Creating updated, scientifically-calibrated mosaics for the RC3 Catalogue

Abstract

The Third Reference Catalogue of Bright Galaxies (RC3) is a reasonably complete listing of 23,010 nearby, large, bright galaxies. By using the final imaging data release (DR10) from the Sloan Digital Sky Survey, we generate scientifically-calibrated FITS mosaics by using the montage program in all SDSS imaging bands for all RC3 galaxies that lie within the survey footprint. We further combine the SDSS g, r, and i band FITS mosaics for these galaxies to create color-composite images. We generalized this software framework to make FITS mosaics and color-composite images by using an arbitrary catalog and imaging data set. Due to positional inaccuracies inherent in the RC3 catalog, we employ a recursive algorithm in our mosaicking pipeline that first determines the correct location for each galaxy, and subsequently applies the mosaicking procedure. As a test of this new software pipeline and to obtain mosaic images of more RC3 galaxies, we applied this pipeline to photographic data taken by the Second Palomar Observatory Sky Survey (POSS-II) with B, R, I plates. We publicly release all generated data, which can be accessed via a web search form, and the software pipeline to enable others to make additional galaxy image mosaics.

1. Introduction

Astronomers have a long history of cataloguing objects for subsequent study, for example the Messier catalog (**REF**) or the New general Catalog (NGC; **REF**) have provided valuable guidance to help astronomers study objects with similar properties. Today, we have entered the era of big data, where large surveys such as the Sloan Digital Sky Survey (SDSS; **REF**) have uniformly surveyed large fractions of the entire sky providing detailed photometric and astrometric information for millions of sources. However, the software pipelines that uniformly process these valuable data are optimized of the more numerous, small sources. As a result, large, nearby, and bright galaxies are essentially treated as contaminants. Yet these galaxies still remain incredibly important, providing detailed insight into the dynamics of galaxies as well as a low redshift comparison sample with which to compare higher redshift galaxies.

One of the more popular catalogs of nearby galaxies is the Third Reference Catalog of Bright Galaxies (RC3;**REF**), which is a catalog containing 23,010 galaxies with an apparent diameter greater than one arcminute at the D_{25} isophotal level, with a total B-band magnitudes brighter than 15.5, and with a redshift not in excess of 15,000 km/s. The overall catalog is supplemented by select galaxies that may only meet one or two of these conditions as well as some compact nearby galaxies. Given the efficacy of this catalog, previous authors have used this sample as the basis for making image mosaics for a large galaxy sample. The first such effort was made by **Cite Hogg and Blanton 2006**, who made color-composite images of select RC3 galaxies by using the SDSS g, r, and i band images from the sixth SDSS data release. A subsequent effort by **Cite EFIGI** (YEAR), dedicated to the study of galaxy morphology, generated a set of 4,458 FITS and color images by using the SDSS DR4 data. This latter effort employed a visual inspection to remove artifacts and galaxies with missing data. Finally, a related effort, known as the NASA-Sloan Atlas [1] has been undertaken to construct and analyze a complete set of galaxies within approximately 200 Megaparsecs (see, e.g., **?)fix the reference**.

Given the SDSS project has published their final imaging data release in SDSS DR10, which in-

[1] http://www.nsatlas.org

cludes the final photometric and astrometric calibrations, we have decided to update these previous efforts to make full, scientifically-calibrated image mosaics as well as color-composite images for all RC3 galaxies that lie within the SDSS footprint. As part of this process, we also generate updates coordinate positions for those galaxies with inaccurate coordinate locations. We further generalize this software framework to generate scientifically calibrated FITS and color-composite images from any suitable survey for an arbitrary galaxy catalog. We test this new pipeline out by also constructing FITS image mosaics and color-composite images for RC3 galaxies that lie within the Second Palomar Observatory Sky Survey (POSS-II; **REF?**). This pipeline can be easily applied to existing data from the two micron all sky survey (2MASS) or to future surveys such as the Dark Energy Survey (DES; **REF?**) or Large Synoptic Survey Telescope (LSST;**REF?**).

In Section 2, we introduce the RC3 catalog and the SDSS and POSS-II data with which we make mosaic images. Section 3 introduces our software pipeline and the relevant algorithmic details that enable us to obtain improved astrometric precision for the catalog galaxies. We discuss the actual construction of the image mosaics for the SDSS and POSS-II data in Section 4, before concluding the paper and discussing the overall project in Section 5. The image mosaics for the SDSS and POSS-II image data are available at the LCDM website[2] and the software pipeline is available from github[3].

