Visible & Infrared Survey Telescope for Astronomy

Data Flow System

Document Title: **VISTA Infra Red Camera Data Reduction Specifications**

Document Number: **VIS-SPE-IOA-20000-0003**

Issue: **0.5**

Date: **2004-04-08**

Document Prepared by:	Peter Bunclark (CASU)	Signature And date:	
Document Approved by:	Mike Irwin (CASU Manager)	Signature And date:	
Document Released by:	Jim Emerson (VDFS Project Leader)	Signature And date:	
Document Reviewed by:	William Sutherland (VISTA Project Scientist)	Signature And date:	

Change Record

Issue	Date	Sections affected	Remarks
0.5	2004-04-08	All	New Document

Notification List

The following people should be notified by email that a new issue of this document is available.

IoA:	W Sutherland
Oxford:	G Dalton
QMUL	J Emerson
ATC	Malcolm Stewart Steven Beard

Table of Contents

This page intentionally left blank.

1 Introduction

1.1 Purpose

This document forms part of the package of documents for the Preliminary Design Review of the Data Flow System for VISTA, the Visible and Infra-Red Survey Telescope for Astronomy. As stated in [AD1] this document defines the detailed data reduction procedures to be applied to raw data frames including input data and algorithms required. This document defines all the algorithms required for the reduction of raw data from the instrument. The Data Flow System pipeline will provide a set of standard functions (e.g. logic and arithmetic on images, manipulation on tables). The algorithms implement the procedures defined in the Calibration Plan [AD3]. A table of contents of the Data Reduction Specifications is provided as specified in [AD1] as required for PDR.

The processing infrastructure which is assumed is outlined in Appendix A.

1.2 Scope

This document describes the VISTA Infra-Red Camera Data Reduction Specifications for the output from the 16 Raytheon VIRGO IR detectors in the Infra Red-Camera for VISTA (VIRCAM). The baseline requirements for calibration are included in the VISTA Infra-Red Camera Data Flow System User Requirements [AD2], and the Calibration Plan is described in [AD3].

1.3 Applicable Documents

The following documents, of the exact issue shown, form part of this document to the extent specified herein. In the event of conflict between the documents referenced herein and the contents of this document, the contents of this document shall be considered as a superseding requirement.

[AD1] *Data Flow for the VLT instruments requirements specification,* VLT-SPE-ESO-19000-1618, issue 1.0, 1999-04-21.
[AD2] *VISTA Infra Red Camera DFS User Requirements,* VIS-SPE-IOA-20000-00001, issue 0.5, 2004-04-08.
[AD3] *VISTA Infra Red Camera DFS Calibration Plan*, VIS-SPE-IOA-20000-00002, issue 0.5, 2004-04-08.

1.4 Reference Documents

The following documents are referenced in this document.
[RD 1] *Data Interface Control Document*, GEN-SPE-ESO-19940-794, issue 2.0, 2002-05-21.
[RD 2] *VISTA IR Camera Software Functional Specification*, VIS-DES-ATC-06081-00001, issue 2.0, 2003-11-12.

[RD 3] *IR Camera Observation Software Design Description*, VIS-DES-ATC-06084-0001, issue 2.0, 2004-04.

[RD 4] *VISTA Science Requirements Document*, VIS-SPE-VSC-00000-0001, issue 2.0, 2000-10-26

[RD 5] *Overview of VISTA IR Camera Data Interface Dictionaries,* VIS-SPE-IOA-20000-0004, 0.1, 2003-11-13

[RD 6] *Definition of the Flexible Image Transport System (FITS),* NOST 100-2.0

1.5 Abbreviations and Acronyms

ADU	Analogue to Digital Unit
DFS	Data Flow System
FITS	Flexible Image Transport System
PDR	Preliminary Design Review
TBC	To Be Confirmed
VDFS	VISTA Data Flow System
VIRCAM	VISTA Infra Red Camera
VISTA	Visible and Infrared Survey Telescope for Astronomy
WFCAM	Wide Field Camera (on UKIRT)
WCS	World-Coordinate System
2MASS	2 Micron All Sky Survey

1.6 Glossary

Confidence Map An integer array, normalized to a median of 100% which is associated with an image. Combined with an estimate of the sky background variance of the image, it assigns a relative weight to each pixel in the image and automatically factors in an exposure map. Bad pixels are assigned a value of 0. It is especially important in image filtering, mosaicing and stacking.

