WFCAM Pipeline Design

Jim Lewis, Simon Hodgkin, Dafydd Evans, & Mike Irwin

23rd October 2002

Applicable Documents

Reference	Document Title	Document Number	Issue and Date
	WFCAM Pipeline science requirements	http://www.jach.hawaii.edu/~adamson/wfplrq.html	
	WFCAM Science archive requirements	http://www.jach.hawaii.edu ~adamson/wfsarq.html	
	The UKIDSS Proposal	http://www.ukidss.org/sciencecase/sciencecase.html	
	UKIDSS technical documents	http://www.ukidss.org/technical/technical.html	
	UKIRT board review PDR document	UKIRTB(01)31	December 2001

Contents

1 Introduction

Automatic pipeline reduction of WFCAM data is predicated on the assumption of a well-defined set of observing protocols that forward the relevant meta-data to the pipeline reduction system. In turn the pipeline has to not only deliver science-quality products but also provide feedback on the health of the camera and on the overall data quality. This leads to several general front- and back-end requirements:

1. All the information that the pipeline needs to process the data and to monitor the state of the system and health of the observations has to be contained within the data file header structure (NDF/FITS) and data file main body.

2. The allowed observing sequences, or MSBs in UKIRT parlance, have to be sufficiently general to satisfy a range of user expectations, but also sufficiently well-defined and self-contained to enable automatic pipeline processing.

3. The summit pipeline must run hands-off in (almost) real-time and provide data quality control and summary information to the Observing Management system, in addition to making available to users first-stage science quality data products.

4. Data products from the UK-based pipelines must satisfy the overall science requirements of the UKIDSS consortium and must provide the necessary input and processing options for the UK science archive.

5. An on-line database describing the observations, and the derived Data Quality Control (DQC) information will be maintained on Hawaii and mirrored in Cambridge, as is done currently with all UKIRT observations. This database will form the knowledge interface that helps define and plan the overall observing strategy.

This design document begins by reviewing the proposed overall philosophy of data processing for WFCAM, proceeds by outlining the plans for the overall design of the MSBs (Minimum Schedulable Blocks), and then discusses the proposed generic pipeline design.

2 WFCAM Data Processing Requirements

In developing a data processing strategy we have had to make various assumptions regarding the detail of the science and data products requirements (since they are still under development) and have of necessity had to allow for uncertainties in the detector properties and the (unknown) properties of the spatial and temporal variation of the sky over 1/2 degree scales. With this in mind we have deliberately designed the pipelines around a selectable modular scheme for maximum flexibility.

In particular, there is an overall requirement for: a summit pipeline to assess data quality on-the-fly, provide feedback to survey planning and progress tools and provide first-pass science products for users at the summit; a basic calibration and science product pipeline to be run in the UK; further processing using an advanced survey product pipeline/toolkit including dealing with the complex issues of asynchronous image processing for deep stacking and optimal multi-passband knowledge extraction; and an active GRID-enabled science archive capable of dealing with and enabling sophisticated datamining queries and of being a cornerstone of the UK Virtual Observatory.

All WFCAM data, including PATT data, is to be processed through the summit and basic pipelines and the calibrated science data made securely available through the on-line WFCAM archive.

2.1 CASU–JAC Memorandum of Understanding

The agreement with JAC for the summit pipeline is based on the following memorandum of understanding:

