

Creating updated coordinates and scientifically-calibrated mosaic images for the RC3 Catalogue

Jung Lin Lee^a, Robert J. Brunner^{b,*}

^a*Astronomy Department, University of California at Berkeley, Berkeley, CA 94720-3411*

^b*Department of Astronomy, University of Illinois, Urbana, IL 61801*

Abstract

The Third Reference Catalogue of Bright Galaxies (RC3) is a reasonably complete listing of 23,011 nearby, large, bright galaxies. We introduce a software pipeline that updates the astrometric positions for an arbitrary catalog and survey imaging data set. Due to positional inaccuracies inherent to the RC3 catalog, we designed a recursive algorithm that first determines the correct location for each galaxy, and subsequently applies the mosaicking procedure. We publicly release data products generated by the pipeline based on the final imaging data release from Sloan Digital Sky Survey (SDSS) DR10 data and photographic plate data taken by the Second Palomar Observatory Sky Survey (POSS-II). These collections of multi-band, science-grade FITS mosaics, color-composite mosaics, and their most up-to-date coordinates can be used for individual or group studies on these RC3 galaxies. Additionally, our open-source software pipeline enable professional astronomers to apply this to their own user-defined catalogs or other pre-existing, possibly outdated, catalogs and easily extend the pipeline for use with imaging data from other surveys.

Keywords: Image mosaics, RC3 catalog, Sloan Digital Sky Survey, Digitized Sky Survey, Montage, SExtractor

1. Introduction

Astronomers have a long history of cataloguing objects for subsequent study, for example the Messier catalog (Messier, 1781) and the New general Catalog (NGC; Dreyer, 1888) 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; York et al., 2000) have uniformly surveyed large fractions of the entire sky, providing detailed photometric and astrometric information for millions of sources. However, the software pipelines that process these valuable data are optimized for the more numerous small sources. As a result, large, nearby, and bright galaxies are essentially treated as

contaminants. Yet these galaxies remain incredibly important, providing detailed insight into the dynamics of galaxies, as well as a low redshift 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; de Vaucouleurs et al., 1991), which is a catalog containing 23,011 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 selected galaxies that may only meet one or two of these conditions as well as some nearby compact galaxies. Given the efficacy of this catalog, previous authors have used this sample as the basis for making image mosaics for a large sample of galaxies.

The first such effort was made by Hogg and Blanton (2006), who made color-composite images of se-

*Principal Corresponding author

Email addresses: dorislee@berkeley.edu (Jung Lin Lee), bigdog@illinois.edu (Robert J. Brunner)

lect RC3 galaxies by using the SDSS g , r , and i band images from the sixth SDSS data release. A subsequent effort by Baillard et al. (2011), 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 separate effort, known as the NASA-Sloan Atlas¹, has been undertaken to construct and analyze a complete set of galaxies within approximately 200 Megaparsecs (Blanton et al., 2011), which is the approximate redshift limit of the RC3 catalog.

Given the SDSS project has published their last imaging data release (DR10; Ahn et al., 2014), which includes 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 DR10 footprint.

Our software pipeline is not just an easy and convenient wrapper around Montage and STIFF but is necessary for successfully mosaicing outdated catalog sources. It is distinguished by its unique relationship of how the positional update algorithm informs the mosaicking process and vice versa. Without the updating algorithm, standard mosaicking procedure will not be able to find and center the mosaic around the source of interest and will instead mosaic an “empty” piece of the sky. Therefore, as a by-product of this mosaicking process, we also generate updated coordinate positions for RC3 galaxies with inaccurate catalog coordinates. Despite the diverse set of studies that incorporates the RC3 catalog mentioned earlier, our study is the only one that uniformly updates the astrometry of all the sources in the RC3 catalog, providing the most accurate astrometry for the RC3 catalog to date.

We further generalize this software framework for use on any arbitrary user-defined catalog using imaging data from any scientifically-calibrated sky survey. We test this new pipeline by also constructing FITS image mosaics and color-composite images for all RC3 galaxies that lie within the Second Palomar Observatory Sky Survey (POSS-II; Reid et al.,

1991). Our pipeline can be easily applied to existing data such as the Two Micron All Sky Survey (2MASS; Skrutskie et al., 2006), or to future surveys such as the Dark Energy Survey (DES; The Dark Energy Survey Collaboration, 2005) or the Large Synoptic Survey Telescope (LSST; Ivezić et al., 2008).

In Section 2, we introduce the RC3 catalog and the SDSS and POSS-II data that we use to generate 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 pipeline performance and evaluate the science-quality for the pipeline results for the SDSS and POSS-II data in Section 4, before concluding the paper and discussing the overall project in Section 5.

2. Data

To construct large, calibrated image mosaics, we need two types of data. First, we need a catalog of galaxies; and second, we need an image data set. A number of different candidate galaxy catalogs exist; we also could follow the example used to develop the NASA-Sloan Atlas and construct a new catalog based on specific physical criteria. We choose, however, to use the RC3 catalog of galaxies (de Vaucouleurs et al., 1995), which we detail in § 2.1. For our imaging data, we actually selected two different surveys: the SDSS and the POSS-II. The SDSS uniformly surveyed a large fraction of the sky with digital detectors, while the POSS-II is an older photographic plate survey that covers nearly the entire sky. We discuss these two imaging surveys in detail at the end of this section. While not directly discussed in this paper, our pipeline approach could easily be extended to other datasets, such as the Two Micron Sky Survey (Skrutskie et al., 2006), the Wide-field Infrared Survey Explorer (Wright et al., 2010), and the Galaxy Evolution Explorer (Martin et al., 2005).

2.1. The RC3 Catalog

The simplest approach to constructing a uniform galaxy catalog is to define a sample by limiting apparent brightness within a survey. Original attempts to accomplish this task date back to the original Harvard Survey of External Galaxies (Shapley and Ames,

¹<http://www.nsatlas.org>

1932), which contained 1,249 objects brighter than 13th 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 (de Vaucouleurs et al., 1976). The original RC3 galaxy catalog compiled by de Vaucouleurs et al. (1995), contains a complete listing of 23,011 galaxies with D_{25} apparent major isophotal diameter greater than one arcminute and with a total B-band magnitude greater than 15.5th magnitude.

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. An update was published a few years after the original RC3 catalog by Corwin et al. (1994). In this project, we use the RC3 catalogue information available through the VizieR Service² as our starting conditions, which contains the updated RC3 data published in 1994. The RC3 data from NASA Extragalactic Database (NED) were last updated from the 1993 version 3.9b of the RC3 catalog. Until 2011, it was incrementally updated with new, published observations and subsequently maintained by Harold G. Corwin³. However, the RC3 positions in NED are non-uniformly updated as only the few RC3 sources that coincidentally overlaps with other newer source catalogs were updated in NED.⁴

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 have been used in astrophysical studies of quasars and X-ray sources (e.g., Walton et al., 2011), and for galaxy morphology and clustering studies (e.g., Best et al., 1996; Knapen et al., 2000). 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). However, despite different levels of precision due to the non-uniform astromet-

	SDSS	POSS-II
Imaging bands	u, g, r, i, z	R_F, B_J, I_N
Sky Coverage (%)	35.28	78.27
Resolution ("/pix)	0.396	1.7
Imaging Technique	CCDs	Photographic Plates

Table 1: A summary of the two surveys processed in this work.

ric updates, many survey catalogs and studies have taken the NED RC3 positions to be the de-facto positions of the RC3. In this project, we present a complete updated set of all the RC3 sources using the DPOSS-II survey data.

2.1.1. Updating the RC3 Astrometry

As described in §3, we use the montage toolkit to reproject and mosaic individual survey FITS images. Initially when we followed the standard mosaicking steps in montage, we obtained many mosaics with off-centered or missing RC3 galaxies. After further review, we realized that this effect was due to the inherent positional inaccuracy in the catalog, which, given the size of a typical RC3 galaxy, was surprising. On further reflection, however, this effect can be understood by the heterogeneous origin and updating of the RC3 catalog.

