

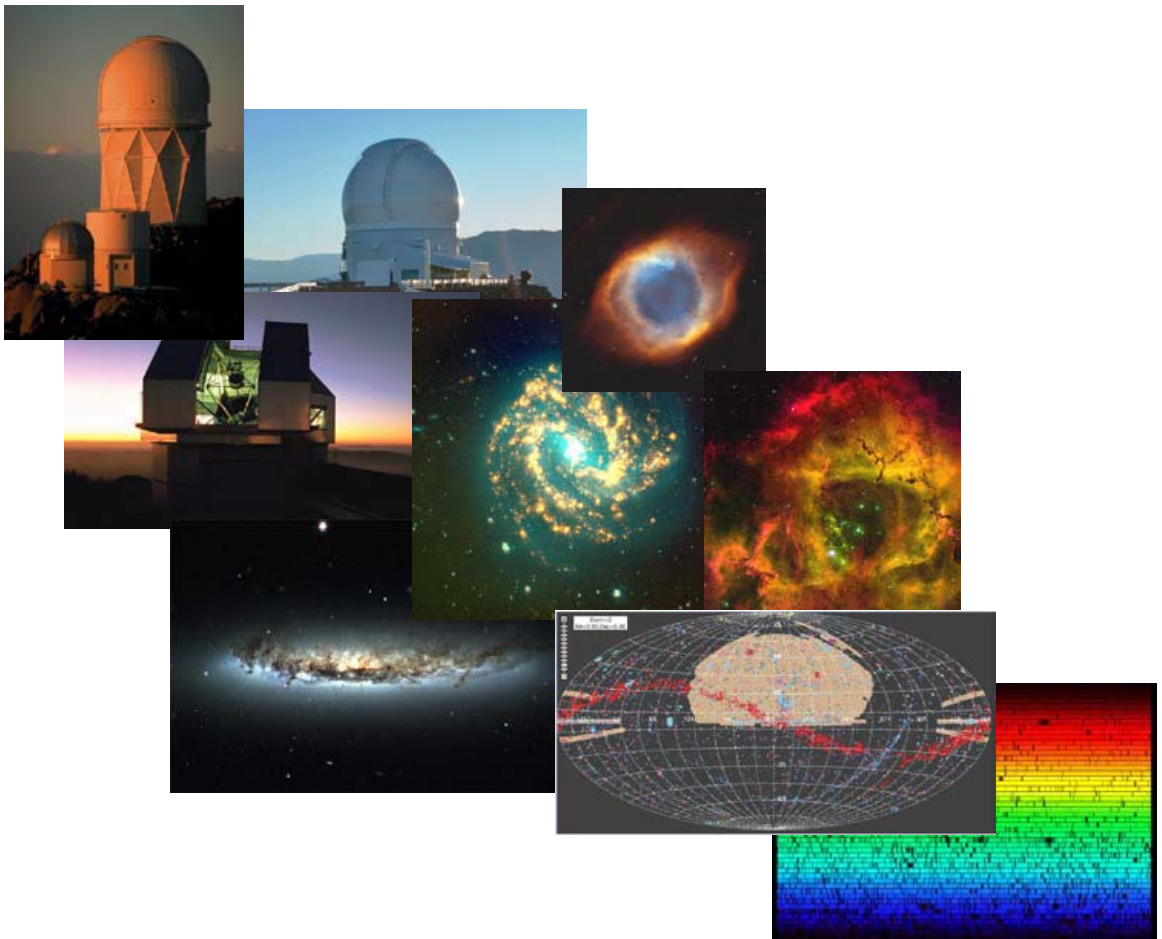


Science Data Management Group

National Optical Astronomy Observatory  
950 N. Cherry Ave.  
Tucson, AZ 85719

# NOAO DATA HANDBOOK

Version 1.1, May 2009



Operated by the Association of Universities for Research in Astronomy (AURA), Inc.  
under cooperative agreement with the National Science Foundation



[www.manaraa.com](http://www.manaraa.com)

المنارة للاستشارات  
AURA

## User Support

For help with accessing or understanding NOAO data products or data services, please email the Help Desk:

Email: [vohelp@noao.edu](mailto:vohelp@noao.edu)

For information, documentation, and other resources available on the Web, see:

<http://nvo.noao.edu/noaonvo/help.shtml>

This manual is available online at:

<http://www.archive.noao.edu/help.html>

## Revision History

Version	Date	Editor(s)
1.0	2007 December	Richard A. Shaw
1.1	2009 May	Richard A. Shaw

Citations to this version of the *NOAO Data Handbook* should read:

Shaw, R. A., ed. 2009, *NOAO Data Handbook* (Version 1.1; Tucson: National Optical Astronomy Observatory)

Send comments or corrections to:

Science Data Management Group  
NOAO  
950 N. Cherry Ave.  
Tucson, AZ 85719  
USA  
Email: [vohelp@noao.edu](mailto:vohelp@noao.edu)



Copyright © 2009, Association of Universities for Research in Astronomy, Inc. All rights reserved.

## *Contents*

Version 1.1, May 2009

<b>Preface</b> .....	<b>1</b>
Handbook Structure	2
Handbook Updates	3
Source Material & Attribution	3
<b>Chapter 1</b>	
<b>Introduction to NOAO Data</b> .....	<b>1-1</b>
1.1 NOAO Data Holdings	1-1
1.1.1 Data Sources	1-2
1.1.2 Processing Levels	1-3
1.1.3 Types of Data Products	1-4
Science Data	1-4
Metadata	1-5
Calibration Data	1-6
Concomitant Data	1-6
Ancillary Data	1-6
1.1.4 Data File Nomenclature	1-7
1.2 NOAO Data Formats	1-8
1.2.1 Imaging Data	1-8
1.2.2 Spectroscopic Data	1-9
1.2.3 Concomitant Data	1-9
1.2.4 Ancillary Data	1-10
1.3 Accessing NOAO Data	1-10
1.3.1 User Registration and Login	1-10
1.3.2 VO Portal Interface	1-12
1.3.3 Obtaining Data via the Portal	1-13
Retrieve Nonproprietary Science Data	1-14
Obtain Data from an Observing Run	1-19
Obtain Calibration Reference Files	1-22
1.4 Data Use Policy	1-25
1.4.1 Proper Citation for NOAO Data	1-25

1.4.2 Referencing Datasets in Journals	1-26
1.5 References & Further Information	1-26
Contributing Authors	1-26
References	1-27
For Further Reading	1-27

## Chapter 2

### Mosaic Cameras..... 2-1

2.1 Instrument Overview	2-1
2.1.1 Instrument Capabilities & Design	2-2
Focal Plane	2-4
Operations	2-6
2.2 Data Products	2-7
2.2.1 Image Formats	2-8
Raw Data	2-8
Calibrated Data	2-10
2.2.2 Header Keywords	2-12
2.2.3 Environmental Data	2-14
2.3 Calibration	2-15
2.3.1 Processing Steps	2-15
Artifact Flagging	2-17
Cross-Talk Correction	2-17
Bias Correction	2-18
Dark Correction	2-18
Pupil Ghost Correction	2-18
Fringe Pattern Correction	2-19
Flat-Field Correction	2-20
World Coordinate System Calibration	2-20
Photometric Calibration	2-21
Image Re-projection	2-22
Image Stacking	2-22
2.3.2 Calibration Reference Files	2-23
Bias Structure	2-24
Dome Flat-Field	2-24
Pupil Ghost	2-25
Fringe Pattern	2-25
Sky Delta-Flat	2-26
2.4 Sources of Error	2-27
2.4.1 Astrometry	2-27
2.4.2 Photometry	2-28
2.4.3 Anomalies	2-29
Metadata	2-29
Detector Performance	2-29
Bias Stability	2-30

Negative Saturated Pixels	2-30
2.5 References & Further Information	2-30
Contributing Authors	2-30
References	2-30
For Further Reading	2-31

## Chapter 3

### NEWFIRM Camera ..... 3-1

3.1 Instrument Overview	3-1
3.1.1 Instrument Capabilities & Design	3-2
Focal Plane	3-3
3.1.2 Operations	3-7
Array Operation	3-7
Observing Sequences	3-9
3.2 Data Products	3-10
3.2.1 Image Formats	3-11
Raw Data	3-11
Calibrated Data	3-13
3.2.2 Header Keywords	3-15
3.2.3 Environmental Data	3-18
3.3 Calibration	3-18
3.3.1 Processing Steps	3-19
Artifact Flagging	3-21
Dark Correction	3-21
Linearity Correction	3-21
Flat-Field Correction	3-22
Sky Subtraction	3-23
Astrometric Calibration	3-24
Photometric Calibration	3-24
Image Re-projection and Stacking	3-26
3.3.2 Calibration Reference Files	3-26
Dark	3-27
Linearity Coefficients	3-27
Flat-Field	3-27
3.4 Sources of Error	3-28
3.4.1 Astrometry	3-28
3.4.2 Photometry	3-29
3.4.3 Anomalies	3-29
Metadata	3-29
Dome Lamp State	3-30
Saturated Pixels	3-30
Image Ghosts	3-30
Background for Narrowband Filters	3-30
3.5 References & Further Information	3-31

Contributing Authors	3-31
References	3-32
For Further Reading	3-32
<b>Glossary</b> .....	<b>Glossary-1</b>

# *Preface*

Version 1.1, May 2009

## In this Preface...

Handbook Structure	2
Handbook Updates	3
Source Material & Attribution	3

The purpose of this *NOAO Data Handbook* is to describe the instruments and archived data products in enough detail that a Virtual Observatory user who is not familiar with them (or who may be new to the optical/infrared wavelength regime) can make informed use of these data products and services for scientific analysis. This *Handbook* is also intended for Principal Investigator teams who obtain data at NOAO or partner facilities and who wish to understand how their data are reduced and calibrated by the applicable pipelines. One of the key aims of this *Handbook* is to provide one-stop shopping for vital background material for each instrument, with references to more detailed material where it exists. It is also intended to be the most up-to-date reference material, and will be central to establishing the pedigree of datasets in the NOAO Science Archive (NSA). However, this *Handbook* is *not* intended to replace instrument manuals, online cookbooks, descriptions of filters, instrument design documents, or detailed technical descriptions of the archive or pipelines. It is also *not* a user guide for data analysis—the suite of available tools for data analysis is already quite rich, and growing. It will, however, draw upon all of these resources in order to provide a self-contained and complete manual for scientific use of the data from all instruments that are archived in the NSA. The complete list of the NSA holdings by instrument is given in Chapter 1.

The first edition of this *Handbook* coincided with the initial deployment of the End-to-End (E2E) data processing system at NOAO. E2E is the system that manages the flow of information and data from the database of approved observations, through data acquisition, transport from the telescope to an NOAO Archive Center, pipeline processing, and retrieval by an end-user. Because E2E was created *after* many instruments were deployed, rather than before, it will be

some time before the Archive is populated with all intended data products from all instruments. Thus, this *Handbook* is necessarily incomplete.

## HANDBOOK STRUCTURE

This *Handbook* is relatively modular in its construction, and is organized conceptually into two parts: an introductory chapter that describes general properties of NOAO data and the means to identify and access data for retrieval from the NSA, and chapters containing reference material for each instrument or class of instruments. These latter chapters are intended to give users a concise overview of each instrument and a good understanding of the metadata and data products that are produced. Instrument chapters are organized identically at the section level and consist of the following topics:

- **Instrument Overview.** Each instrument is described in enough detail to acquaint the user with the basic design and configuration, capabilities, typical operation, and the associated vocabulary.
- **Data Products.** Each of the data products that are produced for an instrument (including file formats) and essential metadata (e.g., header keywords) are described in depth to enable their effective use.
- **Calibration.** A detailed description of the processing that occurs to produce each calibrated data product is presented, along with the mechanisms by which the calibration reference information (e.g., master bias frames, flat-fields, etc.) is created and applied.
- **Sources of Error.** This section describes the accuracies that can be expected of the data products (e.g., astrometric and photometric uncertainties), instrument peculiarities, anomalous conditions that may have occurred at one time or another, known data foibles, and any other limitations to the scientific use of the data products.
- **References and Further Information.** A list of contributing authors, cited references, and pointers to background information is provided at the end of each chapter. Novel uses of the data or useful processing and analysis techniques found in the literature may also be cited.

In addition to the material listed above, this *Handbook* contains a glossary of often used terms and acronyms.

Special notes appear throughout this *Handbook* to convey information of special interest or urgency, particularly for the casual reader. They are the following:




---

Informative notes are denoted with the light-bulb symbol, and generally contain tips and pointers that deserve special attention.

---





Cautionary notes are indicated like this, and indicate potential limitations of the data, the instrument, or the processing software that may affect the use or interpretation of the data products.

---



Warnings of serious consequences are indicated like this, and caution users of problems with the data that could lead to erroneous scientific interpretations, or problems with the software that could lead to errors that may not be apparent to users.

---

## HANDBOOK UPDATES

New data products will be offered in the NOAO Science Archive in the future as new instruments are built and deployed at NOAO and partner facilities, as new processing pipelines are constructed to produce higher levels of data reduction for existing instruments, and as teams of scientists in the community create specialized products as a result of their research endeavors. Undoubtedly, improvements also will be made to the existing processing pipelines, and/or to the characterizations of instrument performance upon which the pipelines depend. When new or improved products are offered, chapters of this *Handbook* will be created or updated as necessary. This *Handbook* has been designed so that each chapter can be independently versioned, and updates will be announced on the NOAO Web pages and in the *NOAO/NSO Newsletter*. Users can download<sup>1</sup> the entire *Handbook* or any chapter of it as needed.

## SOURCE MATERIAL & ATTRIBUTION

The source material for this *Handbook* was drawn from a large number of documents, including instrument manuals, instrument and software design documents, published papers, Web pages, data file headers, and informal conversations with many experts within and outside of NOAO. Often, figures, tables, and even text are taken directly from these sources. In order to keep the style and presentation of this *Handbook* relatively clean, the attribution to the source material is cited the first time it is used in the main body of each chapter.

1. from <http://www.archive.noao.edu/help.html>

The last section of each chapter is devoted to a listing of references, contributing authors, and background references to details that fall outside the scope of this document.

# Chapter 1

## Introduction to NOAO Data

Version 1.1, May 2009

### In This Chapter...

NOAO Data Holdings	1-1
NOAO Data Formats	1-8
Accessing NOAO Data	1-10
Data Use Policy	1-25
References & Further Information	1-26

NOAO instruments routinely produce large quantities of digital data from a variety of science and survey programs directed by principal investigators (PIs). Data from some of these instruments are reduced and calibrated with automated pipelines, and both raw and reduced data products are published through the NOAO Science Archive. This chapter describes the scope of the data holdings in the Archive and the means to access them.

### 1.1 NOAO DATA HOLDINGS

NOAO collects a large quantity and variety of scientific data products, including images, spectra, catalogs, etc., from a large number of instruments on several telescopes that are deployed on three mountain tops on two continents. The primary repository for these data is the NOAO Science Archive (NSA). The data holdings can be categorized in a number of ways, including the source of the data, the type of data product, and the level(s) of processing required to produce them. Each of these views will be described in the following subsections. Since calibration pipelines exist for only a few instruments, most of the current archive holdings consist only of raw data products.

NOAO hosts two science data archives. The NOAO Science Archive for surveys (called Survey Archive) contains data products that were generated by NOAO Survey PI teams; it can be accessed through its custom Web interface.<sup>1</sup>

1. The Survey Archive data products can be accessed at <http://archive.noao.edu/nsa/>

The main NOAO Science Archive (called Archive), which contains all raw data, and calibrated data products for select instruments as produced by NOAO pipelines, is accessed through a Virtual Observatory (VO) compatible portal.<sup>2</sup> The NOAO VO Portal can also access all Survey Archive data products. The remainder of this chapter will describe the data holdings and access methods for the Archive (all data other than PI-produced survey products).



**Warning:** Users of NOAO data products should be aware that, with few exceptions, the data were obtained from observing programs that were designed and carried out by principal investigator (PI) teams, who obtained the data for the purpose of achieving their particular science objectives, and who obtained supporting calibration observations at their discretion. In addition, the quality of the data and the calibrations are strongly affected by the weather and other conditions that prevailed during the observing run.

### 1.1.1 Data Sources

The scope of the Archive data holdings can be measured by the variety of contributing data streams. Data from most instruments that are offered on NOAO facilities (with the exception of some visitor instruments) are captured and stored in the Archive. In addition, most data from partner institutions (i.e., observatories operated by consortia of institutions of which NOAO is a partner) are also captured, though in some cases only data obtained through the NOAO time allocation process may be stored in the Archive. The specific instruments and telescopes that contribute data to the Archive are listed in Table 1.1, along with the highest level of scientific data processing that is currently offered (see next subsection).

**TABLE 1.1: Archive Data Holdings**

Instrument	Type	Observatory/ Location	Highest Level Data Product <sup>a</sup>
Hydra	Optical fiber spectrograph	NOAO/Cerro Tololo	1
ISPI	IR imager	NOAO/Cerro Tololo	1
Mosaic-2 Camera	Optical imager	NOAO/Cerro Tololo	3
R-C Spectrograph	Optical spectrograph	NOAO/Cerro Tololo	1
4-m direct	Optical imager	NOAO/Kitt Peak	1

2. The NOAO VO Portal can be accessed at <http://portal-nvo.noao.edu/>

TABLE 1.1: Archive Data Holdings (Continued)

Instrument	Type	Observatory/ Location	Highest Level Data Product <sup>a</sup>
2.1-m direct	Optical imager	NOAO/Kitt Peak	1
Echelle Spectrograph	Optical spectrograph	NOAO/Kitt Peak	1
FLAMINGOS <sup>b</sup>	IR imaging spectrograph	NOAO/Kitt Peak	1
GoldCam	Optical spectrograph	NOAO/Kitt Peak	1
MARS	Optical spectrograph	NOAO/Kitt Peak	1
Mosaic-1 Camera <sup>c</sup>	Optical imager	NOAO/Kitt Peak	3
NEWFIRM	IR imager	NOAO/Kitt Peak	3
R-C Spectrograph	Optical spectrograph	NOAO/Kitt Peak	1
SQIID	IR imager	NOAO/Kitt Peak	1
0.9-m direct	Optical imager	SMARTS/Cerro Tololo	1
ANDICAM	Optical/IR imager	SMARTS/Cerro Tololo	1
1.5-m direct	Optical imager	SMARTS/Cerro Tololo	1
1.5-m Echelle	Optical fiber spectrograph	SMARTS/Cerro Tololo	1
1.5-m Cass Spectrograph	Optical spectrograph	SMARTS/Cerro Tololo	1
Y4KCam	Optical imager	SMARTS/Cerro Tololo	1
Goodman Spectrograph	Optical imaging spectrograph	SOAR/Cerro Pachón	1
OSIRIS	IR imaging spectrograph	SOAR/Cerro Pachón	1
SOI	Optical imager	SOAR/Cerro Pachón	1
Hydra/Bench	Optical fiber spectrograph	WIYN/Kitt Peak	1
Mini-Mosaic	Optical imager	WIYN/Kitt Peak	1
SparsePak	Optical IFU spectrograph	WIYN/Kitt Peak	1
WHIRC	IR imager	WIYN/Kitt Peak	1

a. See Table 1.2 on page 1-4 for a description of the data product levels.

b. Offered on both the 4-m and 2.1-m telescopes at KPNO.

c. Also offered on the WIYN Observatory 0.9-m telescope.

## 1.1.2 Processing Levels

The basic data files that are generated by the instruments in the observing environment are referred to as *raw* data products. Data from certain instruments have been processed to remove the instrumental signature, and to provide some type of calibration. Such advanced, higher-level science data products, along with concomitant data, are also available from the archive; detailed descriptions of these data products may be found in subsequent chapters, which are specific

to each instrument. The degree to which data have been processed is expressed in terms of *levels*, as described in Table 1.2.

**TABLE 1.2: Levels of Data Processing**

Level	Description	Example Products
1	Raw data, as it is formatted by the data collection software at the telescope	Raw Mosaic camera frames, including overscan; may include approximate WCS predicted from model.
2	Single-frame reduced data	Single Mosaic camera images with instrument signature removed; WCS calibration applied; image possibly re-projected.
3	Combinations of two or more frames	Combined and/or tiled images of a region of sky derived from multiple, spatially overlapping exposures with the same filter. Artifacts and gaps between detectors in a focal-plane array are identified and removed.
4	Measurements from one or more level-3 processed data frames	Star catalogs, including position, magnitudes, colors; identifications of moving objects; time-series of object brightnesses.

Note that not all data products can be usefully described in terms of processing level, including exposures obtained to support calibration and ancillary data that are described in the next subsection.

### 1.1.3 Types of Data Products

The Archive contains a variety of data products that are relevant for a scientist end-user. The following categories of products are part of the vocabulary used throughout this *Handbook* to describe NOAO data. Generally, each type of data product (apart from metadata) is stored in its own file.



At present there are only rudimentary ways of associating reduced science data files with related files such as the original raw data and data quality masks. The value of the header keyword `DTNSANAM` will be identical for raw science data files and products derived directly from them. For stacked images and other products with more than one parent, `DTNSANAM` will correspond to the first raw file that was used in creating the stack. Spatial coincidences of the instrument FoV may be useful, but bear in mind that the WCS is updated from raw to reduced data products.

#### Science Data

Frames of data obtained with an instrument of an astrophysical source are generally referred to as *science* data. Such data include images and spectrograms, and consist primarily of contiguous blocks of pixel values from one or more

detector arrays in the instrument. The science data are packaged with headers of metadata (see below) within a file; these metadata describe essential details of the science data. Science data products may also include high-level products, such as catalogs or time-series, that may or may not contain pixel values.

Science data in raw form are produced from the instruments listed in Table 1.1. Science data in processed form are generated either by dedicated processing pipelines that are part of the NOAO End-to-End (E2E) system, or they are provided by PI teams to the Survey Archive as advanced science products. In the first case, the data products are described in this *Handbook*. In the second case, PI teams provide the documentation necessary to understand the archived data products; links to this documentation are provided on the NSA Survey Archive Web site.<sup>3</sup>

## Metadata

*Metadata*, or data that describe data products, are collected or created at various points in the planning, acquisition, and processing of science data. Most of the metadata that are available to a user are stored in the headers of science data files as keyword–value pairs. Metadata fall into a variety of categories, but perhaps the most practical for the end-user are the following (see Shaw 2007):

**Telescope/Environment.** These metadata include telescope name and configuration information, celestial coordinates of the telescope pointing, airmass/zenith distance, tracking information, ambient temperature and wind speed, ADC use, etc.

**Instrument/Detector.** These metadata are required for generating Level-2 products from Level-1 raw data and include the instrument configuration, detector/amplifier designations, gain settings, operating temperatures, photo-active regions of the image, filter/disperser or other optical elements, and pixel scale.

**Science/Calibration.** These metadata include the World Coordinate System (WCS) mapping, field of view, resolution or delivered image quality, start of the exposure and its duration, effective bandpass and sensitivity of the system, brightness units and dynamic range, background level, and transformation to a standard photometric system.

**Pedigree/Provenance/Quality Assurance.** These metadata include the observer name and proposal identifier, version identifier of the control software, version of the data processing/calibration software, the processing steps performed, calibration reference files used, etc.

3. See the NSA Survey Archive Web site at <http://archive.noao.edu/nsa/>.




---

Critical metadata in raw science data files may be missing or have incorrect values. At present, no mechanism is in place to validate and correct corrupted metadata, apart from the proposal identifier (`PROPID`) and the date of the observation (the *date* portion of `DATE-OBS`). However, certain of the metadata in the headers of calibrated data, notably the WCS information and the filter name, are corrected during the course of pipeline processing. Metadata *added* by the pipeline during processing are not in question.

---

### **Calibration Data**

Calibration data in raw form are often very similar to (or indistinguishable from) science data products. These include exposures obtained for bias (zero duration), dark (finite duration with shutter closed), flat-field (including flats taken of the twilight sky), and standard star calibration. The *Master Calibration* reference files are usually generated by the processing pipeline from exposures obtained during the same observing run as the science data. The methods by which master calibration files are generated and how they are applied during processing is specific to each instrument, and they are discussed in detail in subsequent chapters.

### **Concomitant Data**

Concomitant data quantify the detailed quality of the science data, generally at the pixel level, and are generated during the course of pipeline processing. These data are intimately related to the scientific interpretation of the data products. Examples of concomitant data include data quality masks (DQMs) which codify one or more pathologies that affect (usually) small areas of the image, exposure masks that quantify the exposure level as a function of position in an image mosaic, and variance arrays that quantify the uncertainty associated with pixel values.

### **Ancillary Data**

Most ancillary data are generated contemporaneously with science data, but are not part of the science or engineering data streams. Some ancillary data, such as preview images, are generated by the processing pipeline. Ancillary data are used to establish the context of an observation, or to enable a rough, qualitative assessment of science data quality or applicability. Examples of ancillary data



include observation logs and environmental data such as that from seeing and atmospheric transparency monitors, images from all-sky cameras at or near the observing site, and weather satellite images.



At present no contemporaneous weather data are available from the Archive. However, weather data and images from all-sky cameras are available for KPNO;<sup>a</sup> statistical weather data are also available for CTIO,<sup>b</sup> as well as historical cloud coverage information<sup>c</sup> dating back to 1975.

- a. KPNO weather and sky conditions: [http://www-kpno.kpno.noao.edu/Info/Mtn\\_Weather/](http://www-kpno.kpno.noao.edu/Info/Mtn_Weather/)
- b. CTIO weather: <http://www.ctio.noao.edu/environ/environ.html>
- c. CTIO historical cloud coverage: [http://www.ctio.noao.edu/site/phot/sky\\_conditions.php](http://www.ctio.noao.edu/site/phot/sky_conditions.php)

## 1.1.4 Data File Nomenclature

Raw science data as obtained at the telescope are captured from the observing environment by the Data Transport System, and persisted as FITS files with unique names. Each name consists of a prefix, which is a two-letter code for the observing site (one of “cp,” “ct,” or “kp”), and a numeric serial number that increments with each successive observation (though not necessarily monotonically). Calibrated data files are prefixed with a code for the archive center where they were generated (“tu” for Tucson, “ls” for La Serena), plus a serial number that increases as each calibration data product is made available for Archive ingest. When users obtain data products through the Portal, the filenames are further prepended with a code for the Archive version number (currently, “NSAR3\_”). The files are postpended with the file type, which will also reflect the type of compression applied to the file, e.g., “.fits.gz” for a g-zipped FITS file. *Note that these file names are not related in any way to the names assigned by observers to data files at the telescope.*



Participants in observing programs should note that the filenames they assigned via the data taking system are preserved in the data headers, as the value of the keyword DTACQNAM. For some instruments this name includes the file pathname. After raw data are retrieved from the Archive, it is possible to restore the file names to those assigned in the observing environment using a script, an example of which can be found in the VO Portal contributed tools area.<sup>a</sup>

- a. Available at <http://nvo.noao.edu/noaonvo/contrib.shtml>

## 1.2 NOAO DATA FORMATS

Most data from the Archive will be provided to users in FITS format (FITS 2008), which is the well known data interchange format for astronomy. As most users know, the FITS standard allows for many complex representations of data. The general features of FITS file structures that are used for data in the Archive are described here, but implementation details for each instrument are deferred to subsequent chapters. Note that most data products as published by the Archive have been compressed in one way or another, which minimizes the time required for users to download their data. Depending upon the software that is used to examine or analyze the data, users may need to uncompress or otherwise transform their data after download.




---

Data files from the Archive are stored in one of a few compressed formats. Science images and calibration reference files are currently compressed with *gzip* (see <http://www.gzip.org/>); most systems have a native *gunzip* or equivalent program for decompressing the files after download. FITS tile compression (Pence et al. 2009) is planned for a future release of the Archive. Data quality masks are stored as a pixel list in BINTABLE extensions in a FITS file (this is actually one form of FITS compression). Some software (such as IRAF) treats these files as logical images transparently to the user; other software (e.g., DS9) currently requires that these masks be transformed to FITS images (or image extensions) before they can be imported.