- CASU will write the WFCAM summit pipeline using ORAC-DR to call their algorithm engines. This will use C library modules under perl/XS and use FITS as the primitive transport. JAC will help with any ORAC-DR infrastructure modifications or advice.
- A large subset of this pipeline will run at UKIRT in real time. This will include dark subtraction, flat-fielding, fringe subtraction, interleaving and catalogue primitives. The same precise pipeline modules will run at CASU with whatever additional steps are needed to deliver the UK basic science products.
- JAC will write any additional code primitives required for nightly overview quality monitoring, moving object detection and other potential JAC products. JAC will probably make an additional PC available for this.
- CASU and JAC High-Level Software Group will present a joint document on the WFCAM pipeline for the software CDR.
- JAC conditions on CASU
  - The work will be done respecting the ORAC-DR philosophy.
  - CASU group members will do the work.
  - There will be a single CVS repository at JAC for the ORAC-DR code, and a single CVS repository at CASU for the algorithm code.
  - CASU will allow their ORAC-DR code to be GPL'd and shipped by JAC as appropriate. Also the algorithm code will be made available to anyone who wants to reduce WFCAM data (PATT users) and distributed by JAC.
  - CASU will support data reduction with the summit pipeline by any PATT observers providing they used agreed observing protocols.
  - CASU will follow JAC software requirements, for example, documentation for both ORAC-DR and data processing algorithms.
  - This effort will be run along project lines with CASU and JAC High-level Software group accountable to Adamson and Casali.
- CASU conditions on JAC
  - JAC will make sure that the ATC WFCAM instrument team meets all JAC and CASU header and acquisition requirements.
  - One of the CDR documents will specify contents of FITS headers.

2.2 WFCAM Pipeline Science Requirements

The agreed WFCAM pipeline science requirements document is included as an Appendix.

3 Overview of Strategy

Data will be taken according to a flexible menu of observing protocols (discussed in the next section) to guarantee automatic pipeline processing both at the summit, and in the UK.

Our general philosophy is that all fundamental data products are FITS multiextension files (MEFs) with headers describing the data taking protocols in sufficient detail to trigger the appropriate pipeline processing components, including the generated catalogue binary tables, and that all derived information, data quality control (DQC), photometric and astrometric calibration and processing details, are also incorporated within the FITS headers. These headers thereby provide the basis for ingest into databases for archiving and databases for real time monitoring of survey progress and survey planning.

3.1 Summit Pipeline

- The main purpose of the summit pipeline is to generate near real-time DQC information by processing data on-the-fly using either library master calibration frames or locally derived (from extant MSBs) calibration data and to provide a first-pass science quality product for local users.
- The Data Aquisition System (DAS) will coadd multiple reads/consecutive exposures on the same field, apply non-linearity corrections, as necessary, and assemble each multi-sector read into a coherent whole. This will minimise the data flow system requirements to $\approx$100 Gbytes per night.
- The summit pipeline will run under the control of ORAC-DR and together with the CASU-supplied data processing infrastructure modules, will produce the required calibration and DQC information, the export FITS files and instigate FITS header and data verification.
- The summit pipeline will compute, maintain and update a (time) series of master calibration frames (dark, flats, skies, hot/bad pixel masks) to provide basic 2D instrumental signature removal.
- The summit pipeline will produce object catalogues that will additionally be used to generate DQC measures including: astrometric pointing accuracy; photometric throughput; image shape measures; and sky brightness and noise levels.
- Each detector data array will be processed separately and if tiling is required at the summit an additional processing stage will be added.