DAS Data Acquisition System

Exposure The stored product of many individual **integration**s that have been co-added in the DAS. The sum of the integration times is the exposure time.

Integration A simple snapshot, within the **DAS**, of a specified elapsed time. This elapsed time is known as the integration time.

Jitter (pattern) A pattern of **exposures** at positions each shifted by a small **movement** (<30 arcsec) from the reference position. Unlike a **microstep** the non-integral part of the shifts is any fractional number of pixels. Each position of a jitter pattern can contain a **microstep** pattern.

Mesostep A sequence of **exposures** designed to completely sample across the face of the detectors in medium-sized steps, in order to monitor residual systematics in the photometry.

Microstep (pattern) A pattern of **exposure**s at positions each shifted by a very small **movement** (<3 arcsec) from the reference position. Unlike a **jitter** the non-integral part of the shifts are exact fractions of a pixel, which allows the pixels in the series to be interlaced in an

	effort to increase resolution. A microstep pattern can be contained within each position of a **jitter** pattern.
Movement	A change of position of the telescope that is not large enough to require a new guide star.
OB	Observation Block
Offset	A change of position of the telescope that is not large enough to require a telescope **preset**, but is large enough to require a new guide star.
Pawprint	16 non-contiguous images of the sky produced by VIRCAM with its 16 non-contiguous chips (see Fig 2-2 of [AD3]). The name is from the similarity to the prints made by the padded paw of an animal (the terminology was more appropriate to 4-chip cameras).
Preset	A telescope slew to a new position requiring a reconfiguration of various telescope systems.
Robust Estimate	A statistical estimator that is resilient to small perturbations on the assumed shape of the underlying distribution.
Tile	A filled area of sky fully sampled (filling in the gaps in a pawprint) by combining multiple **pawprint**s. Because of the detector spacing the minimum number of pointed observations (with fixed offsets) required for reasonably uniform coverage is 6, which would expose each piece of sky, away from the edges of the tile, to at least 2 camera pixels. The pipeline does not combine **pawprint**s into tiles.

2 Document Overview

This document provides a description of the data reduction recipes that will be applied to raw VISTA data in order to process it as outlined in the VISTA IR Camera Calibration Plan [AD3]. Readers are referred to that document and the User Requirements Document [AD2] for a description of the camera and detectors.

The requirements are to:
1) remove the two-dimensional instrumental signature from the images;
2) compute and apply photometric and astrometric calibration to the derived products;
3) provide measures for Data-Quality Control.

Section 3   is a description of the recipes required to construct the calibration images and tables from the appropriate calibration observations. All these recipes are concerned with the removal of the 2-D instrumental signature.

Section 4   is a description of the recipes required to measure additional calibration data from the science data frames themselves. These recipes are concerned with both the 2-D instrumental signature removal and the measurement of astrometric and photometric parameters.

Section 5   describes three possible scenarios under which the pipeline may apply calibration data, and illustrates how recipes will be combined to form a data reduction procedure.

Section 6   is a formal list of the high-level data reduction modules which will be used to construct the recipes outlined in sections 3 and 4 and the procedures in section 5

3 Recipes on Calibration Data for Instrument Signature Removal

3.1 Purpose

This section describes what pipeline recipes are executed on the calibration data collected to allow instrumental signatures to be removed. These refer specifically to individual calibration items in [AD3]. Each recipe below is described by the following fields:

- **Purpose:** The task the recipe is to perform
- **Algorithm:** The method the recipe uses to perform the task
- **Modules:** A list of the main data reduction modules used in the recipe. This is *not* a comprehensive list of all the modules that will be needed by the recipe.
- **Prerequisites:** Possible dependencies
- **See Also:** Any further information

3.2 vircam_reset_combine

Purpose: Combine a sequence of reset frames to form a mean frame. Compare to a library reset frame and look for stability of the pedestal and reset structure.