The entries in the RC3 catalog hosted by VizieR are astrometrically tied to the B2000.0 FK4 reference system. Due to the heterogeneous positional updates, the RC3 galaxies are denoted with two different levels of accuracy: HH MM SS.s, DD MM SS for positions that have been updated to an accuracy of approximately 5–8 arcseconds, and HH MM.m, DD MM for galaxies whose positional accuracy remains at approximately 1–2 arcminutes as presented in the original catalog. The 1991 version of the RC3 catalog (de Vaucouleurs et al., 1991) contains 5,492 galaxies that fall in the latter group. As a result, we developed an iterative mosaicking step within our pipeline, as discussed in §3.1.1, to compute updated astrometric coordinates for those galaxies with poor astrometry so that we would have the target galaxy centered on the final mosaic.

2.2. SDSS

The SDSS image data is acquired from the 2.5-m telescope (Gunn et al., 2006) at the Apache Point Observatory in New Mexico. For our mosaics, we

²<http://vizier.u-strasbg.fr/viz-bin/VizieR?source=VII/155>

³<http://haroldcorwin.net/rc3/bugs.rc3>

⁴For example, PGC 2557 was updated with Chandra source catalog (Evans et al., 2010).

use the imaging data from the SDSS Data Release 10. The SDSS imaging camera (Rockosi et al., 2002) was one of the first, large imaging cameras that leveraged Charge Coupled Devices (CCDs) as photon recording instruments. CCD detectors have obvious throughput advantages over photographic plates and also provide a near-linear response to input signals. In addition, the pixel resolution and mean seeing at Apache Point combined to provide a reasonable angular resolution, especially for large, bright galaxies.

The SDSS imaging data is processed by the photo pipeline (Lupton et al., 2002), which is responsible for a number of different image processing procedures, including source detection, deblending, model-fitting, and astrometric and photometric calibration. We quantify the imaging quality of the SDSS data we use by the `clean` flag⁵, which is extracted from the SDSS SkyServer. For extended sources like a galaxy, the `clean` flag is defined from data masks with variables that quantify imaging quality metrics such as the PSF magnitude error, cosmic rays, or undefined profiles from model fits. We extract parameter information by querying the SDSS SkyServer via the SDSS Command Line Query Tool, which retrieves the calibrated, sky-subtracted corrected frames with the calibration meta-data (`fpC`) in bulk from the Science Archive Server.

2.3. POSS-II

The Second Palomar Observatory Sky Survey (POSS-II) was a photographic survey that covered most of the night sky, including the northern region later covered by the SDSS. POSS-II was an improved update to the original National Geographic Society-Palomar Observatory Sky Survey (NGS-POSS or POSS-I; Minkowski and Abell, 1963a), which was one of the first, large area photographic surveys. We acquire FITS images for all three plate types: B_J , R_F , and I_N as required to construct mosaic images from the Space Telescope Science Institute’s Digitized Sky Survey project (DSS; Lasker et al., 1996). These photographic plates were calibrated to the Gunn g , r , i bands by a separate CCD observing campaign, producing the Digitized Palomar Sky Survey data (DPOSS; Gal et al., 2004).

Our software pipeline obtains photographic plate data from POSS-I, POSS-II and UK Schmidt Telescope Survey by using the NASA/IPAC Infrared Science Archive’s (IRSA) Finder Program Interface⁶. Each photographic plate covers $6.5^\circ \times 6.5^\circ$ of the sky with centers spaced approximately 5° apart (thus providing significant overlap between adjacent plates). After lossy compression, each plate is approximately 1.1 GB, although a crop-out of the plate can be retrieved to reduce the download time. Due to the large size of these plates, our pipeline rarely needs to stitch multiple, adjacent fields; instead we simply use the full imaging data to complete the positional update algorithm.

Even though POSS-II has the advantage of a greater sky coverage than SDSS, it suffers from several disadvantages that do impact our photographic-plate based mosaics. First, photographic plates are much less sensitive to incoming photons ($\sim 1\%$ quantum efficiency) as compared to CCDs ($\sim 80\%$), thus the photographic-based mosaics have a reduced clarity. Second, objects detected near the edges of photographic plates suffers can suffer from vignetting, which may result in inaccurate astrometric calibrations being applied to those objects. Finally, photographic plates have a non-linear response to incoming photons, thus the absolute photometric calibration of photographic plates is less accurate, which will affect multi-plate mosaics.

3. Mosaic Pipeline and algorithms

Given the size of the RC3 catalog, we decided to construct a pipeline to automate the construction of the FITS mosaic images and the color-composite images for each RC3 galaxy contained within a given imaging survey like the SDSS. This pipeline is presented as a flowchart in Figure 1, and is implemented in the Python programming language.

Broadly speaking, our pipeline leverages two open-source tools: montage (Jacob et al., 2010) and STIFF (Bertin, 2012). Montage is used to reproject each input image appropriately, and to combine the reprojected images into the final mosaic FITS image. STIFF, on the

⁵<http://skyserver.sdss3.org/dr10/en/help/cooking/general/flags6.aspx>

⁶<http://irsa.ipac.caltech.edu/applications/FinderChart/docs/finderProgramInterface.html>

other hand, is used to make the color-composite images from the mosaicked FITS images generated by montage. While montage can make color-composite images, we found that STIFF was easier to use within our pipelined approach since STIFF automatically determines the relevant color-mapping.

In the rest of this section, we first detail several new software tools we developed to generate more accurate FITS image mosaics for each RC3 galaxy. Next, we provide a more detailed discussion of the specific software components, including specific classes we developed to generalize our approach to support different galaxy catalogs and different photometric surveys. We present a flowchart that highlights the operational steps of our pipeline in Figure 1.

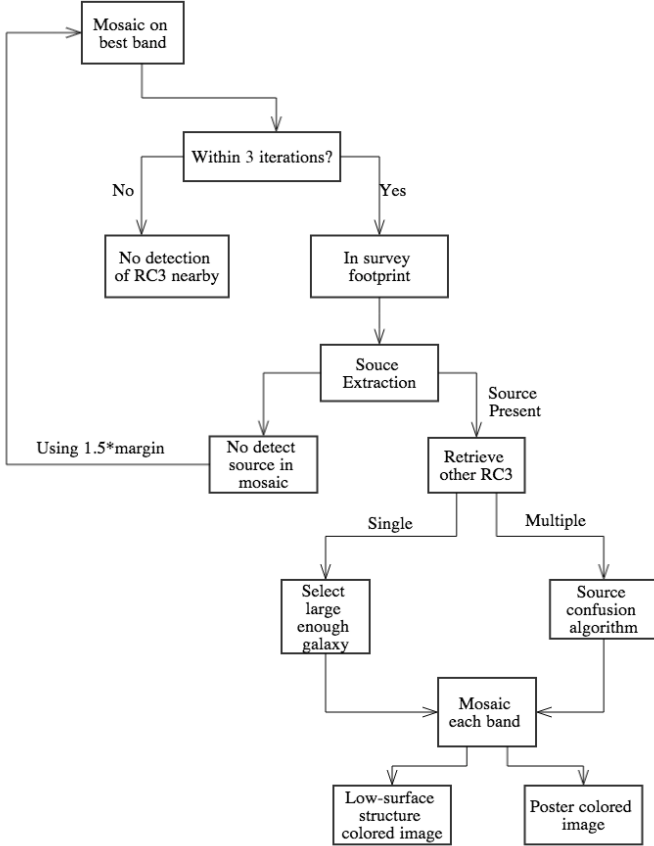


Figure 1: A flowchart detailing the basic steps employed by our pipeline to first construct calibrated FITS image mosaics for each photometric band in an imaging survey of the target galaxy, followed by the construction of color-composite images for the target galaxy.

3.1. Algorithm Enhancements

The Montage software can be accessed via the Virtual Astrophysical Observatory (VAO) Image Mosaic Service at the Infrared Science Archive (IRSA) website to robustly generate mosaic images from multiple surveys, including from both the SDSS and DSS. Given the nature of the RC3 catalog, however, we had to overcome two challenges when including montage as part of our overall mosaicking pipeline. First, the relatively large positional uncertainty present for some galaxies within the RC3 catalog meant a mosaic could be automatically generated by using montage that either had the target galaxy off-center, or worse entirely missing from the mosaic. Second, since galaxies are often in groups or clusters, a positional uncertainty can lead to a mis-identification of the target galaxy within the constructed mosaic. As a result, two components are developed within our pipeline to address these two issues in order to generate a mosaic image centered on the appropriate galaxy.

