fast and slow ones with a division parameter, estimated as the motion of 3" during 20 min (the typical exposure time of a DFBS plate). Extraction of the spectra of asteroids found in DFBS by SkyBoT has been started, and they were grouped into extended (fast asteroids) and star-like (slow asteroids). The further plans include modeling of template spectra of asteroids by means of the star-like spectra, search for new candidate asteroids by similar spectra and comparison with DSS1/DSS2 fields for elimination of the stars, spectral analysis of the asteroid spectra to get some physical parameters, etc.

Another project is the identification of the newly found IR sources from Spitzer Space Telescope (SST). 73 unidentified sources in the Boötes region have been found and classified on the DFBS plates. All available additional data from DSS1/DSS2, other optical and multiwavelength catalogs were used to clarify the nature of these sources.
objects. 51 were found to be known objects from existing catalogs, including 1 QSO, 28 galaxies, and 22 stars. The 22 new objects were classified as 5 candidate QSOs, 10 galaxies (4 AGN candidates and 6 interacting systems), and 7 stars (6 G-M type, and 1 carbon). 8 of the known stars not having spectral classification were classified too.

Any other VO tools will also be applicable for further astrophysical projects when the DFBS is ready in VO standards.

4  CONCLUSION

The electronically-accessible DFBS is the primary result of our project. Use of the digitized 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 utilizes multivariate relations such as relations between magnitudes and widths of the spectra (for separation of stellar and diffuse objects), SED, 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 digitized 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 new research ideas will also originate from the DFBS.

5  ACKNOWLEDGEMENTS


6  REFERENCES