2. Data

To construct large, calibrated image mosaics, we need to types of data. First, we need a catalog galaxies, and second, we need an image data set. A number of different candidate galaxy catalogs exist; we also could follow the example of Blanton et al and construct a new catalog based on specific physical criteria. We choose to use the RC3 catalog of galaxiesde Vaucouleurs et al. (1995), which we discuss in detail in the following subsection. For our imaging data, we actually choose two different survey data products. The first survey is the Sloan Digital Sky Survey**REF**, which uniformly surveyed a large fraction of the sky, and the Second Palomar Observatory Sky Survey , which is an older photographic plate survey that covers nearly the entire sky. We discuss these two imagine surveys in detail at the end of this section. While our pipeline approach can be easily extended to other datasets, such as the Two Micron All Sky Survey2MASS Skrutskie et al., 2006 **Could list GALEX and WISE here as well**, we do not explicitly discuss them in this section.

2.0.1. The RC3 Catalog

The simplest approach to constructing a uniform galaxy catalog is to make a define a galaxy sample by galaxy limiting apparent brightness within a survey. Original attempts to accomplish this task date back to the original Harvard Survey of External Galaxies **Change this to a proper reference:** (H. Shapley and A. Ames 1932), which contained 1249 objects brighter than 13^{th} **What band do you know?** magnitude. Many of these galaxies were subsequently included in the University of Texas monographs in astronomy, which were predecessors to the RC3 catalog .The actual RC3 Catalog is an update to the Original and Second Reference Catalog of Bright Galaxies, which collected data from **is this an exact or approximate number?** 150 different published **is this correct? published, or simply shared data?** sources, which originated from different telescopes with varying accuracy, precision, and coverage. The original RC3 galaxy catalog compiled by de Vaucouleurs et al. (1995), contains a complete listing of 23,010 **why is this number different than the number at vizier?** galaxies with D_{25} apparent major isophotal diameter greater than one arcminute and with a total B-band magnitude greater than 15.5^{th} magnitude.

After the original RC3 catalogue was published, it was incrementally updated with new, published observations and maintained by Harold G. Corwin.[4]. As the imaging data used to update these galaxies came from various different imaging programs, the final, updated catalog is heterogeneous with a non-uniform distribution of updated galaxies. In this project,

[2]http://lcdm.astro.illinois.edu/data/rc3/
[3]http://github.com/ProfessorBrunner/rc3-pipeline/
[4]http://haroldcorwin.net/rc3/bugs.rc3

	SDSS	2MASS	POSS-II
Imaging bands	g,r,i	J,F,K	R, B,IR
Sky Coverage(%)	35.28	99.998	78.27
Resolution(″/pix)	0.396	2.0	1.7
Imaging Technique	CCD	Infared array	Photographic Plates

Table 1: Comparison of survey and telescope specifications

we used the catalogue information available through the VizieR Service, which contains the original RC3 data published in 1991. There is also a 1994 version updated by Corwin available on VizieR.

Since the RC3 catalog is a reasonably complete representation of large, bright nearby galaxies in the extragalactic sky, it remains a popular catalog. Selected galaxies or complete subsets are used in astrophysical studies of quasars and X-ray sources (Walton et al., 2011) or galaxy morphology and properties**Reference?**. The RC3 catalog also serves as a basis for statistical studies in cosmology, for example within the New York University-Value Added Galaxy Catalog (Blanton et al., 2005).

Robert Comment: Stopped here!

2.0.2. Positional Inaccuracy in RC3 Catalog

The principle motivation behind the automation algorithm is to resolve the problem of centering galaxies and finding RC3 galaxy in a given field of view. Initially, using the standard mosaicking steps in Montage, we obtained many images with off-centered or missing galaxies due to the inherent inaccuracy of the positions in the catalog. There is an non-uniform update of the catalog after the catalog was publish, more accurate position was collected from whatever sources and data available. The B2000 coordinates in the FK4 frame of RC3 galaxies are denoted with two different levels of accuracy: HH MM SS.s, DD MM SS for the positions that has been updated with accuracy of about 5-8 arcsec and HH MM.m, DD MM for galaxies whose positional accuracy remain 1-2 arcminutes as in the original catalog. In the 1991 version of the catalog (de Vaucouleurs et al., 1991), there still remains 5492 RC3 galaxies that fall in the latter group.

2.1. Data

2.1.1. SDSS

SDSS images is taken from a 2.5m telescope at the Apache Point Observatory in New Mexico. We use the imaging data from Data release 10 which is part of the SDSS-III . SDSS-III is dedicated to studying galaxy clustering by looking at signatures of Baryon Acoustic Oscillation,structure and history of the galaxy , and the population of exo-planets. SDSS is one of the newest all-sky survey with improved instrumentation. CCD detectors have obvious throughput advantages over photographic plates of POSS-II. In addition, the resolution of SDSS is better than 2MASS and POSS-II. From Figure 4, we see that 2MASS tends to smooth out imaging details of a galaxy (especially at its wings) due to the large tiles that its imaging data is consist of.