3.1.1. Positional Update

We created a class, called RC3, to encapsulate the data and methods relevant to the RC3 sources. Each RC3 instance contains the catalog RC3 coordinate information, any updated coordinate information, radius, and a unique identifier. While we could have adopted a standard numerical identifier scheme, we instead chose to adopt each galaxy’s Catalogue of Principal Galaxies (PGC) number, since this is a unique identifier for each RC3 galaxy that is also already present in the original RC3 catalog. This unique name provides unambiguity for our pipeline, and is also used for naming the generated data products and serving these same products via our website.

To overcome the problem of inaccurate positions discussed in §2.1.1, we developed an approach that finds the target galaxy in the imaging data and uses this position to update the astrometric information appropriately. This algorithm is implemented in the `source_info` method of the RC3 class, and first generates a mosaic with a field of view roughly six times larger than the radius of the target galaxy. Next, SExtractor is used to detect all sources on the newly generated mosaic. The selected mosaic size is generally sufficient to ensure that SExtractor will be able

to determine an accurate background sky level. Of all detected sources, only those with a radius greater than 5.94 arcseconds are retained. This size was empirically selected to eliminate most stellar sources and background noise spikes while still retaining the subset of RC2 galaxies contained in the RC3 catalog that are smaller than one arcminute as described by de Vaucouleurs et al., 1976. Montage has a built-in set of modules that rectifies the background and subtracts the overlapping, neighboring fields in the raw input image tiles by least square fitting solution (Jacob et al., 2010). Since the sole purpose of these initial r band mosaics are to improve the astrometry, we do not rectify the background in these mosaics in order to speed the computation by minimizing data I/O.

At this point, our pipeline is in one of three states depending on whether there are zero, one, or more RC3 candidates detected by SExtractor on the new mosaic. First, if there are zero candidate RC3 galaxies in the new mosaic, we recursively create larger mosaics until either an RC3 candidate galaxy is detected or we exit after three iterations and declare the RC3 source not found. If only one candidate galaxy is detected, the pipeline proceeds directly to the final image generation step. If multiple large galaxies are detected, the pipeline executes the source confusion algorithm as outlined in §3.1.2. At the end of this process if at least one candidate RC3 galaxy is detected, then we generate mosaic images in all bands covered by the survey with Montage’s background rectification procedures turned back on.

Once the mosaic images have been constructed in all bands for the given survey, we select images observed through three different filters and recombine them into a color mosaic image by using STIFF. The three bands selected are the g , r , and i bands for the SDSS, which has five bands, and the B_J , R_F , and I_N bands for the DSS survey, which only has these three bands. At this point, two color images are generated (although this number can be altered if necessary). The first color image generated emphasize low surface brightness structures, such as the halo around a galaxy, or interaction tidal streams. The second color image generated is a poster or publication quality image that uses higher background cuts to ensure a clean, highly contrasted image. We note that to



Figure 2: A visual demonstration, using SDSS image data, of our source confusion algorithm, see Figure 3 for a wider angle view of this same field. (Left) A color-composite image made from mosaicked g , r , and i FITS images. (Right) The regenerated color-composite image made from the original data after the source confusion algorithm has been applied, resulting in a recentering on the target galaxy PGC58.

construct a color image, our pipeline does require at least three bands. Fortunately, most photometric surveys meet this criterion, which is not surprising since the construction of a color-color plot, used to minimize the effects of extinction and identify stellar or extragalactic source populations, require at least three bands (see, e.g., Skrutskie et al., 2006). Since the sensitivity of each image colored filters is survey-dependent, users who wish to extend the pipeline to mosaic images from another survey must empirically test the parameters used in STIFF by visual comparison.

3.1.2. Source Confusion

The second challenge our pipeline was forced to tackle was the source confusion problem. Since large galaxies are often physically located near other galaxies, any astrometric error in the location of our target RC3 galaxy could result in source confusion on a generated mosaic (see Figure 3 for a visual demonstration). To tackle this challenge, we first assumed that any galaxy large enough to cause confusion would also be present in the original RC3 catalog. This assumption is supported by the fact that the RC3 catalog is reasonably complete for galaxies with apparent diameters larger than one arcminute.

The algorithm we developed to overcome this challenge first identifies all RC3 galaxies that may lie within the newly generated mosaic image, and second matches this list to the list of galaxies that were actually detected by SExtractor within the mosaic

image. The list of RC3 galaxies is generated by the `otherRC3` method that is part of the `Server` class described in § 3.2.2. By default, this method queries the VizieR Catalog to obtain this information. Alternatively, this method can be overridden if a survey provides this information directly, like the SDSS, which contains an RC3 table in its SkyServer database.

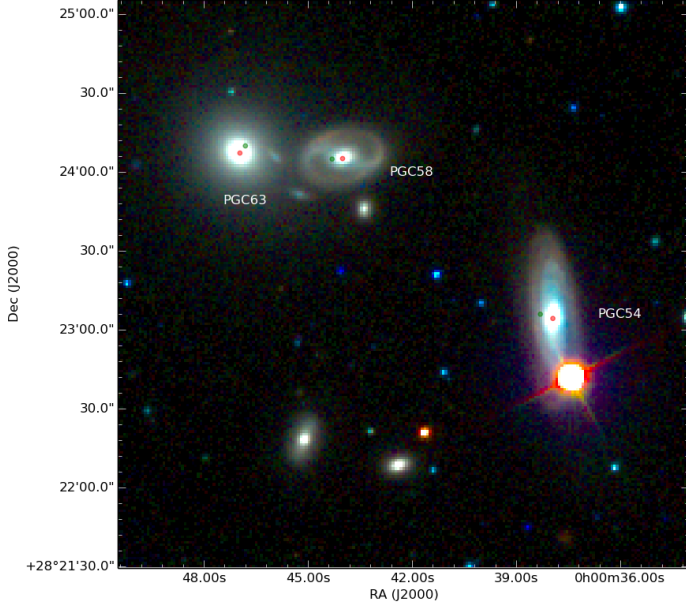


Figure 3: A visual demonstration of the results of the source confusion algorithm for three RC3 Galaxies showing the original RC3 catalog galaxy coordinates (green markers) and our newly calculated coordinates (red markers).

Naively, this cross-match process might appear to be simple; however, since each of these RC3 galaxies may have inaccurate astrometry, the problem becomes non-trivial. To tackle this challenge, we make the reasonable assumption that any positional inaccuracies are due to instrumental or measurement errors that might affect a galaxy’s absolute position but not the relative positions of each galaxy. Thus we identify the n RC3 galaxies that potentially lie within the mosaic image from VizieR, and compute all possible relative positional differences between these candidate galaxies. Next, we compare this list with the positional differences computed between the galaxies detected in the actual image by SExtractor. From this list of cross-matched differences, we identify each RC3 galaxy in the image, correct the overall astrometry for each detected RC3 galaxy, and regenerate a new mosaic centered on the correct loca-

tion). We have verified the accuracy of these assumptions by using the SDSS, where find a success rate of 99.97%. Furthermore, we have found that this approach correctly resolves up to five RC3 sources in a single mosaic.

3.2. Pipeline

To simplify the application of our approach to other catalogs and/or surveys, we developed an automated pipeline to generate the FITS image mosaics and color-composite images. This pipeline is built by using classes that encapsulate specific data that might be relative to a certain catalog, like the RC3, or to a particular survey, like the SDSS. The overall class relationship is presented in Figure 4, which highlights three abstract classes: `Survey`, `Catalog`, and `Server`, which simplify the incorporation of new catalog data or surveys into our pipeline. In the following subsections, we discuss several of these classes in more detail.