Algorithm: Combine the sequence of reset frames into a single mean with rejection. Compare with a library reset frame. Test the difference for a global level, and split the difference frame into small sections and do a robust median estimate in each. Test if the medians are zero to within the noise properties of the image.

Modules: vircam_imcombine

Prerequisites:

See Also:

3.3 vircam_dark_combine

Purpose: Combine a series of dark frames taken with a particular integration and exposure time combination. Compare with a similarly observed master dark frame. Look for variation in the reset anomaly structure and scale. Estimate dark counts in the

new frame and compare with recent measurements.

Algorithm: Combine the sequence of dark frames with rejection. In conjunction with confidence maps, assess the number of rejected pixels to give an indication of the rate of cosmic ray hits and their properties. Work out a robust median in a region that is unaffected by reset anomaly. Compare to previous values to identify any trend. Compare this frame from a library dark frame. Test global median, which should be near zero. Divide the difference frame into sections and do a robust background estimate in each section, which will identify any regions where the structure of the reset anomaly is varying with time.

Modules: vircam_imcombine

Prerequisites:

See Also:

3.4 vircam_dome_flat_combine

Purpose: Combine a series of dome flat exposures and measure the number of bad pixels. Create confidence maps and compare these to previous ones to look for trends.

Algorithm: Combine the dome flat exposures with rejection. Using the background following algorithm and the ratio of dome flats at different illumination levels; create a mask of pixels that fall above or below the local median by a threshold amount. Compare the number of bad pixels on the detector with previous measures to identify any trends.

Modules: vircam_imcombine, vircam_mkconf

Prerequisites:

See Also:

3.5 vircam_detector_noise

Purpose: Measure the detector readout noise and gain

Algorithm: Subtract two dome flat frames and measure the variance of the difference frame σ_f^2. Do the same for the two dark frames to get σ_d^2. If the background means are locally: m_f and m_d the local gain in electrons per ADU is: $\varepsilon = (m_f + m_d)/(\sigma_f^2 - \sigma_d^2)$ and the readout noise in electrons is $\rho = \varepsilon \sigma_d^2 / \sqrt{2}$.

Modules:

Prerequisites: A pair of dome flat frames and a pair of dark frames are required. The dome flats must be taken under similar illumination.

See Also:

3.6 vircam_linearity_analyse

Purpose: Create linearity curves for each detector

Algorithm: The background statistics for a series of dome flats are measured. A polynomial is fitted to the exposure time vs. background mean data. A look-up table of observed counts to linear counts is generated by evaluating the coefficients over the complete data range to calculate a 'time' and then dividing by the linear coefficient to convert the answer to 'linear counts'. Test the result by dividing dome flats at different illumination levels using gradients and relative gain of detectors for comparison.

Modules: vircam_genlincur

Prerequisites: Dome flat observations should cover the full dynamic range of the detectors.

See Also:

3.7 vircam_twilight_combine

Purpose: Create a mean twilight flat-field

Algorithm: Examine a list of twilight flat-field exposures and reject all those that are over or under exposed. Robustly scale and map detector pixel-to-pixel sensitivity variations. Combine the rest with rejection. Normalize each detector's flat-field by the ensemble median level over all the detector's flat-fields. The normalized level in each image is the detector gain correction. Update the confidence map from mean flat properties.

Modules: vircam_imcombine, vircam_mkconf

Prerequisites:

See Also:

3.8 vircam_illumination_analyse

Purpose: Create a map of illumination corrections using a dense grid of secondary standards

Algorithm: Calculate the photometric zero-point for stars on an image with an assumed value of the extinction. Divide the image into sections and calculate the median residual zero-point for each section. The illumination correction for a given place on the image is then a bilinear interpolation over the adjacent sections.