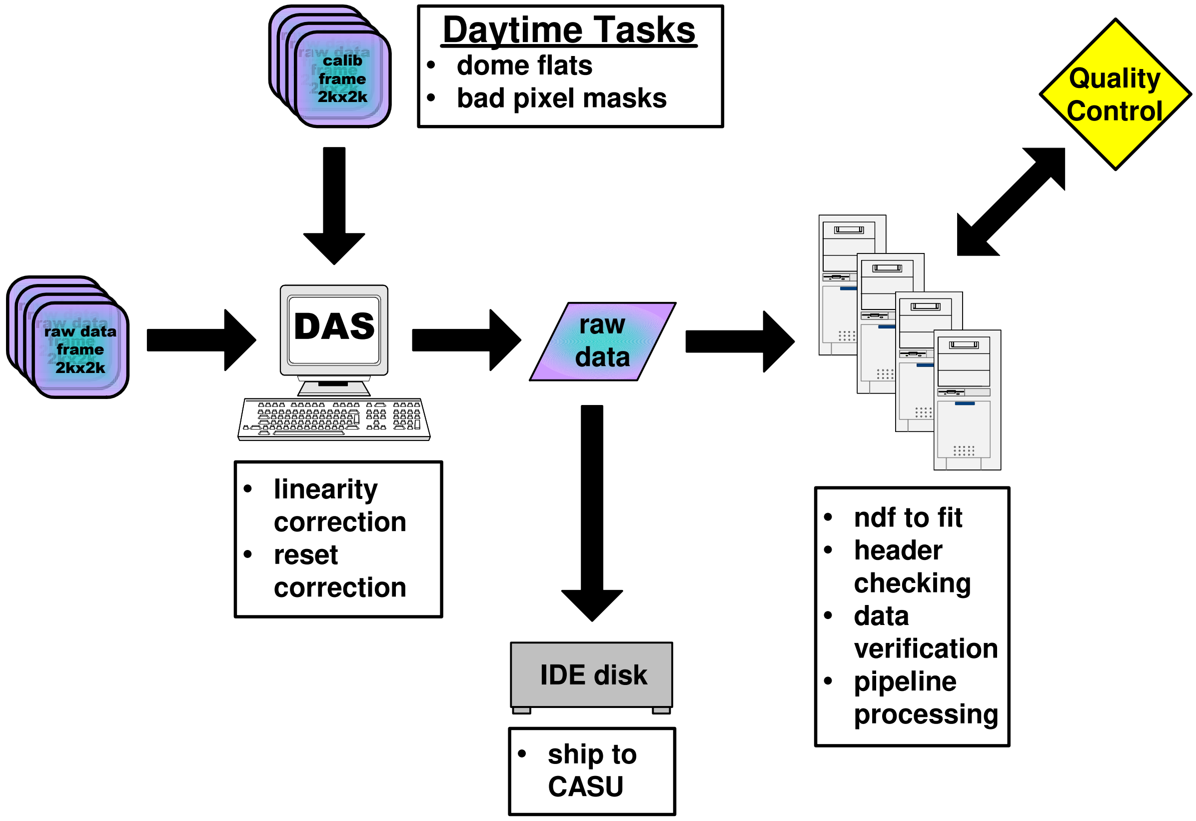

Figure 1: An overview of the summit processing.

3.2 Basic pipeline

- The basic pipeline in Cambridge will be a superset of the summit pipeline such that when the same processing sequence is used identical output products will result and will operate on the MSBs taken during each night to produce instrument signature-free images and to derive object catalogues from them. The object catalogues will be used to provide the first pass astrometric and photometric calibration and together with the images will form the basic science survey products.[1]
- Image data products at this stage will be instrumental signature-corrected and sky variation-corrected: single frames; lossless interleaved superframes; stacked dither (super)frames; tiled contiguous mosaics combining four pointings with any of previous options; and confidence maps for all pipeline image products.
- The catalogue data products will be lists of detected images with an agreed set of parameters (provisional details in Appendix) summarising useful astronomical information and providing the necessary DQC information.[2]
- Data products will be made securely available for WFAU to transfer to Edinburgh from Cambridge directly via the internet with an end-to-end minimum connectivity

[1]Note that the Cambridge basic processing will have access to an entire nights data and may derive different flatfields, and sky correction frames compared to the causal summit pipeline.

[2]Parameters to include: image location, various flux and shape measures, and error estimates for them. For a starting point we will base this on the INT WFC pipeline catalogue products, with selective addition from the SDSS and 2MASS pipelines.