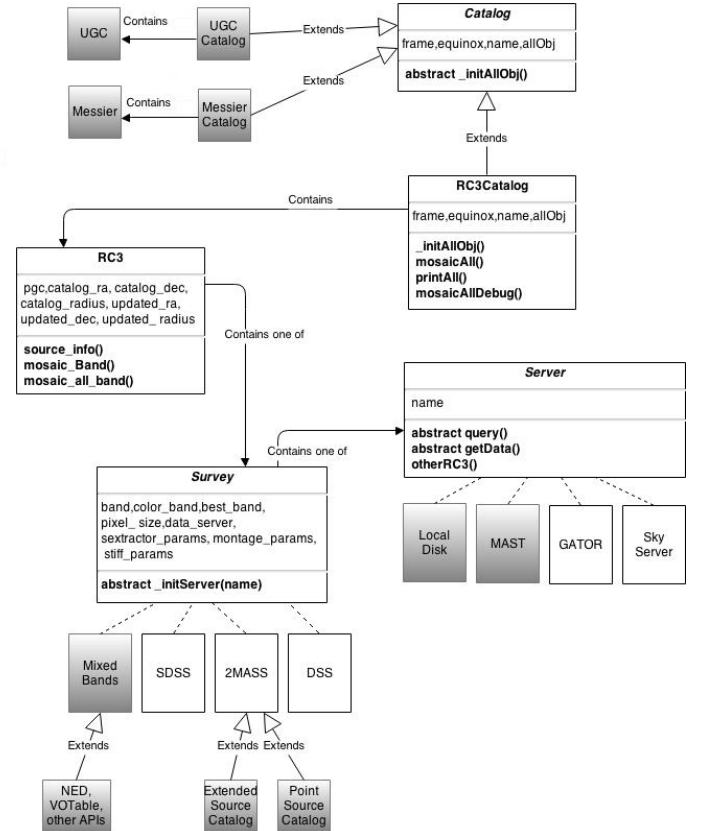


Figure 4: A unified modeling language (UML) class diagram showing the relationships between different classes in the mosaicking pipeline. Grey-filled boxes shows possible extensions of this pipeline.

3.2.1. Survey

The Survey class provides a convenient way of extending the mosaicking pipeline to take inputs from other survey. Additional Survey classes enables the users to obtain multi-band or higher resolution images for the same source object in a catalog. In addition, extending the pipeline to an all-sky survey (such as POSS-II) increases the coverage of sources in the catalog, which leads to a more complete set of updated astrometry. As shown in §4.4.2, our pipeline preserves the photometry of the input images. Therefore, in order to get science-grade, calibrated output mosaics, the survey's input images must be photometrically calibrated. In other words, our pipeline does not perform any additional calibration to the raw images. As this pipeline is intended for use on large, general, sky surveys, the task of calibrating images and tuning telescope-specific parameters are abstracted away by dedicated photometric pipeline used to generated the data products of most sky surveys.

3.2.2. Server

The primary abstract class used by the pipeline is the Server class, which encapsulates the two main tasks of data acquisition: querying imaging data and retrieving the imaging data from the actual server. Many recent surveys provide an application programming interface (API) that enables data access either by using SQL or a customized query mechanism. To actually implement a subclass of the Server class, a mapping between the positional values of galaxies in a particular catalog and the recorded image data for the particular survey must be established. For example, SDSS image frames are uniquely identified by a particular combination of *run*, *camcol*, and *field*, while 2MASS identifies images with a sexagesimal, equatorial position-based source name. For those surveys where this type of mapping has not already been established, such as the imaging data from the POSS-II survey, a subclass can establish a new naming scheme.

In addition to these primary tasks, each Server subclass also must implement functionality to build and execute queries as required by the montage mosaic software. By using a Server class as opposed to a DataObject class enables code reusability across various surveys that can be accessed via common

server tools. For example, this approach allows the pipeline to easily use the IRSA GATOR query service (Alexov et al., 2005) and Astroquery⁷.

3.2.3. Catalog Objects

The current version of the pipeline, which solely generates mosaic images of RC3 galaxies, does not contain an abstract CatalogObject class. As a future extension, this would be a beneficial addition as it would be helpful to have a generic class that provides a similar functionality as the RC3Objects class. These functions contain basic information about the particular object and are survey-independent. They also perform the essential mosaicking features on a per-object basis. Therefore, these functions not only can be used for comparing resulting mosaics from multiple surveys (see, e.g., §5), but these methods can also be conveniently used within the Catalog class.

The final step in the mosaic procedure generates two TIFF color images as described in §3.1.1. When extending the pipeline to other surveys, prototyping work should be done to determine the optimal STIFF parameters in order to generate color-composite images that best capture the details in the telescope-specific imaging by following the guidelines in Bertin (2012).

3.2.4. Catalog

The simplest abstract class contained in the pipeline is the Catalog class, which simply contains a list of objects in a particular catalog. While the use of such an abstract container class might seem superfluous, by using this class, the pipeline is able to cleanly separate the basic mosaic functionality for individual galaxies from the functionality required for an entire catalog. This capability enables a direct study of a single object, enables a processing of all sources listed in the derived Catalog class, or simplifies the debugging process. By using this abstract barrier, we ensure minimal changes are required in the class structure if the pipeline is modified to support imaging data from a new survey.



Figure 5: PGC120 mosaic from POSS-II (top) and SDSS (bottom). The two images are of different scale.

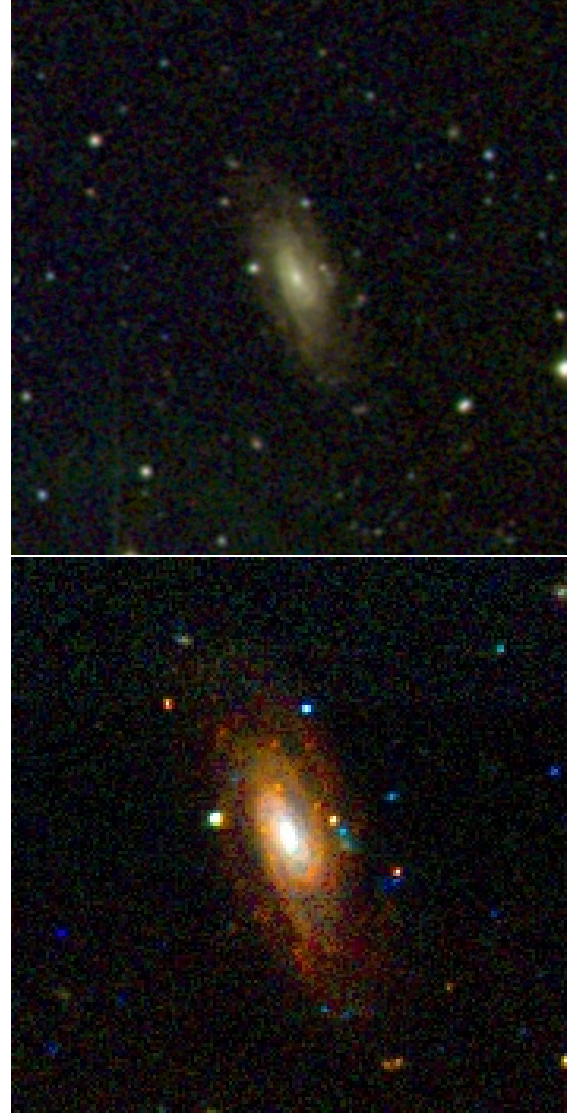


Figure 6: PGC1746 mosaic from POSS-II (top) and SDSS (bottom). The two images are of different scale.

4. Results

4.1. Mosaicking Results

Of the total 23,011 galaxies described in the RC3 catalog, 12,418 lie within the SDSS DR10 footprint. Our automated pipeline successfully generated mosaics for 76.58% of these galaxies. Of these, 4,283 RC3 galaxies were mosaicked by using an updated position that was more than one arcminute away from the recorded RC3 astrometric position. On an 8-core linux server, the pipeline averaged 80 RC3 galaxy mosaics per hour for the SDSS data, which was re-

trieved from a remote SDSS data repository. The final data products, the five band FITS mosaic images and two color composite images for each galaxy, occupy approximately twenty GBs.

The POSS-II data has full sky coverage, thus all 23,011 RC3 galaxies are covered by the POSS-II footprint. Of this full sample, the automated pipeline was successful for all RC3 galaxies, except for 8,431 galaxies that produced a STIFF error (see § ?? for a discussion of this condition). For the POSS-II mosaics, 3,431 RC3 galaxies were mosaicked by using an updated position that was more than one arcminute away from the recorded RC3 astrometric position. Given the larger base images for the POSS-II

⁷<http://astroquery.readthedocs.org/>

survey (~ 1.1 GB each), it is not surprising that our pipeline was slower when creating POSS-II mosaics, averaging about fifty galaxies per hour on the same machine used for the SDSS mosaicking. The finished data products for the POSS-II sample of galaxies occupies approximately 10 GBs.

4.2. Pipeline Performance

The majority of the processing time for our pipeline is in transferring the raw FITS image data from the survey data site to the processing site. However, we have explored techniques to improve the overall performance of the pipeline. First, we accelerate the mosaicking process by performing the positional update on only a single band from the bands available from a given survey. The target band is designated as the `best_band`, and given the results from this single band, the other image FITs mosaics are performed only once per object. For example, for the SDSS we use r-band images, since the r-band filter since transmission has the highest quantum efficiency (Stoughton and Bernardi, 2002).