The SDSS data is pass through the `photo` pipeline, which is responsible for source detection, deblending, model-fitting, photometric calibration, and other image processing procedures. We record imaging quality information on the data product by using the `clean` flag in SkyServer. For galaxy and other extended source, the `clean` flag is defined from data masks with variables describing PSF magnitude error, cosmic rays, undefined profile..etc.[5]The data is queried using SkyServer through a SDSS Command Line Query Tool. Then the the calibrated, sky-subtracted corrected frame with the calibration meta-data (`fpC`) is retrieved in bulk from the Science Archive Server (SAS).

2.1.2. 2MASS

The 2MASS(Skrutskie et al., 2006) is a near-infrared survey in J(1.25 μm), H(1.67 μm), and K_s(2.17 μm) band. Data taken from two 1.3m telescope covered almost the whole sky with overlapping“survey tile” is 6 degree long by 8.5 arcminute stripes. The raw data processed through the 2MASS Production Pipeline System are astrometrically and photometrically calibrated , and the source extracted atlas images are used as inputs for our mosaicking porgram. Since the RC3 catalog contains large galaxies that should not be confused as point sources, we narrowed down our

[5]`http://skyserver.sdss3.org/dr10/en/help/cooking/general/flags6.aspx`

imaging data search to 2MASS's Extended Source Catalogs which contained only 1.6 million objects that are extended with respect to their instantaneous PSF. Some of the RC3 galaxies (larger than 120") is also contained in the 2MASS Large Galaxy Atalas, a subset of the Extended Source Catalog.2MASS data has the advantage of an all-sky coverage, however since each survey tile is large, the imaging resolution is lower. Also, multiband recombined color image for 2MASS is less of a true-color compared to SDSS. When mapping 2MASS's J,H,K band, the relative values show up according to the wavelengths, so the contrast in feature is shown. Since SDSS is an optical survey, it look much better than 2MASS's infrared data since mapping g,r,i band to R,G,B is approximately accurate in its color value.

2.1.3. POSS-II

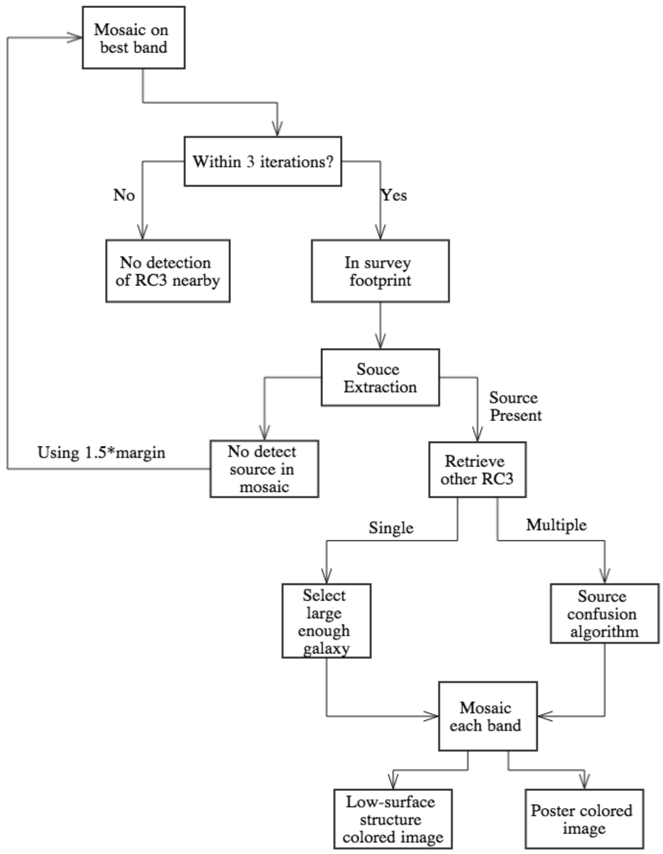
The Second Palomar Observatory Sky Survey (POSS-II) was a photographic survey that also covered most of the northern night sky. It is a update to POSS-I which was one of the first major all-sky photographic survey. We also serve FITS mosaics from POSS-I which contains only the B and R band data taken in the 1950s to serve as comparison. This data can also be joined with another single band survey for RGB mapping to create colored mosaics.

The photographic plates used to create our data product was made accessible through STScI's effort in the Digitized Sky Survey(DSS) project. The photographic plate J,F,N bands was calibrated to the Gunn g, r, i band (Gal et al., 2004). The DPOSS-II data consists a set of photographic plate data from POSS-I,II and UK Schmidt Telescope Survey. The data querying is done through NASA/IPAC Infrared Science Archive's (IRSA) Finder Program Interface. Each photographic plate is 6.5° by 6.5° of the sky with 5° overlap. After lossy compression, each plate is 1.1 GB. A cropout of the plate can be retrieved to speed up downloading time. Due to the large size of these plates, the program rarely has to stitch together multiple fields, only the positional update algorithm is executed.