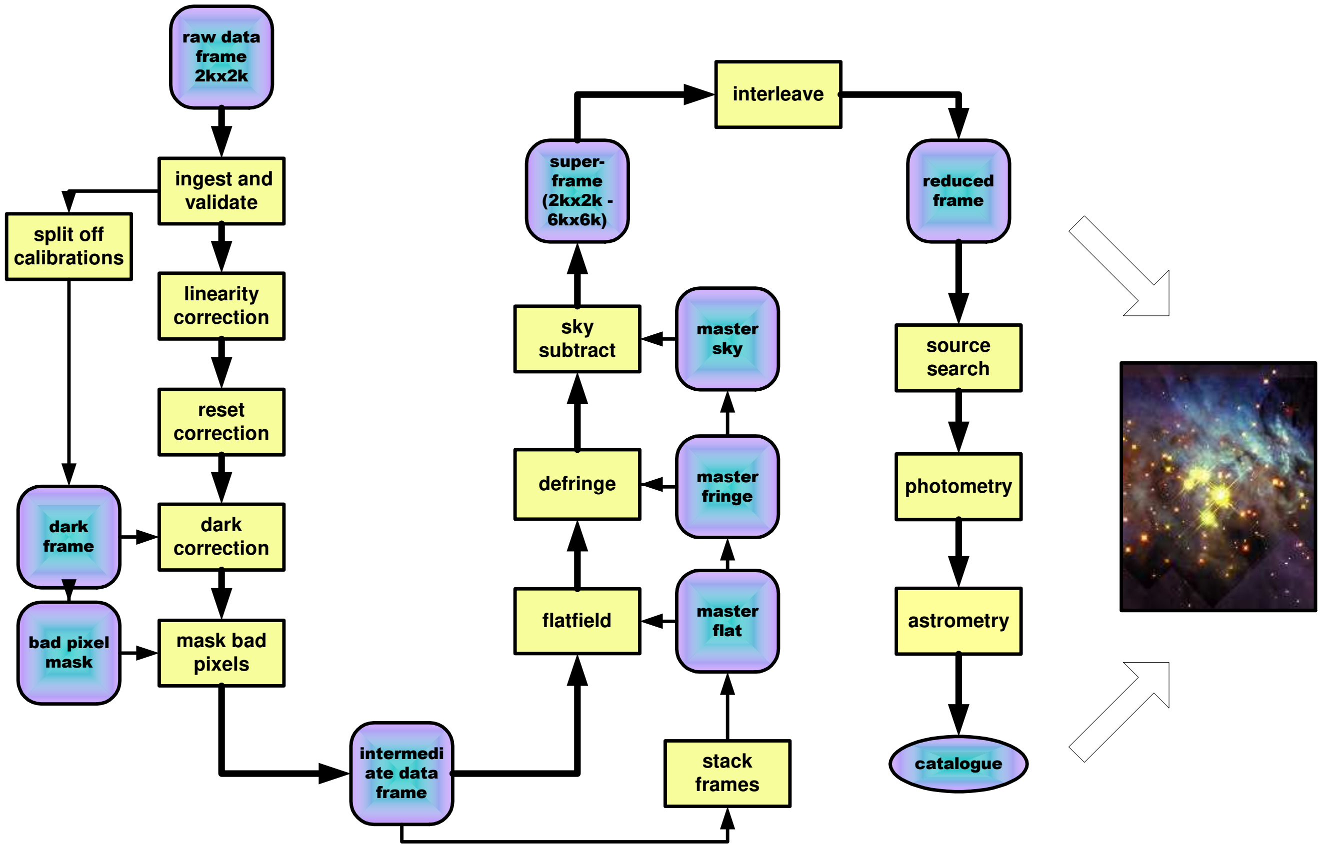

Figure 2: Overview of processing operations required to deliver basic science survey products from WFCAM data.

bandwidth of 100 Mbits/s.

- The raw science data will be held in an online database in Cambridge, both for backup and reprocessing purposes and to provide accredited access to the raw data for external users should they so wish.

A generic flowchart outlining the main processing tasks required for the basic and summit pipelines is shown in figure 2.

3.3 Further and advanced processing

Many of the likely requirements for an advanced dynamic processing system stem from the necessary combinatorial operations that will be required on images/catalogues derived from basic calibrated science survey products. Since the time of aquisition of **all** the requisite data is unpredictable this is distinct from the basic processing and some of the further processing options.

