Visible & Infrared Survey Telescope for Astronomy

Data Flow System

Document Title: **UK VISTA User Requirements**

Document Number: **VDF-SPE-IOA-00001-0002**

Issue: **1.0**

Date: **2004-10-11**

Document Prepared by:	**Will Sutherland** (VISTA Project Scientist)	**Signature And date:**	Will Sutherland Digitally signed by Will Sutherland DN: CN = Will Sutherland, C = GB, O = Cambridge University, OU = Institute of Astronomy Reason: author Date: 2004.10.11 16:16:47 +01'00'
Document Approved by:	**Jim Emerson** (VISTA PI)	**Signature And date:**	
Document Reviewed by:	**Tim Naylor** (VDUC Chair)	**Signature And date:**	
Document Accepted by:	**Mike Irwin** (VDFS Pipeline Manager, CASU)	**Signature And date:**	
	Peredur Williams (VDFS Science Archive Manager, WFAU)	**Signature And date:**	
Document Released by:	**Jim Emerson** (VDFS Project Leader)	**Signature And date:**	

Change Record

Version	Section	Changes
1.0		First official release.

Notification List

The following shall be notified when a new version of this document is issued

Exeter	Tim Naylor
ATC	Malcolm Stewart

Contents

1 Introduction

1.1 Scope of this Document

This document specifies the top-level requirements on the UK pipeline and science archive of the VISTA Data Flow System (VDFS); effectively, it specifies the requirements on the software pipeline which takes the UK copy of the raw VISTA science and calibration data from the VISTA IR Camera (VIRCAM), removes the instrumental signature, produces photometrically and astrometrically calibrated pawprints, combines pawprints at one area to go deeper and/or into contiguous tiles or larger regions, generates object catalogues, and curates this data providing data access and advanced search capabilities to suitably authorised users.

The VDFS project also delivers basic processing software (i.e. derivations of quality control measures, removal of instrumental signature and photometric and astrometric calibration) to ESO. Once delivered this software will be maintained and operated by ESO, who will run it at its sites at Cerro Paranal Observatory and Garching. The requirements for the ESO VDFS deliverables are in a completely separate set of documents [RD04, RD05] which also cover the requirements for observing templates. Where appropriate software delivered to ESO will be reused in the UK VDFS system.

This document is intended to provide only high-level specifications for the overall functionality of the pipeline and archive. Detailed software specifications will be covered by derived documents which should refer back to this document and, where appropriate, and should provide a compliance matrix with respect to the numbered Requirements herein.

1.2 Reference Documents

The following documents are referenced in this document:

RD01	VISTA Technical Specification	VIS-SPE-ATC-00000-0003 Issue 2.0
RD02	VISTA Science Requirements Document	VIS-SPE-VSC-00000-0001 Issue 2.0
RD03	VISTA IR Camera Technical Specification	VIS-SPE-ATC-06000-0004 Issue 2.0
RD04	VISTA Infrared Camera DFS User Requirements	VIS-SPE-IOA-20000-0001 Issue 0.5
RD05	VISTA Infrared Camera Calibration Plan	VIS-SPE-IOA-20000-0002 Issue 0.5
RD06	VISTA Infrared Camera Data Reduction	VIS-SPE-IOA-20000-0003

	Specifications	Issue 0.5
RD07	Table of Proposed Catalog Parameters	VDF-SPE-IOA-00009-0001 Issue 1.0
RD08	VO Standards	TBD
RD09	VISTA Operational Concept Definition Document	VIS-SPE-VSC-00000-0002 Issue 1.0, 2001-03-28
RD10	WFCAM Science Archive Science Requirements Document	Issue 1.2, 2003-01-08

1.3 Acronyms & Abbreviations

BOB	Broker for Observation Blocks
CASU	Cambridge Astronomical Survey Unit
CCD	Charge-coupled device
DAS	Data Acquisition System
ESO	European Southern Observatory
FoV	Field of View
FPA	Focal plane array
FWHM	Full-width half maximum
HOWFS	High-order Wavefront Sensor
IRACE	Infrared Array Control Electronics
LOWFS	Low-order Wavefront Sensor
M1	Primary mirror
M2	Secondary mirror
OB	Observation Block
P2PP	Phase 2 Proposal Preparation
PSF	Point spread function
VDFS	VISTA Data Flow System
VIRCAM	VISTA Infra Red Camera
VISTA	Visible Infrared Survey Telescope for Astronomy
VPO	VISTA Project Office
WFAU	Wide Field Astronomy Unit

1.4 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.

- **Exposure[1]:** The stored product of many individual integrations, that have been co-added in the DAS. Each exposure is associated with an exposure time.
- **Integration[2]:** A simple snapshot, within the DAS, of a specified elapsed time. This elapsed time is known as the integration time. (Usually an integration will comprise a double-correlated sample, i.e. the detector is reset, read once then read a second time after a delay equal to the exposure time. The two readouts are subtracted to give the resulting ``data'' array.)
- **Jitter (pattern):** A pattern of exposures at positions each shifted by a small movement ($\leq$ 30 arcsec) from the reference position. Unlike a microstep the non-integral part of the shifts is any fractional number of pixels (though integer shifts may be preferred for small jitters). 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 to monitor residual systematics in the photometry.
- **Microstep (pattern):** A pattern of exposures at positions each shifted by a very small movement ($\leq$ 3 arcsec) from the reference position. Unlike a jitter the non-integral part of the shifts are specified as 0.5 of a pixel, which allows the pixels in the series to be interleaved 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. (e.g. a jitter or microstep movement).
- **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 the VISTA IR camera, with its 16 non-contiguous chips (see Figure 3). The name is from the similarity to the prints made by the padded paw of an animal (the analogy suits earlier 4-chip cameras better).
- **Preset:** A telescope slew to a new position involving a reconfiguration of the telescope control system and extra housekeeping operations that are not necessary for a movement or an offset.
- **Robust Estimate:** A statistical estimator that is resilient to small perturbations on the assumed shape of the underlying distribution.
- **Tile:** A filled and fully sampled area of sky formed by combining multiple pawprints. 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. For example the set of six pawprints with relative offsets respectively (0,0); (0.475, 0); (0.95,0); (0,0.95); (0.475, 0.95); (0.95,0.95) detector widths in the camera x and y axes (where one detector width = 11.6 arcmin) gives a contiguous rectangle on sky of $1.5°$ (x) $\times$ $1°$ (y) , such that each point on the sky within the tile is covered by **two** or more of the six pawprints (see Fig. 3).

[1] The definition of “integration” and “exposure” here are interchanged relative to WFCAM usage, in order to conform to ESO usage.

[2] See footnote 1

In general, the pawprint will be the preferred unit of processed and calibrated image data, but users will frequently desire data delivered as complete tiles - this combination may be done on-the-fly, see Sec 8 below.

2 Background and Definitions

This Section explains the status of the various reference documents, and sets out the aims and responsibilities of VDFS. For clarity it also lists some related items which are not the responsibility of VDFS.

2.1 Aims

The top-level aim of the Vista Data Flow System is to optimise the science return given the VISTA raw data as input.

The top-level science requirements for the VISTA project as a whole were set out in the VISTA Science Requirements Document [RD02], and the hardware specifications for VISTA are given in the VISTA Technical Specification [RD01]. RD01 was derived to meet the requirements of RD02 as far as possible. RD02 has deliberately not been updated in the light of actual hardware design, to avoid moving the goalposts for the hardware, for example still including all the requirements for the visible camera which will now not be provided unless further funds can be found.

The large surveys to be executed with VISTA will be defined by suitable ESO committees. It is probable that, to a first approximation, these will include Wide, Deep and Very Deep surveys comparable to larger-area versions of the UKIDSS surveys, surveys of the Galactic plane and Magellanic Clouds, plus possibly a shallow ``Atlas'' of the entire Southern hemisphere. These survey specifications may to some extent depend on availability of matching visible data (e.g. from VST or other telescopes), and on early results from UKIDSS. For VDFS design purposes the set of IR surveys defined in the VISTA Operational Concept Definition Document (OCDD) [RD09] should be assumed. A brief summary of these surveys is given in Table 1, which is simply a sketch for VDFS planning purposes and in no way prejudges any aspect of the actual observing programme.

Detailed observing strategy for each such survey remains to be defined, but apart from the obvious wide range in total exposure times, the basic observing sequences of jittering, tiling and calibration are expected to be quite similar for each survey; details are given in Sec 3.2.