Modules: vircam_imcore, vircam_getstds, vircam_illum

Prerequisites:

See Also:

3.9 vircam_mesostep_analyse

Purpose: Create a map of illumination corrections using a mesostep sequence of a standard stars.

Algorithm: Compute the zero-point of standard stars on a grid of each of a series of meso-stepped exposures. Work out the residual zero-point at each position on the detector relative to the zero-point at

the centre. The illumination correction for a given place on the image is then a bilinear interpolation over adjacent mesostep locations.

Modules: vircam_imcore, vircam_getstds
Prerequisites:
See Also:

3.10 *vircam_persistence_analyse*

Purpose: Analyse an image of bright stars and subsequent dark exposures to compute the persistence decay rate and look for any adjacency effects.

Algorithm: Compute the flux and position of bright stars on an image. Look on subsequent dark exposures at the same location and compute the flux. Fit the flux vs. Δt curve to an exponential to work out the characteristic decay constant, τ_0. Examine size of residual structure to quantify adjacency effects.

Modules: vircam_imcore
Prerequisites:
See Also:

3.11 *vircam_crosstalk_analyse*

Purpose: Analyse a series of images to work out the crosstalk matrix for all detector sections.

Algorithm: Analyse a series of images where bright stars have been placed on each channel of each detector in turn. Compute the flux of the bright stars and compute the flux in the expected location of a crosstalk image in each readout channel. The ratio of the two gives the crosstalk factor for each section and detector combination.

Modules: vircam_imcore
Prerequisites:
See Also:

4 Recipes on Calibration Data Derived from Science Data

This section describes the calibration data that will be derived from the science data themselves and how it will be processed. These refer specifically to individual calibration items in [AD3]. Each recipe below is described by the following fields:

- **Purpose:** The task the recipe is to perform
- **Algorithm:** The method the recipe uses to perform the task

- **Modules:** A list of the main data reduction modules used in the recipe. This is *not* a comprehensive list of all the modules that will be needed by the recipe.
- **Prerequisites:** Possible dependencies
- **See Also:** Any further information

4.1 Instrument Signature Removal

4.1.1 vircam_sky_flat_combine

Purpose: Combine sky or object observations to create a mean sky flat and confidence map.

Algorithm: A list of observations is generated from a scan of FITS headers. This is either a list of target images or offset sky images, depending on the type of observations being done. The images are combined with scaling and rejection. The mean image for each detector is divided by the ensemble average background over all the detector images. A confidence map can be generated from the mean flat and ratio of flats taken under different illumination levels.

Modules: vircam_imcombine, vircam_mkconf

Prerequisites:

See Also:

4.1.2 vircam_offset_sky_combine

Purpose: Combine sky or object observations to create a mean sky and fringe frame.

Algorithm: A list of observations is generated from a scan of FITS headers. This is either a list of target images or offset sky images, depending on the type of observations being done. The images are combined with scaling and rejection.

Modules: vircam_imcombine

Prerequisites:

See Also:

4.1.3 vircam_jitter_combine

Purpose: Combine a jitter sequence into a single image.

Algorithm: Detect all the astronomical objects on each frame from a jitter sequence. Cross-correlate the positions to work out the pixel shifts between the frames. Shift and combine the images to form a single image. Weighting using the confidence map and rejection with a suitable threshold ensures an image as free as possible from bad pixels and other artefacts.

Modules: vircam_imcombine, vircam_imcore, vircam_matchxy

Prerequisites:

See Also:

4.1.4 vircam_microstep_interleave

Purpose: Combine a microstep sequence into a single image

Algorithm: The headers of a microstep sequence are examined to find the *x,y* shifts. Images are shifted and interleaved onto an output grid. No pixel rejection is possible.

Modules: vircam_interleaf

Prerequisites:

See Also:

4.2 Astrometric Calibration

4.2.1 vircam_wcs_fit

Purpose: Compute a full WCS for each image

Algorithm: Objects are detected on a frame. These are matched up to astrometric standard stars from a given catalogue (e.g. 2MASS), using the initial WCS that is available in the header. A plate constant solution is obtained in the standard way, using the projection geometry and distortion terms that have been previously determined.