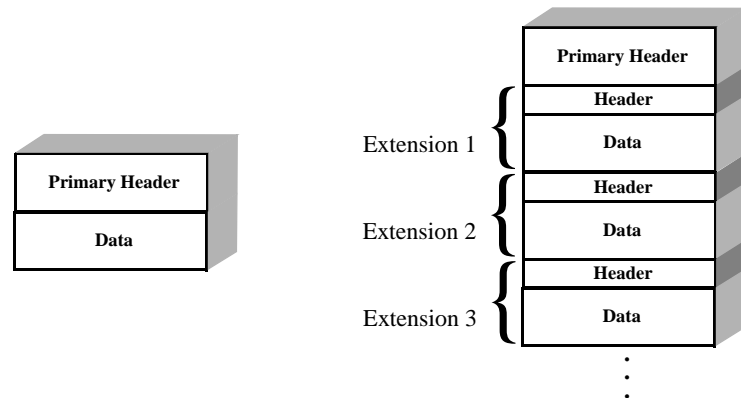
---

### 1.2.1 Imaging Data

In modern imaging cameras it is common to use multiple detectors, arranged in a contiguous array to cover a large fraction of the available focal plane of a telescope focal station. The detectors in these *focal plane arrays* (FPAs) are operated simultaneously, and are read out in parallel. (Some detector electronics allow for parallel read-out of portions of a single detector.) The signal from each of the detectors (or portions thereof) is generally stored in raw form as a separate image, along with the associated metadata; all of the images from a single exposure with the FPA are collected into a single FITS file for convenience. On the other hand, images from cameras with single CCDs, or composite images from, e.g., multiple, overlapping pointings that have been combined into a single image, are stored in simple FITS images in the Primary Header-Data Unit.

A schematic representation of the structure of NOAO FITS files for storing data from FPAs is given in Figure 1.1. Generally, the pixel values for multiple components of an image are stored in image extensions, rather than in the primary header/data block at the beginning of the file. The headers of each image exten-

sion contain metadata that pertain specifically to that image (e.g., the WCS information). Global metadata that apply to all the images (e.g., filter name, exposure duration, etc.) may be stored in the primary header, in which case the keyword `INHERIT = T` will be found in the extension headers.<sup>4</sup>



**FIGURE 1.1:** Schematic of the structure of a simple FITS file that stores a single image (*left*) and a Multi-Extension Format file that stores one or more components of an image (*right*). Metadata that apply to all extensions are often stored in the Primary Header.

## 1.2.2 Spectroscopic Data

The storage format for raw spectrograms is conceptually similar to that for imaging cameras, i.e., pixel values from multiple detectors are stored in simple FITS or FITS MEF format. At present, no Level-2 or higher (i.e., processed) data products for spectra are offered in the Archive.

## 1.2.3 Concomitant Data

Concomitant data are always stored in compressed form, although they are logically equivalent to ordinary FITS images. Since DQMs usually contain only a small fraction of pixels with values different than zero (zero indicates no data quality problem), it is much more efficient to represent the image internally as a list of regions containing particular nonzero values. These lists are stored as tables in FITS BINTABLE extensions: one extension per image, similar to the structure illustrated in the right panel of Figure 1.1. Similarly, exposure maps mostly consist of large regions with a common effective exposure time; these are also stored in BINTABLE extensions.

4. In practice, the most critical metadata in the primary header often also appear in the extension headers because not all data analysis systems support the FITS `INHERIT` convention, which is described at <http://fits.gsfc.nasa.gov/registry/inherit.html>.

## 1.2.4 Ancillary Data

Ancillary data come in a variety of formats, depending on the specific product. At present the only ancillary data offered through the Archive are down-sampled image previews. These data are generated from science or Master Calibration images, and are stored in PNG<sup>5</sup> format. Note, however, that these preview images are encapsulated in a FITS file in extensions of type FOREIGN<sup>6</sup> (Zárate, Seaman & Tody 2007), and must be unpacked before they will be of use to most application software.

## 1.3 ACCESSING NOAO DATA

Access to and use of nonproprietary data is available to everyone, consistent with the *Conditions of Use* policy described in Section 1.4, and requires neither registration nor login. Authorized users who wish to download *proprietary* data may do so only after registering with the Virtual Observatory (this need be done only once) and logging in to the NOAO VO Portal. The VO provides secure *authentication* of users by issuing an electronic security certificate to their browser when they log on to the VO; this mechanism enables access to any service or data in the VO that accepts single sign-on certificates. The Archive makes use of such certificates to provide you with access to all proprietary data to which you are authorized.




---

**The NOAO VO Portal can only be used successfully with compatible browsers, which currently are Mozilla 1.7 or greater, Netscape 9.0 or greater, and Firefox 1.4 or greater. Use with any other browsers will likely result in missing or erroneous functionality that may or may not be apparent to the user!**

---

### 1.3.1 User Registration and Login

PIs wishing to retrieve *proprietary* data from the Archive must complete a one-time registration process with the VO to obtain a user ID and password. Retrieving nonproprietary data does *not* require registration. Investigators who have been awarded observing time on NOAO or partner facilities will receive an invitation via email (from [archive-noreply@noao.edu](mailto:archive-noreply@noao.edu)) to register with the VO. This email will be sent to PIs prior to the actual observing run.

5. <http://www.libpng.org/pub/png/>

6. <http://fits.gsfc.nasa.gov/registry/foreign.html>

Registration is a straightforward process and consists of the following steps:

1. **View the invitation email** from [archive-noreply@noao.edu](mailto:archive-noreply@noao.edu) (with the subject line “NOAO Science Archive Registration Invitation”), and enter the registration URL into your compatible browser.<sup>7</sup>
2. **Complete the Portal registration form.** This form will be prefilled with information that is on file with NOAO: your name, institution, and email. If any of this information is incorrect or out of date, now is the time to fix it. If you have previously registered with the VO, click the [I'm already registered](#) button and skip to step 5. Otherwise, click [Start registration with VO](#), and continue with step 3.
3. **Complete the VO registration form.** Information from the Portal registration form will be transferred to this form. Fill in the remaining information, including your choice for *username* and a *password*. **Note:** *it is important to remember your VO username and password*. Click [Submit](#) to advance to the confirmation phase of the VO registration process.
4. **Confirm your identity.** A confirmation token will be sent to the email address you provided in step 2. Copy this token from the email message into the dialog box on the form.<sup>8</sup> Clicking [Submit](#) will first display a congratulatory page from the VO, and within about 20 seconds your browser will be redirected to the VO login page.
5. **Log in to the VO.** Your Portal registration will not be complete until you perform this step. Click the [Login](#) button after entering your username and password (selected in step 3 for first time registrants). Your browser will automatically be redirected to the Portal **sky** view.

Returning users need only log in to the NOAO VO Portal at <http://portal-nvo.noao.edu/> by selecting **Sign in** from the menu bar, using the username and password they established during the registration process.<sup>9</sup> Users will automatically have access to all proprietary data from all NOAO programs for which they have been authorized. Users who have previously registered with the VO,




---

If an error message appears in your browser indicating that a “Secure Connection Failed” or that “sso.us-vo.org uses an invalid security certificate,” you will need to add a security exception in your browser. Adding such an exception is easy, but the details are browser-specific.

---

but who have not been invited to do so by the NOAO Science Archive via the aforementioned invitation email, will not have secure sign-in privileges with the NOAO VO Portal.<sup>10</sup> They may still use the Portal without logging in, however.

7. Be sure to enable pop-ups, at least for the site <http://portal-nvo.noao.edu/>.

8. Alternatively, you may click the direct URL in your email message **provided that the browser you are using for email is compatible with the Portal**.

9. Attempting to follow the registration link in the invitation email *after you have registered* will result in an error.

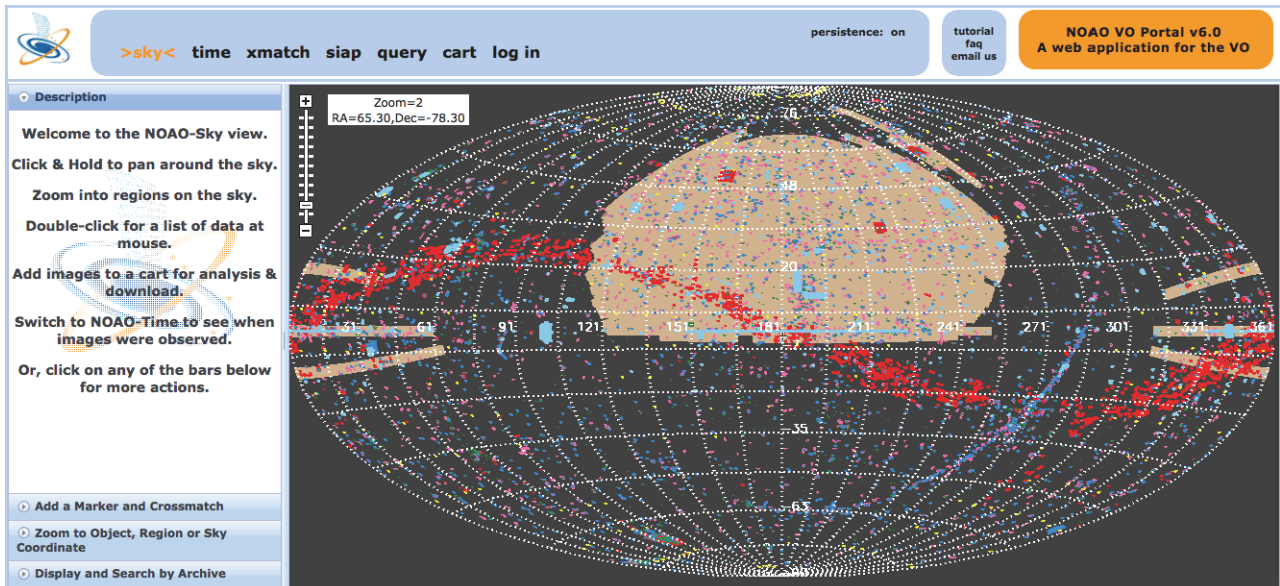
### 1.3.2 VO Portal Interface

The NOAO Science Archive was designed to be compatible with the VO, and users can access NOAO data through a variety of tools and services that are VO-compatible. One such access tool is the NOAO VO Portal, which is composed of a rich set of services that are accessed through a Web interface, as shown in Figure 1.2. The initial view is that of an all-sky Aitoff projection of the celestial sphere, with the footprints of various surveys indicated with shaded regions (in color). The suite of functionality includes:

- panning and zooming on the sky to identify surveys that overlap spatially with regions of interest;
- searching for datasets from selected repositories that are connected to the VO, via a spatial region (e.g., a cone search), epoch of observation, or type of data (image, spectrum, etc.);
- staging data in the VO for retrieval; and
- cross-matching a target location against major VO catalogs.

The NOAO VO Portal is particularly well tuned to retrievals of NOAO data from the Archive. The Portal's browser-based graphical user interface enables the generation of rich queries of the underlying database through mouse actions and simple forms-based user input. The auto-generated SQL (the structured query text) may be saved for later use in advanced queries, or for access to NOAO data through mechanisms outside the browser-based NOAO VO Portal (see "Special Case: Co-I access to proprietary data" on page 1-21).

10. Users who are not NOAO PIs and who wish to have an account for secure sign-in to the NOAO VO Portal may email their request to [vohelp@noao.edu](mailto:vohelp@noao.edu).



**FIGURE 1.2:** Web interface for the NOAO VO Portal, showing an all-sky view of the footprints of major VO holdings that can be accessed (*color shaded areas*). Additional markers will appear for NOAO datasets that are identified in the user's cart.

### 1.3.3 Obtaining Data via the Portal

Obtaining data from the Archive is straightforward and consists of the following steps:

1. **Register as a user of VO resources** (one time only, and only for access to proprietary data).
2. **Access the NOAO VO Portal** at <http://portal-nvo.noao.edu/> (retrieving proprietary data requires login).
3. **Construct a query** for data that meet your science criteria.
4. **Select data products** of interest from the Data Grid, and add them to the cart.
5. **View the cart** contents, and select those data products of interest for retrieval.
6. **Stage the selected data** to the Archive ftp area.
7. **Download the data** to your machine from the ftp area.

The following examples illustrate the above process for some common cases. It is assumed throughout that the user has registered with the VO (if needed), has opened the Portal interface in a compatible browser, and has selected the **query** panel from the top menu bar.

## Retrieve Nonproprietary Science Data

**Construct a Query.** This example describes how to retrieve nonproprietary reduced data from the observing program 2002B-0007, obtained in November 2005. The approach for building the needed query is to use pull-down menus to construct an approximate query, and then make refinements to the search criteria. At the **Query** panel, perform the following steps, as illustrated in Figure 1.3 on page 1-15:

1. Select the **Default Simple Query** view from the tab at the upper left of the Query panel.
2. Choose the following from the pull-down menus at the upper right:
  - Search Type: Search All Data
  - Data Type: Reduced Images
  - Example Queries: Get Reduced Data and Masks by PropID
3. Type the proposal ID into the text box (in this case “2005B-0007”).
4. Enter a Date-Obs string to match any date in the year of the observation (in this case, “2005%”).
5. Select the desired type of reduced data from the pull-down menu (in this case, InstCal).
6. Check the **Public Data** box to match only nonproprietary data.
7. Press the **Update Query** button (note how the text in the left panel changes).
8. Press the **Submit Query** button at the top left.

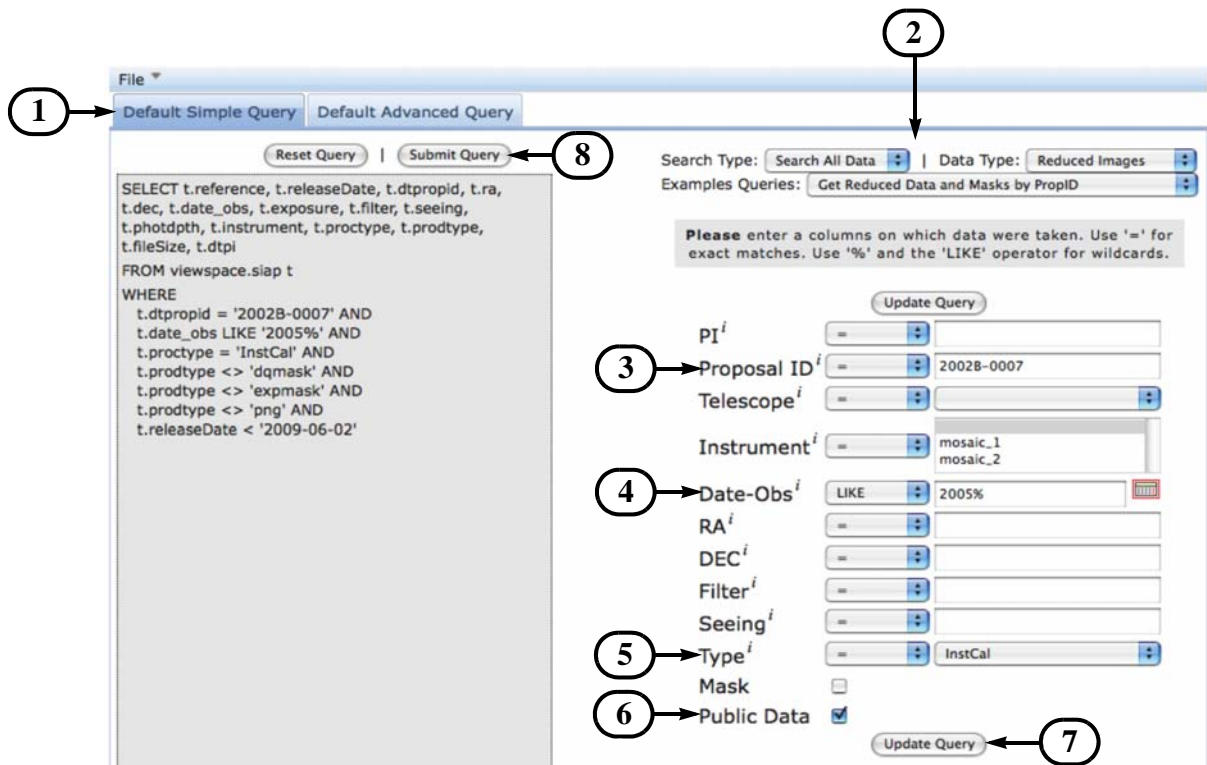



---

It is of course possible to select by other criteria rather than the proposal ID, for instance on a range in RA and Dec (using the “BETWEEN” operator in the pull-down to the left of the text boxes). In this case, one would leave the “Proposal ID” text box blank.

---



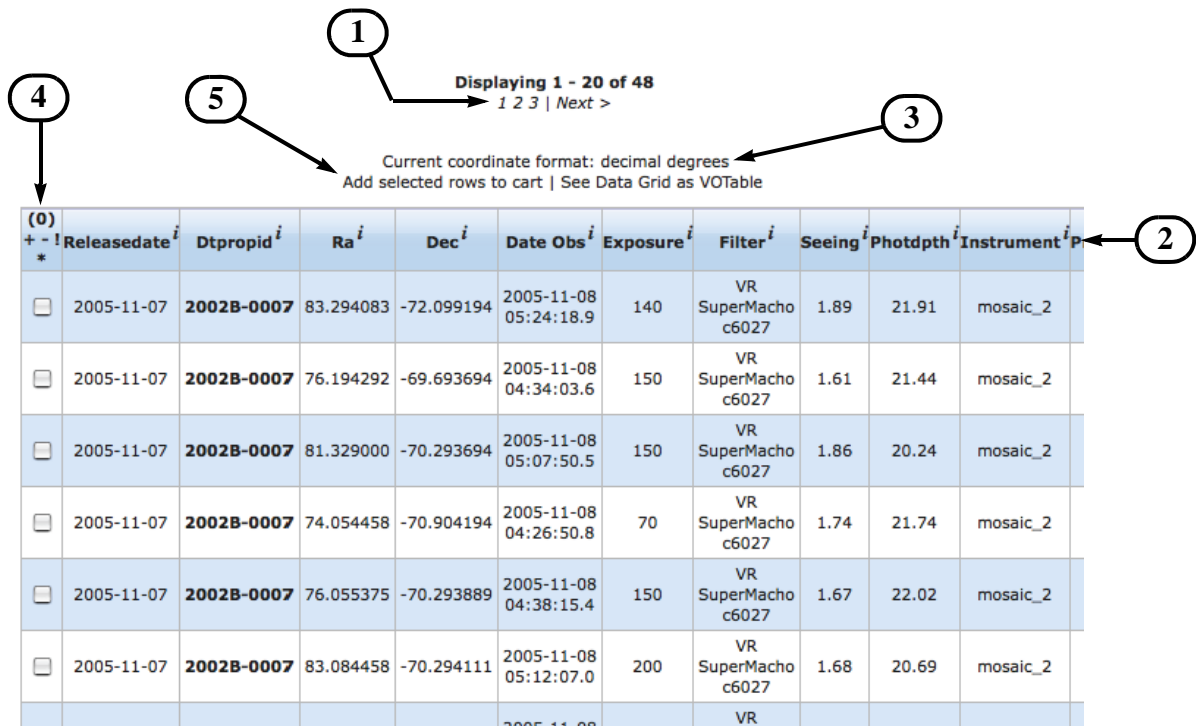


**FIGURE 1.3:** Portal **Default Simple Query** panel, showing the input text boxes (*right*) and the resulting query string (*left*) that can be updated with values in the boxes by pressing the **Update Query** button (*bottom*). Numbered callouts are explained in the text. Pressing the **Submit Query** button (*top*) will send the query to the Archive, the results for which are displayed in the panel described in Figure 1.4 on page 1-16.

The pending query that will be passed to the Archive database is shown on the left portion of the panel in Figure 1.3. The results will appear in the *Data Grid* panel just below the query form, shown in Figure 1.4 on page 1-16; you may want to adjust the Data Grid’s viewable area with the frame boundary widget by dragging upwards the bar that separates the Data Grid from the query panel.

**Select Data.** Each row in the Data Grid corresponds to one unique data product. There are up to 20 rows on each page, and pages can be navigated through links near the text labeled (1) in Figure 1.4. Various attributes of the datasets are given in the header row, labeled (2), and a brief description of the attribute is displayed by clicking the “i” at the upper-right of each column name; clicking the header text of any column sorts the rows by value in that column. The coordinates for the dataset are given by the Right Ascension and Declination columns, which by default are in units of decimal degrees; to view the coordinates in sexagesimal format, click the text labeled (3). You may select datasets of interest by doing one or more of the following actions in the column labeled (4):

- check the box to the left of individual datasets;
- click the plus (+) sign to select all rows on the displayed page (rows will remain checked even if you display another page);
- click the asterisk (\*) to select all rows on all pages;
- click the minus sign (-) to de-select all rows on the displayed page; and/or
- click the exclamation point (!) to de-select all rows on all pages.



**FIGURE 1.4:** Portal Data Grid, showing query results from Figure 1.3. Numbered callouts are explained in the text. Clicking the link labeled (5) will add all *selected* datasets to your cart, the result of which is shown in Figure 1.5 on page 1-17.

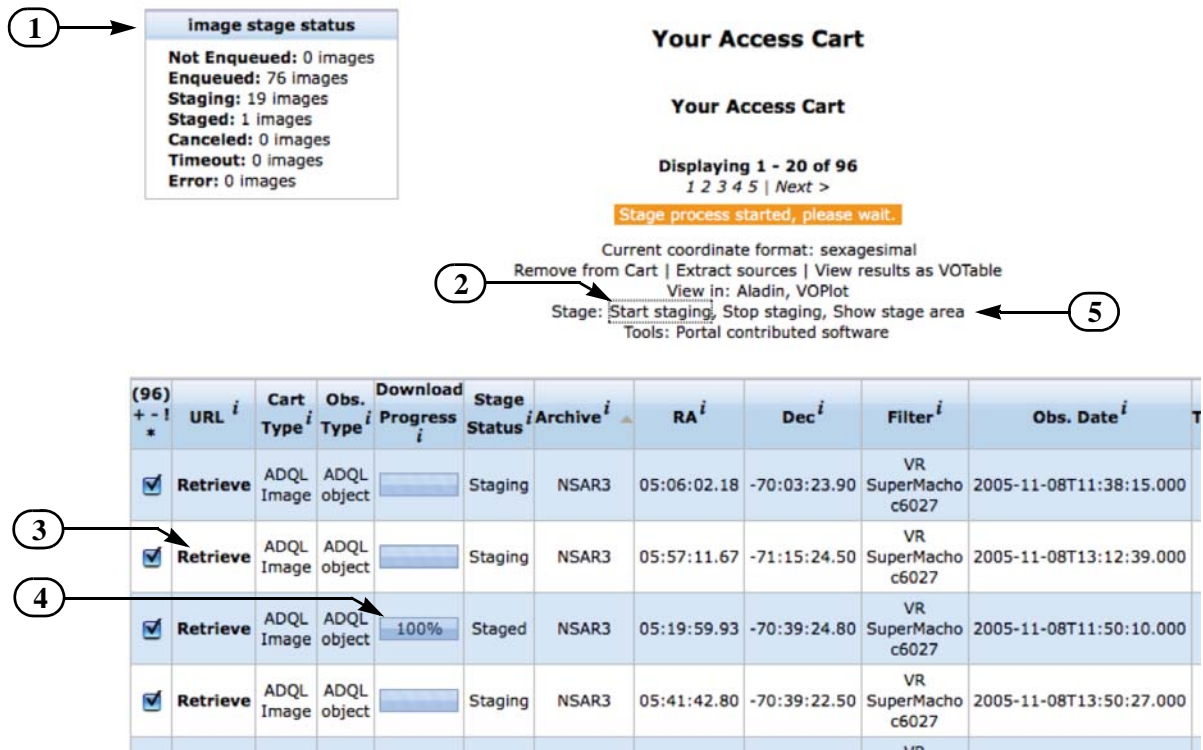
After selecting the datasets of interest, click the link labeled (5) to add them to the cart. At this point, you may either query for additional data (e.g., search for higher-level data products or Master Calibrations from this same dataset) and add them to the cart, or go directly to the **cart** view.

**View the Cart.** Data in the cart may be examined or used for a number of purposes using the Portal, such as viewing the spatial location on the **sky** view. To download the datasets in the cart, select the **cart** entry in the top menu bar.



To actually stage the data for ftp download, select datasets of interest in the cart Data Grid pane using much the same process as described above (i.e., checking boxes in the first column individually or in aggregate), as shown in Figure 1.5.

**Stage the data.** The status (1) of the staging process is updated continuously once the Start staging text (2) is clicked. Alternatively, individual data files may be downloaded directly by clicking the Retrieve text (3) in the **URL** column. Staging progress for individual files is displayed as progress bars (4). To view the stage area (see Figure 1.6 on page 1-18), click the Show stage area text (5). The staging progress is summarized in the **Image stage status** box, with files moving from being *enqueued*, to *staging* (i.e., actively transferring data from mass store to the temporary staging area), to *staged* (transfer complete).

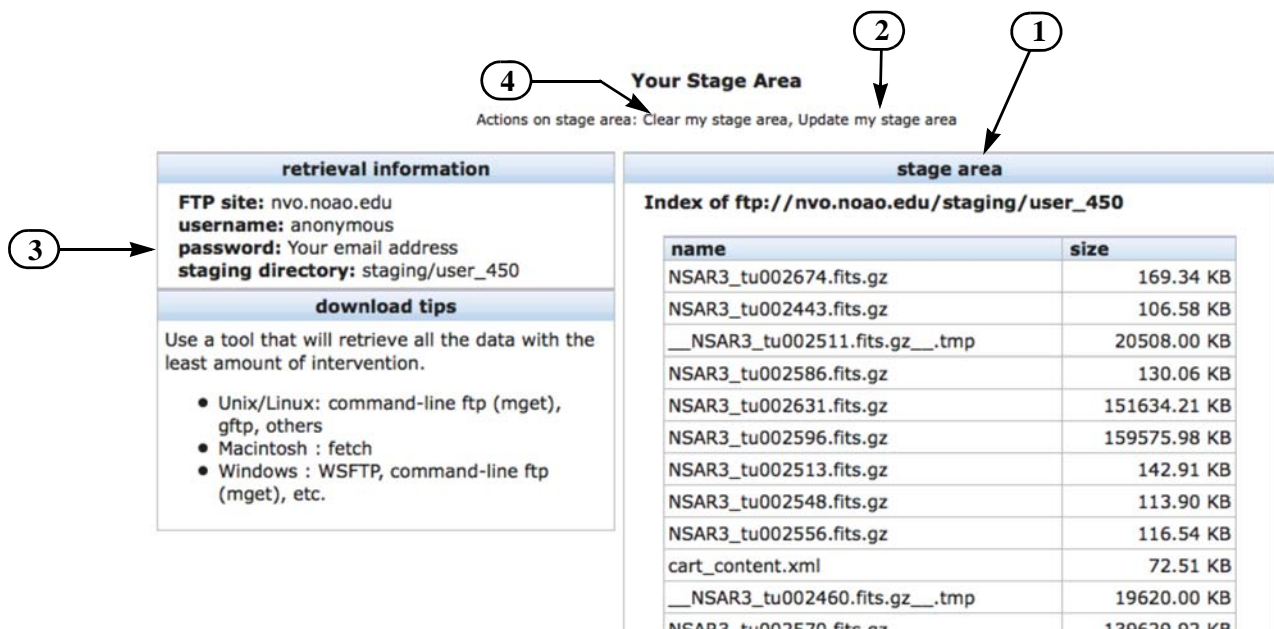


**FIGURE 1.5:** Portal Access Cart. Numbered callouts are explained in the text. The status (1) of the staging process is updated continuously, once the Start staging text is clicked. You may view the stage area (see Figure 1.6 on page 1-18) by clicking the Show stage area text.

If staging takes an unusually long time (perhaps hours), the status may change to “Error.” Problems may have their cause in hardware or software failures in the Archive, network problems, or a staging area that has reached full capacity. The system attempts to recover from such errors, so the status should be monitored for some time before seeking help from the NOAO Help Desk. If staging fails, the affected files are noted in the Error row at the top; the word “Error”

then becomes a link which, if followed, will ask you if these files should be restaged. Note that if you log out of a Portal session before staging completes, you can resume monitoring the progress at any time by logging in again and clicking the [Show stage area](#) text on the **cart** view.

The stage area (Figure 1.6) shows which products have been (or are being) successfully staged for download, as well as other information needed for access. The files that reside in the stage area are listed on the right (1) with the standard file nomenclature (see “Data File Nomenclature” on page 1-7); if staging is not yet completed, the list may be updated manually by clicking the [Update my stage area](#) text (2).



**FIGURE 1.6:** Portal Data Staging Area. Once the data files have all been staged, use a download client (e.g., ftp) to access the ftp server at nvo.noao.edu.