This document does not attempt to specify individual science projects in detail, but to give general requirements on quantities such as photometric and astrometric accuracy, catalogue completeness, archive search capabilities and performance etc, which are common to most of the likely science projects. A sketch of some possible projects is given in Section 11.

Survey name	Area (deg^2)	Y	J	H	K_s	Clear nights
		(5σ)				(exc. overheads)
Very deep	15		23.8	22.5	22.0	55
Deep	100		22.8	21.5	21.0	57
Wide (high-b)	3000	22.0	21.2*	20.0	19.5	100
Wide (plane + MCs)	1500	21.5	20.5*	19.5	19.0	45
Atlas (Goal)	20000		20.2		18.4	150

Table 1 Design reference programme parameters (for illustration only). Magnitude limits are 5σ, point source, Vega scale, assuming 90% of flux in 1.6 arcsec diameter aperture. Nights (excluding overheads) are updated from RD09 using current field of view and throughput estimates; the combination of overheads and bad weather may mean elapsed nights are approximately double the above numbers. Note that required nights will increase significantly if z′ band is not available elsewhere and is done with the VISTA IR Camera; also, wider or deeper surveys may be appropriate given the increases in field of view, quantum efficiency and IR camera time fraction since the baseline DRP was drawn up.
*** 3 epochs for proper motions**

2.2 Responsibilities

Clearly, achieving the science requirements depends ultimately on the whole chain from the sky to final catalogues, including the atmosphere, the telescope, the camera, science and calibration data and VDFS software. In addition there are ``extraneous'' effects such as stray-light, dome thermal emission, interference etc. Some degradations (e.g. poor seeing, poor telescope image quality) cannot be fixed in software. Other effects (e.g. bad pixels, pincushion distortion, chip-to-chip QE variations, persistence, atmospheric transmission variations etc) clearly can be corrected in software, subject to availability of suitable calibration data.

There is a ``grey area'' comprising effects which could in principle be corrected in software, but in practice cannot be, either because the required calibration data cannot be obtained at all, or because the effect is unpredictable or varying on a short timescale and it is infeasible or too inefficient to repeat the calibration with the required frequency. In the event of some effect in a ``grey area'' causing a failure to meet requirements, there is significant potential for conflict as to whether VISTA hardware or VDFS software is at fault. To minimise this, it should be assumed that the calibration data includes that and only that outlined in Sec 4. The VDFS should meet all the requirements in this document as far as reasonably possible given the available science and calibration data. Any degradation in the data quality resulting from hardware, which cannot be calibrated in software given the calibration data listed in Sec 4, shall not be the responsibility of VDFS.

This document intentionally does not describe the detailed steps in data processing, such as reset correction, dark subtraction, flat-fielding, cross-talk correction, stacking of frames, astrometric calibration, etc. The purpose of this document is to summarise as far as possible the probable inputs to VDFS and the desired outputs: the internal details of how the processing is actually done are the responsibility of VDFS to define.

2.3 Requirements, Goals, Comments

In later Sections, there are various Requirements, Goals and Comments, which have status as follows:

- Requirements shall be met. If it is found impossible or unreasonable to meet a Requirement, the VDFS team shall use the Change Control procedure to request a waiver, which may or may not be granted.
- Goals are desirable to be met. VDFS shall attempt to meet Goals on a best-efforts basis.
- Comments are *purely for information*, they are intended to be helpful and to indicate the source of a requirement, but no obligation is expressed or implied. A Requirement or Goal is not modified by the existence or otherwise of an associated Comment. VDFS may as it chooses make use of all, some or none of any information in a Comment. Use of any material in a Comment does not remove the responsibility of the VDFS team to meet the relevant Requirement.

When citing Requirements or Goals in later documents, the format should be as follows: Requirement 4.3/5 refers to the Requirement numbered 5 in Subsection 4.3.

The above assumes VDFS has been awarded sufficient resource to meet the requirements. If the necessary resource cannot be found from within the allocated budget, the necessary resource should be applied for.

2.3.1 Pipeline and Archive division

The VDFS already contains two main parts, the ``Pipeline'' which does processing and image cataloguing from the raw data, and the ``Science Archive'' which curates data products and provides advanced search and access capabilities to suitably authorised science users.

In general this document intentionally does not specify which part of the system has to do which internal steps. Any Requirement should apply to both the Pipeline and Archive considered as one combined black box, and reference to VDFS hereafter should apply to both Pipeline and Archive as far as necessary to meet the requirement. Any grey areas should be resolved within VDFS, if necessary in consultation with the VDFS Project Leader.

3 Telescope & Instrument Overview

This section provides a summary of the key features of the telescope and instrument hardware, for information only.

3.1 General Layout

VISTA is a 4-metre altazimuth 2-mirror telescope with a single focal station (Cassegrain) which can accommodate one of two possible cameras: an IR Camera or (if funded later) a visible camera. The telescope has a fast focal ratio (f/1 primary, f/3.25 at Cassegrain) hence a compact structure. The telescope uses active optics, with 84 independent axial supports controlling the figure of the primary mirror (M1), and a precision hexapod controlling the position of the secondary mirror (M2) in 5 axes.

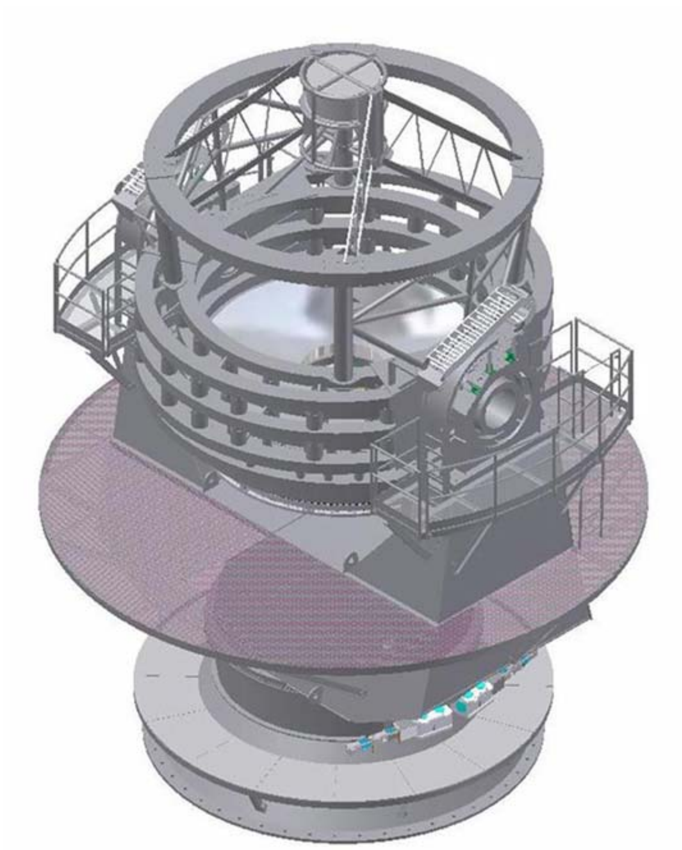

Figure 1: Design of VISTA telescope structure at Final Design Review. The large flat section is level with the dome floor. Camera at Cassegrain and secondary mirror are not shown.

The infrared camera (Figure 2, details in RD03) has a very wide field of 1.65 degrees diameter. It is a novel design with no cold stop, but instead a long cryostat with a cold baffle extending ≈ 2.1m above the focal plane to minimise the detectors' view of warm surfaces. There is a large entrance window (95cm diameter) and 3 corrector lenses, all IR-grade fused silica. The camera contains only one moving part (the filter wheel). There are 16 science detectors, each 2048x2048 pixels, covering 0.59 deg^2 per single integration. The camera also contains fixed autoguiders and wavefront sensors (2 each, using CCDs operating at ≈ 800nm wavelength) to control the telescope tracking and active optics.