Modules: vircam_imcore, vircam_getstds, vircam_matchstds, vircam_platesol

Prerequisites:

See Also:

4.2.2 vircam_distortion_update

Purpose: Compute optical distortion coefficients for the model WCS

Algorithm: The refinement of the projection to be used for the WCS will be determined at commissioning. A good estimate of the distortion coefficients are derived from the optical design, but these will in general be wavelength dependent. Observations of astrometric fields in each wavelength band will be fit in a standard way to determine the best value of the distortion coefficients for each waveband.

Modules: vircam_imcore, vircam_getstds, vircam_matchstds, vircam_platesol

Prerequisites:

See Also:

4.2.3 vircam_catalogue_gen

Purpose: Generate a catalogue of objects on an image

Algorithm: A standard image detection algorithm is used to extract objects from an image. Estimates of the mean sky level, sky noise,

seeing, stellar ellipticity and number of noise objects are derived from the catalogue parameters.

Modules: vircam_imcore
Prerequisites:
See Also:

4.3 Photometric Calibration

4.3.1 vircam_photcal_fit

Purpose: Calculate mean zero-point and extinction coefficient
Algorithm: Instrumental and true magnitudes of stars in a standard field are used in conjunction with repeat observations over several airmass values to compute the mean zero-point and extinction coefficient.
Modules: vircam_imcore, vircam_getstds, vircam_matchstds
Prerequisites:
See Also:

4.3.2 vircam_photcal_apply

Purpose: Apply photometric calibration to object data
Algorithm: Mean zero-point and extinction coefficients are calculated for each filter using 4.3.1. The objects are photometrically calibrated by simply writing these values into the header of the object images and catalogues. This keeps catalogue and image data in observed units, obviating the need to rewrite the data if a better photometric calibration is done in the future.
Modules:
Prerequisites:
See Also:

5 Instrumental Signature Removal for Science Data

Applying calibration information to science data in order to remove the instrumental signature can be done in several ways. The results of laboratory tests on the instrumental detectors and the results of on-sky tests during commissioning will help us to decide on the exact strategy to be used. In the mean time, several scenarios are given below which will aid in understanding some of the data reduction issues. This is not intended to be a comprehensive list of all possible options, but rather a list of some of the most likely ones. Section 5.1 deals with the removal of dark current, reset anomaly, flat-field and sky background. Once this has been done crosstalk and persistence effects can be subtracted out. Combining the jitter and microstep sequences is covered in sections 4.1.3 and 4.1.4. The final step, which is the astrometric and photometric calibration, is done using the recipes defined in 4.2.1 and 4.3.2.

5.1 Dark Current, Reset Anomaly, Flat-field and Sky

5.1.1 Dark Current, Reset Anomaly, Flat-field, Sky All Stable

Assumptions:
1. Flat-field stable over at least the night
2. Reset anomaly is of fixed pattern and scale
3. Dark current stable over at least the night for a given exposure and integration time combination
4. Sky background stable over at least the length of an OB.

Method:
1. Subtract the appropriate mean dark frame to remove dark current and reset anomaly
2. Divide by an appropriate twilight flat-field.
3. Get a sky estimate either from offset sky exposures or the object data and subtract this from each object image.

5.1.2 Dark Current, Reset Anomaly, Flat-field Stable, Sky Unstable

Assumptions:
1. Flat-field stable over at least the night
2. Reset anomaly is of fixed pattern and scale
3. Dark current stable over at least the night for a given exposure and integration time combination
4. Sky background (fringing in particular) varies over the length of an OB.

Method:
1. Subtract the appropriate mean dark frame to remove dark current and reset anomaly
2. Divide by an appropriate twilight flat-field.
3. Get a sky estimate either from offset sky exposures or the object data.
4. Scale the differential sky background data so that spatial features, e.g. fringes, will be removed to the highest degree possible and subtract.

5.1.3 Dark Current and Reset Anomaly Unstable, Flat-field, Sky Stable

Assumptions:
1. Flat-field stable over at least the night
2. Reset anomaly is of fixed pattern and scale for a jitter sequence, but not more.
3. Dark current stable for at least the length of a jitter sequence, but not more.

