Astronomical Data Analysis Software and Systems XIII
ASP Conference Series, Vol. 314, 2004
F. Ochsenbein, M. Allen, and D. Egret, eds.

ImgCutout, an Engine of Instantaneous Astronomical Discovery

María A. Nieto-Santisteban and Alexander S. Szalay

Department of Physics & Astronomy, The Johns Hopkins University, Baltimore, MD 21218, USA. Email: nieto, szalay@pha.jhu.edu

Jim Gray

Microsoft Research, San Francisco, CA 94105, USA. Email: gray@microsoft.com

Abstract. ImgCutout is a Web application that enables professional astronomers and the general public to interactively visualize and explore large, complex astronomical data sets. The application consists of a Web interface that calls a Web service, which accesses SkyServer, a 1 TB SQL Server database containing catalog data for 100 million objects, spectra and images from the Sloan Digital Sky Survey. ImgCutout builds, in real time, color mosaic-images of user-selected regions of the sky, and overlays additional information about astronomical and spatial objects in the database including: boundaries of survey fields and aperture plates, outlines of individual objects and data quality masks, in addition to locations of photometric and spectroscopic objects. The tool can search for lists of known objects, allows new database queries, and provides detailed information about selected objects.

1. Introduction

The Sloan Digital Sky Survey (SDSS) is digitally mapping one-quarter of the sky in 5 spectral bands. SDSS will ultimately obtain 40 TB of raw pixel data, spectra of 1 million objects, and photometry for more than 200 million objects.

Raw survey data from the telescope are processed at Fermilab to produce various image and data files. These data files are then loaded into the SkyServer database at Johns Hopkins University (JHU) after further processing (Thakar, Szalay, & Gray 2004). SkyServer is a relational database that is tuned and heavily indexed to allow fast exploration and analysis of the data. The SkyServer database occupies 1 TB for Data Release One (DR1), but it will grow to about 6 TB for the full survey.

We handle quick database queries interactively SkyServer Web site[1], but queries longer than a minute are handled by the Batch Query System (O'Mullane et al. 2004), which requires registration.

[1] `http://skyserver.sdss.org/`

© Copyright 2004 Astronomical Society of the Pacific. All rights reserved.

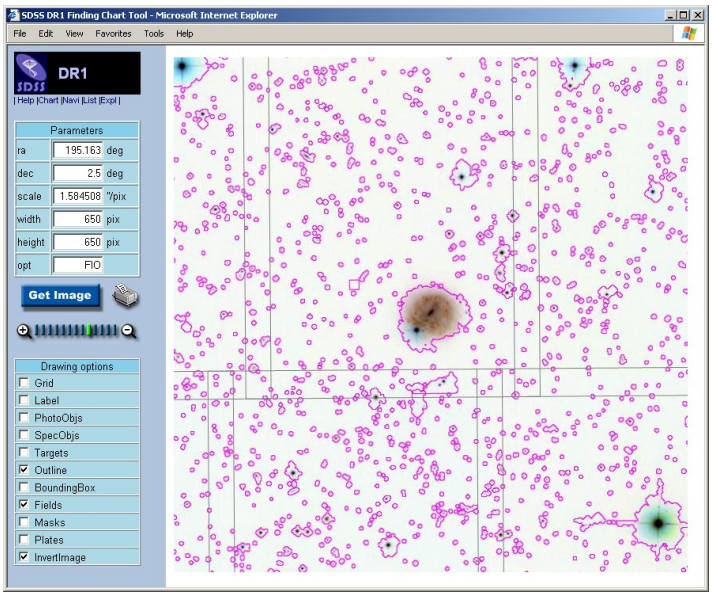

Figure 1. Snapshot of the Finding Chart ImgCutout Web interface

2. ImgCutout Web interface

The SkyServer Web site provides interactive access to images and associated catalog data via Visual Tools, which include the Finding Chart, Navigator, Image List, and Explorer. These interlinked tools exchange information about the current selected object, allowing transparent mobility between data views. All of these visual interfaces use a common underlying SOAP Web service, ImgCutout[2].

2.1. The big picture of the sky

To allow rapid data visualization, images are stored in JPEG format. During the preprocessing stage, FITS images in 3 of the original 5 bands are combined to create a color JPEG image. The JPEGs are stored in the database at 5 zoom levels, along with associated astrometry information. This preprocessing facilitates rapid recovery of the frames and geometric transformations needed to create dynamically a mosaic of any region of interest. Mosaics can be as big as 2048 × 2048 pixels, covering approximately 34 × 34 degrees.

Many interesting features can be overplotted on top of the images. The Field option displays the boundaries of the original FITS images used to construct the mosaic. The Plates checkbox shows the boundaries of the aperture plates used to obtain spectra chosen from a list of potential Targets. Selecting SpecObjs or PhotoObjs highlights objects in the image with spectroscopic or photometric data. Outline and BoundingBox options mark the pixels associated with each object (Figure 1).