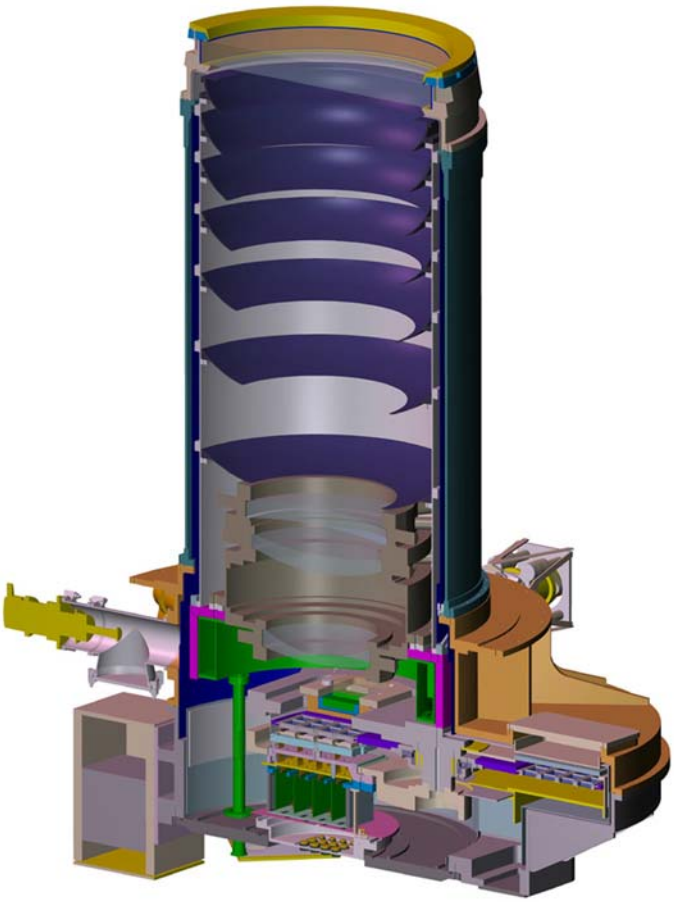

Figure 2 : Cutaway view of IR Camera; note cold baffles and 3 lenses. The bulge at lower right surrounds the filter wheel (note filter exchange hatch). The science detectors are mounted on the light blue plate; the small green box is one AG+LOWFS unit.

Table 2 gives approximate values for the main system parameters.

Telescope	
Telescope mount	Altazimuth
Focal station	Cassegrain
Primary mirror diameter	4.10 metres
Entrance pupil diameter	3.70 metres
Secondary mirror diameter	1.24 metres
Baffle diameter	1.63 metres
System focal length	12.072 metres
M1-M2 distance	2725.73 mm
Mirror coatings	Silver (default) or Al (optional).
IR Camera	
Wavelength range	0.85 - 2.4 μm
FoV (available)	1.65^{o} diameter
FoV (detectors)	0.59 deg^2 instantaneous; 1.5^{o} ×1^{o} tile covered ``twice" after 6 pointings
System image quality	≤ 0.5 arcsec (Goal: 0.4 arcsec).
Detectors	4×4 mosaic of 2048×2048 Raytheon VIRGO
Pixel scale	0.34″ / 20 μm
Controllers	ESO IRACE (1 master + 3 slave)
Readouts	16 per detector (each a ``stripe" of 2048×128 pixels).
Readout time	Approx 1 second.
Filters (mounted)	1 dark + 7 science (Y,J,H, K_s + others TBD)
Auxiliary sensors	2 autoguiders, 2 LOWFS, HOWFS beamsplitter(s).

Table 2: VISTA: baseline system parameters

3.2 Observing considerations

The hardware clearly places certain constraints on operational modes, though not many, since the system has been designed to be fairly flexible and limit the observing parameters as little as reasonably possible. Since there is only one moving part in the camera (the filter wheel), the telescope + instrument configuration is primarily specified by the pointing direction (RA and Dec), the Cass rotator angle and the selected filter. The Cass rotator has a full-range of 540 degrees so the position angle of the focal plane w.r.t. the sky may be chosen freely (the default will probably be the long axis aligned East-West)[3].

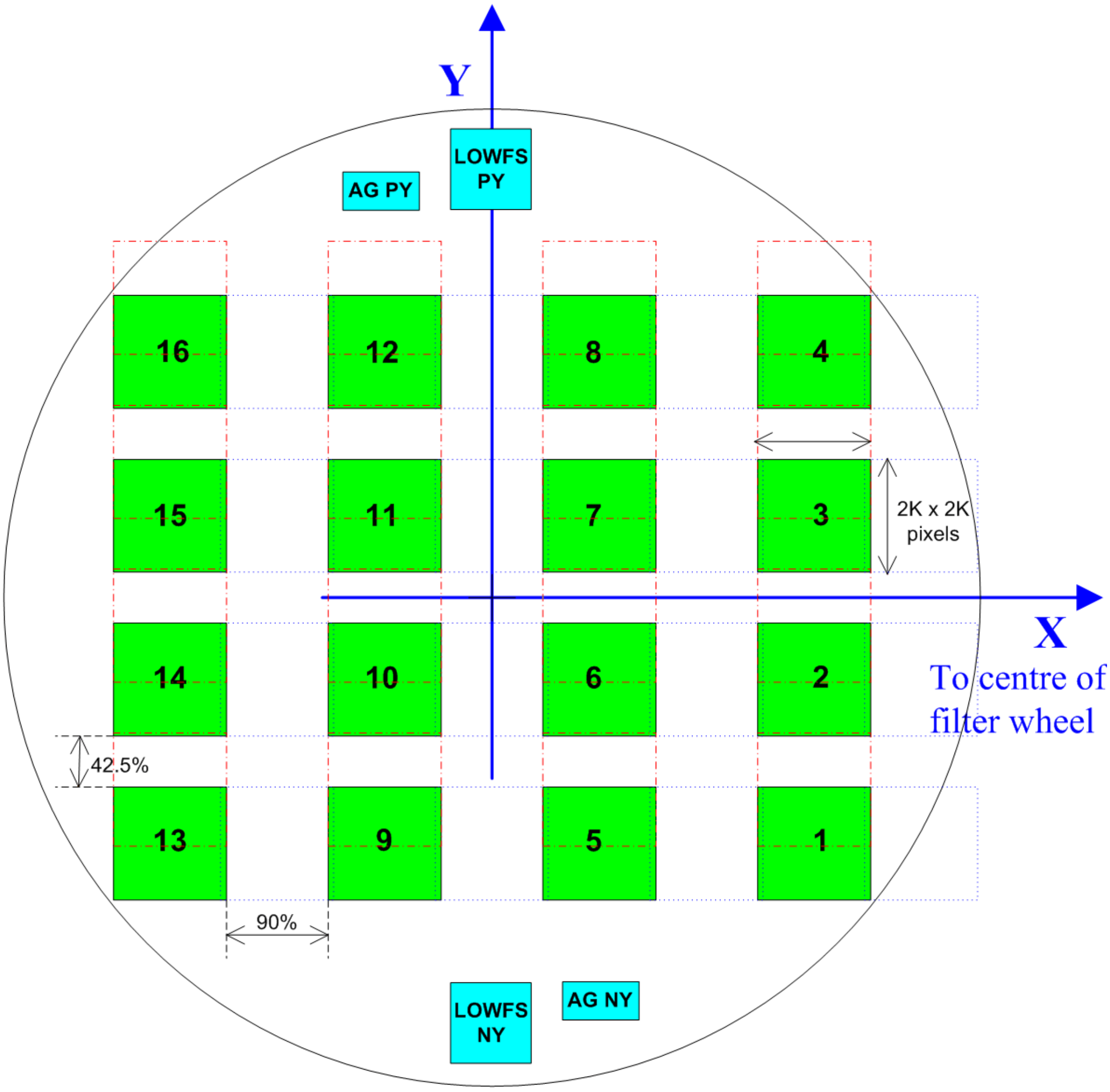

Figure 3: Layout of the 16 science detectors in the focal plane. Circle is 350mm or 1.65 deg diameter. Also shows "effective locations" viewed (via pickoff mirrors) by the autoguiders and LOWFSs (blue).

[3] Note that the FPA, autoguiders and LOWFSs are 180-deg symmetric, so if desired one can observe a field at two camera angles 180^{o} apart while re-using the same guide star and LOWFS stars.

The most obvious constraint is the **layout of detectors** on a 4 ×4 rectangular grid with gaps of 90% of a detector width in x and 42.5% in y (Figure 3). This requires observing 6 ``pawprints'' to give a filled ``tile'', as defined in Sec 1.4 above. The resulting exposure time map is shown in Figure 4. Each pawprint may comprise many exposures at different jitter or microstep positions, as desired for optimal background subtraction and/or sampling. These moves are user-selectable and not ``quantised'', but should be limited in size for the following reason: after tiling moves, the overlaps between adjacent pawprints have a width of ≈ 5% of a detector or 30 arcsec, therefore jitter moves within a pawprint must not exceed ±15 arcsec away from a ``central'' jitter position to avoid underexposed stripes in the final tile. (If a specific science project requires larger jitter moves, these can be delivered by the telescope, but underexposed stripes after tiling may result).