- Many advanced survey products are predicated on the ability to automatically generate detailed well-sampled PSFs as a function of position over each basic product image. Example uses of accurate PSFs are: optional automated PSF fitting to improve s:n of point sources particularly in crowded stellar regions; and adaptive kernel matching of PSFs for difference imaging for transient event detection.

- A related, but functionally separate issue, is how to optimally stack images taken at different times in different seeing conditions. This will involve assessing various stacking techniques where optimality will undoubtedly depend on the science aims for the stacked data and may well delve deeply into (CPU intensive) image restoration techniques.

- Producing contiguous tiled regions and seamless overlap-calibrated large area catalogues. These and the preceding item could be initiated automatically from a database-driven list of survey fields and survey progress eg. is it worth (re)stacking (re)mosaicing (re)calibrating this region ?

- Detection and parameterisation of Low Surface Brightness objects and an investigation of more general purpose image detection methodology that is not just optimised to find faint PSF-like images.

- Matching/merging frames in different passbands and for use in defining master catalogues for any given field for list-driven co-located photometry (cf. SDSS and their ability to measure fluxes and limits to fluxes in a consistent way).

4 Design of Observing Strategy

Flexibility of observing protocols to satisfy different user science requirements and well-defined recipes to enable automatic pipeline processing is the goal. To achieve this we will base the observing strategy on a nested menu system (figure 3) and make each MSB as self-calibrating and self-contained as possible. This is due to ongoing uncertainties in the feasibility of an approach based on master library calibration frames.

The data processing pipeline will be sufficiently modular to cope with a variety of use cases providing agreed observing sequences are followed. These could typically range from quick look exposures, to producing contiguous tiled mosaics using offset skies.

4.1 MSB Definition and Restrictions

A Minimum Schedulable Block (MSB) is a selection from iterated targets, loops, microsteps, dithers, tiling, and/or offset patterns taken at one or more sky positions (where position includes micro and macro offsets). Although this is an open-ended definition, in practice an MSB will be limited by several factors:

- there is bound to be a 'longest acceptable observing time' for any observation block due to possibly rapid seeing, sky fringing and sky brightness variations (temporal and spatial) over the WFCAM 1/2 deg field (this needs to be assessed during commissioning);

- ideally each MSB should be as self-calibrating as possible since there may be, for example, time-dependent flatfielding effects due to thermal emisison or sky colour and emision line variations. This implies that in general sufficient exposure time has to accumulate in each MSB to ensure the data can be properly calibrated;

- offsets in dither and microstep patterns for interleaving should be <20" (TBC) to avoid problems with non-linear astrometric distortions leading to problems stacking/interleaving, and to remain within the throw of the guide star movement – tiling will require non-linear resampling in all cases.

There are two generic modes of NIR observing (notwithstanding the caveats in section 5.2):

- Standard interleaving/dither/tiling pattern (*e.g.* 2x2 + 3x3 + x4) – where the flatfields, sky correction and so on are generated from stacking of all frames. Appropriate for uncrowded fields with no large extended objects.

- Offset sky dither pattern – flatfield and sky estimators generated from offset sky frames. Appropriate for highly crowded regions (*e.g.* Galactic Plane Survey) or fields with large extended objects (*e.g.* Virgo cluster galaxies).

The optional steps required to achieve these are illustrated in figure 3 and are defined more precisely in the next subsection and in the glossary in the appendices.

Briefly in general terms, the lowest level iterator is a loop, where a loop is a combination of N exposures coadded at a given position to give an integration.

This is optionally followed by a combination of microsteps for interleaving and dithers; where a microstep is a small integral + precise sub-pixel shift for interleaving pixels to create better sampled images. These can also be stacked for flatfields and sky estimation but cannot be used to remove bad pixels. A dither is a small offset over, say, a "spiral" grid (*e.g.* 3x3). Combined dither images can be used to remove bad pixels, and other defects and stacking with rejection generates flatfields and sky estimates.