Even though the POSS-II has the advantage of a greater sky coverage than SDSS, objects near the boundary on photographic plates suffers from vignetting patterns which may result in inaccurate astrometry for those objects. Another disadvantage is that photographic plates are much less sensitive to photons($\approx$ 1% quantum efficiency) compared to CCD ($\approx$ 80%). However, this is not a significant issue for the RC3 objects to be seen since RC3 objects are bright, but the sources appear to be dimmer in the images as seen in Figure 5.

3. Mosaic Pipeline and algorithms

3.1. Algorithms



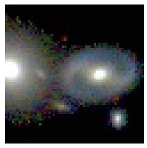
3.1.1. Positional Update

To address the problem of inaccurate positions discussed in subsubsection 2.0.2, we develop an algorithm that first updates the position by finding the galaxy in the imaging data. An `RC3` object contains its The Catalogue of Principal Galaxies(PGC) number, updated and catalog ra,dec,radius for each galaxy in the RC3 catalog. Since every RC3 galaxy has a unique PGC number, we record its PGC number in order to identify the RC3 galaxy of interest for cases of source confusion and naming of data products. The method `source_info` updates the position of individual RC3 objects by first creating a single band mosaic with a field of view 6 times the radius of the galaxy, then SExtractor detects all sources in the mosaic and keeps only the ones with a radius greater than 5.94 arcseconds. This size cut is large enough such that it eliminates most stellar point sources and background noises but retains the subset of RC2 galaxies inside the RC3 Catalog that are smaller than 1 arcminute as described by de Vaucouleurs et al., 1976. Also, if the mosaicking field is chosen correctly, then SExtractor's sky level estimation is fairly accurate.

If there are multiple galaxies in the field of view that satisfy this criteria, then it is passed into the source confusion algorithm to determine the galaxy of interest. Then we record this single source and generate mosaic FITS file for all bands and color images using its updated values. If no sources are detected, then we mosaic the single band FITS using a larger field of view. This is done recursively, while keeping track of the number of iterations in each recursive steps. The process is terminated at the 3rd recursive step and generates a mosaic for all bands. Then we select mosaics from the three imaging bands and recombine them to make two colored mosaic. One emphasizes on low surface structures so that the halos around the galaxy can be seen; the other image is a poster/publication image that uses higher cuts on the background to ensure a clean contrasting image. At least three band passes is necessary for generating a colored image since it enables R,G,B mapping by STIFF. Also, Skrutskie et al., 2006 explains that a minimum number of three bands is required to distinguish the effects of interstellar extinction and stellar and extra-galactic populations, which is why many surveys has three imaging bands.

3.1.2. *Source Confusion*

Large galaxies tends to form gravitationally bound structure in clusters of galaxies. The source confusion algorithm works on the assumption that the only galaxies large enough to cause source confusion also lie in the RC3 catalog, since the catalog is reasonably complete for galaxies having apparent diameters larger than 1 arcminute. First, the program retrieves a list of n number of RC3 galaxies that may lie in the field of view of the generated mosaic. The `otherRC3` method in the `Server` class queries this information using the Vizier Catalog. If this information is provided by a survey's query service, such as the RC3 Table in SDSS's SkyServer, it can be overridden. The goal is to match the results returned by the server with the list of n largest sources detected by SExtractor.



(a) Before

(b) After

Figure 1: The three galaxies' arrangement is shown in Figure 2. Using the source confusion algorithm, we resolve the three sources and generate recentered mosaic for PGC58.

We need to cross-correlate the coordinate by matching together the relative differences between the sources, due to the unreliability of absolute coordinates inherent to the catalog as discussed in section 2.1. We assume that the positional inaccuracy is due to instrumentation and measurement error but retains the galaxies' locations relative to one another. Therefore, we compute all the possible distances between any two RC3 galaxies that lies in our field. Then compare this with the set of differences generated by the SExtractor-generated list. We select n number of galaxies that has the closest values in both list, thereby matching together the galaxies from the catalog with the detected sources. Finally, we mosaic using that galaxy's coordinates as center. Our premises proves to be valid, as the algorithm executed with a success rate of 99.97% in the SDSS run, and is capable of correctly resolving up to five RC3 sources in one example.