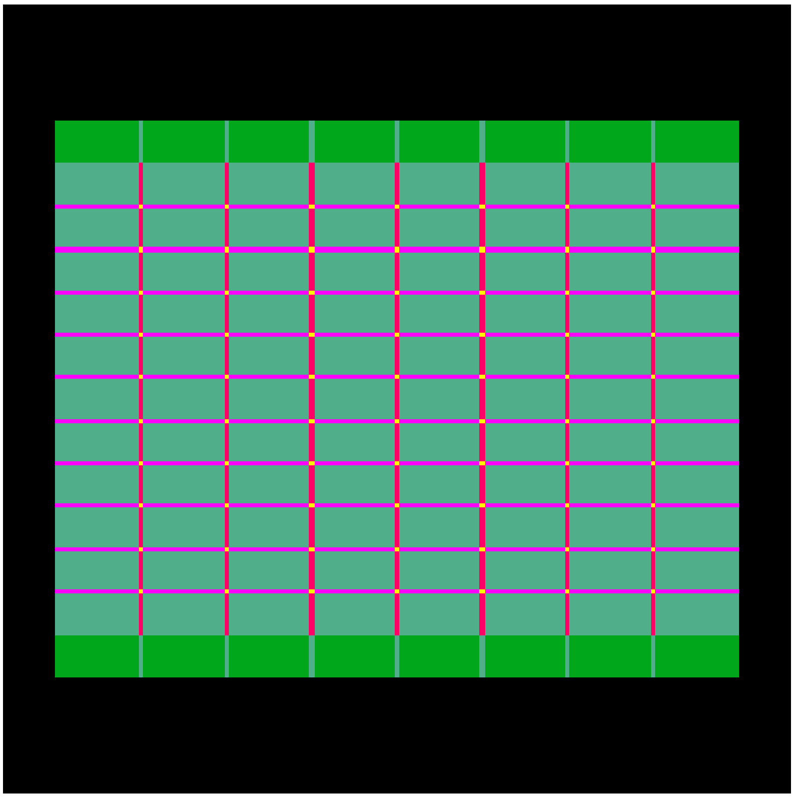

Figure 4 : Exposure time map for a filled tile of 6 pawprints (no jittering). Dark green = 1, light green = 2, magenta = 3, red = 4, yellow = 6 , in units of the single-pawprint exposure time.

The constraint on size of microsteps is more stringent due to field distortion. The radial plate scale in arcsec/pixel is 2.4% smaller at the field corner than the centre, so a microstep of e.g. 10.4 pixels at the field centre corresponds to 10.63 pixels at the field corner. Therefore, to give offsets close to a half-integer number of pixels across the whole field, microsteps must be limited to not more than ~3 arcsec.

Thus, if one wants to observe with both microstepping and a jitter size > 3 arcsec, separate nested jitter and microstep loops are required: this is fine for deep exposures but implies that overheads may preclude microstepping for shallow surveys.

3.3 Active Optics & Guiding

The VISTA active optics system is largely ``transparent'' to science observing: there is a closed-loop low-order wavefront sensor (LOWFS) system controlling 5 axes of M2 (i.e. focus, x+y centering and x+y tilt). The M1 figure is controlled via 84 axial supports in 4 rings: the force pattern is calculated from a combination of a lookup table, occasional tweaks from a high-order wavefront sensor, and optionally measurements of M1 astigmatism and trefoil modes from the LOWFSs.

There are two identical fixed autoguider + LOWFS units at opposite sides of the field of view, each fed by a pickoff mirror and a fixed filter (roughly I-band). There are no patrol mechanisms; large-format detectors and windowed readouts are used so that usable guide and LOWFS stars are available for almost all telescope pointings.

Each autoguider has one frame-transfer CCD chip giving an 8×4 arcmin field; only one autoguider will be in use at a time (whichever has a brighter star). The LOWFSs have a field of view of 8×8 arcmin ; they usually add no ``overhead'' since their exposures should start after and finish before the science integrations. Use of the LOWFSs imposes a minimum time between jitter moves ~ 30sec since they have to complete an exposure with adequate SNR in between consecutive jitter moves. If it is essential to jitter more often than once per 30 sec, this can be done using open-loop M2 control - a slight loss of image quality may result.
Comment: (Alternatively, a software enhancement may be added to enable co-adding two or more 15-sec LOWFS exposures of the same star with relative jitter shifts - this is not currently implemented in the TCS but should be considered as a software enhancement if there's significant demand for ``rapid jittering''. A simple image co-add with a shift by the nearest-integer number of pixels should be fine.)

The one significant overhead from the LOWFSs is after a telescope slew giving a large ($>10^{o}$) change in altitude: in this case there will probably be a need for a 30-sec pause for one LOWFS cycle to complete and update the M2 position before science observing re-starts.

Given the above sensor fields of view, generally a ``jitter'' move ~15 arcsec will re-use the same guide and wavefront sensor stars by simply offsetting the selected readout window in software, whereas a ``tiling'' move of 5-10 arcmin will almost always require different guide + LOWFS stars to be selected after the move; checking and acquiring the new guide star will add a short overhead ~ 1 sec per tiling move.

There are various “graceful degradation” modes in the event of hardware failure of one sensor, unavailability of stars etc; these include reducing the autoguider frame rate below the nominal 10 Hz, and/or operating with one LOWFS and 3-axis M2 control. These are described in more detail in [RD04]. VDFS should assume that each OB has been executed with acceptable active optics and guiding parameters (where the definition of "acceptable" is fixed for each OB but may vary between different OBs).

3.4 Filters

The filter exchange mechanism is a large wheel of 1.37m diameter (actually an annulus with a fixed pillar through the centre) with 8 main positions, 7 science and 1 dark. Filter exchange time is expected to be ~15-45 sec depending on required rotation angle; this is clearly longer than that for a jitter or tiling telescope move, so it is generally more efficient (and gives better sky subtraction) to complete a tile in one filter, then change filter and repeat the tile.

Filters will be ordered in the wheel so that a higher-background filter does not cross the detectors during a filter change. Due to non-uniform temperature across the wheel, a filter change is likely to cause a small warming of the detectors, but this should be corrected by the temperature servo system, so the temperature rise should be < 0.1 K for a few minutes after the change. With a wheel temperature < 110 K, photon emission from the wheel itself should always be negligible.

The wheel also contains ``mini-filters" in the V-shaped sectors in between the large science filters. These mini-filters cover only a small fraction of the science focal plane, and are primarily for engineering use (including one or more beamsplitters giving HOWFS images onto the science detectors). For observing efficiency, it is possible that a row of 3 mini-filters (J, H, K_s) can be provided covering one detector in each band simultaneously: if implemented, this gives the option of taking atmospheric standard stars in all 3 bands in a single exposure (possibly 5 bands if the corners of neighbouring science filters are used also). It is not a requirement for VDFS to handle such data.

3.5 Comparison with WFCAM

Since the VISTA data flow system is an extension of that for WFCAM, it is useful to have an overview of the relevant differences.

1. **Requirement:** The VDFS shall ensure that it can handle, as far as reasonably practical, any differences between VISTA and WFCAM data which may arise from the items listed below.
    a. VISTA is an altazimuth telescope. The effect on data reduction should be minimal, except that diffraction spikes will rotate relative to the detectors during long exposures. There is an extra degree of freedom in that the position angle of the detectors w.r.t the sky is user-selectable - this may be useful for calibration purposes.

b. VISTA has 16 detectors in a 90% / 42.5% spacing, rather than WFCAM with 4 detectors at 94% spacing in each direction. This means that 6 pawprints rather than 4 are the minimum required to give a filled tile, but each sky pixel is covered by at least 2 of the 6 pawprints.
c. VISTA has no cold stop, but a cold baffle instead. This may result in a smooth gradient in scattered thermal radiation across the detectors in K_s band; total intensity of this scattered background should be ~20% of sky and the gradient may be 10% of that, i.e. ~ 2% of sky. This will need to be taken into account in data processing.
(On the plus side, no cold stop means no intermediate focus, so there should be no issue with ``nearly in focus'' warm dust particles.)
d. The VISTA pixel scale is somewhat smaller (0.34" vs 0.40") so the intention is to use microstepping only in the best seeing conditions. 2×2 microstepping should be a user-selectable option for VISTA (pre-selected at the time of OB definition).
e. VISTA has 16 readouts per detector, instead of 32 for WFCAM.
f. VISTA does not have a fast tip-tilt mechanism on the secondary mirror (there is a high-precision hexapod for collimation, but it is `slow'). Therefore jitter moves will be accomplished by moving the entire telescope, so for VISTA there is no hardware upper limit on the jitter size. Due to the very compact VISTA telescope structure, this will not be a serious time overhead (e.g. ~2 sec for a 10 arcsec jitter).
g. Field distortion for VISTA is larger than for WFCAM due to the wider field (0.85% integrated vs 0.4%). This means that jitter steps would only correspond to a near-integer number of pixels across the whole field if this integer were small, ~10. Such small jitter steps would be undesirable for background removal, so rebinning of pixels when coadding frames at different jitter positions is likely to be required for VISTA.
h. The VISTA autoguider(s) will work with floating-point centroiding rather than (as WFCAM) holding the guide star at the corner of 4 pixels, so VISTA does not have any quantisation on the allowed jitter or microstep stepsize.
i. VISTA has a filter wheel rather than WFCAM's ``paddles''. Therefore the VISTA detectors do not receive an increased flux during filter changes.
j. The non-linearity correction for WFCAM will probably be done inside the data acquisition system hardware, while for VISTA it will probably be done during later VDFS processing.