**Download Data.** Once all the data have been staged, they are ready to be downloaded from the Archive ftp site at <ftp://nvo.noao.edu> via a file transfer client of your choice (suggestions are offered in the **download tips** box). The login information for your staging area is given in the **retrieval information** box (3), including the path to your data. The staging area is a limited resource, and older files are deleted periodically to make room for new retrievals. Thus, prompt retrieval of your files is advised. To give yourself room for more downloads, you can delete your old data from the stage area by clicking the [Clear my stage area](#) text (4).




---

**Data in the NOAO staging area have a limited shelf life, and may be deleted after one week if the staging area nears its maximum capacity. Therefore, prompt retrievals are advised.**

---

### ***Obtain Data from an Observing Run***

In the following example, the raw data from the NOAO Survey program 2005B-0045 (PI: C. Miller) will be retrieved using a common type of query, as the PI might be expected to do. These data were obtained with the Mosaic-2 camera at CTIO during 24–30 April 2008.

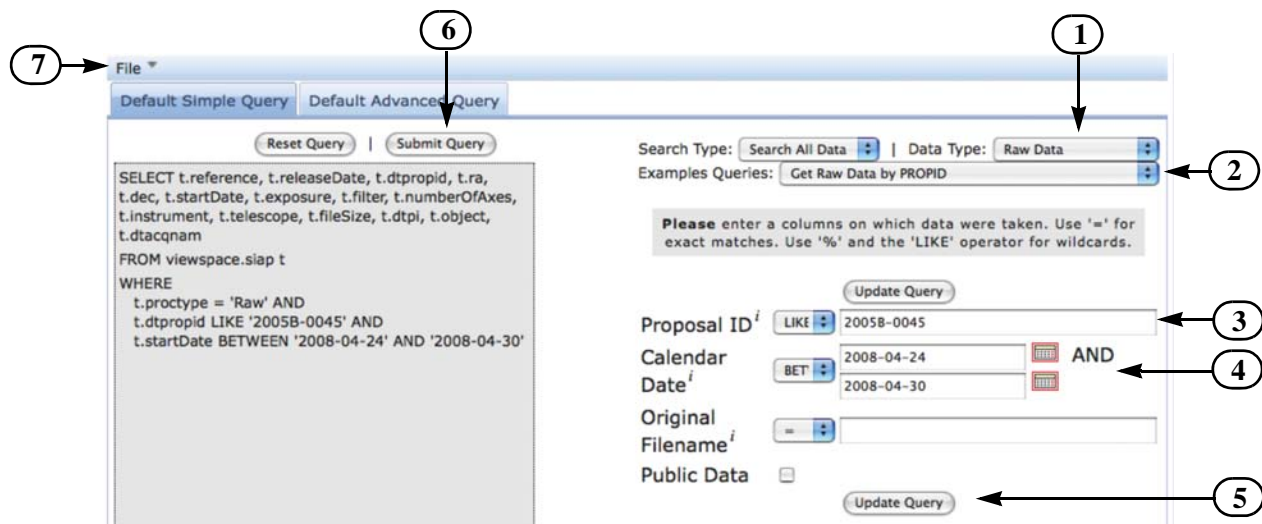



---

Only PIs of observing programs are authorized to retrieve data during the proprietary period, and they must log in to the portal using their assigned password in order to stage their data. Queries for proprietary data from non-PIs will populate the Portal Data Grid, but these data cannot be selected and placed in the cart for retrieval. PIs may transfer authorization to co-investigators (Co-Is) using security certificates (see “Special Case: Co-I access to proprietary data” on page 1-21). To query for nonproprietary data only, check the **Public Data** checkbox.

---

As in the previous example, an approximate query can be generated with the **Default Simple Query** panel, as shown in Figure 1.7 on page 1-20, and refined with the following steps:



**FIGURE 1.7:** A Custom Query to obtain data from an observing run can be generated by modifying the **Get Raw Data by PROPID** example query.

1. Select **Raw Data** from the Data Type pull-down menu at the top-right.
2. Select **Get Raw Data by PROPID** from the Examples Queries pull-down menu at the top-right to build a query string that is similar to what will be needed in this example.
3. Refine the query string by entering the observing proposal number in the Proposal ID text box (in this case, “2005B-0045”).
4. Select the **BETWEEN**<sup>11</sup> operator next to the **Calendar Date** label, and enter the first and last dates of the observations in the text boxes (YYYY-MM-DD format), or use the calendar widgets.
5. Press the **Update Query** button. The refined query will then be displayed in the left portion of the pane.
6. Either press the **Submit Query** button to execute the pending query, or alternatively...
7. Store the query text in a file on your machine by selecting **File**→**Save Query As...** from the pull-down menu at the upper left.

The results of the search will appear in the Data Grid panel, beneath the Query panel. Staging data from the cart and retrieving data proceeds exactly as in the prior example, “Retrieve Nonproprietary Science Data” on page 1-14. Recall that the file names may be restored to those assigned in the observing environment using a script: see “Data File Nomenclature” on page 1-7.

<sup>11</sup>. Note that the **BETWEEN** operator requires that the first argument be strictly less than or equal to the second argument.




---

The Portal currently limits the number of datasets that can match a single query to 1000 files. Although the number of files from even a moderate-length observing run with NEWFIRM will likely exceed this limit, it is necessary to keep the required disk space in the NOAO staging area to a reasonable level when servicing multiple users. Users with large numbers of datasets from lengthy observing runs will need to break their retrievals into individual requests that do not exceed the limit by, e.g., restricting queries to a time window of a few days.

---

**Special Case: Co-I access to proprietary data.** If you are a co-investigator on a program and wish to retrieve the data during the proprietary period, the PI can give you authorization to do so in a relatively secure way *without sharing his/her VO login information*.

1. Download a shell script for command-line file retrieval. Go to <http://nvo.noao.edu/naonvo/contrib.shtml> or click the [Portal contributed software](#) text near the top of the Portal **cart** view. Download the [secure\\_get.sh](#) file to your machine, and ensure that you have permission to run it as an executable.
2. Using the Portal, build and submit a query that selects the data products of interest, bearing in mind the limit on the maximum number of files that can be staged in a single request.
3. In the Data Grid (see Figure 1.4 on page 1-16), right-click the [See Data Grid as VOTable](#) text, and save the file to your local disk (choose XML format, which is the default). This file contains information in VOTable format that is necessary to access the files currently in the Data Grid. You will need to generate more than one of these tables if you require more than one query to specify all the data files of interest.
4. Obtain a VO security certificate file from the PI.
5. Run the `secure_get.sh` script to retrieve the files prior to the certificate's expiration date.

The instructions for running the script are given at the top of the file. The command-line syntax (with a certificate) is:

```
secure_get.sh infile --cert CERT --pass PASS [--dir DIR]
```

where *infile* is the name of the VOTable containing the search results, *CERT* is the name of the certificate file, *PASS* is the password, and (the optional) *DIR* is the path to the destination directory. If you have more than one VOTable of search results, you will need to run the script multiple times.

In order for the PI to provide a secure certificate to a collaborator, he/she must do the following:

1. Log on to the VO at <https://nvologin1.ncsa.uiuc.edu/protected/welcome>.
2. Choose a validity lifetime for the certificate.
3. Choose a Package Key (i.e., a password) for the certificate.
4. Choose PEM format.
5. Click the **Download Certificate** button. A file called “geteec,” which is the certificate file, will appear in your download directory.
6. Send this file to your collaborator, and tell them (preferably in a separate email) the certificate password and expiration date.

## Obtain Calibration Reference Files

On occasion, users may wish to examine in detail the quality of the pipeline instrument signature removal, to recuperate raw data using custom software, or to create custom calibration reference files from other observing runs. In these cases, it is helpful to retrieve either the master calibration reference files or the raw files from which they were created. Retrieving these files is similar to other nonproprietary data retrievals, but the search criteria are sufficiently different that efficient queries require the **Default Advanced Query** panel of the Portal **query** interface. This example illustrates a search for NEWFIRM raw dark files for potential use in calibrating data from another observing run, for which the selection criteria include integration times near 60 s duration and observing dates at KPNO in early May 2009.<sup>12</sup>

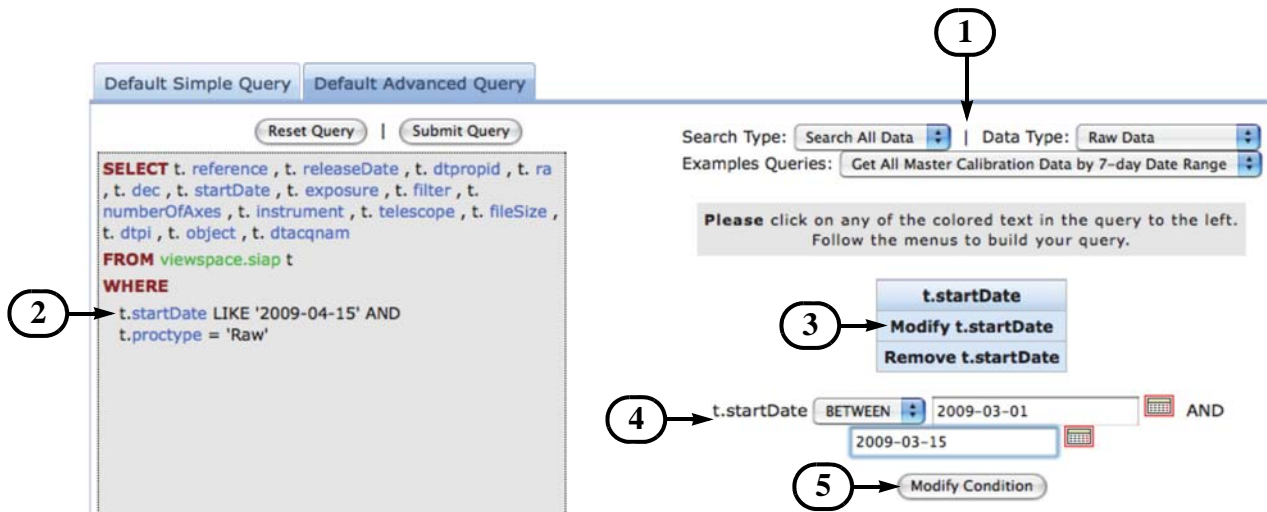
As in prior examples, an approximate query is generated and then refined. But unlike the **Default Simple Query** panel, essential attributes in the advanced query string will be edited by clicking highlighted text<sup>13</sup> and then using custom selection menus and text boxes to define or modify attributes of the query. The approximate initial query, shown in Figure 1.8 on page 1-23, is generated with the following steps:

1. Select the following from the pull-down menus at the upper-right:
  - Search Type: Search All Data
  - Data Type: Raw Data
  - Example Queries: Get All Master Calibration Data by 7-day Date Range
2. Click the blue startDate text (2), which will open the attribute menu.
3. Click the **Modify t.startDate** text (3), which will open widgets to edit the value.
4. Select the BETWEEN operator from the pull-down menu (4) and enter the start and end dates for the search range (YYYY-MM-DD format) or use the calendar widgets.
5. Press the **Modify Condition** button (5) to update the query text with the new date range.

<sup>12</sup>. Consult the KPNO telescope schedule at [http://www.noao.edu/kpno/forms/tel\\_sched/](http://www.noao.edu/kpno/forms/tel_sched/) for dates when NEWFIRM was scheduled.

<sup>13</sup>. See <http://cas.sdss.org/dr7/en/help/howto/search/simplequery.asp> for a tutorial on the similar SDSS Query Builder.



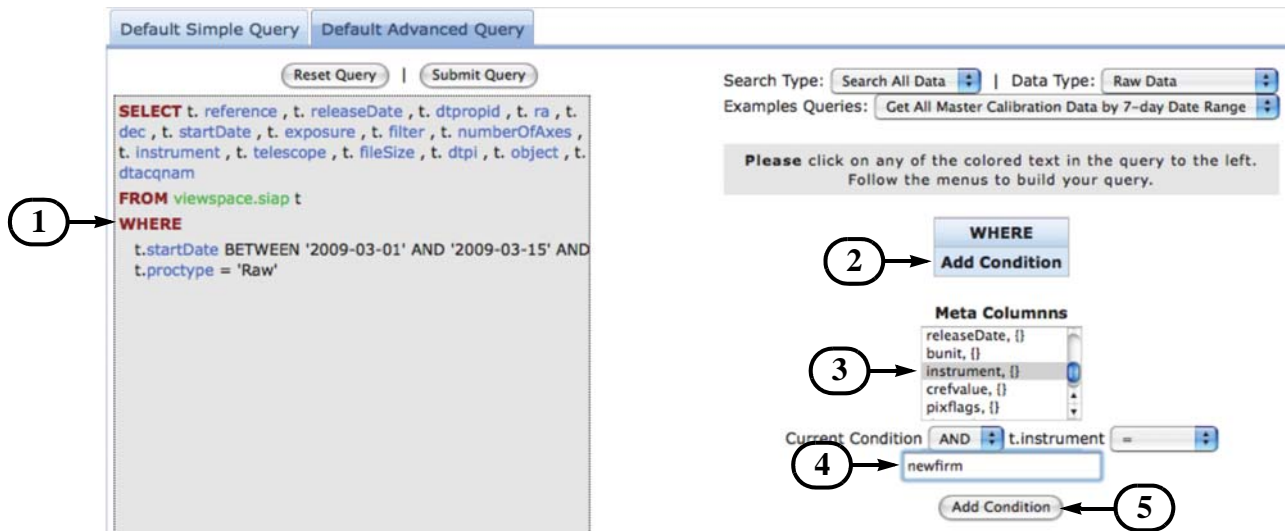


**FIGURE 1.8:** Portal **Default Advanced Query** panel, showing the editable query string (*left*) and input boxes (*right*). After clicking the blue *startDate* text in the query string, it can be updated with values in the boxes by pressing the **Modify Condition** button (*bottom*). Numbered callouts are explained in the text.

The next step is to add a new condition to restrict the query to data from the NEWFIRM instrument. This process is illustrated in Figure 1.9 on page 1-24, and consists of the following steps:

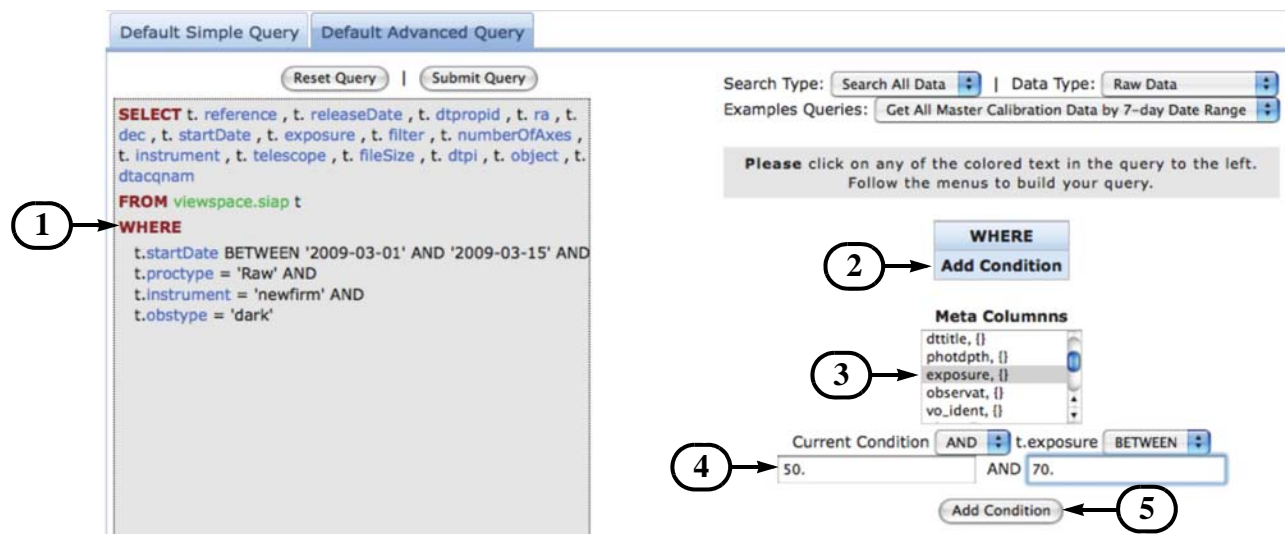
1. Click the **WHERE** text.
2. Click **Add Condition**.
3. Scroll down the Meta Columns selection box, and select instrument.
4. Type “newfirm” into the text box.
5. Press the **Add Condition** button to update the query string.

Now follow the above steps to add a new condition, this time for obstype with the value “dark.” The result of adding this condition is illustrated in Figure 1.10 on page 1-24.



**FIGURE 1.9:** Portal **Default Advanced Query** panel, showing the process for adding a new condition to the query: the instrument name. See text for details.

The final step is to select for exposure times near 60 s, as illustrated in Figure 1.10. Adding this condition is exactly like adding the other conditions, except for choosing exposure in the selection box (3) and selecting a modest range of exposure times (4) **BETWEEN** 50 and 70 s (since an exact match against a floating-point number is unwise). After pressing the **Add Condition** button (5) and then the **Submit Query** button, which will match hundreds of data files, select files of interest in the Data Grid and load them into the cart for staging, as in prior examples.



**FIGURE 1.10:** Portal **Default Advanced Query** panel, showing the process for adding an exposure time as a query condition. See text for details.

Searching for pipeline-produced *Master Calibration* reference files is very similar to the above example, except that the initial default query ([Get All Master Calibration Data by 7-day Date Range](#)) should instead search on the Data Type Master Calibrations.

## 1.4 DATA USE POLICY

There are a number of conditions that apply to the use of data from the NOAO Science Archive; see <http://archive.noao.edu/nsa/condOfUse.html> for details. In general, AURA (or one of the partner observatories), possibly in addition to the PI of the program, assert ownership of the intellectual property that includes the data and derived products (including catalogs, but not including scientific papers), subject to a royalty-free license to the U.S. federal government. AURA policy provides for free use of these data for scientific research, following a limited period during which the investigator of a program has exclusive (proprietary) rights to those science data. Use of the data and derived products also requires that users provide appropriate attribution (see below). Nonproprietary data include the following:

- all data obtained with any program to characterize the instrument health and performance, including CCD bias and dark frames, flat-fields (including dome- and twilight-flats), comparison arc-lamp exposures, and focus exposures;
- all contemporaneous weather data, all-sky camera images, and seeing monitor data;
- all calibration reference frames that are produced by NOAO calibration pipelines, whether generated from on-sky science frames or not (except that metadata that identify the host science program will be redacted);
- all metadata from all exposures as soon as they are archived, including those from science frames, unless specific authorization has been granted by the NOAO Director to withhold these metadata for the duration of the proprietary period; and
- all science data for which the proprietary period has expired.

The applicable proprietary period for science data depends upon the policy of the host observatory, and may be different under some circumstances. The default period for observing programs obtained through the NOAO time allocation process is 18 months from the calendar date of the exposure, except for programs granted “survey” status for which the proprietary period is usually less. Well-justified requests for extensions may be granted by the NOAO Director, although the request should generally be made when the observing proposal is first submitted for consideration.

### 1.4.1 Proper Citation for NOAO Data

Authors of papers that make use of data from the NOAO Science Archive should include one of the following credit statements in their publications:

“National Optical Astronomy Observatory / Association of Universities for Research in Astronomy / National Science Foundation”

The above citation applies to data collected and processed by NOAO. If you use high-level science products that were generated by one of the Survey Project PI teams, the following credit statement should appear in your publication:

“This research draws upon data provided by [*Survey PI*] as distributed by the NOAO Science Archive. NOAO is operated by the Association of Universities for Research in Astronomy (AURA), Inc. under a cooperative agreement with the National Science Foundation.”

The identity of the Survey PI is given on the archive holdings<sup>14</sup> Web page.

## 1.4.2 Referencing Datasets in Journals

The professional journals of the American Astronomical Society have developed a means to reference specific datasets within the manuscripts of submitted papers. The mechanism has been described by Accomazzi, Eichhorn & Rots (2007). It is helpful to the various archive centers to provide this information so that archive usage statistics can be compiled, and so that links to valuable science results from the original datasets can be offered to future archive users. There is now an AASTeX macro that lets authors include Data Set Identifiers in their papers that will provide a link from the electronic edition of the manuscript to the NOAO Science Archive. One of the AASTeX macros is:

```
\dataset[DatasetID]{text}
```

Here, the *DatasetID* is the identifier for your NOAO dataset; this has not been implemented yet in NOAO data headers. Once these identifiers, and the supporting VO services, have been defined for data in the NOAO Science Archive, you will find this information in the value of the DS\_IDENT header keyword.

## 1.5 REFERENCES & FURTHER INFORMATION

### Contributing Authors

People who have contributed to the description of the NOAO Science Archive, the data formats, and the NOAO VO Portal, include Dick Shaw, Chris Miller, Mark Dickinson, Howard Lanning, and Rob Seaman.

<sup>14</sup>. <http://archive.noao.edu/nsa/holdings.html>

## References

- Accomazzi, A., Eichhorn, G., & Rots, A. 2007, in ASP Conf. Ser. 376, ADASS XVI, ed. R. A. Shaw, F. Hill, & D. Bell (San Francisco: ASP), 467
- FITS 2008, *Definition of the Flexible Image Transport System*, IAU FITS Working Group (Version 3.0: Greenbelt, MD: FITS Support Office, NASA/GSFC); available at [http://fits.gsfc.nasa.gov/fits\\_standard.html](http://fits.gsfc.nasa.gov/fits_standard.html)
- Miller, C. J., Gasson, D., & Fuentes, E. 2007, in ASP Conf. Ser. 376, ADASS XVI, ed. R. A. Shaw, F. Hill, & D. Bell (San Francisco: ASP), 625
- Pence, W. D., Seaman, R., & White, R. L. 2009, PASP, 121, 414
- Shaw, R. A. 2007, *Essential Imaging Metadata*, DPP System Technical Report STR2007-01 (Tucson: NOAO)
- Smith, R. C., Dickinson, M., Lowry, S., Miller, C. J., Trueblood, M., & Valdes, F. 2007, in ASP Conf. Ser. 376, ADASS XVI, ed. R. A. Shaw, F. Hill, & D. Bell (San Francisco: ASP), 615
- Zárate, N., Seaman, R., & Tody, D. 2007, in ASP Conf. Ser. 376, ADASS XVI, ed. R. A. Shaw, F. Hill, & D. Bell (San Francisco: ASP), 351

## For Further Reading

Additional documentation on the NOAO VO Portal, including examples and tutorials, is available at <http://nvo.noao.edu/noaonvo/help.shtml>.



## Chapter 2

# Mosaic Cameras

Version 1.1, May 2009

### In This Chapter...

Instrument Overview	2-1
Data Products	2-7
Calibration	2-15
Sources of Error	2-27
References & Further Information	2-30

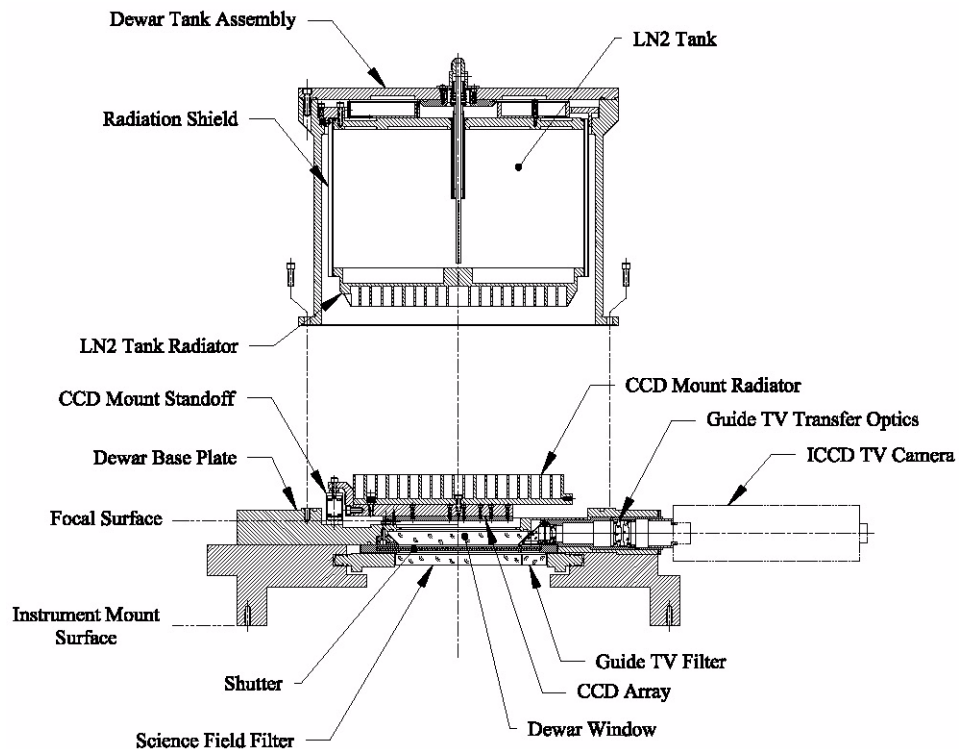
The Mosaic cameras have been in operation at NOAO since the late 1990s, and a good deal of technical literature exists on the instrument design, operation, and performance. The material presented here is drawn primarily from the *NOAO CCD Mosaic Imager User Manual* (Schweiker & Jannuzi 2004), and from the *NOAO Deep Wide-Field Survey Mosaic Data Reductions* notes by Jannuzi, Claver, & Valdes (2003). Interested users should consult these authoritative documents for details of the instrument configuration and operation. Additional resources are cited throughout this chapter and are listed along with other background material in the last section.

### 2.1 INSTRUMENT OVERVIEW

The Mosaic cameras are among the most heavily used instruments offered by NOAO, as measured both by number of nights awarded to principal investigator (PI) programs, and by the volume of data in the NOAO Science Archive. There are actually two Mosaic cameras: **Mosaic-1** is used on both the Mayall 4-m and the WIYN 0.9-m telescopes on Kitt Peak, and **Mosaic-2** is used on the Blanco 4-m telescope at Cerro Tololo. This chapter will cover all usages of both cameras; differences in the instrument design, usage, or other technical details that affect the data will be called out.

## 2.1.1 Instrument Capabilities & Design

The Mosaic cameras are optical/near-IR imagers with a wide field of view (FoV) and a large focal plane array (FPA) containing a total of  $8096 \times 8096$  pixels. A schematic of the instrument is shown in Figure 2.1. The detectors reach peak quantum efficiency (QE) above 85% at  $\sim 6500 \text{ \AA}$ , but have useful sensitivity from the atmospheric cut-off near  $3200 \text{ \AA}$  to  $\sim 1 \text{ \mu m}$ . Both KPNO and CTIO have a large number of filters<sup>1</sup> that are designed to be used with these cameras, up to 14 of which can be held in the *filter track* at one time. The filter sets include standard broad bandpasses, such as Harris *UBV* and Cousins *R*, SDSS *g'r'i'z'*, narrowband filters near  $H\alpha$ ,  $[\text{O III}] \lambda 5007$ , and various other nebular emission lines. Other filters, including filters provided by PIs for their observing programs, have been used occasionally.



**FIGURE 2.1:** Cross-section view of the Mosaic dewar, also showing the arrangement of the CCD array and the filter. Light enters from below in this diagram and passes through the corrector optics (not shown).

The details of the observing configurations for the Mosaic cameras are given in Table 2.1. The cameras are mounted at the Prime focal station of the 4-m telescopes, and at the Cassegrain focus of the 0.9-m, with converging beams as indicated in the Table. The detectors sample the delivered point spread function

1. Filter properties can be viewed at <http://www.noao.edu/kpno/mosaic/filters/filters.html>.



(PSF) very well; image quality is seeing-limited at all wavelengths. The image quality at the 4-m telescopes is good, with little (<10%) focus gradient or PSF variation across the field of view. The pixel scale is slightly variable from the center to the edge of the field of view due to pincushion distortions. The image quality at the 0.9-m telescope is not as good: there is a small focus gradient across the FoV, resulting in degraded images near the corners of the field. Also, the corners of the field are slightly vignetted (~5–10%) by the internal telescope baffle. The image scale at the 0.9-m is relatively constant across the FoV.

**TABLE 2.1: Mosaic Observing Configurations**

	KPNO 4-m	CTIO 4-m	WIYN 0.9-m
Field of view	36' × 36'	36' × 36'	59' × 59'
Pixel scale @center/edge of FoV (arcsec/unbinned pixel)	0.261 / 0.245	0.261 / 0.245	0.425 / 0.420
ADC <sup>a</sup> usage	Recommended	Recommended	N/A
Telescope focal ratio	f3.1	f2.9	f7.9

a. Atmospheric Dispersion Corrector.