Completing the minimum contiguous region on-sky requires four related pointings of the inner iterators to fill the gaps between detectors – a tile.

Optionally, large offsets may be required for flatfield and sky estimation in crowded regions or for large area objects. They could also be used to offset to the next science field if a single field does not meet self-calibration requirements.

Separate calibration observations, or additional targets, optionally make up the top level iterator for the MSB.

4.2 Definition of MSB components

MSB – a Minimum Schedulable Block is a sequence of integrations that makes up a largely self-calibrating and useful series of scientific observations. The component parts of an MSB are defined below and in turn are constrained by the scientific requirements on flatfielding accuracy, astrometric and photometric calibration and sky removal.

pointing – a telescope slew to position with new guide star acquisition. Note that microstep sequences and dither sequences are done within one pointing, *i.e.* within $\approx \pm 10$ arcsec range of the nominal pointing position.

read – a read is the act of physically reading a detector. This generally means the act of digitising the data, and can be either destructive (*i.e.* followed by a reset) or non–destructive

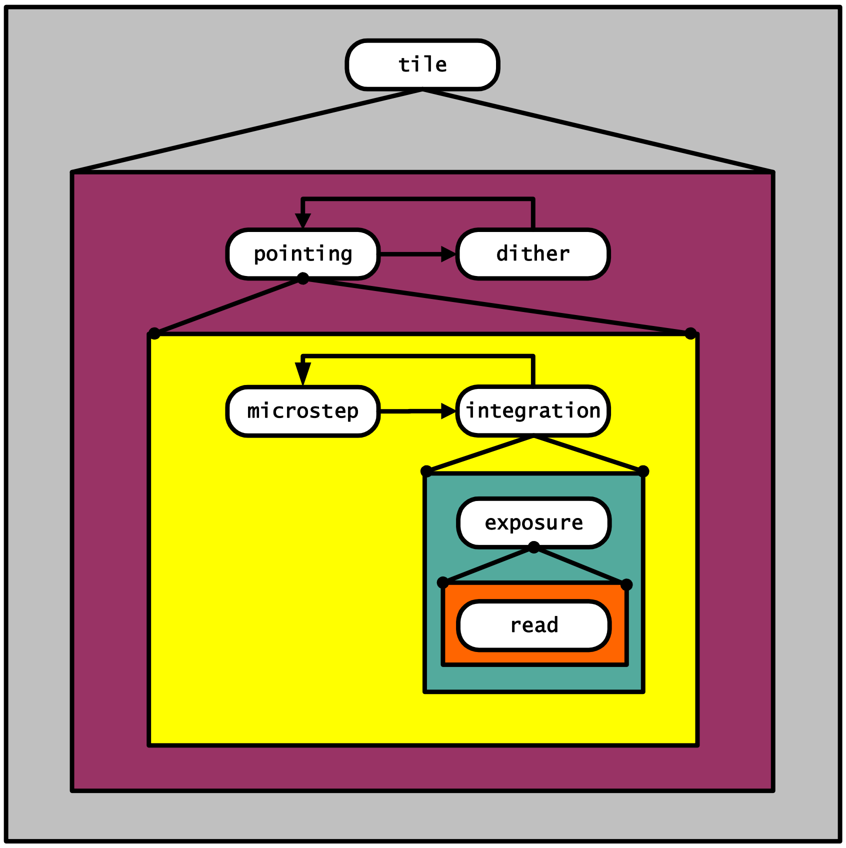

Figure 3: The elements for each target within an MSB. Each component is an optional iteration of a sequence of observations. Thus at the lowest level, an exposure is made up of a sequence of reads. Exposures can be looped over to build up an integration. Microstepping over integrations for interleaving will increase the size of the data frame by 2x2 or 3x3. Interleaving microsteps also include an integral pixel shift (*e.g.* step n+0.5 pixels) so the data can be combined to create flatfields. This forms a pointing. Dithering around pointings can be used to remove bad pixels. Four pointings combined together form a tile. For shallow survey tiles with a 2x2 microstep each detector pixel would see 16 different bits of sky. Offsets to other targets, or for sky estimation, are not included for clarity.