4 Calibration Data Assumptions

This section provides the working assumptions to be used when designing the VDFS. It does not constitute a specification for the actual frequencies of calibration observations.

A detailed Calibration Plan [RD05] is provided as part of the DFS deliverables to ESO. The items below are expected to be consistent with RD05 and shall be assumed for purposes of software design (i.e. how the calibration data is reduced in the UK). In the event that [RD05] changes, or is not consistent with the items below, VDFS shall bring this to the attention of the author of this document.

VDFS software shall so far as possible meet the Requirements given the calibration data below. If an unwanted effect cannot be calibrated out using the items below, VDFS should propose obtaining the necessary calibration data, and take all reasonable measures to see that it is obtained. However if the necessary data cannot be obtained the VDFS shall not be held liable for failure to deal with the effect in software, although it should take all reasonable steps to work around the lack of necessary data.

If any of the following data are not strictly necessary for VDFS to meet the Requirements in this document, then VDFS is at liberty to ignore that data. However, VDFS software should not assume the existence of any calibration data outside the following list, unless it is included in [RD05].

The assumed calibration data shall be as follows:

1. **Reset frames** with the minimum physically possible exposure time and a cold opaque filter in the beam. A set taken weekly, cf [RD05] Sec 4.2.
   Comment: For reset anomaly monitoring. NB these are different from 1-sec dark frames, because the array is read out once for a reset frame, but twice (and differenced) for a 1-sec dark frame.
   Comment: Ideally read out array twice for a reset frame but save only the first read, so the read sequence is identical to normal frames ?
2. **Dark frames** taken with a cold opaque filter in the beam, with exposure times matching the science frames. A set taken daily, cf [RD05] Sec 4.3.
3. **Dome flats** taken with the telescope pointing at the illuminated dome screen: a set taken daily with one (TBC) filter, cf [RD05] Sec 4.4.
   Comment: Primarily for bad pixel monitoring and system health checks.
4. **Twilight flats** on twilight sky, with up to 4 filters per night, cf [RD05] Sec 4.7.
   Comment: limited twilight time probably limits the number of filters.
5. **Linearity frames** comprising dome flats with stable illumination with a suitable range of exposure times: one set taken weekly, cf [RD05] Sec 4.6.
6. **Cross-talk frames** For measurement of detector cross-talk, it may be necessary to observe fields containing a moderate number of bright stars, repeated with various relative offsets. These should be assumed to be taken monthly, cf [RD05] Sec 4.10.
7. **Persistence frames** For measurement of detector persistence, it may be necessary to observe fields containing a moderate number of bright stars, then shifted and a series of further frames taken to measure the decaying signal. Assumed to be taken monthly, cf [RD05] Sec 4.9.

Comment: The persistence and cross-talk frames may optionally be the same if one set of frames are adequate for both purposes.

8. **Astrometric standard frames**: nighttime observations of an astrometric standard field, taken once per week; see [RD05] Sec 6.2.
   Comment: Currently these are not essential since astrometry will use existing standards in science frames, but standard frames can be taken using standard observing templates if later desired.
9. **Photometric standard fields**: nighttime observations of one or more photometric standard fields. It should be assumed that once per hour, one standard field is observed in each of the science filters used during that night. Also, twice per night, two standard fields are observed at airmasses ~1.1 and ~2 respectively. Cf [RD05] Sec 5.2.
   Comment: If the throughput and detector gain are both stable, variation in apparent zero point can also be used as a measure of extinction variation.
10. **Offset sky frames**. When observing very dense star fields (e.g. Galactic bulge), offset sky frames may be taken as necessary to provide background images for sky monitoring and/or subtraction, interleaved in time and taking equal exposure time to the science frames.
    Comment: It is assumed that any science frames requiring offset sky frames will specify this in advance as part of the OB, i.e. no real-time decision is needed.
11. **Mesostep frames.** These comprise a set of frames on a selected region of sky with successive offsets e.g. 0.2, 0.4, …2.0 detector widths in each coordinate. One series taken monthly. Cf [RD05] Sec 4.8.
    Comment: This enables mapping of photometry vs position on the focal plane, which can test (for example) residual gradients due to scattered light, varying aperture correction etc.

VDFS should ensure a list of suitable standard fields is available. If additional calibration data is desired outwith the above list, VDFS may request it, but hardware or operational limitations may not permit it.

5 General Requirements

This section covers general requirements on the overall VDFS system (pipeline and archive), such as data rate, volume, access privileges, bookkeeping and traceability in the data processing.

1 **Requirement:** The VDFS shall process all VISTA science data (including both the ``large surveys'' and the smaller PI-driven programs), except any data taken in ``non-standard'' observing modes outside of those defined in [RD04].

2 **Requirement:** The release, and use, of processed data products shall take account of any proprietary periods that may have been assigned to the data.

3 **Requirement:** The data output by the VDFS science archive shall as far as reasonably possible be compliant with the Virtual Observatory standards given by RD08.
Comment: The ``internal'' data formats within VDFS are not specified, but this applies to data supplied by the Archive to the end users.

4 **Requirement:** The VDFS shall be able to process and archive data at 1.6× the maximum nightly data rate, or 2.5 × the annual average data rate, whichever is the larger.
Comment: To enable keeping up with input data and simultaneously reprocessing ``old'' data with improved algorithms.
Comment: Each exposure is 268.4 MB of pixel data (before Rice-compression). Maximum data rate for which the data handling system is designed is one exposure per 10 sec for 14 hours i.e. 1350 GB/night. One exposure per 15 sec for 12 hours i.e. 773 GB/night is a more realistic maximum in most circumstances, with a typical night probably being closer to 450GB. Assuming 10 nights engineering time per year and 15% loss of the remaining 355 days to weather gives data for about 300 nights a year at 450GB/night so the annual raw data received would then be up to 135 TB.

5 **Requirement:** For reduced data, summary information on the reduction history shall be written in the FITS headers, including (but not necessarily limited to) software version, run time, and filenames for dark, flat-field and similar calibration frames.

6 **Requirement:** The VDFS shall function correctly in the event of one or more partly or wholly non-operational or missing detectors in the VISTA focal plane.
Comment: This implies (for example) that the software should not rely on one ``preferred'' detector and crash if it isn't working, and that ``averaged'' parameters shall exclude non-operating detectors, and that sky coverage maps shall allow for non-operating detectors.
Comment: It is assumed that only those OBs which can tolerate a failed detector will be observed in this event.

7 **Requirement:** The VDFS shall provide means of tracking survey processing progress, e.g. plotting of pawprints / tiles available / reduced / passing quality control.

8 **Requirement:** The VDFS shall provide ``summary statistics'' to enable assessment of data quality for each reduced pawprint or tile, including (but not limited to) the following :

a. Average sky brightness
b. Average PSF FWHM and ellipticity.
c. Sky brightness, PSF FWHM and ellipticity for one location per detector.
d. Photometric zeropoint.
e. Limiting magnitude for point source (5σ)

f. Number of bad pixels (from recent flat field).
g. Dark count (from the relevant dark frame).