Method:
1. Subtract the sky estimate. This is either found from a combination of the offset sky images or from the object frames themselves. This has the effect of removing the reset anomaly, dark current and sky background.
2. Divide by an appropriate twilight flat-field.

6 Data Reduction Modules

Below is a list of the high level data reduction modules that will be needed by the recipes, some of which have been referred to in sections 3, 4 and 5. They are grouped by the type of operation they perform.

Instrumental Signature Removal

vircam_stage1	Linearise, remove dark current and reset anomaly and divide by flat-field
vircam_defringe	Work out optimal scaling factor for fringe removal. Subtract out the fringe frame.
vircam_crosstalk	Remove crosstalk given a crosstalk matrix
vircam_persist	Remove persistence images given a list of data frames and persistence decay constant.
vircam_genlincur	Given a list of exposure times and fluxes, generate a linearity curve
vircam_illum	Given observations of a standard field, calculate an illumination correction table.

Astrometric and Photometric Calibration

vircam_coverage	Work out the coverage in equatorial coordinates of an image
vircam_imcore	Extract objects from images and generate a catalogue
vircam_getstds	Get a list of photometric or astrometric standards from a catalogue that may appear on the current image.
vircam_matchstds	Match a list of standard stars to the x,y positions of objects on a frame.
vircam_platesol	Work out a plate solution for an image given the α, δ, x, y values for objects on the image. Write the solution to the image header.

Image Combination

vircam_imcombine	Combine a list of images. Allow for pixel shifts (jitters), mean or median output, several rejection algorithms, confidence map support (input and output), image biasing and/or scaling before combination.
vircam_interleaf	Combine a list of images from a microstep sequence by interleaving.
vircam_imstack	Stack images onto a resampled output grid. The input header WCSs will determine the mapping of input to output grid.
vircam_matchxy	Work out jitter offsets by cross-correlating the locations of objects on a set of frames.

Miscellaneous

vircam_mkconf	Create a confidence map from flat-field images.

Appendix A.   Infrastructure for pipeline modules

As defined in the Calibration Plan [AD3], VIRCAM data will be ESO-modified standard FITS [RD 6], specifically using the *hierarchical header* proposal. The data will be stored as 32-bit signed integers in Multi-Extension Format (MEF). Because of the universal acceptance of the FITS standard, there are several high-quality FITS manipulation packages available, providing varying levels of rigorously tested functionality which aid the programmer in operating on data files. Below is a list of some of the functionality which is assumed present in the underlying FITS infrastructure:

- Image and image section access:
  The pipeline modules need to be able to specify a subset of the FITS file being operated on; this is specified using the notation:
    fff.fits[1]
    fff.fits[1][100:150,200:237]
  The first example opens the first extension of the file "fff.fits". The second example opens the first extension of the file and gives access to a subset of the data array. This is a very common practice.
- Table subsets and selection
  The ability to open a subset of a FITS table according to some selection criteria. For example, "open this FITS table and give access only to those rows where the object is in a particular search area and whose magnitude is below this threshold".
- Error status passing
  All routines must pass back a status value and must act on a bad input status value.
- Header keywords 'classified'
  Having classes of header keywords helps when building a new file from an old one
- A WCS interface
  Allows access and manipulation of WCS parameters
- A scripting interface
  An interface to a common scripting language like Perl or Python greatly aids prototyping.
- Tile Compression
  This is a compression method applied to image data arrays in FITS files, whereby the array is divided into a number of sections. Each of these sections is compressed using the RICE algorithm and the result is stored as a column element in a binary FITS table. With integer data, the compression is fast, lossless, and typical factors of 4 in compression reduce the disk I/O rate enormously and therefore speed up processing, and also reduce storage and shipping costs. A compressed FITS file is still a valid FITS file and as such can be read by any binary-

FITS table reader (i.e. this is *not* the same as using a compression utility such as 'gzip').

oOo