darks – reference exposures of the same duration as the images in an exposure sequence used for correcting detector artefacts [3]

exposure – an exposure is the result of a sequence of one or more reads, that are used to produce a single output image. The multiple reads are combined within the data aquisition system

loops – a loop is a contiguous sequence of multiple exposures in one place to increase integration time, optionally combined *internally* in the data aquisition system

integration – an integration is the result of a sequence of exposure/loops.

microstep sequence – a sequence of 1x1, 2x2 or 3x3 integrations taken with accurate sub–pixel (including small integral) shifts, combined by interleaving the pixels

[3] current generation NIR detectors require darks of the same exposure time as the observations and the darks may vary as a function of time

dithers – sequence of integrations stepped in "spiral" pattern for use in bad pixel rejection, deep stacking *etc.*, combined using pixel stacking

tile – the four pointings and observing sequences necessary such that a combination of them covers a contiguous area of sky involving all four detectors

offset – a macrostep to another region of sky within an MSB sequence which may be either another target region or an offset region used for sky estimation

mesostep sequence – a type of MSB being a sequence of exposures designed to completely sample across the face of a detector to monitor residual systematics in the photometry

focus run – a type of MSB with a series of exposures for use in providing an estimate of best focus

5 Design of MSB Data Reduction

5.1 Preamble

As stated previously, ideally, an MSB should be self-calibrating, *i.e.* there must be sufficient information in each MSB such that all instrumental signatures can be removed independently of any observations that occur either before or after. While this is a worthwhile goal, in practice there are several steps in a standard infrared reduction procedure where this principle could break down:

Dark correction: Current UKIRT observing practice is generally to take a dark frame with each observing block with the same integration time as the program observations. This frame is then simply subtracted from each of the program frames as the first part of the reduction. There are several difficulties arising from this procedure:

- The dark frame will have cosmic ray hits and one observation is insufficient to allow these to be detected
- The 'dark' exposure is not used only for removing dark current, but also used to correct for a small amount of residual reset bias that is present in the data after normal reset correction. Whether this extra requirement is necessary depends on the currently unknown detailed properties of the WFCAM detectors and read-out system.
- Is it possible to get a statistically valid estimate of the dark current using the sort of short exposure times that are likely to be employed with WFCAM, by using multiple exposures ? If we need to expose for several minutes in order to get good dark frames it would be a significant overhead for each MSB.

The frequency/need for dark correction images will have to be determined once the detectors being used have been characterised.

Image retention: this is potentially a tricky issue, but until we have suitable test results characterising the time decay profile and adjacency effects, is hard to plan for. Correcting for image retention may involve the need to accumulate a persistence map running across MSBs and use this as input to MSB processing.

Flatfielding: The standard procedure at UKIRT is to flatfield images with the night sky observations that comprise the current observing block. While this has several advantages, it may not be a valid way of treating the problem in the presence of sky fringing and other additive effects – see section 5.2 for more details

5.2 Night sky flats and fringes

Deriving flatfields to measure detector pixel-to-pixel gain sensitivity variation and system overall throughput -v- spatial location, is dependent on the ability to produce the equivalent of uniform farfield illumination and to have negligble local non-uniform thermal contribution.

With stable electronics, detector properties and filter positioning, the necessary flatfield correction should be constant over at least a night. However, this is difficult to assess because in practice it is virtually impossible to guarantee uniform illumination. Even on-sky exposures suffer from various extra sources of non-uniform emission that vary with time. In the K-band on UFTI, for example, the flatfield exposures are affected by emission from dust on the window. Although this emission seems constant over time, the sky brightness is not and so the percentage of flux contamination due to dust emission varies over the course of the night. Night sky fringing is another additive component that can cause flatfield exposures to vary over the night. Separating out these additive components is not straightforward, but if it is not done, then the derived flatfield may introduce significant systematic errors.