9 **Requirement:** There shall be ``query-driven'' access to the VDFS Archive, to provide downloads of data for user-selected subsets of data. Examples of search parameters are given in Section 11; the Archive shall facilitate these example queries as far as reasonably possible. Additional requirements are as follows:

a. **Requirement:** There shall be a ``simple'' search interface for basic users and an ``advanced'' interface for expert users.
b. **Requirement:** Queries shall permit any combination of Boolean conditions (AND,OR,XOR and NOT) on arbitrary numbers of parameters including parentheses and nesting.
c. **Requirement:** For clarity, users should be able to define their own parameters as arithmetic or Boolean combinations of catalogue parameters, and search on combinations of these.
d. **Requirement:** Queries must enable searching/selecting by RA/Dec , Galactic or ecliptic coordinates, including circular, square or rectangular windows of user-selectable size.
   Comment: Rectangles necessary for e.g. ``stripe'' surveys such as 2dF.
e. **Requirement:** Queries shall include the facility to ``upload'' files, e.g. user provides a file of ~ 1000 positions and requests a thumbnail image and list of catalogued objects for a specified region around each.
f. **Requirement:** Combined queries must be possible on VISTA data plus the following other catalogues:

- 2MASS
- SDSS
- USNO-B
- FIRST
- IRAS Point Source Catalogue
- ROSAT All-sky survey.
- The UKIDSS surveys.
- VST public surveys.
- The WISE all-sky survey (if available)
- The ASTRO-F (IRIS) all-sky survey (if available).

g. **Requirement:** Query returns for catalogue data shall be able to return either all parameters, or a default ``short'' list limited to the most important ones, or a user-specific selected list (e.g. in a user-specific configuration file).

h. **Requirement:** Query returns shall be able to include a user-selected random-sampled percentage (e.g. 0.1% or 1%) of objects within selected sky area and magnitude limits.
   Comment: For e.g. rare object searches, it will be necessary to know the statistics of a random sample of all objects along with the selected rare objects. Random-sampling reduces the overhead but is unbiased. (Using a smaller area or brighter

magnitude limit is not satisfactory).
Comment: Since the overhead is small, it may be useful to store a 1% random-sample of every catalogue as a distinct item for rapid querying.

10 **Requirement:** When selecting objects by colour, it shall be possible to select using a cut on time-lag between observations in different filters.
Comment: To avoid contamination of high-z quasar samples by variable stars, it may be necessary to reject data where different filters were observed on different nights.

11 **Requirement:** The VDFS Archive shall have capability for user to request arithmetic image processing.
Comment: This should probably be menu-driven as a basic function, followed later by Grid-type system whereby user can upload code to Archive.

12 **Requirement:** The VDFS archive shall comprise both a ``living'' internal database which is updated at regular intervals (at least 6 times per year) with the latest available data, and a small number of data releases issued at suitable intervals. Availability of the database and all releases will respect proprietary data rights as required in 15 below.

13 **Requirement:** Once released, each data release shall remain indefinitely available, and the image and catalogue data shall not be modified (updated meta-data such as improved calibration coefficients, bug lists or bad object lists may be issued, but the actual data shall remain static).
Comment: Typically for each survey the releases may be e.g. at 15%, 40% and 100% of survey completion. Continued access to ``older'' versions is mandatory for cross-checking of results. Data releases do not need to include all data to date, but can be trimmed to provide a sensible coverage map e.g. contiguous regions of sky with at least 70% of the nominal exposure time.

14 **Requirement.** To facilitate follow-up spectroscopy, it shall be possible for a user to upload a list of target positions and a reference-star search radius; then the Archive shall return the updated target position (based on the latest astrometric solution) and a list of reference stars within the search radius.
Comment: It is important that spectroscopic targets and reference stars use the same astrometric solution; in case the astrometric solution has been updated after the user has selected a target list, it is necessary to retrieve updated coordinates for the target objects as well as the reference stars.

15 **Requirement:** The requirements for VDFS Archive security are as follows:

a. The archive shall be capable of limiting data access to suitably authorised subsets of users, e.g. ``any UK user'' for ESO-public surveys, or a more restricted group for PI-driven surveys.

b. Access shall allow for any proprietary periods as necessary. (Coverage maps i.e. planned survey areas and actual telescope pointing history shall be assumed to be world-public immediately.)
c. As far as reasonably practical, it shall be impossible for Archive users to corrupt the data or overload the Archive e.g. with CPU-heavy demands.
d. Releases will require prior authorisation by the VISTA PI to ensure they are agreed by the necessary supervisory bodies, or are consistent with their policies. Details of proprietary periods will be specified to VDFS by the VISTA PI (based on information from the appropriate supervisory committees).

16 **Goal:** The goals for VDFS Archive uptime and downtime, including downtime (if any) during ingestion of new data into the Archive, are as follows:
a. The Archive shall have a downtime percentage including both planned and unplanned downtime of ≤ 3% during working hours and ≤ 6% overall.
b. Planned downtimes during working hours should be limited to the following: Downtimes up to 3 working hours not more than 12 times per year. Downtimes of 3-8 working hours not more than 3 times per year.
c. Recovery from failure of any one piece of VDFS hardware should be possible in not more than 24 hours, assuming resource and manpower availability.

17 **Goal:** The goals for VDFS Archive response time are as follows: On receipt of a user-driven query, the Archive shall estimate the resulting data volume and time to completion, inform the user and proceed as follows :
i. **Requirement:** Short queries (≤ 5min run time and ≤ 50 MB returned data) shall be returned in real time.
ii. **Requirement:** Long queries shall be run in batch mode, the data saved to an FTP site or similar, and the user Emailed when the results are ready.

Comment: As an example, VDFS should be able to return density plots of a colour-magnitude diagram and estimate number of objects selected in a given polyhedral volume of colour-magnitude space.
Comment: As above, running the query first on a 0.1% random sample of objects may be helpful to provide a rapid estimate of run-time for the full query.

6 Astrometric Requirements

The main requirements for astrometric accuracy are taken from Sec 4.5.1 of RD02. These apply for objects sufficiently bright that photon centroiding noise is negligible, and not affected by bad pixels / chip boundaries etc. The requirements are as follows:

1. **Requirement:** Absolute astrometric accuracy ≤ 0.3 arcsec rms for any processed tile or pawprint.
   Comment: Given the availability of catalogues such as 2MASS, UCAC, USNO-B etc, achieving this should be relatively straightforward given Req. 6/2 below.

2. **Requirement:** Differential astrometric accuracy ≤ 0.1 arcsec rms within any processed tile or pawprint.
   Comment: The VISTA detectors are bolted to a rigid Molybdenum plate near-isothermal with the detectors on a stress-free mounting, so the relative detector locations are expected to be stable to ~≤ 1 μm or 0.017 arcsec.
   Comment: One possible issue is the stability of the r^3 ``pincushion" distortion, both in amplitude and in shift i.e. whether its centre moves with respect to the FPA origin.
   Comment: Differential refraction may be important, but numerous standard stars should be available per field.

3. **Requirement:** Differential astrometric accuracy ≤ 0.03 arcsec rms within an area of sky covered by one single detector.
   Comment: Primarily for proper motion measurements after few-year baseline, e.g. for brown dwarfs. There should be several reference stars per chip with good proper motions from UCAC or future surveys (also, could average all other stars and use a Galactic model to give the mean).
   Comment: Intra-pixel sensitivity variation and colour-dependent atmospheric refraction may be an issue at this level.

4. **Goal:** Differential astrometric accuracy ≤ 0.01 arcsec rms within an area of sky covered by one single detector.
   Comment: This goal is challenging. In addition to the above-mentioned intra-pixel and colour effects, since the system is non-telecentric (angle of chief ray up to 7 deg from the normal to the focal plane), small levels of detector tilt or non-flatness can translate into small centroid shifts. This is likely to become relevant at the 0.01 arcsec = 0.58μm level. Also, whether the pixels are truly on a square grid is unclear at this level. Tinney (2003, AJ 126, 975) reaches 0.75μm astrometric accuracy in the IR, so this goal may be achievable.

7 Photometric Requirements

The main photometric accuracy requirements are drawn from Section 4.6.1 of RD02. There are additional related requirements on uniformity of coverage, independence of pixels, and bright objects.

1. **Requirement:** Absolute photometric accuracy ≤ 0.02 mag in J, H, K_s bands, in the VISTA native photometric system corrected to airmass 1.0 .
   Comment: The VISTA filters are specified very similar to the ``Mauna Kea 98" or ``MKO" system (but won't be identical). Correction to other systems or airmass 0.0 is dependent on the object spectral energy distribution: this may introduce additional errors which are outwith the above budget.

2. **Goal:** Absolute photometric accuracy ≤ 0.01 mag in J, H, K_s bands.
   Comment: RD02 contains no specific requirements on differential accuracy in

IR bands, so the differential requirements are the same as the absolute ones.