3.2. *Pipeline*

The motivation behind building a class hierarchy is to establish a mosaicking pipeline for any given catalog and survey data. We identify the essential components required for mosaicking program to work. The relationship between these essential steps are re-

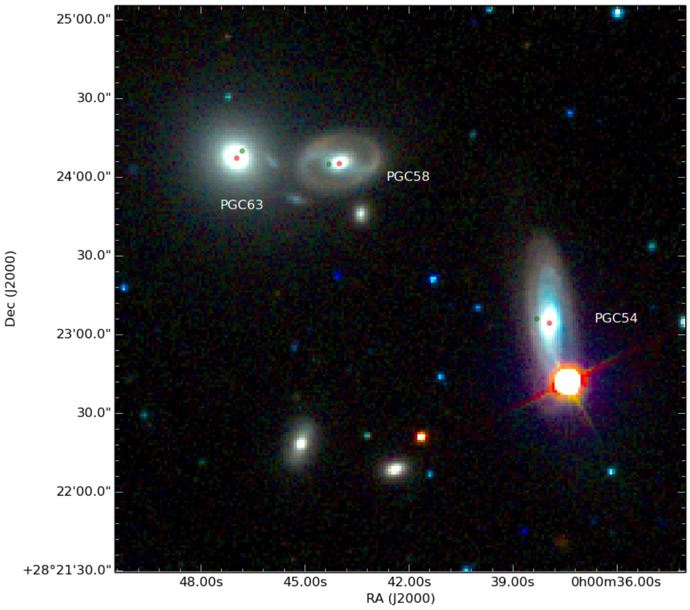


Figure 2: Three source confused RC3 Galaxies. Green marker denotes coordinates found in the original RC3 catalog, red marker denotes coordinates updated by the algorithm

flected in our design of the abstract `Survey`, `Catalog`, and `Server` classes.

3.2.1. *Server*

Data acquisition is composed of two main tasks: querying imaging data and retrieving data from server. Most surveys have an API that enables data access by SQL or customized query. To make the `Server` class work, we need to figure out how to convert positional values to record-keeping parameters dependent on the survey's telescope. For example, SDSS image frames are uniquely identified by a particular combination of run, camcol, and field and 2MASS refers to this with its sexagesimal, equatorial position-based source name. In the case where the survey has no special name referring to the imaging frame, such as in the case of POSS-II data, the user can invent their own naming scheme as placeholders for naming the raw imaging data. In addition to these two essential functions, each `Server` class should also implement the query builder methods that are used inside the main mosaicking program. Our choice of using a server instead of a data object enables code reusability across various surveys that uses common server tools , such as IPAC's GATOR query service

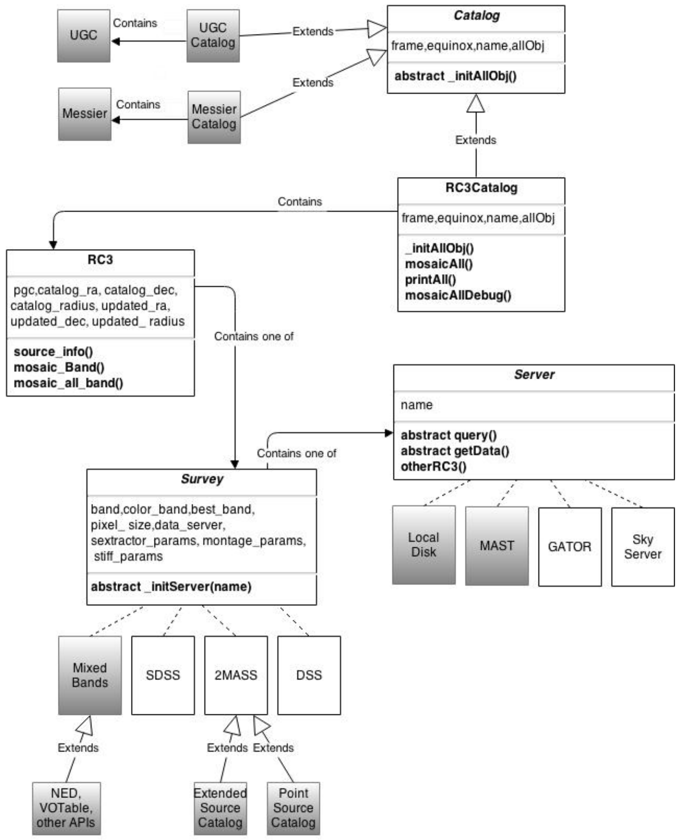


Figure 3: UML Class diagram showing the relationships in the mosaicking pipeline. Grey-filled boxes shows possible extensions of this pipeline

(Alexov et al., 2005) or Astroquery.

3.2.2. *Catalog Objects*

For the purpose of mosaicking RC3 galaxies, we did not create a `CatalogObject` class, although it would be helpful to have a generic class similar to the functions in `RC3Objects` as an extension to the pipeline. They contain basic information about the particular object and are survey-independent. They perform the essential mosaicking features on a per-object basis so that they not only can be used for comparing resulting mosaics from multiple surveys (Figure 4) but these methods can also be conveniently used in the `Catalog` class. The final step in the mosaic procedure generates two TIFF color images as described in subsubsection 3.1.1. When extending the pipeline to other surveys, investigators can adjust the STIFF parameters to better suit the telescope-specific imaging details following the guidelines in Bertin, 2012.