Since data transfer dominates the overall processing time, we do not employ traditional parallel programming techniques (we could, however, employ embarrassingly parallel techniques since mosaic images are constructed independently). Thus, even though Montage’s modular design enables its performance to scale with the number of processors (Jacob et al., 2010), our pipeline would not be accelerated by using Message Passing Interface (MPI). Other factors that affect the runtime include the sky coverage of a particular survey and the response speed of query from a survey archive.

One technique that could be employed to accelerate our pipeline would be local image caching, or alternatively the capability of executing our pipeline within a survey archive. We have designed our pipeline by using a class hierarchy that enables subclassing of the `Server` class to specify the location of a survey’s raw FITS images. However, downloading an entire survey’s imaging data set for this purpose is unlikely to be beneficial (if done solely for this task), since that would likely take significantly more time than simply downloading the required input raw FITs imaged required for the mosaics when running the pipeline.

4.3. Technical Details

Due to the recent rise in the popularity of the Python programming language, we developed our pipeline by using Python 2.7.6. As a result, our pipeline’s dependencies are widely supported, which simplifies the extension of our work to future datasets. The majority of the mosaicking actions are done by calls to the Montage API along with the AstroPy Montage wrapper⁸ developed by Robitaille (2014). The final color-composite images are generated by using Astromatic’s STIFF v.2.4 (Bertin, 2012).

Our combination of these two programs allows us to make use of their best features. Montage excels at creating scientifically-calibrated images that retains the astrometry and photometry of input sources during the image reprojection. Our choice was also aided by the fact that the efficacy of montage was already demonstrated by the montage developers by using SDSS and POSS-II image data sets (Jacob et al., 2010). On the other hand, STIFF provides the flexibility of adjusting a number of parameters to optimize the appearance of the final color-composite image, and also automatically estimates the upper and lower limits for the dynamic range of the final color image by using statistics derived from a pixel histogram.

Other major steps in our pipeline that required custom code development include query construction, query result processing, source extraction on the mosaicked image, and the development of a web-accessible database to facilitate access to our data products. We use the SDSS Command line query tool written by Budavari (2004) to submit SQL queries to SkyServer and we use Astroquery as a more general archive query tool and to find other RC3 galaxies that lie within a field by using the VizieR catalog database. For IRSA and most web databases, querying consists of building a URL string, submitting the query, and parsing the resulting raw text or XML file returned by `wget`. Source extraction was performed by using SExtractor v.2.19.5 (Bertin and Arnouts, 1996) with standard processing parameters. Our web-accessible database was created by using the Python `sqlite3` module, and the web search interface was written in PHP with interacting HTML elements.

⁸<http://www.astropy.org/montage-wrapper/>

4.4. Science Quality of Output Mosaic

We chose to only conduct this on the SDSS survey and not the DSS since, as discussed in Sec. 2.3, we are only cropping the digitized plate images and not mosaicking these images, the astrometry and photometry must be preserved along with FITS header.

4.4.1. Astrometric Preservation

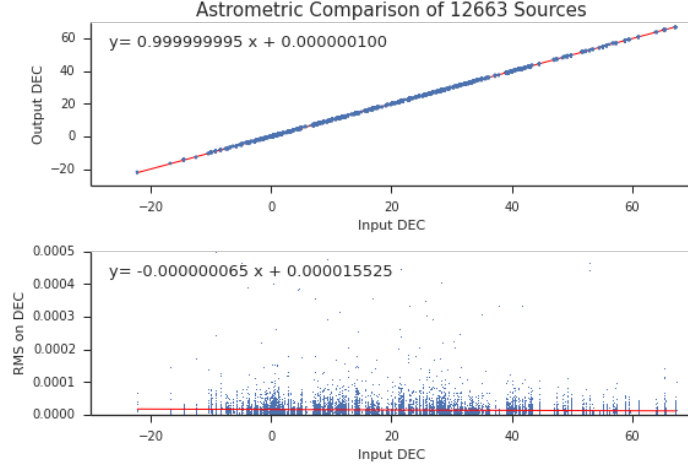


Figure 7: Top: The plots show that astrometry is preserved in the output mosaic images for 12663 detected sources. Bottom: There are no systematic offsets in the residual plots.

To test whether the pipeline introduce errors in the astrometry, we mosaiced a random set of RC3 galaxy with a variety of sizes and uniformly distributed around the sky through the pipeline and match them with the detected sources in the input fields found by SExtractor. Fig.7 shows the one-to-one correlation between the input and output DEC. A similar trend was observed between the input and output RA. Further, we find that there is no systematic offset in the input and output coordinates down to a order of magnitude of 10^{-5} degree. An offset at this level is negligible and probably due to the floating point arithmetic error inherent to SExtractor.

4.4.2. Photometric Preservation

We use Montage’s default for handling image coaddition which takes the area-weighted average of the input fluxes to compute the output pixel intensity. There are no image resampling artifacts introduced as part of this process. Montage does not perform PSF matching across different imaging fields since

that would require telescope-specific detail about the survey. However, since we are using calibrated SDSS data and mosaicking neighboring fields, the PSF variability should not too significant. So we expect that the photometry should be preserved in the input and output images.

We conduct a similar analysis as described in §4.4.1 by comparing the SExtractor-computed values for corrected isophotal magnitudes (MAG ISOCORR). In our analysis, we filter out two types of outliers that appear to have large differences in the input and output magnitude but are in fact artifacts of how SExtractor is used to compute the magnitude values here, as illustrated in Fig. 8.

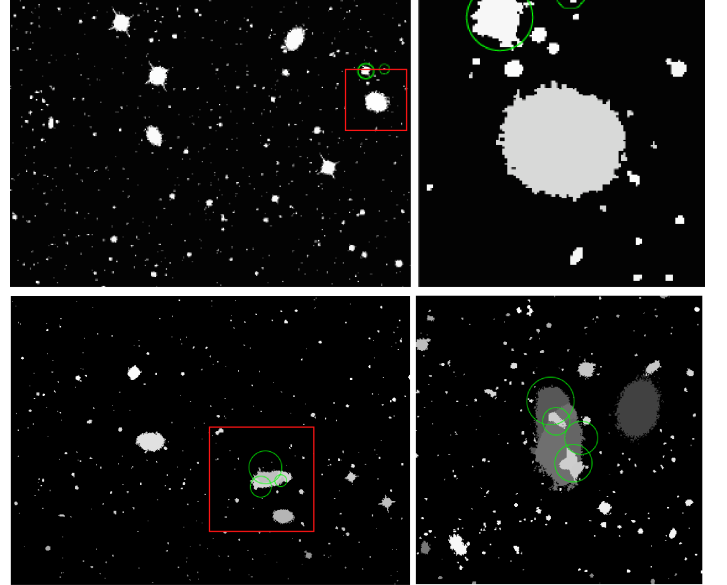


Figure 8: The left panels show the check image from the SDSS input image field. The SExtractor-detected sources are circled in green. The red bounding box shows the boundaries on the check image of the output mosaic in the right panels. Top: PGC 5095 contains two sources that are severely “cut off” in the output image. Bottom: The input and output image shows that the debblending on PGC 1921 is performed differently.

We eliminate objects that lie at the boundary of the image fields since sources that are cut off on the edges will contribute less flux to the SExtractor-detected source, resulting in a large mismatch between the input and output magnitudes. Additionally, it is a well-known problem that SExtractor have trouble with debblending and detection on large, extended sources, such as the RC3 galaxies. One possible reason for this is that even though the SExtractor debblending

thresholds are set as the same, the input image fields are often much larger than the output mosaic, so the different sampling of sky level may lead to a different estimation of the background. Since detection and deblending hinges on the estimated sky level, a single extended source may be broken up differently in the input and output image, resulting in large RMS deviations. Therefore, we design our analysis algorithm to automatically reject wrongly-deblended sources that lies too close to an RC3 source.