3. **Requirement:** Absolute photometric accuracy ≤ 0.03 mag in Y, z′ bands, in the VISTA native photometric system corrected to airmass 1.0 .
   Comment: RD02 contains no spec on these bands, but we need one here. A slight relaxation w.r.t. J, H, K_s requirement appears reasonable since filter leaks/moonlight/detector QE roll off may be relatively more important .

4. **Goal:** As an occasional option to get accurate photometry of bright objects, VISTA may observe with a deliberate defocus ~ few arcsec. VDFS should be able to handle this as necessary.

8 Tiling, Stacking, Microstepping

This section summarises the requirements on combining multiple jittered and/or microstepped and/or tiled frames into pawprints, tiles etc.

RD05 indicates that a short ``menu'' of default jitter and microstep patterns will be provided for the main surveys. (Science programmes may specify their own custom jitter patterns if there is a good reason to do so, but the menu should cover the majority of cases).

1. **Requirement:** The VDFS shall be able to combine jittered frames into a single pawprint, generating object-clipped background maps, and eliminating cosmic rays and other spurious transient events from the image stacks.

2. **Requirement:** When stacking multiple jittered or microstepped frames into a pawprint or tile, the image stacking shall save basic information (e.g. rms of pixel values after cosmic-ray clipping) to enable flagging of suspected variable objects.

3. **Requirement:** The image stacking shall provide a choice of options for pixel interpolation, including but not necessarily limited to the following:
   a. Nearest-neighbour
   b. Bilinear interpolation
   c. Lanczos-2 or Lanczos-3.
   d. Keys spline interpolation.
   e. The ``Drizzle'' algorithm.

   Comment: Each of the above schemes has advantages and limitations - which is best may well depend on the specific survey aims.

4. **Requirement:** The image stacking shall be able, when requested, to provide non-interlaced data but interpolated onto smaller output pixels.
   Comment: This may be a good alternative to microstepping: it avoids the ``hedgehog PSF'' problem in variable seeing, and output pixels e.g. 0.7 × the

physical pixel size can be used, giving less data volume than 2 ×2 microstepping. This is likely to be a good compromise for most purposes; the drawback is that pixel noise becomes non-independent.

5. **Requirement:** The VDFS shall be capable of handling interlaced data taken with either 2×1 or 2×2 microstepping. (NB larger numbers e.g. 3×3 are not required).
   Comment: 2 ×1 microstepping (i.e. shifts of (0,0) and (n+0.5, n+0.5) pixels) gives an ``effective'' 0.24 arcsec sampling interval on a 45-deg rotated grid, and may be a useful compromise with less overheads and data volume than 2×2 microstepping.

6. **Requirement:** The VDFS shall be capable of combining a set of (default six) pawprints (with the same filter) into a single tile on a common pixel grid. There shall be a choice of two methods for dealing with pawprint overlaps within a tile:
   a. ``Trimming'' the pawprints to non-overlapping regions ≈ 95% of a detector width, so that every sky pixel internal to the tile is covered by exactly 2 pawprints (apart from bad pixels).
   b. Stacking all pawprints to give extra exposure time in the pawprint overlap strips.

   Comment: This is possibly best dealt with by archiving individual pawprints and combining pawprints into tiles on-the-fly in response to Archive requests, otherwise proliferation of files may result.

9 Variable Objects

It is not a requirement for VDFS to process data for ``specialised'' monitoring programmes (e.g. supernova surveys, microlensing surveys, killer-asteroid surveys etc) which require rapid response and specialised software pipelines. However, basic source variability should be considered for the reasons below. The actual Requirements/Goal(s) are in other sections (see cross-references) since they apply to logically distinct parts of the VDFS.

1. It is desirable to flag probable variables when co-adding data from multiple epochs (e.g. different nights) into a single image. See Requirement 8/2.
2. Time-lags between observing a field in different filters can result in spurious colour outliers caused by variable objects. Minimising such time-lags should be done at the OB preparation and scheduling level, but if and when significant time-lags have occurred, then VDFS queries should be able to reject the resulting subsets of the data from rare-object searches. See Requirement 5/10.
3. Basic (non-optimised) variable-object searches can be handled by chopping the time series into discrete epochs, processing each epoch as an independent dataset, and rearranging the resulting object catalogues into lightcurves. See Requirement 10/12.

4. A more sophisticated approach is to create a ``master'' object catalogue from one epoch of good seeing, and for other epochs do PSF or aperture photometry at fixed positions corresponding to those of the master epoch. See Goal 10/13 .

10 Object Catalogues

The VDFS shall generate object catalogues from the photometrically and astrometrically calibrated tiles or pawprints. The requirements are as follows:

1. **Requirement:** The catalogues shall include the photometric and astrometric calibrations described in Sections 6 and 7.

2. **Requirement:** The catalogued object parameters shall include (but not necessarily be limited to) the list as specified for the WFCAM data in RD07.

3. **Requirement:** There shall be the following options for cataloguing when a field has been observed in multiple bands:
   a. Separate ``independent'' catalogues per band.
   b. ``Matched'' catalogues where an object list is detected in one ``master'' passband, and then photometry is generated in matched apertures in all bands (so that consistent colours are given for extended objects).

4. **Requirement:** Moving objects (e.g. asteroids) shall as far as possible be eliminated from the ``standard'' catalogues and (if possible) stored in a separate moving-object file.

5. **Requirement:** Known asteroids which appear in the standard VDFS catalogues shall be flagged as such.

6. **Requirement:** The catalogue 10σ point-source completeness should be at least 99.5%: i.e., for real point sources with true flux above the 10σ level, and unaffected by crowding / bad pixels / ghosts / diffraction spikes etc, at least 99.5% should appear in the output catalogue with measured fluxes above 5σ.
   Comment: No more than 0.5% of nominally high-significance objects to be missing without clear reason. (Clearly, faint objects can scatter below the detection limit due to photon noise).

7. **Requirement:** The catalogue 8σ extended source completeness should be at least 90%: i.e., for faint sources with isophotal area ≥ 4 arcsec^2 at an isophote of 4× (in linear units) the 1σ per 1 arcsec^2 sky noise, at least 90% shall appear in the output catalogue.
   Comment: This is intended to cover completeness for faint-galaxy detection. Note that flux level and completeness are lower than the stellar case, since faint galaxy studies tend to push closer to the detection limit.

8. **Requirement:** The catalogue 10σ point-source reliability should be ≥ 99.9%: i.e., for objects detected in a stack of many jitter positions from a single band, with measured PSF-fitted flux ≥ 10σ and classified as ``stellar'' and good PSF fit and good quality flags, not more than 0.1% shall be spurious (i.e. not a real astronomical object or blend of objects).
   Comment: Rare-object searches require very low levels of spurious bright single-band detections. If the above is met, spurious objects at matching locations in 2 or more bands should be extremely rare .
   Comment: Real galaxies misclassified as stars etc are inevitable, and clearly don't count as spurious above. Cosmic rays, airplane/satellite tracks, bad pixels etc do count as spurious, but can be removed by the jitter combination. Ghosts are much larger than the PSF. Persistence effects from saturated stars may need software clipping.

9. **Goal:** Catalogue 10σ point source reliability (defined as above) ≥ 99.99%.

10. **Requirement:** VDFS shall provide a simple completeness estimate for each tile as a function of object magnitude and intrinsic half-light angular radius.
    Comment: This may be a standard tabulated function which is appropriately scaled according to image statistics e.g. exposure time, seeing and sky background level.

11. **Goal:** VDFS shall provide a facility for a full Monte-Carlo completeness estimate, i.e. facility for user to insert simulated objects into data frames and VDFS processes these in the same way as real data to measure completeness.
    Comment: While this is the ideal method, it is appreciated that this may be challenging to implement, and for many purposes the simpler method in (10) above is likely to be adequate.

12. **Requirement:** VDFS shall be able to generate basic variability data by choosing a time spacing (e.g. 12 hours), using a friends-of-friends algorithm to group a series of frames on one pawprint and one filter within this time interval into ``epochs'' (so each frame belongs to exactly one epoch), and then reduce each epoch as an independent unit of data to the level of one object catalogue per epoch. .
    Comment: Further processing, e.g. matching objects detected at different epochs and rearranging into lightcurves, is outwith the responsibility of VDFS.

13. **Goal:** As a goal, it should be possible to group epochs as above, construct a ``master'' catalogue from one selected epoch, and perform PSF photometry at all other epochs using the fixed positions from the master catalogue.