An atmospheric dispersion corrector (ADC) is available on the 4-m telescopes for the Mosaic cameras (see Jacoby, et al. 1998 for details). The effect of atmospheric dispersion is greatest for very wide filters, and especially in the blue (stretching a stellar image by 0.9 arcsec in the *U* filter at a zenith distance of 60°, or 2 airmasses). Because of the wide FoV of the Mosaic cameras, differential atmospheric refraction within the FoV is noticeable even at the zenith.

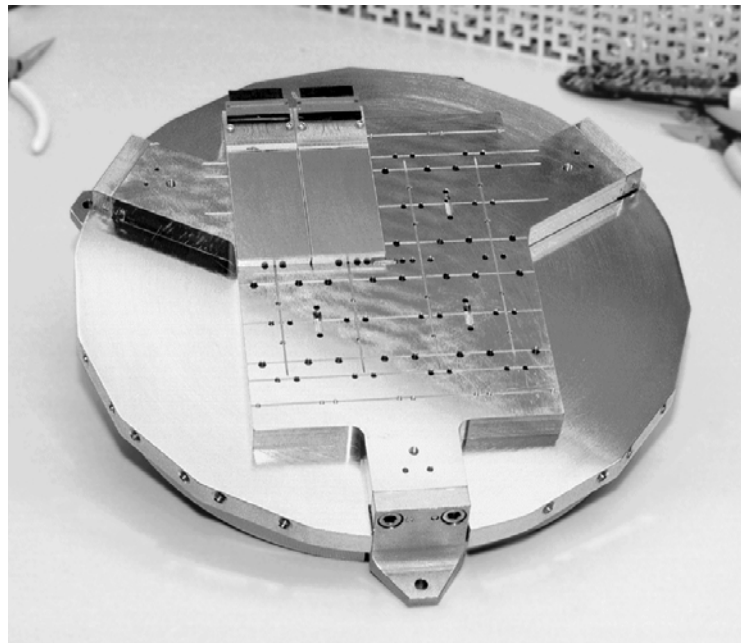


Note that the ADC is used at the discretion of the observer, so that subsequent users of the 4-m Mosaic data will need to examine the value of the ADC header keyword to determine if any given image was obtained with the ADC. Since no ADC is available at the WIYN 0.9-m telescope, images obtained there at high airmass will suffer significant image degradation.

The shutter for the Mosaic cameras consists of a pair of opposing sliding blades, which serves to regulate the exposure time. The motion of the shutter blades is along columns, and the time-of-flight is approximately 23 ms. The accuracy of the shutter timing is good, but not perfect: one experiment showed that a 1 s exposure illuminated the FPA for 0.97 s at the top of the array, and 0.96 s at the bottom. A shutter shading correction is in principle possible, but most Mosaic science data are obtained with exposures greatly exceeding 1 s so that the correction would be negligible.

## Focal Plane

The focal plane of the Mosaic cameras is populated with an array of eight SITE CCD detectors, arranged in a  $4 \times 2$  mosaic, as shown in Figure 2.2. Each of the thinned CCD detectors has two amplifiers to provide the capability of reading the arrays using either or both amplifiers. The Mosaic-2 camera is generally operated with all 16 amplifiers. The Mosaic-1 camera does not have working dual amplifiers on all CCDs, and so can only be used with one amplifier per CCD. All the detectors have very few cosmetic defects.



**FIGURE 2.2:** Mounting for the Mosaic-1 detectors, showing two of the eight 2048 x 4096 pixel CCD arrays in position.

The detector properties, or ranges of properties, are given in Table 2.2. The linearity is good, but the dynamic range is modest by modern standards and varies from detector to detector. The charge transfer efficiency (CTE) when the detectors were first installed was excellent (0.999995 or better), and is not believed to have degraded significantly since then. The pixels may be binned at read-out by integral factors. Observers have selected this option with some regularity to reduce the read-out time (e.g., for observations of variable phenomena or moving targets that require a rapid cadence), or to reduce the read noise per binned pixel (compared to post-read down-sampling). The penalty, of course, is the possibility of under-sampled images and reduced spatial resolution, depending upon the delivered image quality at the time of the observations. Finally, the QE varies somewhat from detector to detector, both on average and as a function of wavelength. (See the Mosaic instrument manuals for details.) Therefore the

transformation from instrumental magnitudes to standard systems depends upon the detector. See “Photometric Calibration” on page 2-21 and “Photometry” on page 2-28.

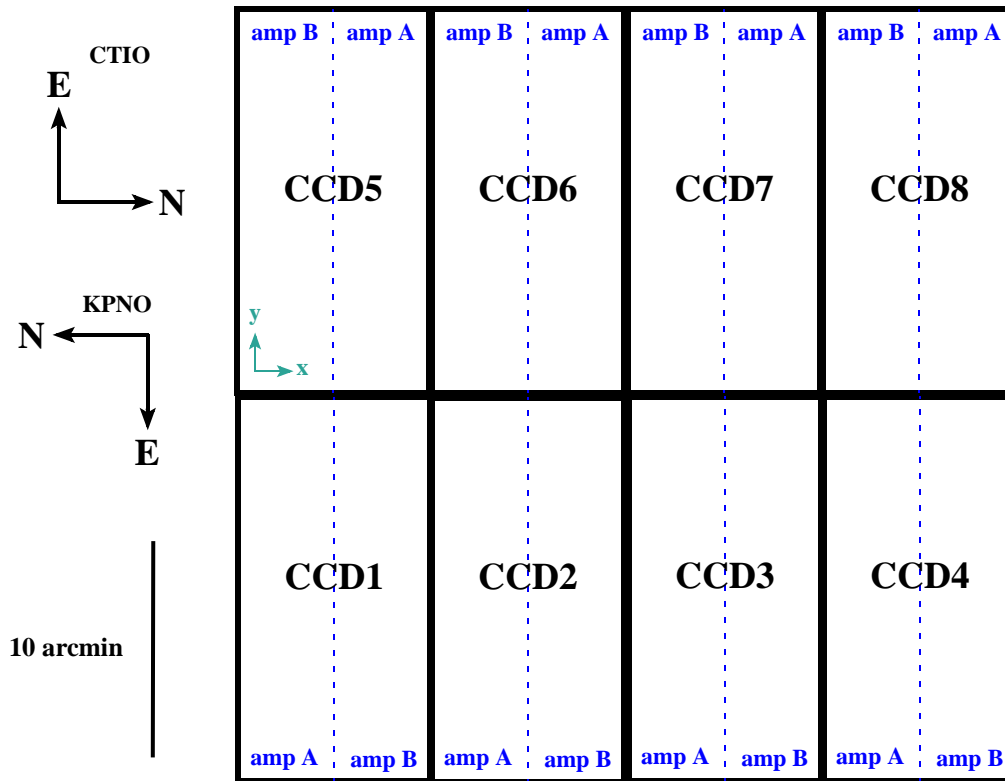
**TABLE 2.2: CCD Array Characteristics**

Array Dimensions	
Axis 1:	2048 pixels
Axis 2:	4096 pixels
Pixel Size	15 $\mu\text{m}$ square
Binning factors <sup>a</sup>	1, 2, 3, 4
Gain	1.8 to 3.3 $e^-/\text{ADU}$
Read noise	5.0 to 9.0 $e^-$
Max. linear count	21,000 to 45,000 $e^-$
CCD Gaps	
in rows:	50 unbinned pixels
in columns:	35 unbinned pixels
Exposure overheads <sup>b</sup>	
binned 1 $\times$ 1	2:34 (min:sec)
binned 2 $\times$ 2	1:06 (min:sec)

a. Configurable by observer, independently for each image axis.

b. Times dominated by detector read-out, assuming a single amplifier.

Figure 2.3 shows the arrangement of the FPA on the sky, which differs between hemispheres, using the same nomenclature as is found in the science data extension headers. Note that the raw image array coordinates are remapped from the detector coordinates, such that the origin is always in the lower-left corner when oriented as in the figure. For reduced data, the images are transposed so that North is up and East is to the left.



**FIGURE 2.3:** Orientations on the sky and spatial footprint of the FPA for both Mosaic cameras, with CCD detector designations indicated. Two amplifiers on each chip may be used for parallel read-out (CTIO Mosaic-2 only) over areas bounded by dotted lines. Image coordinate axes are also indicated (small arrows); pixel coordinate origin is in the lower-left corner of each detector array (or sub-array).

### Operations

Most data obtained with the Mosaic cameras have a binning of  $1 \times 1$  or  $2 \times 2$  pixels. Many observers obtained multiple exposures of their fields using the same filter in order to reject cosmic rays; often this was combined with a sequence of small spatial dithers in order to observe in the gaps between the CCDs in the FPA, and to enable the construction of continuous regions of sky free of gaps and detector artifacts. A few observing programs construct sequences of (slightly) overlapping images to map large regions of sky from the component tiles. Finally, most observing programs obtain calibration frames, such as bias (zero) frames and dome flats, in order to facilitate calibration. See “Calibration” on page 2-15 for a detailed discussion of how the calibration frames are used in the data reduction process.

## 2.2 DATA PRODUCTS

This section describes the content and format of the various data products that are produced for the Mosaic cameras. Most of the products are generated during the course of calibration processing, the details of which are discussed in the next subsection. The data products can be distinguished by the combination of the PROCTYPE, PRODTYPE, and OBSTYPE keywords in the primary header; the possible values are summarized in Table 2.3. The processing level (see Table 1.2, “Levels of Data Processing,” on page 1-4) at which the product is generated is listed in column 5 (Proc. Level).

**TABLE 2.3: Data Product Content Type**

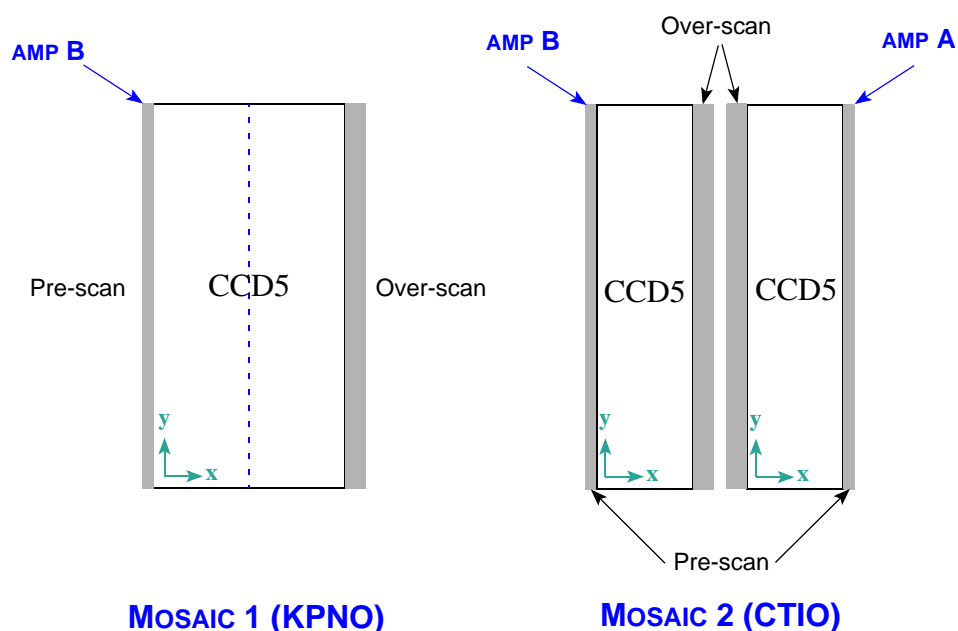
PROCTYPE	PRODTYPE	OBSTYPE	Extensions	Proc. Level	Description
Raw	image	object   zero  dark   dome flat	IMAGE * N <sub>ext</sub>	1	Raw data as obtained at the telescope, with additional metadata included in the header. Note: OBSTYPE value unreliable in raw data.
InstCal	image	object	IMAGE * N <sub>ext</sub>	2	Calibrated, single-frame reduced image with instrument signature removed, WCS and rough photometric calibrations applied
InstCal	dqmask	object	BINTABLE * N <sub>ext</sub>	2	Data quality mask for InstCal
InstCal	png	object	FOREIGN*2	2	Down-sampled preview for InstCal
Resampled	image	object	[none]	2	Calibrated, re-projected image
Resampled	dqmask	object	BINTABLE	2	Data quality mask for resampled image
Resampled	png	object	FOREIGN*2	2	Preview for resampled image
Stack	image	object	[primary]	3	Stack of 2 or more overlapping images
Stack	dqmask	object	BINTABLE	3	Data quality mask for stacked image
Stack	expmap	object	BINTABLE	3	Exposure map for stacked image
Stack	png	object	FOREIGN*2	3	Preview for stacked image
MasterCal	image	zero	IMAGE * N <sub>ext</sub>	2	Bias structure image template
MasterCal	png	zero	FOREIGN*2	2	Preview for bias structure
MasterCal	image	dome flat	IMAGE * N <sub>ext</sub>	2	Dome flat-field template image
MasterCal	png	dome flat	FOREIGN*2	2	Preview for dome flat
MasterCal	image	sky flat	IMAGE * N <sub>ext</sub>	2	Sky delta-flat-field image
MasterCal	png	sky flat	FOREIGN*2	2	Preview for sky flat
MasterCal	image	pupil	IMAGE * N <sub>ext</sub>	2	Pupil ghost template image; generated only for Mosaic-1 data obtained at KPNO
MasterCal	png	pupil	FOREIGN*2	2	Preview for pupil ghost template
MasterCal	image	fringe	IMAGE * N <sub>ext</sub>	2	Fringe pattern template image
MasterCal	png	fringe	FOREIGN*2	2	Preview for fringe pattern

## 2.2.1 Image Formats

The image data from the Mosaic cameras is stored either in FITS multi-extension files (MEFs), the general structure of which was described in Chapter 1, or, for some reduced data, in simple FITS files with no extensions. Although most images are stored in extensions, the detailed arrangement of the image portions among the extensions differs depending upon whether the data are raw (unprocessed) or reduced, and upon the operational practice for each Mosaic camera.

### *Raw Data*

Raw data from the MOSAIC cameras consists of 16-bit unsigned integers, and includes virtual over-scan along each row at the beginning (pre-scan: 24 pixels) and end (over-scan: 48 or 64 pixels) of the CCD read-outs, which is stored with the image pixels as shown in Figure 2.4. Note that the coordinate origin for all images is in the lower-left corner of the read-out section (for the convenience of image display), rather than at the location of the read-out amplifier. While each detector in both Mosaic cameras has two read-out amplifiers, in practice only the Mosaic-2 camera uses both amplifiers in parallel to read both halves of the arrays simultaneously. The output from each amplifier (including the over-scan regions) is stored in a separate image extension in the FITS MEF file (see Chapter 1); thus, there are as many image extensions in the raw science file as the total number of amplifiers used to read out all detectors in the focal plane. The size and location of the photo-active regions and the over-scan are given in Table 2.4. Note that the size of the region used for determining the bias level is smaller than the size of the over-scan region, and that the pre-scan area is not used to determine the bias level.



**FIGURE 2.4:** Schematic of the image arrays just after read-out for detector 5 of the Mosaic camera used at Kitt Peak (*left*) and Cerro Tololo (*right*). Virtual over-scan regions are indicated in grey, but are expanded horizontally by a factor of 4 for clarity. Amplifier locations are indicated, as is the orientation and approximate origin of the pixel coordinate axes (small arrows).

**TABLE 2.4: Raw Data and Over-scan Regions**

Raster Dimensions	Detectors	Amp	Photo-Active Data Section	Bias Section (Columns)
<b>Mosaic-1</b>				
2136 × 4096	1–4	B	[65:2112, 1:4096]	[1:50]
	5–8	B	[25:2072, 1:4096]	[2087:2136]
<b>Mosaic-2</b>				
1112 × 4096	1–4	A	[25:1048, 1:4096]	[1063:1112]
		B	[65:1088, 1:4096]	[1:50]
	5–8	A	[65:1088, 1:4096]	[1:50]
		B	[25:1048, 1:4096]	[1063:1112]

## Calibrated Data

The Archive contains data products that are produced with the Mosaic calibration pipeline. The specific calibrated science data products are listed in Table 2.3 on page 2-7 and are described in more detail below. Each image has an associated *data quality mask* (DQM) and a *preview image*, which are described at the end of this subsection. The science images are stored as 16-bit scaled integers, which is a mild form of compression, when the sigma of the sky pixels can be sampled at more than 30:1 with a dynamic range over all non-flagged pixels. That is, the integer digitization has at least 30 values within 1-sigma of the sky distribution, and the 32K values span all good (i.e., non-flagged) pixel values in the calibrated data. If these conditions cannot be met, then the images are stored as 32-bit floating-point values.

**InstCal.** These images have been processed to remove instrumental signature, and have been astrometrically and photometrically calibrated (see “Calibration” on page 2-15), although there are some cases where the calibration can fail (see “Sources of Error” on page 2-27). The data are organized almost identically to that for the raw science data, except for the following:

- The over-scan regions have been trimmed from the image arrays.
- Image portions for a given detector that were read by different amplifiers have been joined. Thus, even science files from the Mosaic-2 camera contain 8 image extensions, one for each detector.
- The resulting images in each extension are all  $2048 \times 4096$  pixels.

**Resampled.** These images are the result of geometrically rectifying InstCal images, where each array has been re-projected to a common tangent-point on the sky, with pixels aligned to a common grid with uniform scale. The pixels have been transposed, so that when displayed, images will appear with North “up” (i.e., declination increases along Axis 2) and East “left” (i.e., Right Ascension decreases along Axis 1). The images are stored in the primary header-data unit within the file. The tangent-point is selected from a grid with roughly  $1^\circ$  steps in RA, and exactly  $1^\circ$  in Dec. The aim is to maximize the probability that exposures from different nights and processing runs may be combined without requiring further interpolation. Within a single night, all overlapping exposures, such as from dither sequences or serendipitous overlapping pointings, are assigned a common tangent point even if some elements of a dither would have been assigned to different grid points when considered independently. The resampled images are approximately  $8526 \times 8642$  pixels; the actual size depends on the degree of differential atmospheric refraction and the WCS tangent point selected for a particular image.

**Stacked.** If two or more observations of a given target are obtained on the same night using the same filter and have a sufficient degree of spatial overlap, these images are combined using an average with outlier rejection to remove



detector blemishes, gaps between the detectors, and artifacts such as image persistence and cosmic rays. The result is a union of the spatial footprints of the stack, which is stored in a simple FITS image (i.e., not in image extensions), with nearly the same pixel scale as the raw images. In general, these images will be larger—sometimes very much larger—than the resampled images because the area of the sky that is mapped can be significantly larger than the instrument FoV.




---

The exposure duration for stacked images, as recorded in the `EXPTIME` keyword, refers to the sum of all exposure durations of all images used to create the stack. The exposure depth and noise properties of a stack of dithered images is a discontinuous and possibly complicated function of position in the image. Use the exposure map to track the detailed exposure depth at the pixel level.

---

**Master Calibration.** Reference files are created during the course of pipeline processing, such as bias structure, flat-fields, etc. These files are used in pipeline processing to remove instrumental signatures from the science data. (See “Calibration Reference Files” on page 2-23.) These reference files are 32-bit floating-point images, stored as MEF files with as many extensions as the corresponding raw images.

**Concomitant Data.** All reduced images are accompanied by data quality masks (DQM); stacked images are accompanied in addition by exposure maps. The DQMs and exposure maps are logically images, but they are stored in compressed form as FITS binary tables, with one table per extension in the science image. Some utilities (such as the IRAF `XImtool`) can display and work with these files directly, but other utilities may require that the files be translated to FITS images before they can be used. The pixel values in the DQMs are non-zero when affected by detector pathologies and image artifacts such as bleed trails, saturation, and cosmic rays; the values are zero otherwise. Table 2.5 lists mask values that are applicable to *nonstacked* images.

**TABLE 2.5: Data Quality Mask Values**

Value	Meaning
0	No problem
1, 6	Bad/compromised pixel identified in static DQM; generally indicates known detector blemishes
2	[Not used]
3	[Not used]
4	Saturated (plus padding)

**TABLE 2.5: Data Quality Mask Values (Continued)**

Value	Meaning
5	Bleed trail (plus padding); indicates possibly compromised pixels
7, 8	Transient artifacts (includes cosmic rays)

Most *stacked* images are created from dithered frames, so that most of the pathologies affecting single frames will not be applicable. DQMs for stacked images have the following values: “1” for areas of the re-projected image with no data, “2” for areas with no data *after rejection*, and zero otherwise.

The exposure maps are images whose values are the cumulative exposure duration at each pixel, which can be a complicated function of position. The value of the `EXPTIME` keyword in a stacked image is the sum of the exposure durations for all images that contributed to the stack.

**Ancillary Files.** The preview images are down-sampled versions of the files they accompany and are stored in Portable Network Graphics<sup>2</sup> (PNG) format. There are two samplings for each file, which are intended mainly for quick display of the results of pipeline processing on Web pages. Note that these files have been embedded in FITS files as extensions of type `FOREIGN` so that most users will need to extract them with some tool before viewing.

## 2.2.2 Header Keywords

A wide variety of metadata are recorded in the headers of the science frames. Users should review these headers (and the extension headers) to familiarize themselves with the content. The more critical metadata are described in this subsection. Table 2.6 lists metadata by the keyword name, the header unit in which the keyword will be found (Primary or Extension), the point in the data processing where the keyword is introduced (or where the value is updated), and the meaning of the keyword (or group of keywords, if they are related). Some of the keywords are indexed by image axis, meaning they come in pairs, as indicated by the suffixes *i* and *j*.

2. <http://www.libpng.org/pub/png/>

TABLE 2.6: Important Image Keywords

Keyword Name	HDU	Origin	Meaning
<b>Telescope</b>			
ADC	P	R	Identification of the Atmospheric Dispersion Corrector (ADC)
ADCSTAT	P	R	ADC tracking status
AIRMASS	P	R	Atmospheric pathlength for target at observation start
OBSERVAT	P	R	Observatory that operates this telescope
TELESCOP	P	R	Telescope used to obtain these data
TELDEC	P	R	Declination for the telescope position on the sky in degrees
TELRA	P	R	Right ascension for the telescope position on the sky in hours
<b>Instrument/Detector Configuration</b>			
CCDSUM	E	R	CCD binning factors along Axis1 and Axis2
DETECTOR	E	L2	Detector designation
FILTER	P	R	Filter name/designation
GAIN	E	R, U	Detector effective gain, in $e^-/ADU$
INSTRUME	P	L2	Instrument name
OGAIN	E	L2	Original detector gain setting for the exposure, in $e^-/ADU$ (same as GAIN for raw images)
<b>Time</b>			
DATE-OBS	P	R	Date and time of observation start
EXPTIME	P	R	Effective exposure duration in seconds
MJD-OBS	P	R	Time of observation start in MJD
TIME-OBS	P	R	Time of observation start
TIMESYS	P	R	The principal time system for all time-related keywords. Always UTC.
<b>World Coordinates</b>			
$CDi_j$	E	R, U	Transformation matrix from pixel to intermediate world coordinates; $CDi_i$ is the pixel scale for axis $i$
$CRPIXi$	E	R, U	Location of the reference point along axis $i$ in units of pixels
$CRVALi$	E	R, U	Value of the world coordinate at the reference point for axis $i$ in degrees
$CTYPEi$	E	R	Name of the coordinate represented in axis $i$
DEC	P	R, U	Declination for the center of the detector FoV in degrees
EQUINOX	E	R	Equinox in years for the celestial coordinate system in which the positions are expressed
$NAXISi$	E	R	Number of pixels along axis $i$
$PIXSCALi$	E	R	Pixel scale along axis $i$ in arcsec/pixel

TABLE 2.6: Important Image Keywords (Continued)

Keyword Name	HDU	Origin	Meaning
RA	P	R, U	Right ascension for the center of the detector FoV in hours
RADESYS	E	R, U	Name of the reference system in which the world coordinates are expressed
WATi_nnn	E	R, U	IRAF-specific description of the nonlinear portion of the transformation from detector to world coordinates for axis <i>i</i> . This character string contains coefficients for a polynomial; the length of the string is such that it must continue for <i>nnn</i> FITS header records.
<b>Calibration</b>			
BLDPROC	E	L2	Bleed trail processing parameters (bleed threshold & grow radius)
BUNIT	E	L2, L3	Brightness units, normally “electron/s” for calibrated images
MAGZERO	E	L2	Magnitude corresponding to one count in the image
OBSTYPE	P	R, U	Type of target observed (see Table 2.3 on page 2-7)
PHOTBW	P	L2	RMS width of bandpass (Å)
PHOTCLAM	P	L2	Central wavelength of bandpass (Å)
PHOTDPTH	P	L2	Photometric depth of the exposure. (See “Photometric Calibration” on page 2-21.)
PHOTFWHM	P	L2	FWHM of bandpass (Å), i.e., width measured at 50% of peak transmission
PIPELINE	P	L2, U	Pipeline name
PLVER	P	L2, U	Pipeline version identifier
PROCTYPE	P	L2, U	Product type (see Table 2.3 on page 2-7)
PRODTYPE	P	L2, U	Product data description (image   mask   expmap)
SATPROC	E	L2	Saturation processing parameters (saturation threshold & grow radius)
SEEING	P	L2	Average FWHM of point sources in arcseconds
SKYBG	P	L2	Brightness level of background averaged over all CCDs in ADU
SKYBG1	E	L2	Brightness level of background in single CCD in ADU
SKYNOISE	P	L2	RMS noise in the background level in ADU

### 2.2.3 Environmental Data

At present no environmental data are accessible from the Archive, although data from all-sky cameras and seeing monitors is planned. A terse summary of sky conditions at CTIO<sup>3</sup> and at KPNO<sup>4</sup> are available on the Web.

3. [http://www.ctio.noao.edu/site/phot/sky\\_conditions.php](http://www.ctio.noao.edu/site/phot/sky_conditions.php)

4. [http://www-kpno.kpno.noao.edu/Info/Mtn\\_Weather/](http://www-kpno.kpno.noao.edu/Info/Mtn_Weather/)

## 2.3 CALIBRATION

The current generation of pipeline processing produces Level-2 products, i.e., images where the instrumental signature has been removed and geometric and photometric calibrations have been applied; and Level-3 products, where spatially overlapping images in the same filter (and that have been obtained within the same observing run) have been stacked. The pipeline uses both calibration exposures, such as bias frames and dome flats, as well as portions of science exposures (i.e., areas free of defects, artifacts, and astrophysical objects) to construct the calibration reference files that are used as input to the processing. For simplicity, the subsections below will first describe the processing of the science images once the calibration reference files are available; then the processes for constructing the calibration reference files will be described with the overall flow as context. The actual sequence of processing in the pipeline software differs somewhat in detail, in part to optimize the performance of the parallel processing environment.




---

The calibration pipeline for Mosaic reductions is very much tuned to maximize the use of exposures within the observing run in which they were obtained. As such, the quality of the data depends to a large degree on the quality of the calibration exposures that were obtained by the observer. However, under some circumstances, the pipeline will attempt to use calibration data from prior observing runs if the separation in time is not too great; the specific rules are described in the following sections.