In our final analysis of 12663 sources, the filters rejected 427 outliers that lied on the edges of the image and 94 outliers that lied in the proximity of a RC3 galaxy. We are confident that the filter is correctly removing these artifacts rather than actual matched data since we found that after applying the first filter, the number of outlier at bright magnitudes (-6~-10) with RMS from 0.5~1.5 decreased, and applying the second filter decreased the number of outlier at faint magnitudes (-4~0) with RMS from 1.5~3.0, while still preserving correctly-deblended RC3 sources, as expected. As shown in Fig.9, the

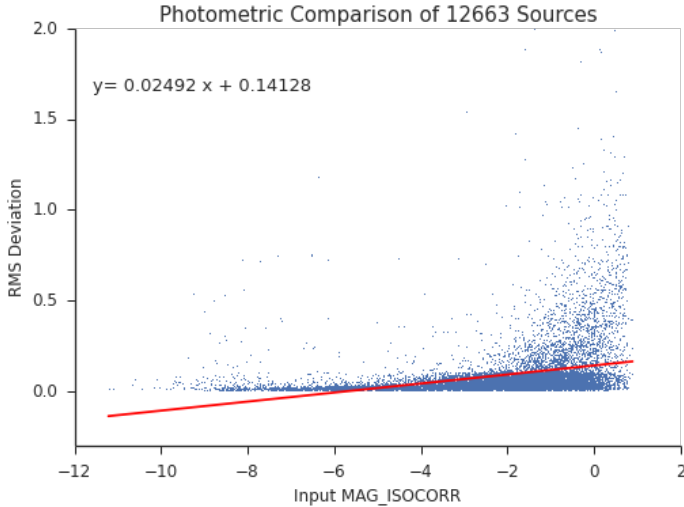


Figure 9: RMS deviations showing differences between the magnitudes of detected sources in input fields and output mosaic.

spread of the RMS corresponds to the well-known trend that the error on the magnitude increases exponentially at faint magnitudes. At low S/N regime, small changes in the background level will be comparable to low values of source flux, so the uncertainty S/N is strongly affected. Since there are more faint sources in each mosaic image than bright sources,

the linear fit on the RMS is positively biased by the faint-magnitude uncertainty resulting in a non zero slope as expected. Possible sources of error includes different background level sampled by SExtractor resulting in different detection threshold on the corresponding image, wrongly matched sources due to cluttered star fields and other badly deblended non-RC3 extended sources. Additionally, the average RMS (~ 0.086) for the whole dataset is comparable to the mean error on the MAG ISOCORR measurement returned by SExtractor (0.088 on the input and 0.058 on the output). Fig. 10 again shows that the deviation for source with $\text{RMS} < 0.25$ can be largely accounted by the error bars inherent to SExtractor's computation, and the number of outliers beyond that is scarce.

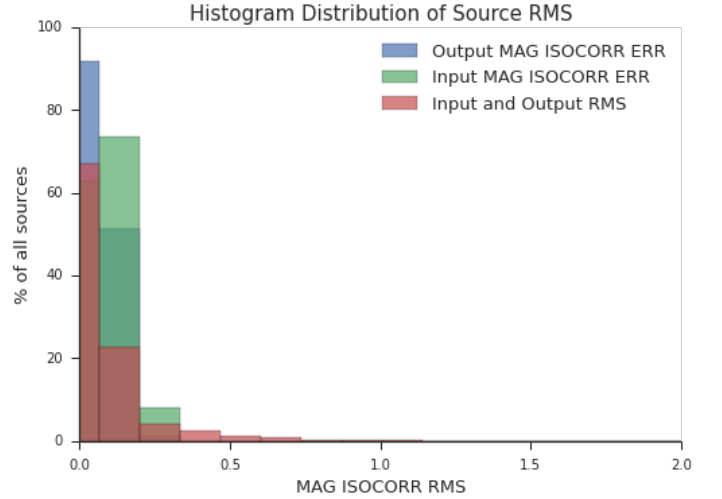


Figure 10: Histogram of RMS shows the distribution of outliers with larger RMS.

5. Conclusion

By design, wide-area sky surveys are used to answer fundamental questions regarding the formation and evolution of large-scale structure and of the cosmological history of the universe. But these data can be used for many other purposes. The mosaicking pipeline described in this paper, for example, provides a convenient way to generate mosaic images for specific sources from an archived set of sky survey images. Furthermore, this pipeline can be easily adapted to work with future data to create scientifically-calibrated FITS mosaics as well as color-composite images. The source code and documentation for the

pipeline described in this paper can be found in the project repository⁹. In addition, we have provided documentation on GitHub that will guide other investigators to adapt the pipeline for alternative imaging data sets.

To ensure that the generated mosaic images are centered on the target source, we implemented an algorithm that automatically determines the correct source astrometry, updates the source catalog appropriately, and generates mosaic images centered on the newly updated coordinate location. Finally, to demonstrate the efficacy of this new pipeline, we generated FITS image mosaics and color-composite images of galaxies in the RC3 catalog by using the SDSS and POSS-II data, along with the most-up-to date set of positional values for all the RC3 galaxies. All of these data products are publicly released and accessible via a searchable web form on the Laboratory for Cosmological Data Mining website¹⁰.

By developing this new pipeline, we can generate image mosaics using newly-obtained data, which will enable a more complete sky coverage for a given source catalog or the construction of potentially higher resolution images or images in other wavelengths. In addition, the pipeline simplifies the extension of this work to either user-defined catalogs or to other published catalogs, such as the Messier Catalog or the New General Catalog. Furthermore, a specific scientific inquiry may require the construction of a user-defined catalog by imposing selection criteria to study certain types of objects. To accomplish this task, a user simply needs to generate a text file containing source positions, source radii, and unique identifier for each source, which can subsequently be used as input to the pipeline.

The scientifically calibrated FITS image mosaics generated by the pipeline for a specific galaxy catalog can be used both for individual source studies, as well as ensemble studies of source populations. In addition, the multi-band FITS mosaic images can be combined with calibrated mosaic images at different wavelengths into new, multi-band color images

by using mosaic images from other surveys. This can be especially beneficial when a survey has only one or two bands, such as the POSS-I (Minkowski and Abell, 1963b) or GALEX (Martin et al., 2005), as long as the image sizes are adjusted to match. Multiband images of the same area of the sky are useful for extracting a wealth of science information about a particular source. The scientific value of multi-band images are evident in areas such as the Stripe 82, which many surveys have chosen to maximize their overlapping area with in order to obtain such a collection of multi-band images on the targeted sources. The generated FITS mosaic images can also be used as inputs to existing tools such as Astrometry.net (Lang et al., 2010) or SExtractor (Bertin and Arnouts, 1996) for subsequent processing or to tools like ds9 (Smithsonian Astrophysical Observatory, 2000), Alladin (Bonnarel et al., 2000), or APLpy (Robitaille and Bressert, 2012) for scientific visualization.

Finally, the SDSS and POSS-II mosaic images that we have generated and publicly released via the LCDM website may prove be useful during the commissioning and data management processing of new surveys such as the DES or LSST. With the growth in the size of imaging cameras on new and existing telescopes, large, nearby galaxies, such as those in the RC3 catalog, need to be properly identified and possibly masked in order to minimize the loss of imaging details from saturated CCDs when imaging its vicinity. The updated RC3 coordinates indicate regions of the sky that may be affected by these large galaxies, and the FITS mosaic images can be used to model-fit the galaxy’s shape and light distribution. Alternatively, this same information can be used to place spectroscopic fibers on RC3 catalog sources.