3.2.3. Catalog

The `Catalog` class contains a list of objects that lies inside the catalog. Although the use of the `Catalog` container class seems unnecessary, it enables a clear separation between basic mosaicking functionality for individual galaxies and for the entire catalog. This can be used to study some particular object, executed on all objects in a `Catalog`, or simply used for debugging purposes. This abstraction barrier ensures minimal changes to the code in both classes when an investigator decides to input data from a new survey in the future.

4. Results

4.1. Result from SDSS and DPOSS run

There are a total of 12512 RC3 galaxies within the SDSS footprint. The automated mosaicking pipeline was successful for 90.22% of the galaxies. 2446 RC3 galaxies is mosaicked using an updated position that was more than 1 arcminute off the values recorded in the RC3 catalog. On a 8-core machine, the program processes on average about 80 RC3 objects per hour using SDSS data. The finished data product occupies 14.7 GB of diskspace.

(# More stats on this after DSS finishes its run) For the POSS-II run, there are a total of [*] galaxies within the POSS-II footprint. The automated mosaicking pipeline was successful for[*] % of the galaxies. 2446 RC3 galaxies is mosaicked using an updated position that was more than 1 arcminute off the values recorded in the RC3 catalog. On the same machine that did the SDSS run, the program processes on average about 18.6 galaxies per hour using POSS-II data. The finished data product occupies [*]GB of diskspace.

4.2. Performance

We accelerate the mosaicking process by performing the recursive algorithm on only a single band FITS file designated as `best_band`, then mosacking all bands only once per object. For SDSS, we use the image obtained by the r band filter since transmission curve in Stoughton and Bernardi, 2002 shows that r band has the highest quantum efficiency. For

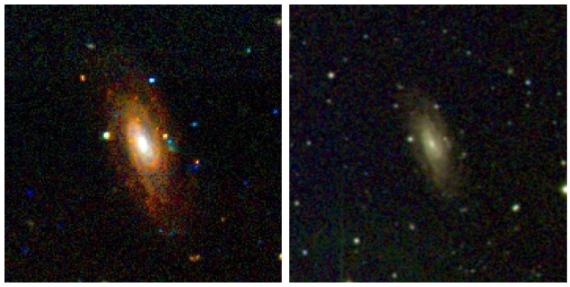


Figure 5: PGC1746 mosaic using SDSS and DSS

2MASS, The K_s band was chosen to be the `best_band` since it has the highest transmission as shown in Figure 2 in Skrutskie et al., 2006.

Most of the computation time is spent on downloading the raw FITS files from the survey's specific server. Therefore, even though Montage's modular design enables its performance to scale with number of processor (Jacob et al., 2010), the process would not be significantly sped up by the use of Message Passing Interface (MPI). Therefore, the speed depends more on the data transfer rate between the host site and local machine rather than the number of cores in the machine. This average runtime would also depend on the sky coverage of the particular survey and the speed of query.

This process can be significantly sped up if the investigator already has imaging data stored locally on a disk or have access to running the mosaic program alongside the survey's datacenter. We have designed our class hierarchy such that this can be written as a subclass of `Server` with user-defined details of where the images are stored. However, it is not worthwhile to download the whole survey's imaging data set for this purpose since that would probably take more time than

4.3. Technical Details

Due to the recent rise in popularity of Python usage in astronomy, we wrote the program in Python 2.7.6 so that the program's dependencies are more widely supported for extending its usage on future datasets. Most of the mosaic steps are done using IPAC's Montage Jacob et al., 2010 mosaic engine along with the AstroPy Montage wrapper[6] and the

[6] `http://www.astropy.org/montage-wrapper/`

Figure 4: PGC120 mosaic using SDSS, POSS-II, 2MASS data

final colored image is generated using Astromatic's STIFF v.2.4 Bertin, 2012. Our choice of the two program makes use of the best features from each programs. Montage creates scientifically-calibrated images by retaining the astrometry and photometry of input sources during image reprojection step.STIFF provides the flexibility of adjusting many variables for the final colored image, as well as automatically estimating upper and lower cuts on the dynamic range using statistics derived from a pixel histogram. Most of the program interface requires building the URL query and parsing the resulting raw text or XML files downloaded by wget. Source extraction is done using SExtractor v.2.19.5 Bertin and Arnouts, 1996. The resulting web search database was created using the Python sqlite3 module and the web search interface was written in PHP with interacting HTML elements.