---




---

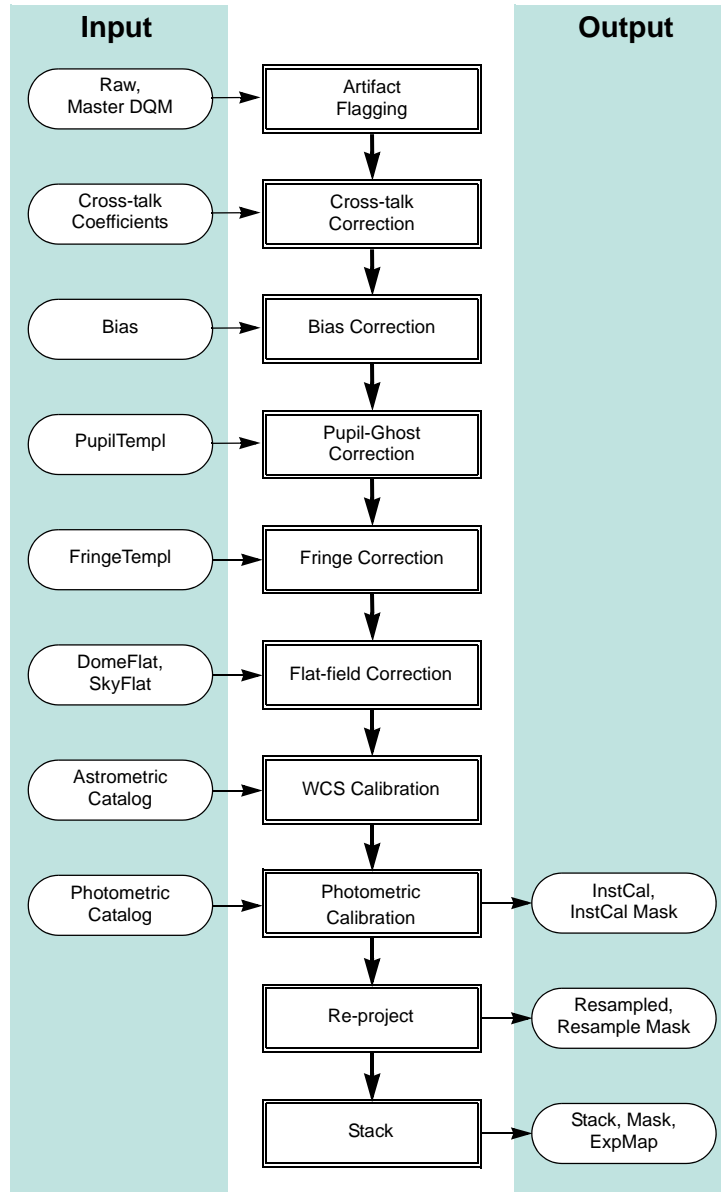
This pipeline *does not* process Mosaic data that were obtained with the WIYN 0.9-m telescope. This pipeline *only* processes data obtained with a CCD binning of  $1 \times 1$  (using either 4-m telescope).

---

### 2.3.1 Processing Steps

The flow of the science data through the pipeline is shown in Figure 2.5. Each step of the processing, indicated by the boxes in the center of the figure, is described in detail in the following subsections. As explained below, some steps are not performed for data where the correction in question is not needed. Inputs to the processing include the raw science frames, calibration reference files (see “Calibration Reference Files” on page 2-23), and photometric and astrometric catalogs. Outputs include the various reduced science images, plus their associ-

ated mask files (see “Image Formats” on page 2-8). Intermediate products that are produced during the course of pipeline processing, but that are not archived, are not shown. The processing is segregated by night, filter, and further by overlapping observing sequences when deriving stacked frames.



**FIGURE 2.5:** Flow of data taken with a common filter through pipeline processing and calibration to produce Level-2 and Level-3 data products. Note that pupil ghost correction only applies to Mosaic-1 data from the 4-m telescope. External catalogs, and data products defined in Table 2.3, are shown as inputs or outputs of the processing.

## Artifact Flagging

Various image artifacts, including detector defects marked in the Master DQM reference file, are flagged in the DQM for each science image. In addition, pixels with values above the saturation threshold for the parent CCD are flagged at the beginning of the processing. Further artifact flagging occurs later in the processing, including *transients* (e.g., moving objects and cosmic rays) which are detected during the final image stack. Note that the regions flagged using the saturation and bleed trail thresholds are grown by up to a few adjacent pixels to minimize the risk of compromised data near artifacts. The final data quality mask for each InstCal image indicates all detected artifacts noted above; these values are propagated to the Resampled images, which generally increases the number of affected pixels to include areas adjacent to the pixels in the original geometry. All flagged pixels are excluded from contributing to the final, *stacked* image; if this process results in areas with no data, such pixels are so flagged (see “Concomitant Data” on page 2-11 for details).




---

Note that all pixels in calibrated science frames (InstCal, Resampled, and Stacked) that are flagged in the DQM with a value other than zero will be interpolated over. The 1-dimensional linear interpolation is performed along the shortest dimension of the region over which to interpolate. This process is intended to avoid introducing “ringing” artifacts during the down-stream resampling of the images, as well as to mitigate scaling problems when using image display software. The affected pixels retain their flags in the DQM, however.

---

## Cross-Talk Correction

A small but significant amount of *cross-talk* occurs between the readout electronics of pairs of CCDs in the FPA. The effect is to introduce a small fraction of the signal from one CCD into the signal chain of the CCD that shares the same electronics box, such that “ghosts” of bright objects appear in the paired CCD. This is an additive effect of ~0.1%, and is most noticeable in objects that are at or near saturation. The proportionality coefficients differ from one electronics box to another, but are relatively stable over time unless significant maintenance has been performed on the instrument. The processing pipeline applies pre-calibrated proportionality coefficients and subtracts the cross-talk signature from the images.

### ***Bias Correction***

This correction is applied to remove the (additive) electronic bias that is present in the signal chain. The bias is, for the most part, composed of a constant pedestal, but it has structure that is related to the electronic stability of the bias during read-out of the detector. The processing pipeline removes the bias contribution in a two-step process. In the first step, the post-scan region (see Table 2.4 on page 2-9) is averaged along the serial direction, and this average value for each image line is subtracted. The virtual over-scan regions (i.e., the pre-scan and the post-scan: see Figure 2.4 on page 2-9) are then removed from the image. In the second step, the master bias image is subtracted from each science image to remove higher-order structure.

Following the bias correction, the pixels are scaled by the gain factor for the appropriate CCD and normalized to unit exposure time. The brightness units, as expressed in the value of the BUNIT keyword, are electrons  $s^{-1}$  (or equivalently, detected photons  $s^{-1}$ ) for calibrated images.

### ***Dark Correction***

The dark current, i.e., the signal introduced by thermal electrons in the detector with the instrument shutter closed, is extremely low for the Mosaic CCDs, so **no dark correction is presently applied in the processing pipeline**. Were a dark correction to be applied, it would occur after bias correction and would be applied by subtracting a dark master frame from the science frames, after scaling by the ratio of the exposure times.

### ***Pupil Ghost Correction***

The pupil ghost is an artifact that is evident in data from the Mosaic-1 camera when used at the 4-m telescope. The artifact is caused by light that reflects off the filter, off the back surface of the corrector, and returns through the filters to the detector. For Mosaic-1, the effect is evident in raw data as an image of the telescope pupil near the intersection of detectors 2, 3, 6, and 7 (see Figure 2.6 on page 2-20). The amplitude of the effect is greatest where the coating of the corrector is least effective (in the extremes of red and blue), particularly in the  $U$ ,  $I$ , and  $z$  bands, and in narrowband filters of similar wavelength. The pupil ghost template is scaled to the pattern in the science frame and subtracted.






---

The pupil ghost correction is only applied to data obtained with the Mosaic-1 camera on the 4-m telescope on Kitt Peak, and only for exposures taken through broadband or narrowband filters with central wavelengths near *U*, *B*, and *I*. No pupil ghost is evident in Mosaic-1 data when used at the WIYN 0.9-m telescope, or in Mosaic-2 data on the CTIO 4-m telescope.

---

### ***Fringe Pattern Correction***

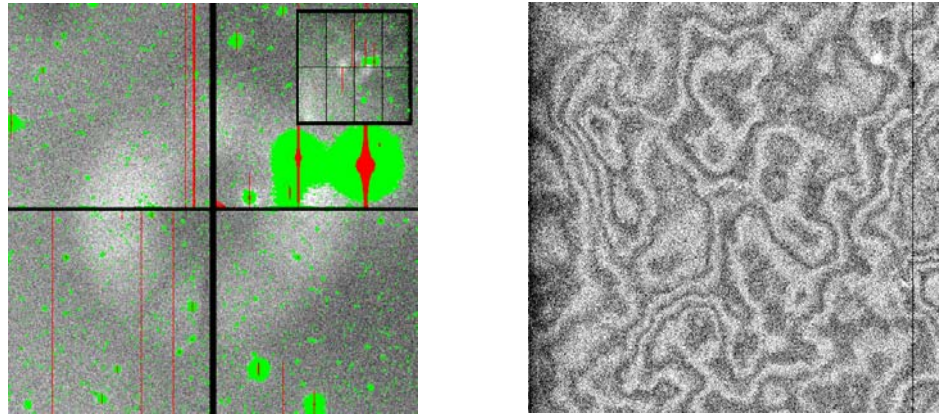
A fringe pattern is evident in data from both Mosaic cameras at certain (mostly red) wavelengths. The pattern occurs because of interference between the incident, nearly monochromatic light from night sky emission lines (from both air glow and reflected city lights) and light that is reflected internally between layers of the CCD substrate. The details of the fringe pattern depend mostly upon the spatial variation in thickness of the top layer in the substrate, but also depend upon a number of other factors including the wavelength(s) of the incident emission lines, the composition of the substrate, the temperature of the CCD, and the focal ratio of the incident beam. (See Walsh, Pirzkal & Pasquali, 2002 for a useful discussion of this phenomenon in the ACS camera on the Hubble Space Telescope.) The amplitude of the fringe pattern background varies with time and telescope pointing. Since the amplitude of the fringes can be large compared to the mean sky background, it is important to remove it where possible. An example of fringing over a portion of an image can be seen in Figure 2.6. To remove the effect, the fringe template is scaled to the science image and subtracted.




---

The fringe correction is only applied to exposures taken through narrowband filters, or through broadband filters with central wavelengths greater than 6290 Å.

---



**FIGURE 2.6:** Example of the pupil ghost pattern (*left*, with bright targets and artifacts masked) near the intersection of the central 4 detectors, and the fringe pattern (*right*) in a 512 x 512 pixel region.

### ***Flat-Field Correction***

The flat-field correction removes the variations in the pixel-to-pixel response of the detector. Ideally the flat-field is derived for each filter from blank sky (i.e., the sum over all science images of the background regions), in order to match the illumination of the pixels in science images as well as the color of the background. However, if there is insufficient illumination of the background pixels it may not be possible to construct a sky flat of sufficient quality (see “Sky Delta-Flat” on page 2-26). In any case, the flat-field correction is performed by dividing each science frame by the flat-field templates for the corresponding filter: first by the dome flat, and then by the sky flat (if available).

### ***World Coordinate System Calibration***

The WCS calibration for science images is described by a two-dimensional polynomial (the function type and coefficients are found in the header) of a tangent-plane projection of stellar coordinates to the image pixel grid. The lower-order terms relate to the location of the reference pixel on the sky, the plate scale, and the rotation of the image, while the higher-order terms relate to non-linear effects of the optical system (such as image distortion) and differential atmospheric refraction. The form and order of the function varies with the filter and with the airmass of the observation. The initial values for all but the zero-point terms are taken from prior, full WCS solutions of calibration images where the astrometry of the stars in the field is known with very high accuracy; the initial value for the telescope pointing is taken from the TELRA and TELDEC header keywords, which are supplied by the telescope control system.

A full WCS solution is determined for each CCD by associating centroids of stars in the science images with objects in an astrometric catalog. Since the number of suitable stars in any given image (and their distribution within the image) can vary considerably, depending upon such details as the exposure time, ambient seeing, and clouds, the input list to the WCS solution includes roughly 100 additional pseudo-stars, which are merely computed points from the trial WCS solution. Thus, if the number of genuine stars in the image is small compared to the pseudo-stars, the high-order terms in the WCS solution will be dominated by the trial solution. If instead the number of genuine stars is large, they will dominate the WCS solution. In all cases, only the genuine stars contribute to determining the zero-point. The WCS solution, while robust, may fail if the telescope pointing is in error by several arcminutes or more (which is rare). See “Sources of Error” on page 2-27 for details.

### Photometric Calibration

An estimate is made of the magnitude zero-point of each science image by comparing the instrumental magnitude of each field star to their published magnitudes in nearby passbands in a reference catalog, applying color transformations as necessary. Currently, the reference photometric catalog is USNO-B1.0 (see Monet, et al. 2003). Note that the result of the photometric calibration is to populate the science header with keywords—the pixel values remain unchanged, and have units of detected photons  $s^{-1}$ .

One quantity of use, the photometric depth of the exposure, is defined as:

$$-\frac{2.5}{2.3026} \log\left(\frac{3.988 \cdot DIQ \cdot \sigma}{\sqrt{A}}\right) + m_{zero}$$

In physical terms, the photometric depth is the faintest point-source that can result in an  $5\text{-}\sigma$  detection above the sky background, in units of magnitudes. In the equation above,  $DIQ$  is the delivered image quality (expressed as the value of the header keyword `SEEING` in the image header) in units of arcseconds,  $\sigma$  is the noise in the background in ADU,  $A$  is the area of a pixel in  $\text{arcsec}^2$ , and  $m_{zero}$  is the magnitude zero-point. This quantity is stored in the header of the calibrated images as the value of the `PHOTDPTH` keyword.

At present the photometric calibration is not adequate to support science, and is intended solely to provide a rough estimate of exposure depth to support queries from the Virtual Observatory. See “Sources of Error” on page 2-27 for details.




---

The magnitudes for the USNO-B1.0 catalog were for the most part derived from photographic source material and have high uncertainties. In addition, the mapping from filters used with the Mosaic cameras to USNO-B bandpasses, and then the color transformation from USNO-B to the SDSS filter set, introduces substantial (but uncharacterized) additional uncertainties. The resulting magnitude zero-point and photometric depth given in the headers of Mosaic science data are uncertain by an unknown amount that may approach 0.5 mag.

---

### ***Image Re-projection***

A single, re-projected Mosaic image is constructed by resampling, using a *sinc* interpolation, the component science image sections. This form of the science data product is easier to use for image stacking and inter-comparisons, without the need to handle the distortion mapping. The astrometric system is a tangent-plane projection with a uniform sampling of  $0.25 \text{ arcsec pixel}^{-1}$ , oriented with North “up” (i.e., Declination increases along Axis 2) and East “to the left” (i.e., Right Ascension decreases along Axis 1). The tangent point is selected from a grid on the sky with roughly  $1^\circ$  steps in RA, and exactly  $1^\circ$  in Dec. This selection makes it easy to stack images that are part of a sequence of small spatial dithers.

### ***Image Stacking***

The spatial footprint of images obtained during a single observing run using the same filter are compared, and those that have at least a 10% spatial overlap are combined into an image stack, which is an image containing the union of the spatial extent of all the constituent images. By construction, the pixel grids of the constituent images are aligned (see “Image Re-projection” above), so the first approximation to the output stack is the median of all non-masked pixels in the stack, after scaling to a common exposure time (if needed) and adjusting for the average sky background in each input image. This approximation is used to identify outliers in the constituent images (to eliminate cosmic rays and hot pixels), which are marked in the mask files. Ultimately the stack is constructed from the mean of the constituent images, after masking bad pixels, scaling by the exposure times, adjusting for the backgrounds, and weighting the input inversely by the seeing for each constituent image. The background level in the stack is the mean of the input image backgrounds, scaled by the exposure times.

## 2.3.2 Calibration Reference Files

The calibration reference files that are used in the pipeline processing are constructed for each observing run where a sufficient number of appropriate exposures exist. For all but the bias structure file, the exposures that are used for calibration are affected by multiple instrumental signatures. It is important to distinguish between the additive backgrounds and the multiplicative sensitivity variations in order to avoid biasing the photometric accuracy of the processed science frames. Thus, the path to creating the reference files is an iterative one, and involves isolating the various effects, storing a characterization of the effect in a template image, and using these templates to construct down-stream reference files.

The background includes two artificial sources: a pupil ghost, and a fringe pattern that is caused by the night sky emission lines. The nature of these effects and that of the flat-field was explained above; key properties of these effects are summarized in Table 2.7. Each effect differs somewhat in its spatial scale, amplitude, color dependence, and stability; these differences are in fact what makes it possible to separate out their respective contributions. Note that the background effects may not be apparent at all in some science images.

**TABLE 2.7: Background Characteristics**

Effect	Contribution	Spatial Scale (pix)	Amplitude	Color Dependence	Notes
Pupil Ghost	Additive	2300	10%	High <sup>a</sup> @blue & red	Spatially stable per filter; amplitude depends on field brightness.
Fringe Pattern	Additive	10 s to 100 s	1–10% of sky	Strong	Spatially stable per filter; amplitude variable in minutes to hours.
Flat-Field	Multiplicative	2	~1–2% <sup>b</sup>	Moderate	Stable spatially & temporally, apart from shadows of dust particles.

a. Effect is much smaller for Mosaic-2, where the corrector has a superior AR coating.

b. Localized variations of >10% are common near detector blemishes or dust particles.

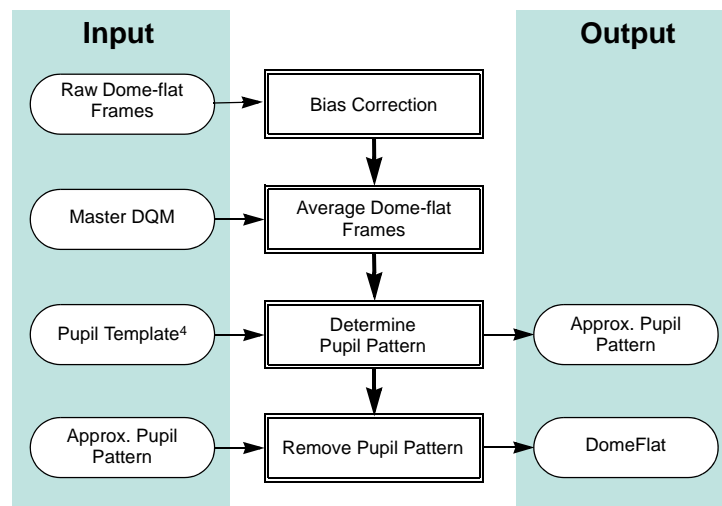
With the exception of the bias and dome-flat images, all calibration reference files are generated from the bias- and cross-talk-corrected science frames themselves. Masks are then generated for each science frame that mark the locations of artifacts (e.g., bad columns, bleed trails, and cosmic rays) and astrophysical objects in the field. The masked science frames for each filter are then combined to yield master sky frames in order to characterize the background effects, as explained below.

## Bias Structure

The master bias structure file is constructed very simply by applying the bias over-scan correction to each of several zero-second exposures, trimming the over-scan regions, and averaging the result (with rejection of outliers). Once constructed, both steps in the bias correction can be applied to all images.

## Dome Flat-Field

Dome flats, or exposures of an illuminated screen affixed to the interior of the telescope enclosure, are used to form the first approximation to the true flat-field. The process for constructing the master dome flat is illustrated in Figure 2.7. After the bias correction is applied, the images for each filter are averaged (with rejection of outliers and masking of bad columns). The next step<sup>5</sup> is to apply a template mask (from a library) with the same shape and spatial extent as the pupil ghost, so that a ratio may be determined between the portions of the dome flat-field image that are affected by the pupil ghost and those that are not. The ratio is applied to the pupil mask to generate a ghost template image that is divided out of the averaged dome flat-field image for each filter. This leaves the dome-flat template images, sans pupil ghost.



**FIGURE 2.7:** Process for characterizing the dome flat-field for each filter. Input dome-flat frames are uncalibrated. Pupil pattern characterization/removal applies to Mosaic-1 data obtained at the 4-m telescope only.

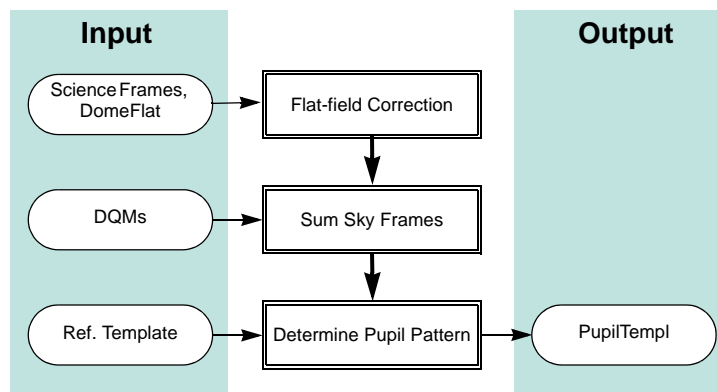
For some observing runs there are few if any dome-flat exposures. In this case, if there are exposures of the twilight sky (a.k.a., twilight-flat exposures) avail-

5. For Mosaic-1 data obtained at the 4-m telescope only.

able, these are processed in much the same way as described above and are used in place of the dome-flats.

### ***Pupil Ghost***

An accurate pupil ghost pattern is created<sup>6</sup> for each affected filter from the science images themselves, as illustrated in Figure 2.8. The science images are first corrected for cross-talk and bias, before applying the dome flat. The non-masked portions of the images are averaged to produce a sky frame. The same pupil template mask mentioned above is used to create a much improved pupil template for each filter.

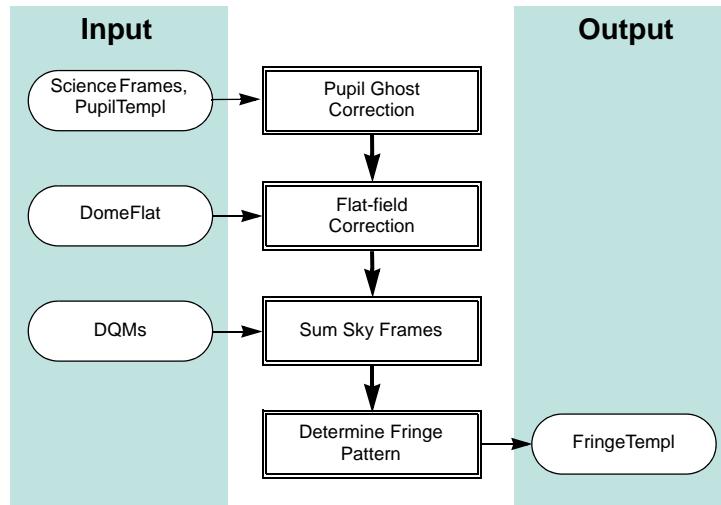


**FIGURE 2.8:** Process for characterizing pupil ghost pattern for each filter. Input science frames have had cross-talk and bias corrections applied.

### ***Fringe Pattern***

Having characterized the additive pupil ghost pattern, the science images are processed once again to scale and remove this pattern from each individual science image, and to correct the results by the dome flat. As illustrated in Figure 2.9, the frames are masked and combined, leaving a sky frame for each filter with the fringe pattern signature. This pattern can be characterized by smoothing the sky frame with a smooth, two-dimensional function with a width comparable to the fringe period, leaving a fringe template for each filter. Although in practice some or all of the fringe pattern within the pupil ghost template region is corrected by the pupil ghost pattern, this is not of concern because both terms are additive.

6. For Mosaic-1 data obtained at the 4-m telescope only.

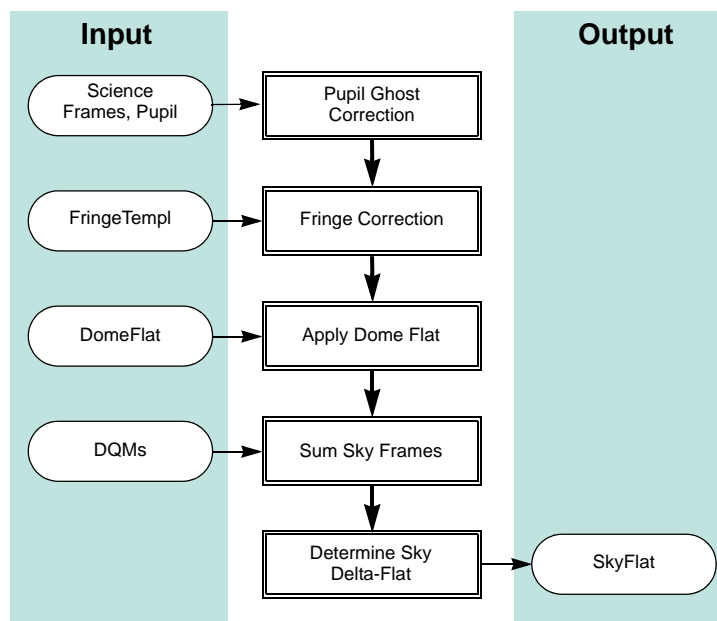


**FIGURE 2.9:** Process for characterizing the fringe pattern for each filter. Input science frames have had cross-talk and bias corrections applied.

### ***Sky Delta-Flat***

Once the pupil ghost and fringe pattern templates have been determined, an improvement to the flat-field can be generated from the accumulated night-sky background for each filter, provided there are enough counts. A flat-field generated in this way will be more closely matched to the science frames in the illumination pattern (i.e., large-scale, artificial gradients can be minimized), and in the color of the incident light (for which the effect is greatest for objects that are at or below the sky brightness). The processing of science data to determine the sky flat-field is shown in Figure 2.10. After correcting for cross-talk and bias, the pupil ghost and the fringe pattern are each scaled and subtracted from the science images, and the dome flat is applied. The results are masked and averaged to form a sky delta-flat template. The effect of applying this correction at full resolution is to use the sky as the flat-field, rather than the dome flats.





**FIGURE 2.10:** Process for characterizing the sky-flat for each filter. Input science frames have had cross-talk and bias corrections applied.

To be of sufficient quality, the sky flat-field must be composed of at least five (masked) input images, and have a signal-to-noise ratio of at least 50. If there are insufficient counts in the pixel-to-pixel sky delta-flat template, the sky delta-flat is smoothed by a function with a width of 255 pixels. In this case the dome-flat corrects for pixel-to-pixel variations in sensitivity, but the sky delta-flat corrects the dome-flat for low-order illumination effects.

## 2.4 SOURCES OF ERROR

This section describes the accuracy that can be expected for the major data products, the major sources of error, instrument foibles, and other noteworthy problems and issues with Mosaic data.

### 2.4.1 Astrometry

The astrometric accuracy for the image coordinate system is fundamentally limited by the accuracy of the astrometric catalog used to determine the WCS solution, which for the present is limited to USNO-B1.0. This catalog has the advantage of full-sky coverage, and a faint limiting magnitude that provides a

good overlap in the brightness range covered by most Mosaic exposures. Typical accuracies obtained for the WCS solution (0.25 arcsec measured in RMS deviations from the fit) are comparable to the accuracy of the catalog. The geometric stability of the Mosaic cameras is good, although rotations of the field of  $\sim 1^\circ$  are common after the camera has been reinstalled for an observing block.

There are two known circumstances where the pipeline WCS solution may fail entirely. The first is if the trial WCS is missing entirely from the raw data header, which is rare. The second case is if the telescope pointing as recorded in the RA and DEC (or TELRA and TELDEC) keywords is too far from the actual pointing, in which case the WCS solution will not converge.

## 2.4.2 Photometry

The photometric stability of the Mosaic cameras is relatively good, and with care and good flat-fielding, it is possible to achieve photometric accuracies of 1%. However, **at present the photometric calibration provided by the pipeline processing is not adequate to support science** and is intended solely to provide a rough estimate of exposure depth to support queries from the Virtual Observatory. While the systematic uncertainty may be large, within a sequence of observations covering the same region of the sky, the relative uncertainties are significantly smaller and provide a useful diagnostic on the photometric stability of the sky.

The photometric performance for any given image depends mostly on a number of external factors, including the sky conditions and reflectivity of the telescope optics (both of which are time dependent). Representative photometric zero-points and color terms for some standard, broadband filters are given in Table 2.8 for the Mosaic-2 camera. The equations used to fit the standard star photometry are of the form:

$$u = U + A_0 + A_1X + A_2(U - B)$$

$$b = B + B_0 + B_1X + B_2(B - V)$$

$$v = V + C_0 + C_1X + C_2(B - V)$$

$$r = R + D_0 + D_1X + D_2(V - R)$$

$$i = I + E_0 + E_1X + E_2(V - I)$$

where capital letters are the standard magnitudes, lower-case letters are instrumental magnitudes ( $25 - 2.5\log(ADU \cdot s^{-1})$ ),  $X$  is the airmass, and the first and last subscripted coefficients are given in the table.