11 Science Examples

The Requirements in previous sections are quite specific in quantitative terms but do not give much flavour of the likely range of science usage. To give some more specific science-driven examples, below we provide some example science-driven usage examples as general guidelines for likely usage of the science archive, for representative purposes only.

Note that some considerations of wide applicability to many of the potential science aims are as follows:

Near-IR magnitudes are an excellent tool for selecting samples of objects with minimal biases against cool, high-redshift , dusty and/or obscured and/or non-star-forming objects. The converse however is that near-IR colours alone provide limited leverage for object classification except in a few special cases; increasing the wavelength baseline is usually necessary. Therefore, a large fraction of VISTA science will involve matching objects against catalogues or pixel data from other wavelengths. Most commonly this involves visible bands (e.g. from VST, Megacam, or VISTA's visible camera if/when available) but also from a range of other wavelengths from X-ray, UV (GALEX), mid-IR (Spitzer,WISE), far-IR (IRIS), submm (Herschel, APEX, SCUBA-2) and radio surveys. Followup spectroscopy will also be important for many projects. This leads to the following considerations for VDFS:

- Ability to deliver matched-aperture photometry at many bands is important (cf Requirement 10/3). It is not possible to do this "perfectly" if the PSF varies between bands, but if the aperture is large enough to enclose most of the flux, the PSF effect is small.
- Flux upper limits are also important – the reddest objects in a sample (aka "dropouts") are often of special interest, and in many cases these will be detected at high significance in redder bands but almost undetectable in a bluer band in a practical integration time. If an object with high significance at one band is undetected at some other band, then a useful 2σ upper limit should be provided when available. If the non-detection is due to data problems (e.g. missing data, dead detector, bad pixels, nearby bright star, cosmic ray etc) rather than intrinsic faintness, this should be flagged. (NB: asteroids are a potentially serious contaminant here, hence Requirement 10/5).
- A major driver for VISTA surveys is to discover new interesting objects for follow-up with VLT, ALMA and other facilities. In many cases visible and/or NIR spectroscopy will be important, with a variety of instruments. This is outside the remit of VDFS, but the Archive should facilitate it. Most such cases will require accurate astrometric offsets from fiducial stars (e.g. for multi-object spectrographs) or offset stars (e.g. where a target is very red and thus invisible on a typical CCD acquisition camera). See Requirement 5/14 above.
- VISTA's source catalogues may contain a few billion stars and ~ 200 million galaxies, with the majority in the wider/shallower surveys. (The Deep and Very Deep surveys may contain "only" a few million objects). While these are large numbers, Moore's law suggests that distribution of entire source catalogues on one hard-disk will be feasible by 2007, but this is probably not

the case for pixel data. Therefore, if resources are limited, Archive requests involving serving or computations on pixel data should take precedence over complex processor-intensive operations on the source catalogues.

- Many statistical-based projects require completeness maps in addition to a source list; i.e the detection probability as a function of magnitude and colour as a function of position on the sky. Relatively coarse bins (e.g. 1 arcmin) are generally sufficient, so the data volume is modest, but working out the completeness function may require adding simulated objects to pixel data and re-running source detection algorithms.

The following list of usage examples is in arbitrary order with no prioritisation implied:

1. **z ~ 7 Quasars:** Select a sample of candidate quasars at z ~ 7 from a wide-area survey. Candidates shall be selected with good significance, stellar PSF and good quality flags at both Y and J bands, with very red z-J colours (including objects with upper limits in z-band placing them definitely in the selection box).

2. **Weak lensing tomography**. For a catalogue of galaxies selected at i-band with Petrosian magnitudes, provide J,H,Ks magnitudes in matched apertures. (NB actual weak lensing shape measurements require visible data and specialised processing outwith the scope of VDFS, but near-IR magnitudes need to be provided for robust photometric redshift estimation).

3. **Baryon wiggles with FMOS/KAOS.** Select objects in VISTA Deep survey at J-band using Petrosian magnitudes. For all objects above appropriate magnitude limit (e.g. corresponding to 10σ in median survey conditions, 7σ in poor conditions) provide fluxes in matched apertures in other visible+NIR bands. Also provide completeness map, e.g. estimated detection probability at given magnitude limit averaged in 1 arcmin pixels across survey area, and fiducial stars for subsequent massive redshift survey. Targets will be selected using photo-z's to preferentially select higher-redshift objects.

4. **Galactic plane map.** Using object catalogue from Galactic Plane survey, Archive returns map of star counts at 1 arcmin resolution binned into 0.2 magnitude bins and 0.05 mag J-Ks colour bins; can be searched for star clusters and analysed to produce dust map.

5. **Magellanic Cloud survey.** From VISTA + VST Magellanic Cloud survey, return star counts binned to 1 arcmin sky pixels, 0.1 mag bins and 0.05 colour bins. User fits clump-giant location and shape to produce reddening distribution.

6. **SWIRE galaxies.** User uploads source catalogue detected by Spitzer SWIRE survey at 5σ at 3.5μm; Archive returns VISTA Very Deep survey J,H,Ks

photometry in 2 arcsec apertures at given positions including upper limits or missing-data flag, and including time-tags.

7. **APEX/SCUBA2 objects.** User uploads list of submm sources detected by APEX or SCUBA-2 including elliptical position error contour. Within each 95% confidence error ellipse, Archive returns multi-band NIR+visible object list.

8. **z ~ 8 quasars.** From VISTA wide-area survey, select sources with good significance, stellar PSF and good quality flags at both J and H (or Ks) bands, with very red Y-J colours including Y upper limits. Archive returns both candidate list and (processed) thumbnails. User inspects thumbnails to weed out obvious defects, then requests pre-stacking image thumbnails to confirm good candidates. (Known L and T dwarfs do not get redder than Y-J ~ 1.5, so anything substantially redder than this in high-latitude sky is a good z ~ 8 candidate. Compact ellipticals at z~1.6 are a potential contaminant, but should be bright in Ks.)

9. **Brown dwarfs.** VISTA's Wide area survey and Atlas may each contain ~ 50 times more brown dwarfs than 2MASS. These require colour selection using i,z,J or z,Y,J diagrams, so user searches for good-quality point sources in appropriate region of 2 colour space.

10. **Y dwarfs:** (brown dwarfs cooler than T spectral type). Select candidates from VISTA Wide-area survey with stellar PSF and near-IR colours appropriate to these.

11. **Ultracool white dwarfs.** For VISTA wide-area survey with 2nd epoch J-band data, produce reduced proper motion diagram (i-J, J – 5 $\log_{10} \mu$) and select objects in region appropriate to old He atmosphere WDs.

12. **SZ clusters.** From a list of cluster candidates detected via the Sunyaev-Zeldovich effect (e.g. from South Pole or Atacama telescopes), take 2 arcmin radius circle and return multicolour object catalogue and images for user to estimate photometric redshifts.

13. **Identification of WISE objects.** Take source list from WISE (a space-based all-sky mid-IR survey at 3.5, 4.8, 12, 25 μm, planned launch ~ 2008), select matching objects in VISTA Atlas and provide J,Ks photometry. Numerous science applications including LIRGs, star-forming regions. VISTA J-band non-detections may include luminous quasars at z > 10, if they exist.

14. **Identification of IRIS objects:** Take source list from IRIS selected at 75 μm with elliptical position error contours. Return object list from VISTA Atlas or Wide-area survey within error ellipses, sorted in order of likelihood. (Actual identifications may be non-unique due to the significant position errors of

IRIS).

15. **Pre-main sequence stars**. Select sources in Galactic cluster, return z,J,H,Ks colours to search for excess due to possible disk.

16. **Star cluster mass function**. Provide luminosity function for Galactic cluster, using a comparison region to account for background contamination. Completeness function also required, e.g. using simulated objects added to real images.

17. **Low surface brightness galaxies**. There are two aspects: i) take list of LSB galaxies detected in a visible survey, return pixel data for user to fit visible-NIR colours for star formation history. (ii) Detection in NIR. This is challenging due to the high sky brightness; user will need to upload specialised code which the Archive then runs on the Deep survey pixel data.

18. **K-band microlensing.** A specialised project to take short exposures nightly of a tile near the Galactic Center, to measure optical depth close to the GC where extinction is too high for optical searches. This requires sophisticated difference image processing, possibly in real time. Processing steps beyond standard reduction of tiles are outwith the scope of VDFS.

19. **Supernovae**. Repeated imaging of a tile at z or Y band with VISTA to search for supernovae at z ~ 0.8 to 1. This requires real-time processing which will have to run at Paranal, hence outwith the scope of UK DFS.

oOo