Acknowledgements

The work was supported by the Google Summer of Code Program. RJB would also like to acknowledge support from the National Science Foundation Grant No. AST-1313415. 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

⁹<http://github.com/ProfessorBrunner/rc3-pipeline/>

¹⁰<http://lcdm.astro.illinois.edu/data/rc3/search.html>

the NASA/IPAC Infrared Science Archive. This research made use of Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration et al., 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 Astrofísica 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

- Ahn, C. P., Alexandroff, R., Allende Prieto, C., Anders, F., Anderson, S. F., Anderton, T., Andrews, B. H., Aubourg, É., Bailey, S., Bastien, F. A., et al., Apr. 2014. The Tenth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-III Apache Point Observatory Galactic Evolution Experiment. *The Astrophysical Journal*, Supplement 211, 17.
- 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.
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., Greenfield, P., Droettboom, M., Bray, E., Aldcroft, T., Davis, M., Ginsburg, A., Price-Whelan, A. M., Kerzendorf, W. E., Conley, A., Crighton, N., Barbary, K., Muna, D., Ferguson, H., Grollier, F., Parikh, M. M., Nair, P. H., Unther, H. M., Deil, C., Woillez, J., Conseil, S., Kramer, R., Turner, J. E. H., Singer, L., Fox, R., Weaver, B. A., Zabalza, V., Edwards, Z. I., Azalee Bostroem, K., Burke, D. J., Casey, A. R., Crawford, S. M., Dencheva, N., Ely, J., Jenness, T., Labrie, K., Lim, P. L., Pierfederici, F., Pontzen, A., Ptak, A., Refsdal, B., Servillat, M., Streicher, O., Oct. 2013. Astropy: A community Python package for astronomy. *Astronomy & Astrophysics* 558, A33.
- 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.
- Best, J. S., Charlton, J. C., Mayer-Kress, G., Jan. 1996. Analysis of Galaxy Morphology and Evolution Using the Pointwise Dimension. *The Astrophysical Journal* 456, 55.
- Blanton, M. R., Kazin, E., Muna, D., Weaver, B. A., Price-Whelan, A., Jul. 2011. Improved Background Subtraction for the Sloan Digital Sky Survey Images. *The Astronomical Journal* 142, 31.
- 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.
- Bonnarel, F., Fernique, P., Bienaymé, O., Egret, D., Genova, F., Louys, M., Ochsenbein, F., Wenger, M., Bartlett, J. G., Apr. 2000. The ALADIN interactive sky atlas. A reference tool for identification of astronomical sources. *Astronomy & Astrophysics Supplement* 143, 33–40.
- Budavari, T., 2004. Sdss command line query tool. <http://cas.sdss.org/dr7/sp/help/download/sqlcl/>, Last Accessed: 2014-09-22.
- Corwin, Jr., H. G., Buta, R. J., de Vaucouleurs, G., Dec. 1994. Corrections and additions to the third reference catalogue of bright galaxies. *The Astronomical Journal* 108, 2128–2144.
- 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.
- Dreyer, J. L. E., 1888. A New General Catalogue of Nebulae and Clusters of Stars, being the Catalogue of the late Sir John F. W. Herschel, Bart, revised, corrected, and enlarged. *Memoirs of the Royal Astronomical Society* 49, 1.
- Evans, I. N., Primini, F. A., Glotfelty, K. J., Anderson, C. S., Bonaventura, N. R., Chen, J. C., Davis, J. E., Doe, S. M., Evans, J. D., Fabbiano, G., Galle, E. C., Gibbs, II, D. G., Grier, J. D., Hain, R. M., Hall, D. M., Harbo, P. N., (He-

- len He, X., Houck, J. C., Karovska, M., Kashyap, V. L., Lauer, J., McCollough, M. L., McDowell, J. C., Miller, J. B., Mitschang, A. W., Morgan, D. L., Mossman, A. E., Nichols, J. S., Nowak, M. A., Plummer, D. A., Refsdal, B. L., Rots, A. H., Siemiginowska, A., Sundheim, B. A., Tippetts, M. S., Van Stone, D. W., Winkelman, S. L., Zografou, P., Jul. 2010. The Chandra Source Catalog. *The Astrophysical Journal*, Supplement 189, 37–82.
- 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.
- Gunn, J. E., Sigmund, W. A., Mannery, E. J., Owen, R. E., Hull, C. L., Leger, R. F., Carey, L. N., Knapp, G. R., York, D. G., Boroski, W. N., Kent, S. M., Lupton, R. H., Rockosi, C. M., Evans, M. L., Waddell, P., Anderson, J. E., Annis, J., Barentine, J. C., Bartoszek, L. M., Bastian, S., Bracker, S. B., Brewington, H. J., Briegel, C. I., Brinkmann, J., Brown, Y. J., Carr, M. A., Czarapata, P. C., Drennan, C. C., Dombeck, T., Federwitz, G. R., Gillespie, B. A., Gonzales, C., Hansen, S. U., Harvanek, M., Hayes, J., Jordan, W., Kinney, E., Klaene, M., Kleinman, S. J., Kron, R. G., Kresinski, J., Lee, G., Limmongkol, S., Lindenmeyer, C. W., Long, D. C., Loomis, C. L., McGehee, P. M., Mantsch, P. M., Neilsen, Jr., E. H., Neswold, R. M., Newman, P. R., Nitta, A., Peoples, Jr., J., Pier, J. R., Prieto, P. S., Prosapio, A., Rivetta, C., Schneider, D. P., Snedden, S., Wang, S.-i., Apr. 2006. The 2.5 m Telescope of the Sloan Digital Sky Survey. *The Astronomical Journal* 131, 2332–2359.
- Hogg, D. W., Blanton, M. R., 2006. Sdss images of selected rc3 galaxies. <http://cosmo.nyu.edu/hogg/rc3/>, Last Accessed: 2014-09-17.
- Ivezic, Z., Tyson, J. A., Abel, B., Acosta, E., Allsman, R., AlSayyad, Y., Anderson, S. F., Andrew, J., Angel, R., Angeli, G., Ansari, R., Antilogus, P., Arndt, K. T., Astier, P., Aubourg, E., Axelrod, T., Bard, D. J., Barr, J. D., Barrau, A., Bartlett, J. G., Bauman, B. J., Beaumont, S., Becker, A. C., Becla, J., Beldica, C., Bellavia, S., Blanc, G., Blandford, R. D., Bloom, J. S., Bogart, J., Borne, K., Bosch, J. F., Boutigny, D., Brandt, W. N., Brown, M. E., Bullock, J. S., Burchat, P., Burke, D. L., Cagnoli, G., Calabrese, D., Chandrasekharan, S., Chesley, S., Cheu, E. C., Chiang, J., Claver, C. F., Connolly, A. J., Cook, K. H., Cooray, A., Covey, K. R., Cribbs, C., Cui, W., Cutri, R., Daubard, G., Daues, G., Delgado, F., Digel, S., Doherty, P., Dubois, R., Dubois-Felsmann, G. P., Durech, J., Eracleous, M., Ferguson, H., Frank, J., Freeman, M., Gangler, E., Gawiser, E., Geary, J. C., Gee, P., Geha, M., Gibson, R. R., Gilmore, D. K., Glanzman, T., Goodenow, I., Gressler, W. J., Gris, P., Guyonnet, A., Hascall, P. A., Haupt, J., Hernandez, F., Hogan, C., Huang, D., Huffer, M. E., Innes, W. R., Jacoby, S. H., Jain, B., Jee, J., Jernigan, J. G., Jevremovic, D., Johns, K., Jones, R. L., Juramy-Gilles, C., Juric, M., Kahn, S. M., Kalirai, J. S., Kallivayalil, N., Kalmbach, B., Kantor, J. P., Kasliwal, M. M., Kessler, R., Kirkby, D., Knox, L., Kotov, I., Krabbendam, V. L., Krughoff, S., Kubanek, P., Kuczewski, J., Kulkarni, S., Lambert, R., Le Guillou, L., Levine, D., Liang, M., Lim, K., Lintott, C., Lupton, R. H., Mahabal, A., Marshall, P., Marshall, S., May, M., McKercher, R., Migliore, M., Miller, M., Mills, D. J., Monet, D. G., Moniez, M., Neill, D. R., Nief, J., Nomerotski, A., Nordby, M., O'Connor, P., Oliver, J., Olivier, S. S., Olsen, K., Ortiz, S., Owen, R. E., Pain, R., Peterson, J. R., Petry, C. E., Pierfederici, F., Pietrowicz, S., Pike, R., Pinto, P. A., Plante, R., Plate, S., Price, P. A., Prouza, M., Radeka, V., Rajagopal, J., Rasmussen, A., Regnault, N., Ridgway, S. T., Ritz, S., Rosing, W., Roucelle, C., Rumore, M. R., Russo, S., Saha, A., Sassolas, B., Schalk, T. L., Schindler, R. H., Schneider, D. P., Schumacher, G., Sebag, J., Sembroski, G. H., Sepala, L. G., Shipsey, I., Silvestri, N., Smith, J. A., Smith, R. C., Strauss, M. A., Stubbs, C. W., Sweeney, D., Szalay, A., Takacs, P., Thaler, J. J., Van Berg, R., Vanden Berk, D., Vetter, K., Virieux, F., Xin, B., Walkowicz, L., Walter, C. W., Wang, D. L., Warner, M., Willman, B., Wittman, D., Wolff, S. C., Wood-Vasey, W. M., Yoachim, P., Zhan, H., for the LSST Collaboration, May 2008. LSST: from Science Drivers to Reference Design and Anticipated Data Products. ArXiv e-prints.
- 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.
- Knapen, J. H., Shlosman, I., Peletier, R. F., Jan. 2000. A Subarcsecond Resolution Near-Infrared Study of Seyfert and “Normal” Galaxies. II. Morphology. *The Astrophysical Journal* 529, 93–100.
- 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.
- Lasker, B. M., Doggett, J., McLean, B., Sturch, C., Djorgovski, S., de Carvalho, R. R., Reid, I. N., 1996. The Palomar-ST ScI Digitized Sky Survey (POSS-II): Preliminary Data Availability. In: Jacoby, G. H., Barnes, J. (Eds.), *Astronomical Data Analysis Software and Systems V*. Vol. 101 of *Astronomical Society of the Pacific Conference Series*. p. 88.
- Lupton, R. H., Ivezic, Z., Gunn, J. E., Knapp, G., Strauss, M. A., Yasuda, N., Dec. 2002. SDSS Imaging Pipelines. In: Tyson, J. A., Wolff, S. (Eds.), *Survey and Other Telescope Technologies and Discoveries*. Vol. 4836 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. pp. 350–356.
- Martin, D. C., Fanston, J., Schiminovich, D., Morrissey, P., Friedman, P. G., Barlow, T. A., Conrow, T., Grange, R., Jelinsky, P. N., Milliard, B., Sigmund, O. H. W., Bianchi, L., Byun, Y.-I., Donas, J., Forster, K., Heckman, T. M., Lee, Y.-W., Madore, B. F., Malina, R. F., Neff, S. G., Rich, R. M., Small, T., Surber, F., Szalay, A. S., Welsh, B., Wyder, T. K., Jan. 2005. The Galaxy Evolution Explorer: A Space Ultraviolet Survey Mission. *The Astrophysical Journal*, Letters 619,