4.4. Known Errors

Even though a series of exception handling and error prevention mechanisms were put in place, there are still errors in the data product that we produced. Error flags were put inside the web search database denoting the errors described below.

```
Error Information
0 = no error
1 = mosaicAll error
2 = stiff error
3 = strange error
4 = Montage image reprojection failure
5= mSubImage failure
```

1. mosaicAll error is a general error that is raised when the program breaks when mosaicking all bands.
2. STIFF enforces that the mosaic FITS used for RGB mapping but must be exactly the same dimensions in order to generate a colored image. Sometimes no resulting color images produced because 3 colored band not the same size. This happens particularly more often in SDSS's g band and in DPOSS. This also happens more often in DPOSS data since the photographic plates are non uniform in size.
3. In the source confusion algorithm, we store the galaxies that pass the size cut inside a radius with PGC as key and position as values. This error flag is raised when SQL region search does not included the galaxy itself.
4. `mProjectExec` is the Montage procedure that creates the reprojected image from the raw FITS files. Sometimes reprojected images are not created even when Montages' debug statement clearly shows that the reprojection was successful and table and header files are corrupted. This results in an error in later mosaicking steps. We have implemented error prevention mechanism to ensure that mosaic procedures terminates correctly in such cases and wrote the problematic galaxy into failed_projection, which can be examined later.
5. `mSubImage` crops a finished mosaic to the specified box frame. Montage throws an error when the program tries to crop the image with specified boundary lying outside the image field.

We typically use a box length of two times the radius. If it doesn't work, another attempt for a smaller box size is made or the error is recorded.

Montage's background rectification module is not used in creating the FITS mosaics. Attempts were made in implementing these procedure, however it is sometimes stuck in the step where the background model is applied to the projected images and produces huge difference images files (≈50GB) for reason that we have not figured out yet. This should not affect the astrometric quality of the data product. Sometimes scanlines may be seen as brightness differences in the SDSS mosaics, but the effect is much more severe and clear visible difference is mostly observed in the color images for 2MASS which manifests as distinct boundaries, and will likely affect the photometric quality of the resulting 2MASS mosaics.

5. Conclusion

Wide-area sky surveys are used to answer fundamental questions regarding large-scale structure and the cosmological history of the universe. The mosaicking pipeline described in this paper provides a convenient way to generate mosaics for sky survey imaging data for a given catalog, or on a per-image basis. This can be easily adapted for use on future data sets to create the scientifically-calibrated FITS mosaics as well as two colored images. We describe an algorithm that automatically updates the catalog using its inherently inaccurate positions and centers the galaxy around the newly updated coordinate for generating the mosaic. We generate a set of data product that resulted from running the RC3 catalog through the mosaicking pipeline using SDSS and POSS-II data and made this publicly available through a searchable web form.

Feeding in the new data enables a greater sky coverage and higher resolution images of objects inside the catalog, which can result in better astrometric and photometric results. In addition, the pipeline can also be extended for other astronomical catalog such as the Messier Catalog or NGC, as well as user-defined catalogs. An investigator may define their own catalog by imposing selection limits to study a type of objects of interest. A text file containing positions, radius, and unique identifier for each object can be used as input to the pipeline. The scientifically calibrated FITS images can be used as for individual studies, as well as used for improved astrometry in updating the catalog position values. The multi-band FITS mosaics can also be combined with the calibrated mosaics of different wavelengths into multi-band color images with mosaics from single/double band surveys such as POSS-I and GALEX , if their image sizes are adjusted to match.

The SDSS and POSS-II data product available on the LCDM website may also be useful during the target selection and commissioning stage of new surveying telescope such as the LSST and DECam. With the trend of future telescope having larger focal plane,, large nearby galaxies covered in the RC3 catalog needs to be properly masked to prevent the loss of imaging details from saturated CCDs. The updated RC3 coordinates tells us which regions of the sky may be affected by these large galaxies, and the FITS files can be used to model-fit the galaxy's shape and light distribution. This information can also be used in selecting fiber location on spectroscopic targets that belong to the RC3 catalog. The FITS images can be used along with existing tools for scientific analysis, such as Astrometry.net(Lang et al., 2010), SExtractor, APLpy(Robitaille and Bressert, 2012). The program is designed such that other APIs such as HEASARC's SkyView, Aladin, and NED can easily be latched on.

We have provided documentation on GitHub that guides investigators through the steps to adapting the pipeline for future imaging data.The source code and documentation for the pipeline described in this paper can be found in the project repository: `https://github.com/ProfessorBrunner/rc3-sdss`. The data product from the SDSS and POSS-II run is available via a search form at `http://lcdm.astro.illinois.edu/search.html`.

Acknowledgements

The work was supported by Google Summer of Code Program. We thank Harold G. Corwin, Jr. for helpful discussion that helped this work. This research made use of Montage, funded by the National Aeronautics and Space Administration's Earth Science Technology Office, Computation Technologies

Project, under Cooperative Agreement Number NCC5-626 between NASA and the California Institute of Technology. Montage is maintained by the NASA/IPAC Infrared Science Archive. This research made use of Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration, 2013).