[2] `http://skyservice.pha.jhu.edu/dr1/ImgCutout/ImgCutout.asmx`

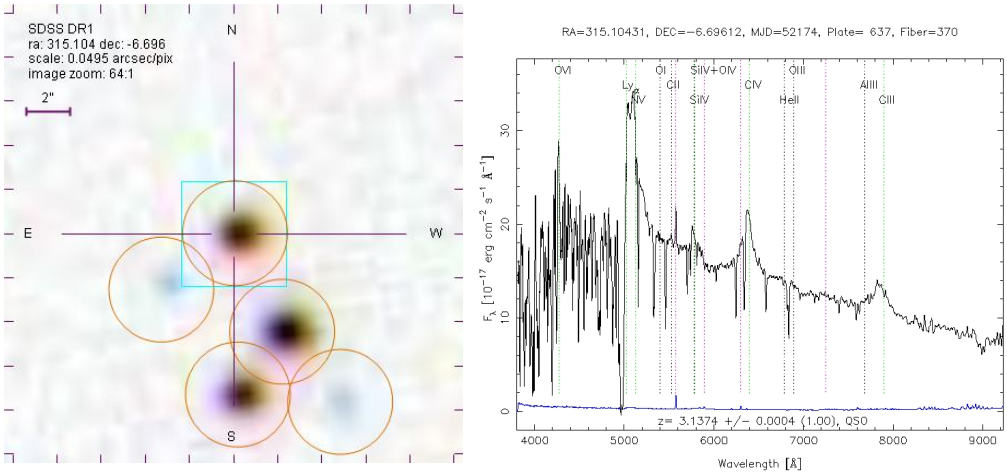

Figure 2. Possible lensed quasar and spectrum.

2.2. Data quality

Due to bad weather, bright stars, satellite trails, or other anomalies, 10–15% of the raw survey data are substandard. The SkyServer Mask option allows easy identification of regions with poor data quality. This visualization tool helps guide data analysis and lets users decide quickly whether particular data are useful for their purposes.

2.3. Viewing known sources. Making new discoveries

Using the Image List tool, users can simultaneously view thumbnails of SDSS images for objects specified by a list of coordinates. The thumbnails provide a visual index for objects specified by the user or returned by an SQL query of the database. More detailed data views can be interactively selected from the Image List tool allowing creation of Finding Charts, for example.

The most interesting feature of this tool is that users can generate a list of coordinates and associated thumbnails by querying the database directly using SQL. For example, visual inspection of the thumbnails generated in response to the following search for high-redshift quasars:

```
SELECT specObjId, ra, dec
FROM SpecObj
WHERE specClass = 3 and z > 3
```

yields a potential lensed quasar. Selecting the thumbnail brings up a more detailed image in the Navigator. Inverting the colormap improves the visibility of faint background objects. Selecting PhotoObjs shows 5 closely spaced objects present in the photometric catalog. The Explore tool then allows detailed investigation of colors, morphological type, and spectrum (see Figure 2). Based on similar colors for some of the objects, followup observations were requested to test whether or not this particular group of sources is indeed a lensed quasar.

3. Design and Technologies

The user interfaces described above are all layered on the ImgCutout Web service for SDSS data in the SkyServer database, but the same principles would work equally well for other surveys. Interactive tools must respond quickly to be successful, so JPEG images and database indexing should be considered for any generic image cutout service. In general terms, the system is constructed as follows:

1. Combine and convert FITS into JPEG.
2. Store JPEGs with several zoom levels in the database.
3. Store astrometry for frames and objects.
4. Store other relevant information (masks, plates, etc).
5. Construct a spatial index, HTM2 (Fekete, Szalay, & Gray 2004).

The database provides images and object data to an image cutout service in response to SQL queries. The cutout service assembles the images into mosaics and displays the mosaics with the selected overlays. In the specific case of the SDSS, we perform these tasks using the following technologies:

- Database: Microsoft SQL Server 2000.
- HTM2 libraries: implemented with C++ and C#.
- Web interface: ASP pages and Java Scripts.
- Web service: implemented in .Net and C#.

The ImgCutout Web service can be accessed in various ways:

- Via our Web site[3].
- Through other Web sites using HTTP protocol[4].
- Using the SOAP protocol. A detailed tutorial[5] is available on the IVOA Web site about how to implement clients in a number of languages and on different platforms.
- Using the Virtual Observatory (VO) Simple Image Access Protocol (SIAP) service for SDSS data.

All of the ImgCutout software, a 1.3 GB subset of the SkyServer DR1 database, and the SkyServer Web site code and content are available for download in the MySkyServer site[6]. MySkyServer database and software are useful for experimenting with queries and seeing how the tools are implemented.

References

Fekete, G., Szalay, A., & Gray, J. 2004, this volume, 289

O'Mullane, W., Gray, J., Li, N., Budavari, T., Nieto-Santisteban, M., & Szalay, A. 2004, this volume, 372

Thakar, A., Szalay, A., & Gray, J. 2004, this volume, 38

---

[3] `http://skyserver.sdss.org/`

[4] `http://skyserver.sdss.org/dr1/en/tools/chart/chart.asp`

[5] `http://www.ivoa.net/twiki/bin/view/IVOA/WebgridTutorial`

[6] `http://www.skyserver.org/myskyserver/`