TABLE 2.8: Representative Photometric Coefficients for Mosaic-2 Camera

Filt.	CCD 1	CCD 2	CCD 3	CCD 4	CCD 5	CCD 6	CCD 7	CCD 8
<b>Zero-Point</b>								
<i>U</i>	2.217±0.007	2.220±0.005	2.231±0.005	2.254±0.008	2.234±0.010	2.215±0.005	2.216±0.005	2.227±0.009
<i>B</i>	-0.107±0.008	-0.120±0.008	-0.115±0.008	-0.111±0.009	-0.111±0.008	-0.101±0.008	-0.106±0.008	-0.104±0.008
<i>V</i>	-0.448±0.008	-0.472±0.004	-0.472±0.004	-0.476±0.011	-0.475±0.004	-0.471±0.006	-0.472±0.005	-0.475±0.007
<i>R</i>	-0.636±0.004	-0.633±0.004	-0.648±0.004	-0.632±0.004	-0.628±0.004	-0.616±0.004	-0.629±0.005	-0.624±0.006
<i>I</i>	-0.089±0.010	-0.118±0.008	-0.076±0.010	-0.094±0.008	-0.094±0.004	-0.123±0.007	-0.113±0.011	-0.088±0.010
<b>Color Term</b>								
<i>U</i>	0.027±0.019	0.028±0.015	0.021±0.015	0.018±0.019	-0.025±0.024	0.019±0.016	0.015±0.016	-0.002±0.021
<i>B</i>	-0.069±0.010	-0.062±0.010	-0.058±0.010	-0.055±0.011	-0.056±0.010	-0.079±0.011	-0.081±0.010	-0.068±0.010
<i>V</i>	0.038±0.009	0.044±0.005	0.042±0.005	0.060±0.012	0.063±0.005	0.041±0.007	0.039±0.006	0.056±0.008
<i>R</i>	-0.025±0.008	-0.037±0.008	-0.022±0.008	-0.032±0.008	-0.028±0.008	-0.045±0.008	-0.035±0.010	-0.029±0.011
<i>I</i>	0.009±0.009	-0.002±0.007	-0.008±0.009	0.006±0.007	0.003±0.004	-0.002±0.006	0.003±0.011	0.011±0.010

### 2.4.3 Anomalies

#### *Metadata*

Critical metadata in raw science data files may sometimes be missing or have incorrect values. However, certain of the metadata in the headers of calibrated data, notably the WCS information and the filter name, are corrected during the course of pipeline processing (unless the calibration fails). Metadata *added* by the pipeline during processing are not in question.

#### *Detector Performance*

In September 2003, the CCD 3 in the Mosaic-2 camera failed, and was replaced in April 2004. Between these dates, the image extensions for this detector are still present in the data files, but contain no meaningful data.

The CCD array controllers for Mosaic-1 are known to suffer data drop-outs on occasion. The most recent documented cases occurred intermittently during January–May 2009, with array 2. This anomaly manifests itself in the data as a constant value everywhere in the affected array.

### ***Bias Stability***

It is possible to improve upon the bias correction by fitting the averaged over-scan vector with a polynomial (of order  $\sim 100$ ). This approach would have the effect of smoothing the correction to reduce the effect of noise in the over-scan region. Unfortunately, the bias level is known on occasion to suddenly change (or jump) during read-out. This makes the line-by-line approach to the over-scan correction more robust.

### ***Negative Saturated Pixels***

Saturated pixels (where the accumulated charge exceeds the full-well, e.g., in the cores of over-exposed stars) will have negative values when the binning factor (given in the CCDSUM keyword) is  $2 \times 2$  or higher.

## ***2.5 REFERENCES & FURTHER INFORMATION***

### ***Contributing Authors***

A large number of people have contributed to the technical knowledge of the Mosaic cameras. Extensive material has been contributed by George Jacoby (lead author of the first version of the Mosaic Instrument Manual), and by Heidi Schweiker and Buell Jannuzi (lead authors of Version 3 of the Instrument Manual). Other substantial contributors were: Taft Armandroff, Todd Boroson, Jim De Veny, Steve Heathcote, Tod Lauer, Bob Marshall, Phil Massey, Gary Muller, Knut Olsen, Rich Reed, Frank Valdes, David Vaughnn, and Dick Shaw. Documentation on the reduction of Mosaic data is also fairly extensive, much of which was authored (see below) by Frank Valdes (in IRAF documentation); Frank Valdes, Rob Swaters, and Tracy Huard (in pipeline documentation); and Heidi Schweiker, Buell Jannuzi, and Frank Valdes (in the form of the NDWFS data reduction notes).

### ***References***

Jacoby, G. H., Laing, M., Vaughnn, D., Reed, R., & Armandroff, T., 1998, *A New Wide-Field Corrector for the Kitt Peak Mayall 4-m Telescope*, SPIE 3355, 721

Jannuzi, B., Claver, J., & Valdes, F. 2003, *The NOAO Deep Wide-Field Survey Mosaic Data Reduction Notes* (Tucson: NOAO)<sup>7</sup>

7. Available at <http://www.noao.edu/noao/naodeep/ReductionOpt/frames.html/>

- Massey, P., Armandroff, T., De Veny, J., Claver, C., Harmer, C., Jacoby, G., Schoening, W., & Silva, D. 2002, *Direct Imaging Manual for Kitt Peak* (Tucson: NOAO)
- Monet, et al. 2003, *AJ*, 125, 984
- Muller, G. P., Reed, R., Armandroff, T., Boroson, T., & Jacoby, G. 1998, *What is Better than an 8 K x 8 K Mosaic?*, *SPIE*, 3355, 487
- Schweiker, H., & Jannuzi, B. T. 2004, *NOAO CCD Mosaic Imager User Manual* (Version 3.0: Tucson: NOAO)
- Walsh, J. R., Pirzkal, N., & Pasquali, A. 2002, *Modelling the Fringing of the ACS CCD Detectors*, in 2002 *HST Calibration Workshop*, ed. S. Arribas, A. Koekemoer, & B. Whitmore (Baltimore: STScI)
- Wolfe, T., Reed, R., Blouke, M., Borson, T., Armandroff, T., & Jacoby, G. H. 1998, *CCD Detector Upgrade for NOAO's 8192 by 8192 MOSAIC*, *SPIE* 3355, 487

### **For Further Reading**

A document repository<sup>8</sup> for the Mosaic-1 camera is available on the Web. A general discussion of data reduction techniques for CCD mosaics was published by Frank Valdes, who also implemented these techniques in IRAF, including most of the applications that are used in the current Mosaic pipeline. Additional, more recent papers on the Mosaic pipeline appeared in the proceedings of the ADASS XVI conference.

Swaters, R. A. & Valdes, F. G., 2007, *The NOAO High-Performance Pipeline System: Mosaic Camera Pipeline*, in ASP Conf. Ser. 376, ed. R. A. Shaw, F. Hill, & D. J. Bell (San Francisco: ASP), 269

Valdes, F. G. 2002, *The Reduction of CCD Mosaic Data*,<sup>9</sup> in *Automated Data Analysis in Astronomy*, ed. H. P. Singh, R. A. Gupta, & C. A. L. Bailer-Jones (New Delhi: Narosa Pub. House), 309

Valdes, F. G., & Swaters, R. A. 2007, *The NOAO High-Performance Pipeline System: Mosaic Camera Pipeline Algorithms*, in ASP Conf. Ser. 376, ed. R. A. Shaw, F. Hill, & D. J. Bell (San Francisco: ASP), 273

Various PI teams have used the Mosaic cameras to conduct major surveys. One team, lead by Phil Massey (Lowell Observatory), has published extensive

8. Mosaic-1 document repository: <http://www.noao.edu/kpno/mosaic/mosaic.html>

9. Available electronically at <http://iraf.noao.edu/projects/ccdmosaic/reductions/adaa/valdes2.pdf>

notes<sup>10</sup> on their data reduction procedures. The following journal papers describe the calibration techniques and science results from this survey.

Massey, P., Olsen, K. A. G., Hodge, P. W., Strong, S. B., Jacoby, G. H., Schlingman, W., & Smith, R. C., 2006, AJ, 131, 2478

Massey, P., Olsen, K. A. G., Hodge, P. W., Jacoby, G. H., McNeill, R. T., Smith, R. C., & Strong, S. B. 2007, AJ, 133, 2393

---

10. Available at <http://www.lowell.edu/users/massey/lgsurvey/splog2.html>

## Chapter 3

# *NEWFIRM Camera*

Version 1.0, May 2009

### In This Chapter...

<b>Instrument Overview</b>	<b>3-1</b>
<b>Data Products</b>	<b>3-10</b>
<b>Calibration</b>	<b>3-18</b>
<b>Sources of Error</b>	<b>3-28</b>
<b>References &amp; Further Information</b>	<b>3-31</b>

The NEWFIRM camera was placed in routine operation at KPNO in late 2007, and definitive technical and performance information is still being assembled. The material presented here is drawn primarily from the *NEWFIRM Quick Guide for Proposal Preparation* (Probst 2008), and from the *NEWFIRM Quick Reduce Pipeline and Data Analysis Tools* notes by Dickinson, et al. (2008). Interested users should consult these documents for details of the instrument description and pipeline operation. Additional resources<sup>1</sup> are cited throughout this chapter, and are listed along with other background material in the last section.

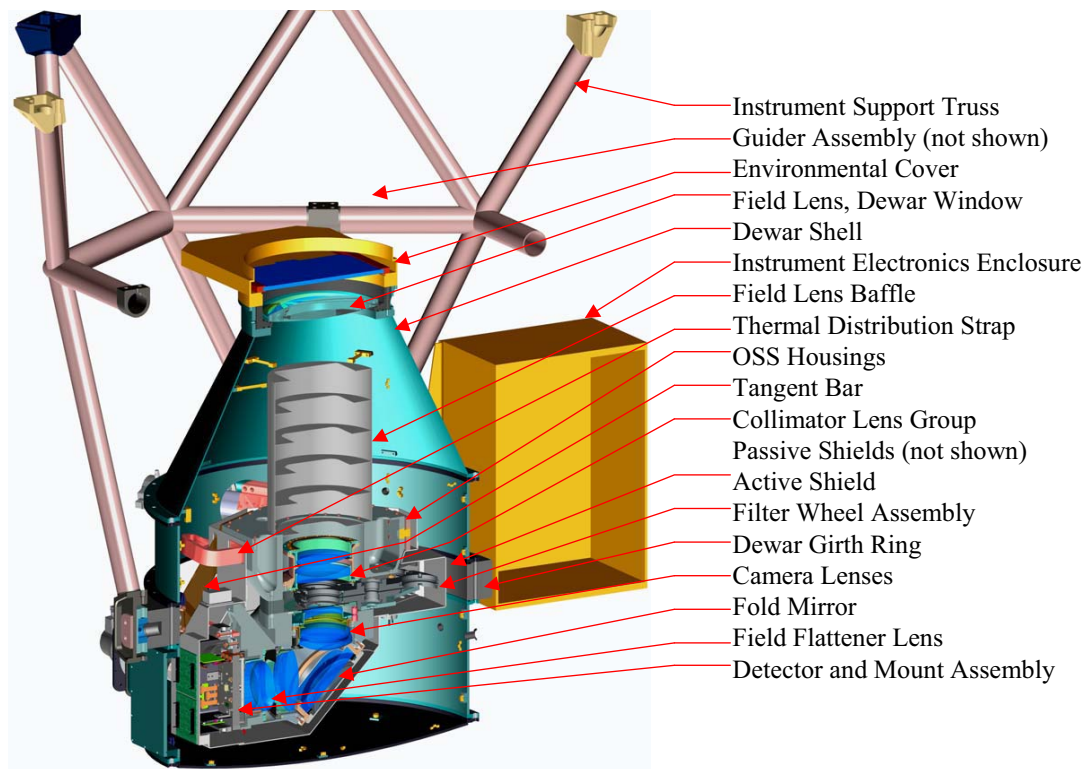
### 3.1 INSTRUMENT OVERVIEW

The NOAO Extremely Wide-Field Infrared Mosaic (NEWFIRM) camera has rapidly become one of the most heavily used instruments offered by NOAO because of its excellent infrared (IR) sensitivity and very large field of view. NEWFIRM is planned to be used at both the Mayall 4-m on Kitt Peak and the Blanco 4-m telescope on Cerro Tololo, in alternating campaigns. This chapter describes both usages of the camera.

1. See the NEWFIRM instrument Web site at <http://www.noao.edu/ets/newfirm/>

### 3.1.1 Instrument Capabilities & Design

The NEWFIRM camera is a near-IR imager with a wide field of view (FoV) and a large focal-plane array (FPA) containing four InSb detectors with a total of  $4096 \times 4096$  pixels. A schematic of the instrument is shown in Figure 3.1. The detectors have quantum efficiency (QE) near 90% over the wavelength range 0.8–4.5  $\mu\text{m}$ , with the exception of one detector (SN019), where a thicker anti-reflection coating degrades the QE by  $\sim 10\%$  shortward of  $\sim 2.5 \mu\text{m}$ . There are two filter wheels, each with 7 slots, that can hold a subset of the available filters for NEWFIRM. The filter sets<sup>2</sup> include standard broad bandpasses, such as  $JHK_s$ , mediumband filters, and a few narrowband filters including [Fe II] 1.64  $\mu\text{m}$ , H<sub>2</sub> 2.12  $\mu\text{m}$  and Br  $\gamma$  2.17  $\mu\text{m}$ . Since the detector has high QE well beyond the bandpasses of the current filter set, blocking filters are sometimes used to minimize red leak.



**FIGURE 3.1:** Cross-section view of the NEWFIRM camera. Light enters from above in this diagram, and passes through various optical elements (see labels).

The details of the observing configurations for NEWFIRM are given in Table 3.1. The camera is mounted at the Cassegrain focal station of the 4-m tele-

2. The filter properties can be viewed in the online Users Manual at <http://www.noao.edu/ets/newfirm/>.



scopes, with a converging beam as indicated in the table. The pixel scale samples the delivered point spread function (PSF) well, which is seeing-limited at all wavelengths. The image quality at the 4-m telescopes is good, with little (<10%) focus gradient or PSF variation across the field of view. The pixel scale is slightly variable from the center to the edge of the FoV due to pincushion distortions. There is very little vignetting.

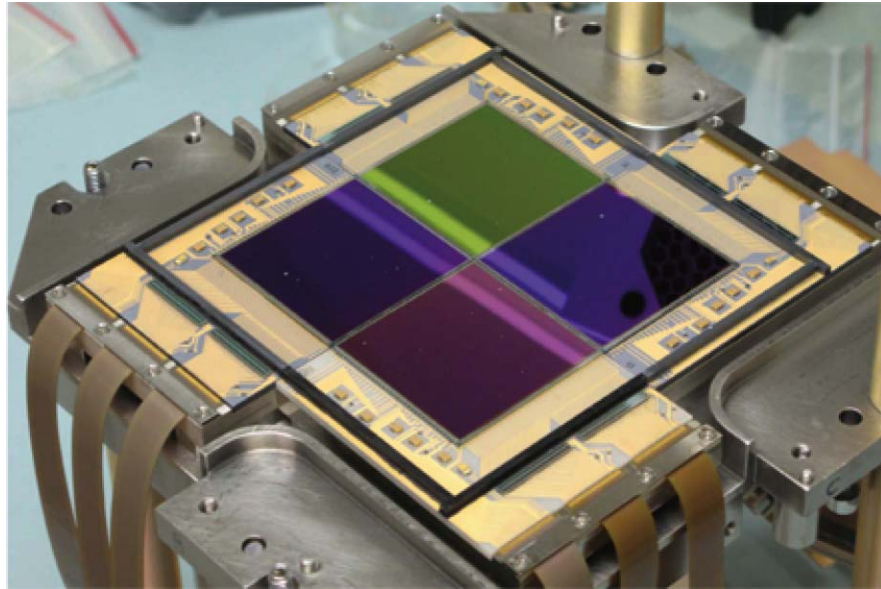
**TABLE 3.1: NEWFIRM Observing Configuration**

	<b>KPNO 4-m</b>
Field of view	27' × 27'
Pixel scale of FPA	0.40 arcsec pixel <sup>-1</sup>
Inter-array gaps	35 arcsec, or 88 pixels
Delivered image quality	0.8 arcsec in excellent seeing conditions
Telescope focal ratio	f3.1

It is important to note that there is no mechanical shutter for the NEWFIRM camera; rather, the exposure duration is regulated by the difference between the initial and final nondestructive reads of the array. This has important implications for the linearity and measured counts, particularly at high count rates; see “Array Operation” on page 3-7 for details.

### ***Focal Plane***

The focal plane of the NEWFIRM Camera is populated with an array of four Orion InSb detectors, arranged in a 2 × 2 mosaic, as shown in Figure 3.2. Each of the detectors has 64 amplifiers to provide high-speed, parallel readout of the entire array as fast as 1.4 s. Representative detector properties are given in Table 3.2; most values do not differ significantly from array to array. Compared to present-day CCD detectors for the optical band, the sensitivity and dynamic range of this generation of large-format IR detectors are excellent, but the array cosmetics, noise, and linearity are inferior. The response of the NEWFIRM Orion detectors is linear to within about 6% for exposures less than 80% of the full-well. The QE varies somewhat from detector to detector, both on average and as a function of wavelength, in part due to differences in the anti-reflection coatings. (See Merrill 2008 for details.) Therefore, the transformation from instrumental magnitudes to standard systems depends upon the detector. See “Photometric Calibration” on page 3-24 and “Photometry” on page 3-29.

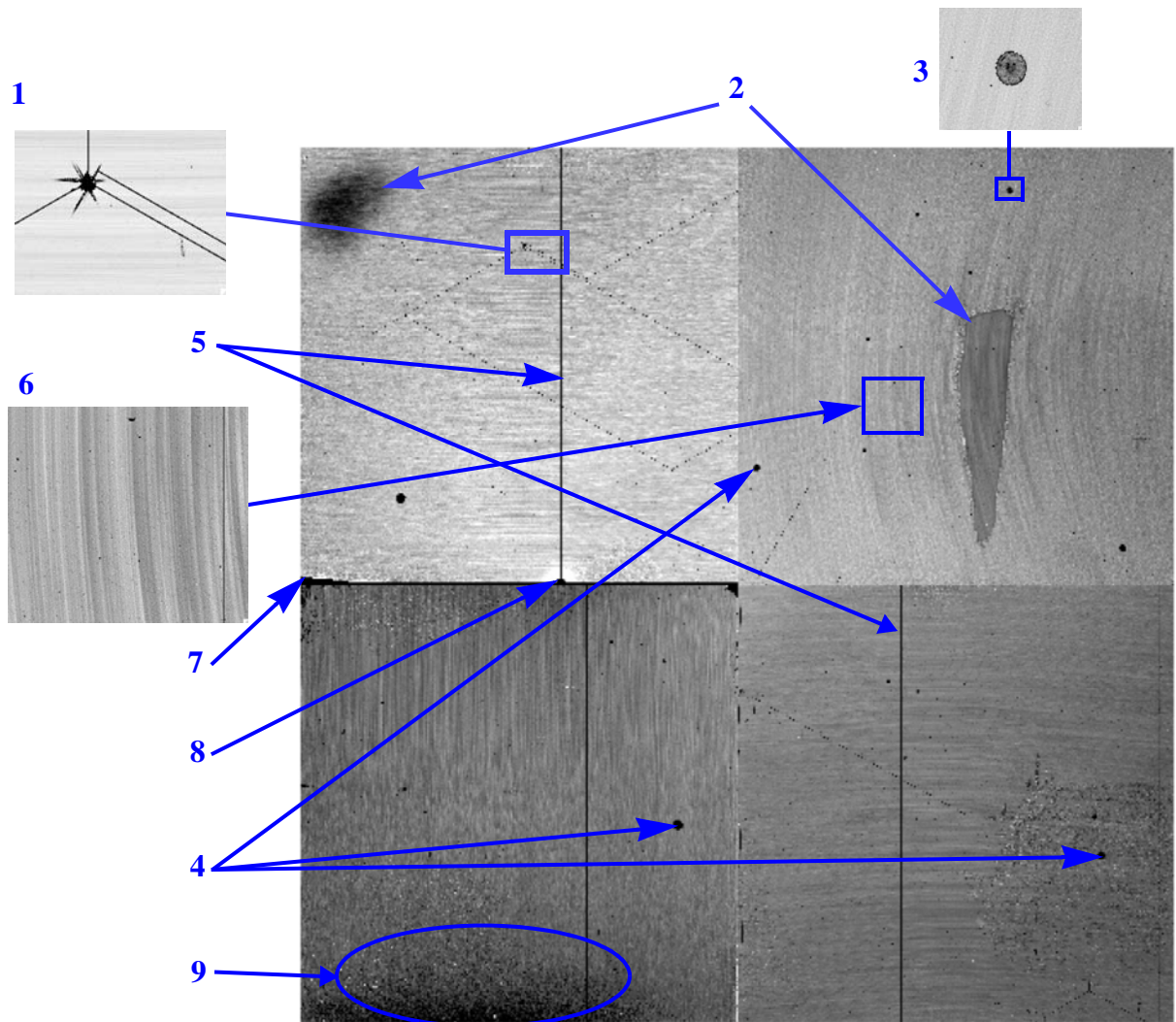


**FIGURE 3.2:** Mounting for the NEWFIRM focal plane array, showing the four 2048 x 2048 pixel Orion InSb detectors.

**TABLE 3.2: Orion InSb Array Characteristics**

Array Dimensions	2048 × 2048
Pixel Size	24 μm square
Gain	8 $e^-$ /ADU
Read noise, RMS	35 $e^-$ (1 Fowler sample)
Dark current	0.17 $e^-$ / s
Well capacity	~100,000 $e^-$
Nonlinearity	~6% at 10,000 ADU
Read Time (Dig. Avg. = 4, Fowler samples = 1)	1.4 s

The flat-field image shown in Figure 3.3 illuminates the most significant cosmetic defects in the photo-active areas of the arrays. The spatial extent of these defects (along with the inter-chip gap) determines the minimum separation of the spatial dithers that are required during observing to mask their effect on the final, stacked images. A brief description of the defect types, numbered as in the figure, follows:

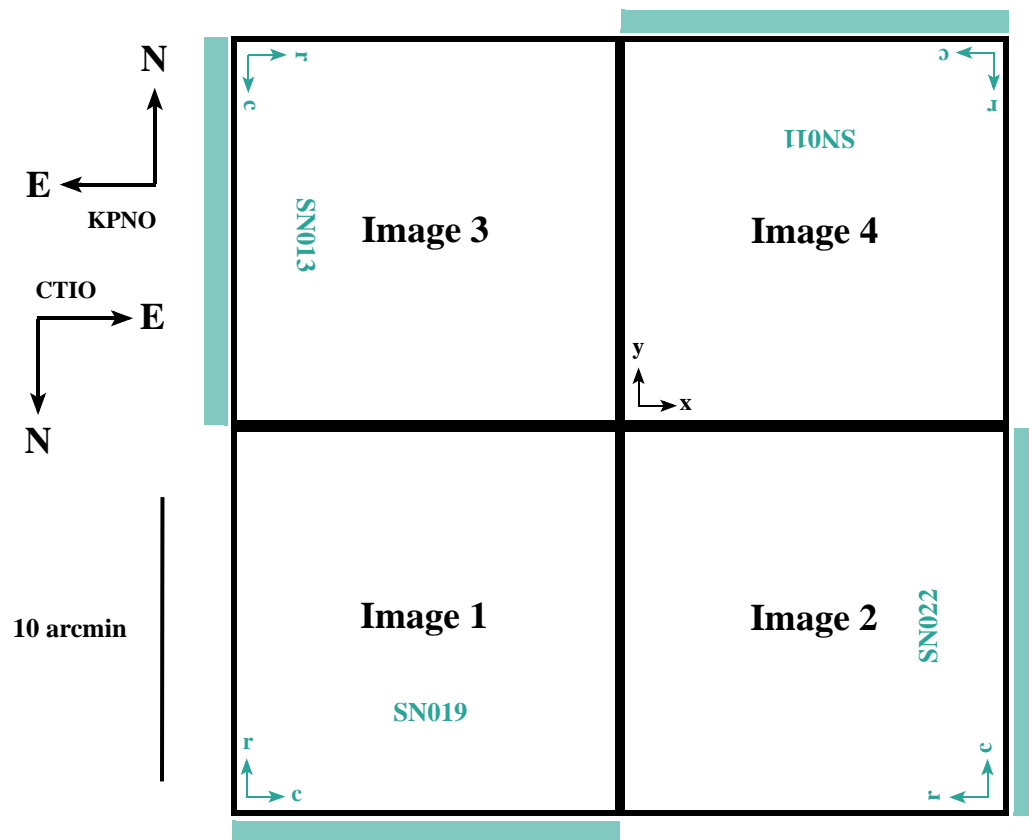


**FIGURE 3.3:** *H*-band Flat-field image showing all 4 arrays abutted, indicating examples of various cosmetic defects in the NEWFIRM focal plane array. See text for descriptions.

1. Diagonal cracks in the InSb substrate, where there is zero sensitivity.
2. Broad regions of slightly reduced sensitivity and higher dark current.
3. Small regions of reduced sensitivity due to contaminants on the arrays.
4. Photo-Emissive Defects (PED), which result from shorts in the bonded readout electronics, and which have been passivated at the expense of a number of roughly circular areas as large as ~50 pixels diameter with zero sensitivity.
5. Inoperable rows or columns, which are often but not always associated with a PED.
6. Residual patterns (with few-percent amplitude) of substrate crystal formation (i.e., “growth rings”), which are easily removed with the flat-field correction.

7. Regions at the detector edges or corners that have debonded from the readout electronics, resulting in zero sensitivity.
8. Damaged region on detector, with surrounding area of high dark current and nonlinear response that is masked in pipeline processing.
9. Extended regions of elevated dark count and pixel nonlinearity.

Figure 3.4 shows the arrangement of the NEWFIRM focal plane array on the sky, using the same nomenclature as is found in the FITS file extension headers.<sup>3</sup> Note that the raw image array coordinates are remapped from the detector coordinates, such that the image origin is always in the lower-left corner when oriented as in the figure.



**FIGURE 3.4:** Orientations on the sky and spatial footprint of the focal plane array for the NEWFIRM camera, with FITS image extension and IR detector designations indicated. Coordinate  $(x, y)$  axes originate in the lower-left corner for all images (black arrows). The detector coordinate  $(row, column)$  origin is indicated for each array (teal arrows). Location of the 64 output amplifiers for each array is also indicated (teal bars). Gap between adjacent sensors is 35 arcsec.

3. The detectors and the orientation of the image arrays were changed during the course of instrument commissioning; this diagram reflects the final configuration for data obtained after semester 2007B.

### 3.1.2 Operations

An understanding of NEWFIRM data products requires a basic knowledge of how data are obtained during an observing run, which in turn requires knowing both how the arrays are operated and the strategies used to optimize the spatial coverage of the intended region of interest on the sky.

#### *Array Operation*

There are a few key points to bear in mind when trying to understand data from NEWFIRM. First, the exposure duration is not regulated with a mechanical shutter, as is done with CCD cameras, but is instead controlled by rapid, precisely timed, nondestructive reads of the detector array. The strategy for operating the NEWFIRM detectors is called *correlated double-sampling* (CDS), where the pixel values in the output image are in reality a *difference* between a final and an initial read of each pixel. The integration time is constant across the array, but the interval over which a given pixel integrates depends on its position on the array. This differencing also masks the fact that the pixels are accumulating counts during the finite time required to reset the array and read it out; this “hidden” flux affects the linearity correction at high backgrounds. The total elapsed time from the array reset to the last read of the last detector pixel depends upon four parameters that are set by the observer, as shown in Table 3.3. Each combination of these readout parameters changes the readout

**TABLE 3.3: NEWFIRM Array Readout Parameters**

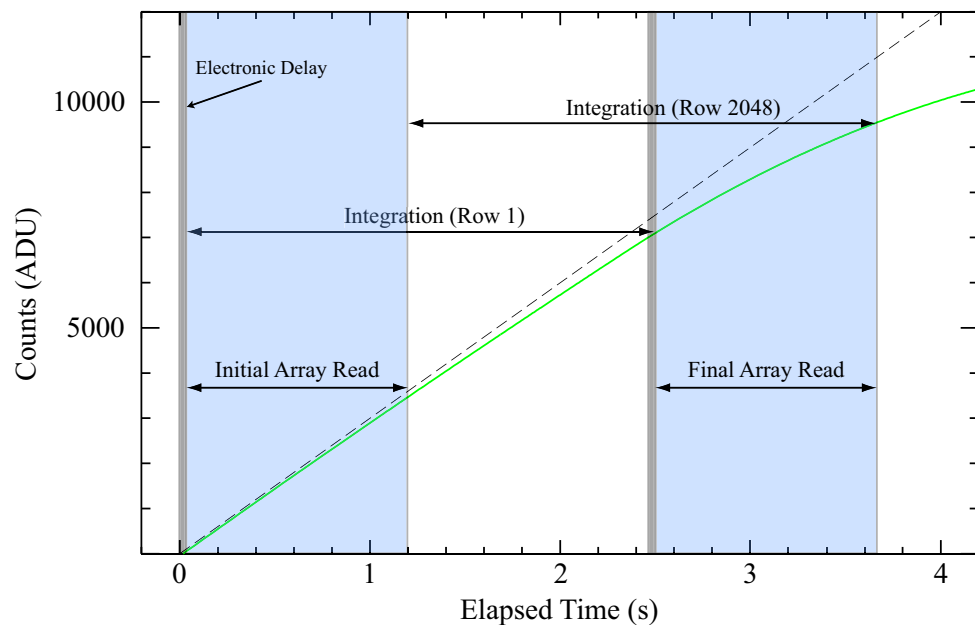
Parameter	Default Value	Description
Exposure Time	—	Duration of exposure, which is the elapsed time between the first sampling of the initial readout of the array and the first sampling of the final readout of the array.
Digital Averages	4	Number of back-to-back samples of each pixel during the course of a single array read; samples are <b>averaged</b> to determine the raw pixel value.
Fowler Samples	1	Number of times the entire array is read just after reset, and again at the end of the integration. Resulting raw image is the difference between the <b>sum</b> of the last readouts and the <b>sum</b> of the initial readouts. Using $N > 1$ is appropriate when the S/N ratio is dominated by read noise, rather than the background.
Co-adds	1	Number of integrations taken in sequence and <b>summed</b> (on-board) into a single output image.

timing (see Table 3.4), the effects of which can be quite significant: at high count rates; the main effect is in the linearity correction as illustrated in Figure 3.5; at low count rates the amplitude and structure of the dark “pedestal”

is affected by the readout timing. This change in readout timing is the reason that dark frames must match the science frames with respect to the integration time and the array operation parameters in the above table; otherwise the dark frames will not accurately represent the accumulated dark current during a science exposure.