- L1–L6.
- Messier, C., 1781. Catalogue des Nébuleuses & des amas d'Étoiles. Connaissance des Temps for 1784, 227–267.
- Minkowski, R. L., Abell, G. O., 1963a. The National Geographic Society-Palomar Observatory Sky Survey. the University of Chicago Press, p. 481.
- Minkowski, R. L., Abell, G. O., 1963b. The National Geographic Society-Palomar Observatory Sky Survey. the University of Chicago Press, p. 481.
- Reid, I. N., Brewer, C., Brucato, R. J., McKinley, W. R., Maury, A., Mendenhall, D., Mould, J. R., Mueller, J., Neugebauer, G., Phinney, J., Sargent, W. L. W., Schombert, J., Thicksten, R., Jul. 1991. The second Palomar Sky Survey. Publications of the Astronomical Society of the Pacific 103, 661–674.
- Robitaille, T., 2014. The montage-wrapper github repository. <https://github.com/astropy/montage-wrapper>, Last Accessed: 2014-09-22.
- Robitaille, T., Bressert, E., 2012. APLpy: Astronomical Plotting Library in Python. <http://aplpy.github.io>, Last Accessed: 2014-08-15.
- Rockosi, C. M., Gunn, J. E., Carr, M. A., Sekiguchi, M., Ivezić, Z., Munn, J. A., Dec. 2002. Sloan Digital Sky Survey imaging camera: design and performance. In: Tyson, J. A., Wolff, S. (Eds.), Survey and Other Telescope Technologies and Discoveries. Vol. 4836 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. pp. 180–188.
- Shapley, H., Ames, A., 1932. A survey of the external galaxies brighter than the thirteenth magnitude. Annals of Harvard College Observatory 88, 41–76.
- 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.
- Smithsonian Astrophysical Observatory, 2000. SAOImage DS9: A utility for displaying astronomical images in the X11 window environment. <http://ds9.si.edu>, last Accessed: 2014-10-15.
- Stoughton, C. Lupton, R. H., Bernardi, M. e. a., Jan. 2002. Sloan Digital Sky Survey: Early Data Release 123, 485–548.
- The Dark Energy Survey Collaboration, Oct. 2005. The Dark Energy Survey. ArXiv Astrophysics e-prints.
- Walton, D. J., Roberts, T. P., Mateos, S., Heard, V., Sep. 2011. 2XMM ultraluminous X-ray source candidates in nearby galaxies. Monthly Notices of the Royal Astronomical Society 416, 1844–1861.
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., Ressler, M. E., Cutri, R. M., Jarrett, T., Kirkpatrick, J. D., Padgett, D., McMillan, R. S., Skrutskie, M., Stanford, S. A., Cohen, M., Walker, R. G., Mather, J. C., Leisawitz, D., Gautier, III, T. N., McLean, I., Benford, D., Lonsdale, C. J., Blain, A., Mendez, B., Irace, W. R., Duval, V., Liu, F., Royer, D., Heinrichsen, I., Howard, J., Shannon, M., Kendall, M., Walsh, A. L., Larsen, M., Cardon, J. G., Schick, S., Schwalm, M., Abid, M., Fabinsky, B., Naes, L., Tsai, C.-W., Dec. 2010. The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance. The Astronomical Journal 140, 1868–1881.
- York, D. G., Adelman, J., Anderson, Jr., J. E., Anderson, S. F., Annis, J., Bahcall, N. A., Bakken, J. A., Barkhouser, R., Bastian, S., Berman, E., Boroski, W. N., Bracker, S., Briegel, C., Briggs, J. W., Brinkmann, J., Brunner, R., Burles, S., Carey, L., Carr, M. A., Castander, F. J., Chen, B., Colestock, P. L., Connolly, A. J., Crocker, J. H., Csabai, I., Czarapata, P. C., Davis, J. E., Doi, M., Dombeck, T., Eisenstein, D., Ellman, N., Elms, B. R., Evans, M. L., Fan, X., Federwitz, G. R., Fiscelli, L., Friedman, S., Frieman, J. A., Fukugita, M., Gillespie, B., Gunn, J. E., Gurbani, V. K., de Haas, E., Haldeman, M., Harris, F. H., Hayes, J., Heckman, T. M., Hennessy, G. S., Hindsley, R. B., Holm, S., Holmgren, D. J., Huang, C.-h., Hull, C., Husby, D., Ichikawa, S.-I., Ichikawa, T., Ivezić, Ž., Kent, S., Kim, R. S. J., Kinney, E., Klaene, M., Kleinman, A. N., Kleinman, S., Knapp, G. R., Korienek, J., Kron, R. G., Kunszt, P. Z., Lamb, D. Q., Lee, B., Leger, R. F., Limmongkol, S., Lindenmeyer, C., Long, D. C., Loomis, C., Loveday, J., Lucinio, R., Lupton, R. H., MacKinnon, B., Mannery, E. J., Mantsch, P. M., Margon, B., McGehee, P., McKay, T. A., Meiksin, A., Merelli, A., Monet, D. G., Munn, J. A., Narayanan, V. K., Nash, T., Nielsen, E., Neswold, R., Newberg, H. J., Nichol, R. C., Nicinski, T., Nonino, M., Okada, N., Okamura, S., Ostriker, J. P., Owen, R., Pauls, A. G., Peoples, J., Peterson, R. L., Petravick, D., Pier, J. R., Pope, A., Pordes, R., Prosapio, A., Rechenmacher, R., Quinn, T. R., Richards, G. T., Richmond, M. W., Rivetta, C. H., Rockosi, C. M., Ruthmansdorfer, K., Sandford, D., Schlegel, D. J., Schneider, D. P., Sekiguchi, M., Sergey, G., Shimasaku, K., Siegmund, W. A., Smee, S., Smith, J. A., Snedden, S., Stone, R., Stoughton, C., Strauss, M. A., Stubbs, C., SubbaRao, M., Szalay, A. S., Szapudi, I., Szokoly, G. P., Thakar, A. R., Tremonti, C., Tucker, D. L., Uomoto, A., Vanden Berk, D., Vogeley, M. S., Waddell, P., Wang, S.-i., Watanabe, M., Weinberg, D. H., Yanny, B., Yasuda, N., SDSS Collaboration, Sep. 2000. The Sloan Digital Sky Survey: Technical Summary. The Astronomical Journal 120, 1579–1587.