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is http://www.sdss3.org/. SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, University of Cambridge, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofisica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University

References

References

Alexov, A., Berriman, G. B., Chiu, N.-M., Good, J. C., Jarrett, T. H., Kong, M., Laity, A. C., Monkewitz, S. M., Tahir-Kheli, N. D., Norton, S. W., Zhang, A., Dec. 2005. The NASA/IPAC Infrared Science Archive (IRSA): The Demo. In: Shopbell, P., Britton, M., Ebert, R. (Eds.), Astronomical Data Analysis Software and Systems XIV. Vol. 347 of Astronomical Society of the Pacific Conference Series. p. 7.

Baillard, A., Bertin, E., de Lapparent, V., Fouqué, P., Arnouts, S., Mellier, Y., Pelló, R., Leborgne, J.-F., Prugniel, P., Makarov, D., Makarova, L., McCracken, H. J., Bijaoui, A., Tasca, L., Aug. 2011. The EFIGI catalogue of 4458 nearby galaxies with detailed morphology 532, A74.

Bertin, E., Sep. 2012. Displaying Digital Deep Sky Images. In: Ballester, P., Egret, D., Lorente, N. P. F. (Eds.), Astronomical Data Analysis Software and Systems XXI. Vol. 461 of Astronomical Society of the Pacific Conference Series. p. 263.

Bertin, E., Arnouts, S., Jun. 1996. SExtractor: Software for source extraction. Astronomy and Astrophysics Supplement 117, 393–404.

Blanton, M. R., Schlegel, D. J., Strauss, M. A., Brinkmann, J., Finkbeiner, D., Fukugita, M., Gunn, J. E., Hogg, D. W., Ivezić, Ž., Knapp, G. R., Lupton, R. H., Munn, J. A., Schneider, D. P., Tegmark, M., Zehavi, I., Jun. 2005. New York University Value-Added Galaxy Catalog: A Galaxy Catalog Based on New Public Surveys. The Astronomical Journal 129, 2562–2578.

de Vaucouleurs, G., de Vaucouleurs, A., Corwin, H. G., Buta, R. J., Paturel, G., Fouque, P., Feb. 1995. Third Reference Cat. of Bright Galaxies (RC3) (de Vaucouleurs+ 1991). VizieR Online Data Catalog 7155, 0.

de Vaucouleurs, G., de Vaucouleurs, A., Corwin, Jr., H. G., Buta, R. J., Paturel, G., Fouqué, P., 1991. Third Reference Catalogue of Bright Galaxies. Volume I: Explanations and references. Volume II: Data for galaxies between 0^h and 12^h. Volume III: Data for galaxies between 12^h and 24^h.

de Vaucouleurs, G., de Vaucouleurs, A., Corwin, J. R., 1976. Second reference catalogue of bright galaxies. In: Second reference catalogue of bright galaxies, 1976, Austin: University of Texas Press. p. 0.

Gal, R. R., de Carvalho, R. R., Odewahn, S. C., Djorgovski, S. G., Mahabal, A., Brunner, R. J., Lopes, P. A. A., Dec. 2004. The Digitized Second Palomar Observatory Sky Survey (DPOSS). II. Photometric Calibration. The Astronomical Journal 128, 3082–3091.

Jacob, J. C., Katz, D. S., Berriman, G. B., Good, J., Laity, A. C., Deelman, E., Kesselman, C., Singh, G., Su, M.-H., Prince, T. A., Williams, R., May 2010. Montage: a grid portal and software toolkit for science-grade astronomical image mosaicking. ArXiv e-prints.

Lang, D., Hogg, D. W., Mierle, K., Blanton, M., Roweis, S., May 2010. Astrometry.net: Blind Astrometric Calibration of Arbitrary Astronomical Images. The Astronomical Journal, Volume 139, Issue 5, pp. 1782-1800 (2010) 139, 1782–1800.

Lupton, R., Gunn, J. E., Ivezić, Z., Knapp, G. R., Kent, S., 2001. The SDSS Imaging Pipelines. In: Harnden, Jr., F. R., Primini, F. A., Payne, H. E. (Eds.), Astronomical Data Analysis Software and Systems X. Vol. 238 of Astronomical Society of the Pacific Conference Series. p. 269.

Robitaille, T., Bressert, E., Aug. 2012. APLpy: Astronomical Plotting Library in Python. Astrophysics Source Code Library.

Skrutskie, M. F., Cutri, R. M., Stiening, R., Weinberg, M. D., Schneider, Feb. 2006. The Two Micron All Sky Survey (2MASS). The Astronomical Journal 131, 1163–1183.

Stoughton, C. Lupton, R. H., Bernardi, M. e. a., Jan. 2002. Sloan Digital Sky Survey: Early Data Release 123, 485–548.

Walton, D. J., Roberts, T. P., Mateos, S., Heard, V., Sep. 2011. 2XMM ultraluminous X-ray source candidates in nearby galaxies. MNRAS 416, 1844–1861.