Note that raw data frames obtained with **N Fowler Samples** will, in effect, contain the **sum** of the N readouts, rather than the average. Raw data frames obtained with **M Co-Adds** will also be recorded as the **sum** of M images, although this is perhaps the expected behavior. Images calibrated by the NEWFIRM pipeline are normalized to unit exposure time.



**FIGURE 3.5:** Schematic of the detector response (*green curve*) to a very high incident flux of  $3000 \text{ ADU s}^{-1}$  (*dashed line*) as a function of time. After the array is reset the counts begin accumulating from zero. In *correlated double-sampling*, after a brief delay the array begins reading out, which takes 1.16 s when  $\text{DigAvg} = 4$  and one Fowler sample (*blue shaded regions*); following a programmed delay to achieve the integration time of 2.5 s, the detector is read again. Although the integration time is constant across the array, the interval over which a given pixel integrates depends on its position in the array.

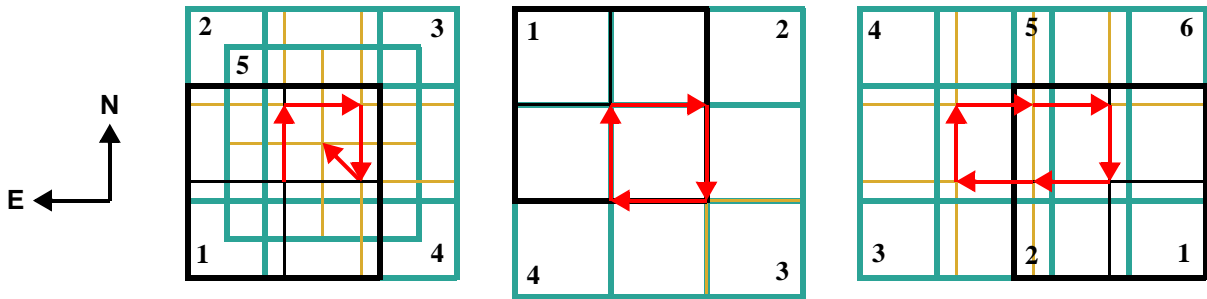
TABLE 3.4: Array Readout Times

		Number of Fowler Samples		
		1	4	8
No. Digital Avgs.	1	0.565	1.52	2.753
	2	0.894	2.270	4.103
	4	1.195	3.014	5.439
	8	1.785	4.092	8.101
	16	2.965	7.448	13.425

### Observing Sequences

Essentially, all NEWFIRM observers obtain multiple exposures of their science fields per filter; usually in a sequence of small spatial dithers about a reference position in order to obtain data in the gaps between the detectors in the FPA, and to ameliorate the effects of detector artifacts (see Figure 3.3 on page 3-5). Many observing programs also obtain sequences of (slightly) overlapping images, which enables the mapping of large regions of sky from the component images. These sequences are enabled at the telescope with a set of standard scripts, the parameters of which specify the observing pattern on the sky and the relative offsets between pointings within the pattern; these sequence parameters are written to the FITS headers of the science images.

A few of the pre-scripted observing patterns are shown in Figure 3.6, where the footprint of the NEWFIRM FPA is shown as a series of overlapping “window panes.” While the magnitude of most offsets is set by individual observers via a script, the “5PX” and “Random” patterns are most often used with small offsets to *Dither*, the “4Q” pattern positions a source of modest angular extent in the center of each detector, and the “RAxDec” pattern is used either with small offsets to dither, or with large offsets to *Map* an extended region of sky. It is also possible to obtain separate “sky” (or background reference) frames at a very large offset from an angularly very extended target (see the NEWFIRM operations manual for details). Note that these patterns can be combined: e.g., a *Dither* sequence may be executed at each position in the *Map* sequence to mosaic a large area of sky free of data gaps. The calibration pipeline will create a dithered stack of the entire map provided that the offset between map centers is no larger than 20 arcmin.



**FIGURE 3.6:** Schematic of three types of offset observing sequence patterns: a five-point “X” (5PX, *left*), four quadrants (4Q, *center*), and a 3 x 2 rectangle (RA x Dec, *right*) of six positions. Translation of the center of the FoV from the initial exposure in the sequence (*black footprint*) is indicated (*red arrows*).

Finally, most observing programs obtain calibration frames, such as dark frames and dome flats, using scripted sequences in order to facilitate automated pipeline processing, since the pipeline is able to recognize calibration frames from the sequence identifier. See “Calibration” on page 3-18 for a detailed discussion of how the calibration exposures are used in the data reduction process.

### 3.2 DATA PRODUCTS

This section describes the content and format of the various data products that are produced for the NEWFIRM camera. Most of the products are generated during the course of calibration processing, the details of which are discussed in the next section. The data products can be distinguished by the combination of the PROCTYPE, PRODTYPE, and OBSTYPE keywords in the primary header; the possible values are summarized in Table 3.5. The processing level (see Table 1.2, “Levels of Data Processing,” on page 1-4) at which the product is generated is listed in column 5 (Proc. Level).

**TABLE 3.5: NEWFIRM Data Product Types and Contents**

PROCTYPE	PRODTYPE	OBSTYPE	Extensions	Proc. Level	Description
Raw	image	object	IMAGE x 4	1	Raw data as obtained at the telescope, with some additional metadata included in the header
InstCal	image	object	IMAGE x 4	2	Calibrated, sky-subtracted, single-frame image
InstCal	mask	object	BINTABLE x 4	2	Bad-pixel mask for InstCal
InstCal	png	object	FOREIGN	2	Preview for InstCal
Resampled	image	object	IMAGE	2	Calibrated, sky-subtracted, re-projected image



**TABLE 3.5: NEWFIRM Data Product Types and Contents (Continued)**

PROCTYPE	PRODTYPE	OBSTYPE	Extensions	Proc. Level	Description
Resampled	mask	object	BINTABLE	2	Data quality mask for resampled image
Resampled	png	object	FOREIGN	2	Preview of resampled image
Stack	image	object	[None]	3	Stack of multiple, overlapping, calibrated frames with sky subtracted
Stack	mask	object	BINTABLE	3	Data quality mask for calibrated, stacked image
Stack	expmap	object	BINTABLE	3	Exposure map for calibrated, stacked image
Stack	png	object	FOREIGN*2	3	Preview for stacked image
MasterCal	image	dark	IMAGE x 4	2	Master dark calibration image
MasterCal	png	dark	FOREIGN*2	2	Preview for dark
MasterCal	image	dome flat	IMAGE x 4	2	Master dome flat-field calibration image
MasterCal	png	dome flat	FORIEGN*2	2	Preview for dome flat

### 3.2.1 Image Formats

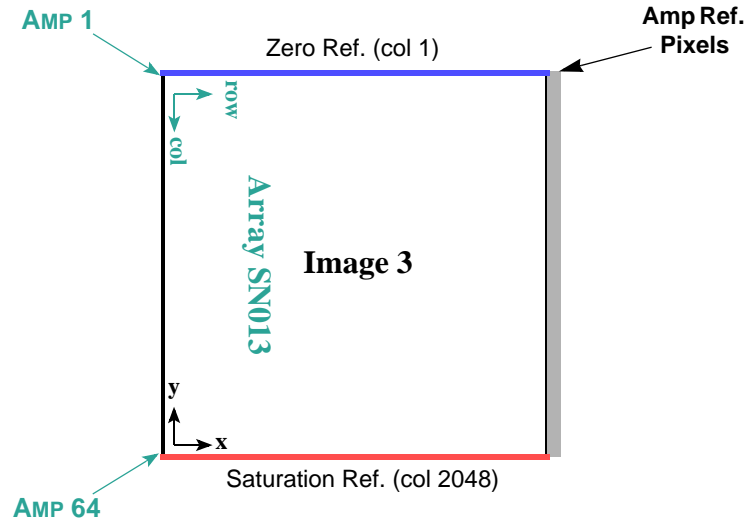
The science image data from the NEWFIRM camera is stored either in FITS multi-extension files (MEFs), the general structure of which was described in Chapter 1, or in simple FITS files with no extensions. The detailed arrangement of the image portions among the extensions differs, depending upon whether the data are raw (unprocessed) or reduced.

#### *Raw Data*

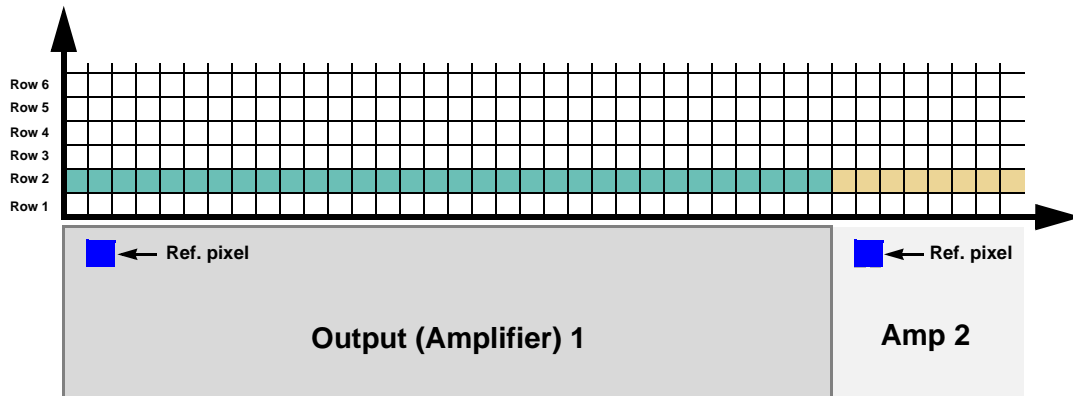
Raw data from the NEWFIRM camera are organized by detector, with the output from each array stored as 32-bit integers in a separate image extension in the FITS MEF file (see Chapter 1). Thus, there are as many image extensions in the raw science file as the number of arrays in the focal plane. The coordinate origin for all images is in the lower-left corner of the readout section, rather than at the location of the first readout amplifier (see Figure 3.4 on page 3-6). This approach means that individual array images, displayed as image extensions, all have the same orientation on the sky.

The Orion detectors contain reference pixels that enable tracking the electronic stability during readout. There are actually three varieties of reference pixels on each detector, as illustrated in Figure 3.7 and Figure 3.8. The first detector column consists of pixels that electronically simulate the effect of unilluminated pixels, and the last column consists of pixels that simulate saturated pixels. There is, in addition, one reference pixel for each of the 64 parallel readout amplifiers. That reference pixel is read once for each row during the course of a

readout. The 64 amplifier reference pixels are stored at the end of each image row, as illustrated in Figure 3.7. At present, the reference pixels are not used in pipeline processing.



**FIGURE 3.7:** Schematic of the image arrays just after readout for detector SN013 of the NEWFIRM camera. Note that the image coordinates are rotated relative to the detector array. Location of the first and last of the 64 output amplifiers are indicated, as are the zero (*blue*) and saturated (*red*) detector reference columns. The photo-active region of the detector is indicated in white, while the amplifier reference pixel values, shown in grey, are stored in the highest columns of the image.



**FIGURE 3.8:** Schematic of a portion of the detector array illustrating the parallel readout of the second row of the IR arrays. For each row in the array, the first amplifier reads its reference pixel (fictitious locations shown in blue), followed by the first 32 pixels of the row; the second amplifier simultaneously reads its reference pixel, followed by the next 32 pixels in the row, and so on. The process is repeated for each row of the array. The values of the array reference pixels for the 64 amplifiers are stored in columns 2049:2112 of the raw image.

TABLE 3.6: Raw Data and Reference Pixel Regions

Raster Dimensions	Image	Detector	Photo-Active Data Section	Amp Reference Pixel Section
2112 × 2048	1	SN019	[1:2048, 2:2047]	[2049:2112, 1:2048]
	2	SN022	[1:2048, 2:2047]	[2049:2112, 1:2048]
	3	SN013	[2:2047, 1:2048]	[2049:2112, 1:2048]
	4	SN011	[2:2047, 1:2048]	[2049:2112, 1:2048]

### Calibrated Data

The Archive contains data products that are produced with the NEWFIRM calibration pipeline. The specific calibrated science data products are given in Table 3.5 on page 3-10 and are described in more detail below. Each image has an associated *data quality mask* (DQM) and a *preview image*, which are described at the end of this subsection. The science images are stored as 16-bit scaled integers, which is a mild form of compression, when the sigma of the sky pixels can be sampled at more than 30:1 with a dynamic range over all non-flagged pixels. That is, the integer digitization has at least 30 values within 1-sigma of the sky distribution, and the 32K values span all good (i.e., non-flagged) pixel values in the calibrated data. If these conditions cannot be met, then the images are stored as 32-bit floating-point values.

**InstCal.** These images have been processed to remove instrumental signature and have been astrometrically and photometrically calibrated (see “Calibration” on page 3-18), although there are some cases where the calibration can fail (see “Sources of Error” on page 3-28). The pixel values have been normalized to unit exposure time (i.e., ADU s<sup>-1</sup>). The data are organized in the FITS file almost identically to that for the raw science data, except for the following:

- The reference pixel regions have been trimmed from the image arrays.
- The resulting images in each extension are all 2046 × 2046 pixels containing most of the photo-active regions of the detectors.

**Resampled.** These images are the result of geometrically rectifying InstCal images, where each array has been re-projected to a common tangent-point on the sky, with pixels aligned to a common grid with uniform scale. The resampled image is stored in the primary header-data unit within the file. Using a common tangent-point (which is selected from a grid with roughly 1° steps in RA, and exactly 1° in Dec) facilitates stacking overlapping images within an observing sequence. The *reduced, re-projected* images are approximately 4092 × 4092 pixels; the actual size depends on the WCS tangent-point selected

for a particular image. Both the single-frame and the re-projected images are accompanied by DQMs.

**Stacked.** When two or more observations of a given target are obtained on the same night using the same filter and have a sufficient degree of spatial overlap (which is generally the case), these images are combined using an average with outlier rejection to remove detector blemishes, gaps between the detectors, and artifacts such as image persistence and cosmic rays. The result is a union of the spatial footprints of the stack, which is stored in a simple FITS image (i.e., not in image extensions), with nearly the same pixel scale as the raw images. These images will, in general, be larger—sometimes very much larger—than  $4092 \times 4092$  pixels because the area of the sky that is mapped can be significantly larger than the instrument FoV.




---

The exposure duration for stacked images, as recorded in the `EXPTIME` keyword, refers to the sum of all exposure durations of all images used to create the stack. The exposure depth and noise properties of a stack, particularly from a *Map* sequence, is a discontinuous and possibly complicated function of position in the image. Use the exposure map to track the detailed exposure depth at the pixel level.

---

**Master Calibration.** Reference files are created during the course of pipeline processing: namely, darks and flat-fields. These files are used in pipeline processing to remove instrumental signatures from the science data. The Linearity Correction file is created by NOAO scientists from engineering data. See “Calibration Reference Files” on page 3-26. These reference files are 32-bit floating-point images, stored as MEF files with 4 extensions, matching the raw images.

**Concomitant Data.** All reduced images are accompanied by data quality masks (DQM); stacked images are in addition accompanied by exposure maps. The DQMs and exposure maps are logically images, but they are stored in compressed form as FITS binary tables, with one table per extension in the science image. Some utilities (such as the IRAF `XImtool`) can display and work with these files directly, but other utilities may require that the files be translated to FITS images before they can be used. The pixel values in the DQMs are non-zero when affected by detector pathologies and image artifacts such as bad rows or columns, saturation, and regions of passivated pixels; the values are zero otherwise. Table 3.7 on page 3-15 lists mask values that are applicable to *non-stacked* images.

**TABLE 3.7: Data Quality Mask Values**

Value	Meaning
0	No problem
1	Bad/invalid pixel identified in the master data quality mask; generally indicates known detector blemishes
2	Bad linearity correction
3	Saturated
4	[Not used]
5	Bad/missing background subtraction
6	Affected by detector persistence
7	[Not used]
8	Transient artifact

Most stacked images are created from dithered frames, so that most of the pathologies affecting single frames will be obviated. DQMs for *stacked* images have the following values: “1” for areas of the re-projected image with no data, “2” for areas with no data *after rejection*, and zero otherwise.

The exposure maps are images whose values are the cumulative exposure duration at each pixel, in seconds, which can be a complicated function of position for *Map* sequences (see “Observing Sequences” on page 3-9). The value of the EXPTIME keyword in a stacked image is the sum of the exposure durations for all images that contributed to the stack.

**Ancillary Files.** The preview images are down-sampled versions of the files they accompany and are stored in Portable Network Graphics<sup>4</sup> (PNG) format. There are two samplings for each file, which are intended mainly for quick display of the results of pipeline processing on Web pages. Note that these files have been embedded in FITS files as extensions of type FOREIGN so that most users will need to extract them with some tool before viewing.

### 3.2.2 Header Keywords

A wide variety of metadata are recorded in the headers of the science frames. Users should browse these headers (and the extension headers) to familiarize themselves with the content. The more critical metadata are described in this subsection. Table 3.8 lists metadata by the FITS keyword name, the header unit

4. <http://www.libpng.org/pub/png/>

in which the keyword will be found (Primary or Extension), the point in the data processing where the keyword is introduced (or where the value is updated), and the meaning of the keyword (or group of keywords, if they are related). Some of the keywords are indexed by image axis, meaning they come in pairs, as indicated by the suffixes *i* and *j*.



The NEWFIRM processing pipeline is still evolving, and not all intended keywords were implemented prior to the start of data archiving in June 2009. Some keywords may be populated, but are not correctly updated by the pipeline.

**TABLE 3.8: Important Science Image Keywords**

Keyword Name	HDU	Origin	Meaning
<b>Telescope</b>			
AIRMASS	P	R	Atmospheric pathlength for target at observation start
OBSERVAT	P	R	Observatory that operates this telescope
TELESCOP	P	R	Telescope used to obtain these data
TELDEC	P	R	Declination for the telescope position on the sky in degrees
TELRA	P	R	Right ascension for the telescope position on the sky in hh:mm:ss.s
<b>Instrument/Detector Configuration</b>			
DETECTOR	E	R	Identifier of Orion detector in focal plane
FILTER	P	R	Filter name/designation
GAIN	E	R, U	Detector effective gain, in $e^-/ADU$
INSTRUME	P	R	Instrument name
NOCFSMPL	P	R	Number of Fowler samples
NOCDGAVG	P	R	Number of digital averages per pixel during readout
NOCCOADD	P	R	Number of on-board co-adds performed for this image
<b>Image Sequences</b>			
NOCDHS	P	R	Name of the image sequence script
NOCDITER	P	R	Dither iteration count
NOCDPAT	P	R	Name of dither pattern
NOCDPOS	P	R	Dither position
NOCDREP	P	R	Dither pattern repetition count
NOCDROF	P	R	Dither offset in RA in arcsec
NOCDDOF	P	R	Dither offset in Dec in arcsec

TABLE 3.8: Important Science Image Keywords (Continued)

Keyword Name	HDU	Origin	Meaning
NOCMITER	P	R	Map pattern iteration count
NOCMPAT	P	R	Map pattern
NOCMPOS	P	R	Map position
NOCMREP	P	R	Map repetition count
NOCMROF	P	R	Map offset in RA in arcmin
NOCMDOF	P	R	Map offset in Dec in arcmin
NOCNO	P	R	Observation number in this sequence
NOCSKY	P	R	Sky offset modulus
NOCTOT	P	R	Total number of exposures in sequence
<b>Time</b>			
DATE-OBS	P	R	Date and time of observation start
EXPCOADD	P	R	Duration of a single exposure in a co-add sequence, in seconds
EXPTIME	P	R	Effective exposure duration, in seconds. For stacked images the effective exposure duration can be a complicated function of position in the image.
MJD-OBS	P	R	Time of observation start in MJD
TIME-OBS	P	R	Time of observation start
TIMESYS	P	R	The principal time system for all time-related keywords. Always UTC.
<b>World Coordinates</b>			
CDi_j	E	R, U	Transformation matrix from pixel to intermediate world coordinates; CDi_i is the pixel scale for axis i
CRPIXi	E	R, U	Location of the reference point along axis i in units of pixels
CRVALi	E	R, U	Value of the world coordinate at the reference point for axis i in degrees
CTYPEi	E	R	Name of the coordinate represented in axis i
DEC	P	R, U	Declination for the center of the FoV in degrees
EQUINOX	E	R	Equinox in years for the celestial coordinate system in which the positions are expressed
NAXISi	E	R	Number of pixels along axis i
PIXSCALEi	E	R	Pixel scale along axis i in arcsec/pixel
RA	P	R, U	Right ascension for the center of the FoV in hh:mm:ss.s
RADESYS	E	R, U	Name of the reference system in which the world coordinates are expressed
WATI_nnn	E	R, U	IRAF-specific description of the nonlinear portion of the transformation from detector to world coordinates for axis i. This character string contains coefficients for a polynomial; the length of the string is such that it must continue for nnn FITS header records.
<b>Calibration</b>			
BUNIT	E	L2, U	Brightness units of image (adu   adu/s   electron/s)

TABLE 3.8: Important Science Image Keywords (Continued)

Keyword Name	HDU	Origin	Meaning
DARKINFO	P	L2	Extent to which Dark MasterCal file matches the EXPTIME and the array sampling parameters of this science image
MAGZERO	E	L2	Magnitude corresponding to one count in the image
OBSTYPE	R	R	Type of target observed (object   dark   dome flat)
PIPELINE	P	L2	Pipeline name
PLVER	P	L2	Pipeline version identifier
PROCTYPE	P	L2, U	Product type (see Table 3.5 on page 3-10)
PRODTYPE	P	L2, U	Product data description (image   mask   expmap)
PHOTCLAM	P	L2	Central wavelength of bandpass (Å)
PHOTBW	P	L2	RMS width of bandpass (Å)
PHOTDPTH	P	L2	Photometric depth of the exposure. (See “Photometric Calibration” on page 3-24.)
PHOTFWHM	P	L2	FWHM of bandpass (Å), i.e., width measured at 50% of peak transmission
SEEING	P	L2	Average FWHM of point sources, in arcseconds
SKYBG	P	L2	Brightness level of sky background averaged over all arrays, in BUNIT
SKYBG1	E	L2	Brightness level of sky background in single array, in BUNIT
SKYNOISE	P	L2	RMS noise in the background level, in ADU

### 3.2.3 Environmental Data

At present no environmental data are accessible from the Archive, although data from all-sky cameras, seeing monitors, and weather conditions at KPNO are available on the Web.<sup>5</sup>

## 3.3 CALIBRATION

The current generation of pipeline processing produces Level-2 products, or images where the instrumental signature has been removed and geometric and photometric calibrations have been applied; and Level-3 products, where spatially overlapping images in the same filter (and that have been obtained within the same observing run) have been stacked. The pipeline uses calibration exposures, such as darks and dome flats taken during an observing run, to construct the calibration reference files that are used as input to the processing. For simplicity, the subsections below will first describe the processing of the science

5. [http://www-kpno.kpno.noao.edu/Info/Mtn\\_Weather/allsky/kpasca.html](http://www-kpno.kpno.noao.edu/Info/Mtn_Weather/allsky/kpasca.html)



images once the calibration reference files are available; then the processes for constructing the calibration reference files will be described with the overall flow as context. The actual sequence of processing in the pipeline software differs somewhat in detail, in part to optimize the performance of the parallel processing environment.



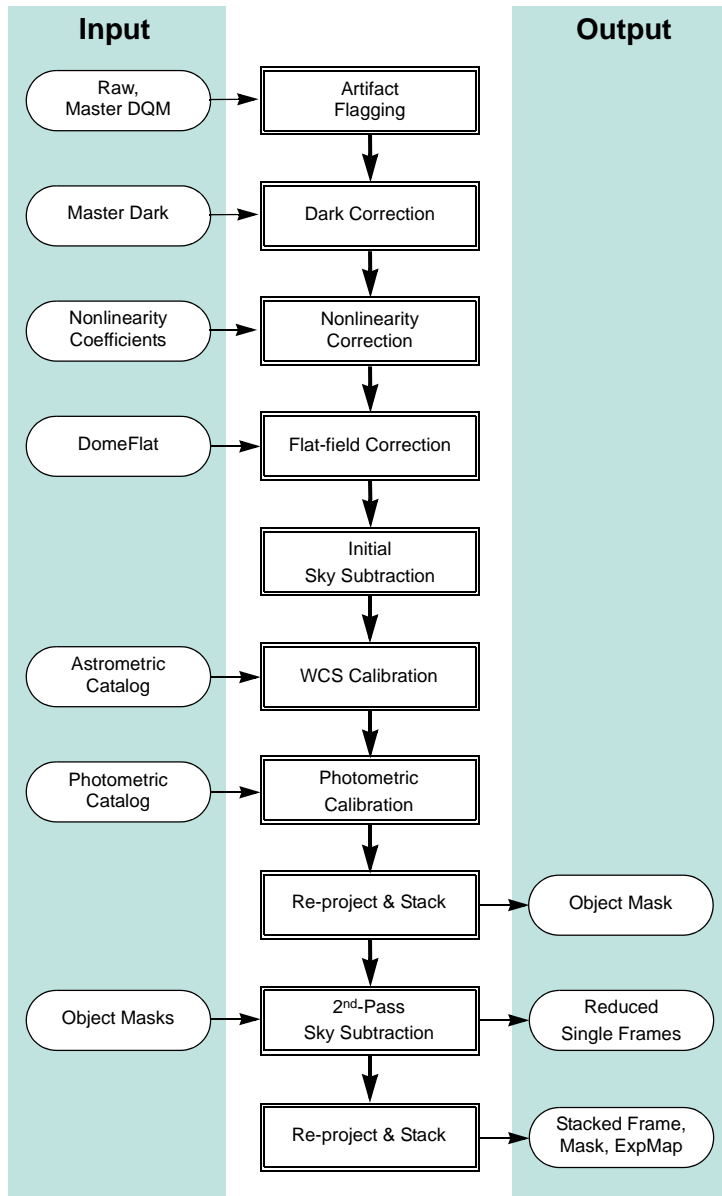

---

The calibration pipeline for NEWFIRM reductions is very much tuned to maximize the use of exposures within the observing run in which they were obtained. As such, the quality of the data depends to a large degree on the quality of the calibration exposures that were obtained by the observer. However, under some circumstances the pipeline will attempt to use calibration data from prior observing runs if the separation in time is not too great.

---

### 3.3.1 Processing Steps

The flow of the science data through the pipeline is shown in Figure 3.9. Each step of the processing, indicated by the boxes in the center of the figure, is described in detail in the following subsections. Inputs to the processing include the raw science frames, calibration reference files (see “Calibration Reference Files” on page 3-26), and photometric and astrometric catalogs. Outputs include the various reduced science images, plus their associated mask files (see “Image Formats” on page 3-11). Intermediate products that are produced during the course of pipeline processing, but that are not archived, are not shown. The processing is segregated by night, filter, and further by overlapping observing sequences when deriving stacked frames.



**FIGURE 3.9:** Flow of data through pipeline processing and calibration to produce Level-2 and Level-3 data products. External catalogs, and data products defined in Table 3.5, are shown as inputs or outputs of the processing. Intermediate products are not shown.

## Artifact Flagging

Various image artifacts, including detector blemishes as described in the section on “Focal Plane” on page 3-3, are marked in the DQM of each science image. In addition, pixels with values above 8000 DN (times the product of the number of Fowler samples and Co-adds; see Table 3.3) are marked as likely to cause persistence on the next 5 images in an observing sequence. Further artifact flagging occurs later in the processing, including *saturation*, which is detected during linearity correction, and *transients* (e.g., moving objects and cosmic rays), which are detected during the final image stack. The DQM for each InstCal image indicates all detected artifacts noted above. All flagged pixels are excluded from contributing to the final, *stacked* image; if this process results in areas with no data, such pixels are so flagged (see “Concomitant Data” on page 3-14 for details).




---

Note that all pixels in calibrated science frames (InstCal, Resampled, and Stacked) that are flagged in the DQM with a value other than zero are replaced by a linear interpolation using adjacent pixels. The one-dimensional linear interpolation is performed along the shortest dimension of the region over which to interpolate. This process is intended to avoid introducing “ringing” artifacts during the down-stream resampling of the images, as well as to mitigate scaling problems when using image display software.

---

## Dark Correction

The dark current, i.e., the signal introduced by thermal electrons in the detector without external illumination, is very low in the Orion arrays, and is typically  $0.3 e^-/s/\text{pixel}$ . However, the amplitude and structure of the dark frames is dominated by electronic bias effects, which can be a strong function of the number of on-chip co-additions, digital averages, and Fowler samples (see “Array Operation” on page 3-7), as well as the exposure time if the integrations are very short. Thus, the dark correction is applied to science frames using the Master dark frame that matches all of these parameters as closely as possible. A keyword, DARKINFO, is inserted into each science header to indicate the extent to which the master dark file matches the exposure duration and sampling type of the science exposure.

## Linearity Correction

The response of the Orion infrared detectors is not perfectly linear to incident radiation, but generally can be well characterized and corrected (see Dickinson

2009). A schematic of the detector nonlinearity is shown in Figure 3.5 on page 3-8. The linearity correction for NEWFIRM is complicated by the fact that the images as recorded by the instrument do not include all the counts that accumulated in each pixel. This effect is most severe at high incident count rates and results from the method used to record the data, as described in “Array Operation” on page 3-7. In essence, counts that accumulate between the pixel reset and the first (nondestructive) read ( $t_r$ ) are not reflected in the final difference image (i.e., what is recorded), but they must be accounted for when assessing where the accumulated counts fall on the linearity curve and for evaluating detector saturation. Thus, the magnitude of the linearity correction for a given pixel is dependent on a number of factors, including the number of co-adds, the incident count rate, the exposure duration, and location on the array. The linearity function is of the form  $\Lambda(n) = 1 + s \cdot r_0 t$ , where  $r_0$  is the true incident count rate over total time  $t$  (which is the sum of the integration time and  $t_r$ ), and  $s$  is negative. The final pixel value is the product of the measured counts and a correction factor (with a value  $>1.0$ ) that takes into account the effects noted above.

Dickinson (2009) points out that the correlated double-sample (CDS) readout technique used for the NEWFIRM detectors means that detector saturation cannot be defined simply in terms of a fixed threshold of measured counts. The counts in a CDS readout will saturate at different apparent levels, depending on the intrinsic count rate, the exposure duration,  $t_r$  (see Figure 3.5 on page 3-8), and the number of co-adds and Fowler samples. It is possible to define a threshold in the context of the linearity correction, but this approach will fail at the highest incident count rates. In this regime, the affected pixels may be near saturation by the time the first readout occurs, so that the difference between the saturated second read and the initial read (i.e., the measured counts) actually declines with increasing count rate.




---

At very high count rates, identifying a saturation condition over broad areas of the detector (as opposed to within the cores of bright stars) is difficult to do correctly within the pipeline.

---

### **Flat-Field Correction**

The flat-field correction removes the variations in the pixel-to-pixel response of the detector. In the current version of the pipeline, the flat-field is derived for each filter solely from dome-flats (i.e., images of the illuminated dome flat-field screen). The flat-field correction is performed by dividing each science frame by the flat-field calibration reference frame for the corresponding filter. The pixel values are then normalized to unit exposure time. The brightness units, as

expressed in the value of the BUNIT keyword, are  $\text{ADU s}^{-1}$  for calibrated images.

## Sky Subtraction

In ground-based IR imaging, flux from the background dominates the detected signal in broadband filters from most astrophysical sources and can exceed the flux from the faintest stars and galaxies by a few orders of magnitude. The sources of the background flux include the night sky (from  $\text{OH}^-$  emission, moonlight, and a thermal component at  $K$ -band), heat sources in the telescope dome, the telescope optics, and dark current. The molecular emission in the night sky can vary significantly on few-minute timescales and can vary spatially across the large NEWFIRM FoV. Thus, highly accurate sky subtraction is essential to good photometry, and this is facilitated by relatively short exposure durations (except for narrowband filters) and spatial dithering between exposures. *It is at this stage of the pipeline that the corrections to single images generally depend upon properties derived from multiple exposures in an observing sequence.*

Determination of the sky background for individual images is a two-pass process. In the **first pass** the sky value for each pixel in the output image is determined by computing a running median with a window of up to 9 exposures in the observing sequence, excluding masked pixels in the input images and rejecting outliers to minimize the influence of transients and astrophysical sources in the stack. Note that the input images are not geometrically registered in this step: the implicit assumption is that the images are dithered, with offsets that are larger than the inter-chip gaps and detector blemishes. For sequences of 3 or fewer images, initial sky subtraction is performed by subtracting the image closest in time (i.e., pair-wise sky subtraction). For single images (not in a sequence) a smooth, low-order polynomial is fit to the background, with rejection of pixel values that are well above the background (i.e., from sources). Finally, if offset sky exposures (i.e., exposures of relatively blank sky some distance from an extended target) were *obtained as a part of the observing sequence*, these will be combined with masking and outlier rejection and used for sky subtraction of the science target. Naturally, offset sky exposures that are




---

Note that at present the pipeline does not provide a header keyword to indicate which type of sky subtraction was actually performed for any given reduced science image.

---

dithered will result in improved sky subtraction. The sky is then subtracted from each of the science images, which are then re-projected to a common geometric

frame and combined using a median with outlier rejection. The purpose of this stacked image is to create a mask for all the astrophysical sources in the field.

In the **second pass**, the source mask is unprojected to the geometry of each original (raw) image frame. The sky is redetermined using the above prescription, except that the astrophysical sources are in addition masked from the input images. The average value of the sky in units of ADUs is written to the extension headers (SKYBG1), and the average over the four arrays is written to the Primary header (SKYBG), for each calibrated science image, including the final stack.

### ***Astrometric Calibration***

The astrometric, or World Coordinate System (WCS) calibration for science images is described by a two-dimensional polynomial (the function type and coefficients are found in the header) of a tangent-plane projection of stellar coordinates to the image pixel grid. The lower-order terms relate to the location of the reference pixel on the sky, the plate scale, and the rotation of the image, while the higher-order terms relate to nonlinear effects of the optical system such as image distortion. The form and order of the function varies with the filter and to a small degree with the airmass of the observation. The initial values for all but the zero-point terms are taken from prior, full WCS solutions of calibration images where the astrometry of the stars in the field is known with very high accuracy; the initial value for the pointing is taken from the TELRA and TELDEC header keywords, which are supplied by the telescope control system.

A full WCS solution is determined for each array by associating centroids of stars in the science images with objects in the 2MASS catalog. Since the number of suitable stars in any given image (and their distribution within the image) can vary considerably, depending upon such details as the exposure time, ambient seeing, and clouds, the input list to the WCS solution includes roughly 100 additional pseudo-stars, which are merely computed points from the trial WCS solution. Thus, if the number of genuine stars in the image is small compared to the pseudo-stars, the high-order terms in the WCS solution will be dominated by the trial solution. If instead the number of genuine stars is large, they will dominate the WCS solution. In all cases, only the genuine stars contribute to determining the zero-point.

### ***Photometric Calibration***

An estimate is made of the magnitude zero-point of each science image by comparing the instrumental magnitude of each field star to their published magnitudes in overlapping passbands in a reference catalog, applying color transformations as necessary. Currently, the 2MASS reference photometric catalog is used (see Skrutskie 2006). Note that the result of the photometric calibra-

tion is to populate the science header with keywords—the pixel values remain unchanged, and have units of  $\text{ADU s}^{-1}$ .

One quantity of use, the photometric depth of the exposure, is defined as:

$$-\frac{2.5}{2.3026} \log\left(\frac{3.988 \cdot DIQ \cdot \sigma}{\sqrt{A}}\right) + m_{zero}$$

In physical terms, the photometric depth is the faintest point-source that can result in an  $5\text{-}\sigma$  detection above the sky background, in units of magnitudes. In the equation above,  $DIQ$  is the delivered image quality (expressed as the value of the header keyword `SEEING` in the image header) in units of arcseconds,  $\sigma$  is the noise in the background in  $\text{ADU}$ ,  $A$  is the area of a pixel in  $\text{arcsec}^2$ , and  $m_{zero}$  is the magnitude zero-point. This quantity is stored in the header of the calibrated images as the value of the `PHOTDPTH` keyword.




---

Note the importance of the exposure time and the mean sky values that were determined during pipeline processing. These values, plus the detector gain, are stored in FITS header keywords and must be used to reconstruct the original source counts and background (in units of detected photons) when computing the statistical errors on measurements such as source magnitudes.

---

At present the photometric calibration is only approximate (i.e., not necessarily adequate to support science), and is most useful for the broadband  $JHK_s$  filters. For the other filters, the `MAGZERO` keyword may be useful as a rough estimate of exposure depth to support queries from the Virtual Observatory and as a measure of sky transparency over an image sequence. See “Sources of Error” on page 3-28 for details.




---

**While the magnitudes for the 2MASS catalog are of high quality, the mapping from filters used with the NEWFIRM camera to 2MASS bandpasses, and then the color transformation from 2MASS to the nonbroadband filter set, introduces potentially significant (but uncharacterized) additional uncertainties. The resulting magnitude zero-point and photometric depth given in the headers of NEWFIRM science data are uncertain by an unknown amount that could approach 0.5 mag.**

---

## Image Re-projection and Stacking

A stacked NEWFIRM image is constructed by resampling, using a *sinc* interpolation, the component science images to a common geometric grid. Note that image blemishes and other artifacts are interpolated over in this step, so as not to cause “ringing” in the resampled images. This form of the science data product is simple to use for image stacking and transient detection, without the need to handle the distortion mapping. The astrometric system is a tangent-plane projection with a uniform sampling of  $0.4 \text{ arcsec pixel}^{-1}$ , oriented with North “up” (i.e., Declination increases along Axis 2) and East “to the left” (i.e., Right Ascension decreases along Axis 1). The tangent point is selected from a grid on the sky with roughly  $1^\circ$  steps in RA, and exactly  $1^\circ$  in Dec. This selection makes it easy to stack images that are part of a sequence of small spatial dithers.

When stacking, the pipeline scales the input images to a common magnitude zero-point and sky transparency. The input images are weighted by the seeing and sky brightness. The final, stacked image is a weighted average of the input images, after masking and outlier rejection.




---

**It is common during the course of customized IR image reduction to reject images from the stack that suffer from poor image quality or unusual background levels. However, no automated mechanism for image rejection has yet been implemented in the pipeline.**

---

### 3.3.2 Calibration Reference Files

The calibration reference files that are used in the pipeline processing are constructed for each observing run where a sufficient number of appropriate exposures exist. For the master flat-field reference file, the exposures that are used for calibration are affected by multiple instrumental signatures. It is important to distinguish between the additive backgrounds and the multiplicative linearity and sensitivity variations in order to avoid biasing the photometric accuracy of the processed science frames.

Master data quality masks are used to generate a DQM for each science frame that marks the locations of artifacts (e.g., inoperable rows/columns, PED dead spots, and high-valued pixels that indicate possible saturation or persistence in subsequent images). When the final, stacked calibration frames are created for each filter these artifacts are masked from the averages.



## Dark

The master dark file is constructed very simply by averaging the frames after rejecting the minimum and maximum values (at least 5 dark images are required to create the reference file). Once constructed, the dark correction will be applied to all images, matching as closely as possible the exposure time, number of co-adds, and sampling type. A keyword, DARKINFO, inserted into each science header indicates the extent to which the master dark file matches the science exposure.




---

Since the response of the detector depends upon the rate and method of sampling the array, it is important that the master dark calibration frame correspond to the observations in exposure duration, number of co-additions, and sampling type. Significant deviations in any of these parameters will compromise the accuracy of the dark correction.

---

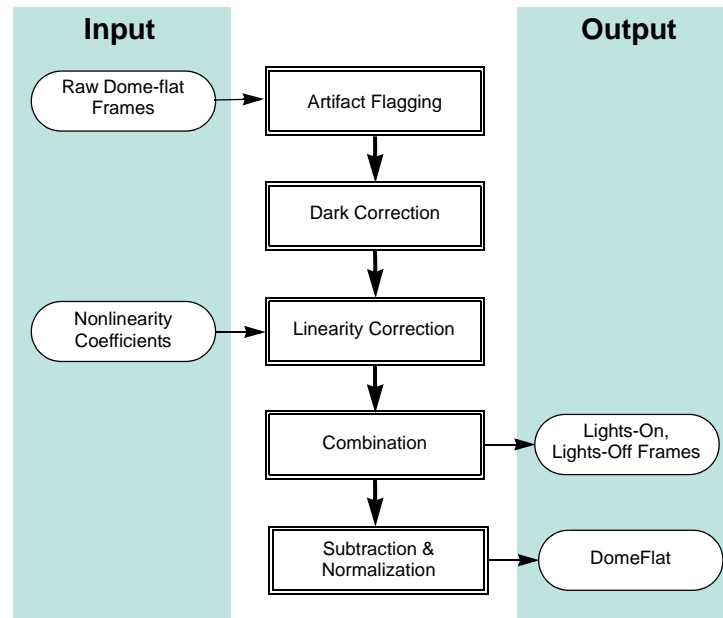
## Linearity Coefficients

The linearity coefficients are derived from engineering observations that consist of a series of dome-flat exposures that were taken with a range of exposure durations such that the full dynamic range of the detectors is sampled. The intensity of the dome flat-field lamp is monitored during this sequence, and a pixel-by-pixel mapping of the measured counts as a function of illumination level may be derived. Note, however, that at the brighter illumination levels it is necessary to correct for the counts that accumulate during the “reset-to-first read” interval, as discussed in “Array Operation” on page 3-7. In the end, the pixel-to-pixel linearity curves are characterized with simple polynomial coefficients and stored in the master linearity calibration reference file. The coefficients are believed to be temporally stable and are provided to the pipeline processing system by NOAO Science staff (see Dickinson 2009).

## Flat-Field

Dome flats, or exposures of an illuminated screen affixed to the interior of the telescope enclosure, are used to construct the flat-field reference file. The process for constructing the master dome flat, illustrated in Figure 3.10, begins with flagging artifacts, followed by the dark and linearity corrections on each frame. The images for each filter are then combined (with rejection of outliers and artifact rejection), separately for the illuminated and also for unilluminated frames. These “lights-on” and “lights-off” frames are then subtracted from one

another to eliminate the contribution from thermal emission even when the flat-field lights are off.<sup>6</sup> The result is then normalized to an average of 1.0 over all four detectors, i.e., the ratio of the sensitivity between detectors is preserved in the flat-field, and hence is removed when applied to the science images.



**FIGURE 3.10:** Process for characterizing the dome flat-field for each filter. Input dome-flat frames are uncalibrated and consist of alternating sequences with and without flat-field lamp illumination; these are combined separately and subtracted from one another.

Note that twilight- and sky-flats are neither generated nor used at present in the calibration pipeline.

## 3.4 SOURCES OF ERROR

This section describes the accuracy that can be expected for the major data products, the major sources of error, instrument foibles, and other noteworthy problems and issues with NEWFIRM data.

### 3.4.1 Astrometry

The astrometric accuracy for the image coordinate system is fundamentally limited by the science image exposure depth and the accuracy of the astrometric

6. The effect is most pronounced for *K*-band exposures.

catalog used to determine the WCS solution, which is presently that of the 2MASS catalog. This catalog has the advantage of full-sky coverage, and a limiting magnitude that provides a good overlap in the brightness range covered by most NEWFIRM exposures. Typical accuracies obtained for the WCS solution (0.2 arcsec measured in RMS deviations from the fit) are somewhat larger than the accuracy of the catalog. The geometric stability of the NEWFIRM camera has not been quantified.

### 3.4.2 Photometry

The photometric stability of the NEWFIRM camera has not been rigorously examined, but with care and good flat-fielding, it should be possible to achieve photometric accuracies of a few percent or better. However, **at present the photometric calibration provided by the pipeline processing is most adequate for science only with the J, H, and K<sub>s</sub> filters**, and otherwise is intended mainly to provide an estimate of exposure depth to support queries from the Virtual Observatory. While the systematic uncertainty may be large, within a sequence of observations covering the same region of the sky, the relative uncertainties are significantly smaller and provide a useful diagnostic of the photometric stability of the sky.

The photometric performance for any given image depends mostly on a number of external factors, especially the variable sky emission and the delivered seeing. In stacked images, the photometric depth varies discontinuously across the image, so that photometric measurements of targets that span these discontinuities may suffer some additional uncertainty.

Spatial structure in the sky background is tracked implicitly by the pixel-by-pixel robust median clipping algorithm used during the two-pass sky subtraction step. However, if the actual spatial structure of the sky background is not well characterized with this technique then, photometry of extended sources may not be accurate.

### 3.4.3 Anomalies

#### *Metadata*

Critical metadata in raw science data files may sometimes be missing or have incorrect values. However, certain of the metadata in the headers of calibrated data, notably the WCS information and the filter name, are corrected during the course of pipeline processing (unless the calibration fails). Metadata *added* by the pipeline during processing are not in question.

### ***Dome Lamp State***

The status of the dome flat-field lamps as recorded in the header keywords is unreliable, and at present is correct only by accident. The pipeline uses a simple algorithm to distinguish between images that have been illuminated by the lamp versus those that are not. The difference is obvious except in the  $K_s$ -band.

### ***Saturated Pixels***

Pixel saturation (where the accumulated charge exceeds the full-well, e.g., in the cores of over-exposed stars) is not easy to characterize in extended regions illuminated at a high count rate. Saturated regions in flat-field exposures, for example, may not always be flagged accurately in the DQM file.

### ***Image Ghosts***

Bright stars can produce three different pupil ghosts, the effects of which may be difficult to remove. Their characteristics in the  $J$ -band filter are the following:

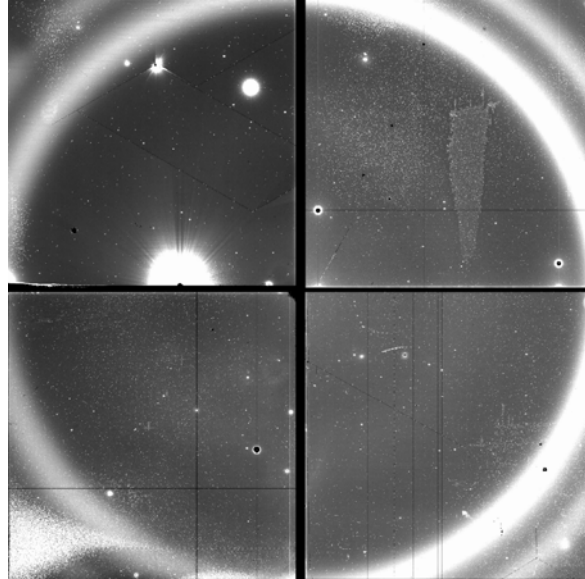
- A roughly 300 pixel diameter ghost, about 17 mag/arcsec<sup>2</sup> fainter than the integrated magnitude of the source star. The ghost is offset radially by hundreds of pixels with respect to the field center; the size of the offset depends upon the distance of the source from the field center.
- A roughly 100 pixel diameter ghost, about 14 mag/arcsec<sup>2</sup> fainter than the integrated magnitude of the source star, and nearly centered on it. This ghost could give rise to the appearance of false nebulosity around bright stars in stacked images.
- A compact, star-like ghost of 4 pixels diameter, which is 8.5 mag fainter than the source star and offset radially from it by hundreds of pixels, in a field-dependent manner. This ghost might appear in stacked images as bogus point sources near bright stars. But since the radial position of this ghost is field-dependent, it may be rejected from the stack if the dither pattern is sufficiently large.

### ***Background for Narrowband Filters***

Images of astrophysical targets obtained with the ultra-narrow filters (1.056  $\mu\text{m}$  and 1.063  $\mu\text{m}$ ) contain ring-like features in the background, as shown in Figure 3.11. These ultra-narrowband filters are designed to capture the designated emission line feature with high transmission over the entire field of view at zero redshift. Owing to the convergent beam of the optical system, the filter bandpass shifts in wavelength with radial position in the FoV: at field center, the designated line falls near the blue edge of the high transmission window; at field corner, near the red edge. The “ring” features are due to the presence of OH emission lines in the night sky background near the bandpass of the narrowband filter. The bandpass shifts slightly as a function of position (distance from the

field center, in particular), and thus different amounts of the OH emission pass through the filter as a function of radius from the field center, leading to the rings.

The ring features are an additive background that varies in time. It is easily characterized, although the current version of the pipeline processing does not remove it explicitly.



**FIGURE 3.11:** Image of the night sky with the 1.056  $\mu\text{m}$  ultra-narrowband filter showing the “ring” feature in the background, which results from night-sky OH emission at the edge of the filter bandpass.

## 3.5 REFERENCES & FURTHER INFORMATION

### **Contributing Authors**

Several people have contributed to the technical knowledge of the NEWFIRM camera. Extensive material has been contributed by Ron Probst (NEWFIRM Project Scientist) and Mike Merrill (Telescope Scientist and detector expert); by the pipeline development team of Mark Dickinson, Frank Valdes, Rob Swaters, and Tracy Huard; and by Dick Shaw. Documentation on the reduction of NEWFIRM data with the IRAF IRRED package is also fairly extensive, much of which was authored by Frank Valdes.

## References

- Dickinson, M., Huard, T., Swaters, R., & Valdes, F. G. 2008, *NEWFIRM Quick Reduce Pipeline and Data Analysis Tools* (Tucson: NOAO)
- Dickinson, M. 2009, *NEWFIRM Linearity Calibration and Correction* (Tucson: NOAO)
- Fowler, A. M., & Gatley, I. 1990, *ApJ*, 353, L33
- Hoffman, A. W., Corrales, E., Love, P. J., Rosbeck, J., Merrill, M., Fowler, A., & McMurtry, C. 2008, *2K x 2K InSb for Astronomy*, SPIE, 5499, 59
- Merrill, K. M. 2008, *Orion SCA Laboratory Test Results Summary* (Tucson: NOAO)
- Probst, R. 2007, *NEWFIRM Quick Guide for Proposal Preparation* (Tucson: NOAO)
- Probst, R. 2008, *NEWFIRM Start-Up, Data Taking, and Error-Recovery Procedures* (Version 3.2; Tucson: NOAO)
- Probst, R. G., George, J. R., Daly, P. N., Don, K., & Ellis, M. 2008, *First Light with NEWFIRM*, SPIE, 7014, 93
- Skrutskie, M. F., et al. 2006, *The Two Micron All-Sky Survey (2MASS)*, *AJ*, 131, 1163

## For Further Reading

A document repository<sup>7</sup> for NEWFIRM is available on the Web, where many of the above-referenced papers may be found. The *JHK<sub>s</sub>* filter bandpasses follow the Mauna Kea Observatory prescription:

Tokunaga, A. T., Simons, D. A., & Vacca, W. D. 2002, *PASP*, 114, 180

The reduction of NEWFIRM data outside the automated pipeline, including a general discussion of IRAF data reduction techniques for IR mosaics, may be found in:

Dickinson, M., & Valdes, F. G. 2009, *A Guide to NEWFIRM Data Reduction with IRAF* (Tucson: NOAO)

Additional papers on the NOAO pipeline processing system generally and the NEWFIRM pipeline in particular have appeared in the proceedings of the ADASS conference series.

7. Available at <http://www.noao.edu/ets/newfirm/>

Scott, D., Pierfederici, F., Swaters, R. A., Thomas, B., & Valdes, F. G. 2007, *The NOAO High-Performance Pipeline System: Mosaic Architecture Overview*, in ASP Conf. Ser. 376, ed. R. A. Shaw, F. Hill, & D. J. Bell (San Francisco: ASP), 265

Swaters, R. A., Valdes, F., & Dickinson, M. E., 2009, *The NOAO NEWFIRM Pipeline*, in ASP Conf. Ser., ed. D. Bohlender, P. Dowler, & D. Durand (San Francisco: ASP), in press<sup>8</sup>

---

8. Available at <http://arxiv.org/abs/0902.1458>





## *Glossary*

---

Version 1.1, May 2009

The following is a Glossary of common terms and acronyms found throughout this *Handbook*.

2MASS	Two-Micron All Sky Survey
ADC	Atmospheric Dispersion Corrector
ADU	Analog-to-Digital Unit (see DN)
AURA	Association of Universities for Research in Astronomy
CCD	Charge-coupled device, a common detector for the visible waveband
Bias	The nearly constant pedestal voltage, seen as an offset from zero DN, of a raw CCD frame. Also refers to a zero-second exposure used to determine the spatial structure of the bias level in the CCD
CTIO	Cerro Tololo Inter-American Observatory, located near La Serena, Chile and operated by NOAO/AURA for the NSF
Digital Average	During the read-out of an infrared detector array, two or more sequential reads of a single pixel are averaged to obtain the final value before the next pixel is read
Dither	An observing sequence of spatially overlapping exposures (with small relative offsets) intended to image a continuous region of sky free of interdetector gaps and other array defects
DN	Digital Number, or the quantized value of a pixel prior to conversion to electrons (or detected photons) via the detector gain. Synonymous with ADU.
DQM	Data quality mask
E2E	End-to-End system: The system that manages the flow of information and data from the database of approved observations, through data acquisition, transport from the telescope to an archive center, pipeline processing, and retrieval by an end-user
FITS	Flexible Image Transport System, the standard for data interchange in astronomy
Flat-field	Generally refers to an image of a uniform source in the relevant coordinates (spatial or wavelength/frequency), which is used to calibrate variations in the detector pixel-to-pixel sensitivity

FoV	Field of view (generally, for an imaging device or a telescope)
Fowler Sample	During the read-out of an infrared detector, a Fowler Sample refers to multiple, sequential reads of the entire array; this procedure is performed both at the start and at the end of an exposure, and the mean difference is retained
FPA	Focal plane array: Multiple, abutting detectors mounted in the focal plane of a camera, arranged in an array
IR	Infrared portion of the spectrum, which for a ground-based observatory is generally restricted to the near-IR range: 0.9–2.5 $\mu\text{m}$
KPNO	Kitt Peak National Observatory, located near Tucson, Arizona, USA and operated by NOAO/AURA for the NSF
Map	An observing sequence of spatially adjacent exposures (usually with some overlap) intended to image a region of sky that is significantly larger than the camera FoV
MEF	Multi-Extension FITS file, i.e., a file containing multiple data structures such as tables or binary images
NEWFIRM	NOAO Extremely Wide-Field Infrared Mosaic camera
NOAO	National Optical Astronomy Observatory
NSA	NOAO Science Archive
NSF	National Science Foundation
PI	Principal investigator, generally of an approved observing program
PSF	Points spread function
SMARTS	Small and Moderate Aperture Research Telescope System, located at CTIO near La Serena, Chile, and operated by a consortium of several universities and NOAO
SOAR	SOuthern Astrophysical Research telescope, located on Cerro Pachón near La Serena, Chile, and operated by a consortium of the Brazilian Ministry of Science, NOAO, the University of North Carolina, and Michigan State University
SOI	SOAR Optical Imager
VO	Virtual Observatory: Generally refers to the global, interconnected data repositories, supporting services, and tools that have been built by IVOA member organizations of the International Virtual Observatory Alliance
WCS	World Coordinate System, generally refers to the mathematical mapping from detector coordinates to a physical coordinate system, such as position in Right Ascension and Declination, as defined by FITS conventions
WIYN	A consortium of the University of Wisconsin, Indiana University, Yale University, and NOAO operating an observatory on Kitt Peak