There are several possibilities for aquiring the necessary information for deriving flatfields:

Dome/internal flats:

Advantages – can be done during the day, no problem with insufficient flux, dust emission should be neglible, no fringing, can be used for linearity measures as "light" source can be made very stable;

Disadvantages – virtually impossible to get uniform illumination.

Twilight sky flats:

Advantages – sky background level may be bright enough to make fringes and dust emission unimportant;

Disadvantages – only a limited amount of time per night in order to acquire them, the assumption that flatfield properties do not change may not be true.

Night sky flats:

Advantages – constantly updating, so potential variability in the flatfield properties is taken into account, the source function is the same colour as night sky in the observation frames (by definition);

Disadvantages – the presence of fringing and dust emission in some bands.

In practice a combination of these measurements will be required to calibrate WFCAM, with daytime measures being used for non-linearity updates and sky measures used to derive the rest.

We can get some idea of the likely fringing and dust emission levels by using existing UFTI frames. K-band UFTI images show readily visible dust emission at the level of 25 photons/s/pixel for the brighter dust spots. This appears to be constant through the night (which is to be expected as the camera should be at a constant temperature). The mean UFTI K-band night sky background level is roughly 10× the dust emission level, but does however change with time, which is why a flat derived from one part of the night does not work well for another part of the night. For twilight sky, selected to have, say, a count rate 10× higher, the dust contribution drops to less than 1% of the sky.

The fringing level is a more difficult problem to predict as it is highly detector-specific. UFTI has fringing present at a level ±3% of sky in the H band (negligible in J and K). The fringe pattern seems quite stable over a night, which means that fringe subtraction works well at the sky subtraction stage. However, with the larger detector area and larger sky coverage of WFCAM there is no guarantee that this will remain so, indeed wide field optical mosaic cameras show variable fringe patterns as a function of sky conditions.

Although some estimate of the fringing can be made from laboratory tests before WFCAM goes on the telescope, most of this will have to wait until the instrument is commissioned for a final answer.

5.3 Processing in the Data Aquisition System

At present there are several basic reduction steps that will be done at the telescope in the DAS before data are written to disk. In the case of the first two items, these are done in order to cut the data rate by a large factor:

Reset correction: The reset image (*i.e.* bias frame) will be removed in situ which eliminates the need to write a separate image for each data frame.

Read combinations: Data from all reads that comprise an exposure will be combined together to form a single frame.

Linearity correction: By and large the data obtained from NIR arrays are strongly non-linear. The linearity curve can be determined through observations of a stable "light source" for a range of exposure times. This can be done during the day, and it remains to be determined how often this will be necessary. Because each detector will be read out in (32) sections with each section having its own electronics, it is very likely that a separate linearity correction will be needed for each section, since in our experience most of the non-linearity is generated in the electronics. The linearity correction will be applied on-the-fly in the DAS.

5.4 Main reduction steps

In what follows we deal with each step in the approximate location they occur in the pipeline (see figure 2).

Dark correction: as mentioned before this is not a completely resolved issue. A mean dark frame should come from several dark observations either from within the present MSB or from earlier observations, suitably combined with rejection to get rid of any cosmic ray hits. In theory, dark correction involves scaling the mean dark frame to the exposure time of the program frame and then subtracting from the latter. In practice if other artefacts are present, dark frames of the same exposure time as the indivudal science exposures will be necessary.

Flatfield correction: see section 5.2 for the issues involved. Because WFCAM is a multi-detector camera, internal calibration will also involve accounting for the variation in mean gain from chip to chip. If the sky "on average" uniformly illuminates each detector then the variation in the "mean" counts on the flatfield map for each chip is a measure of the variation of the mean gain. Flatfielding is usually defined such that the "mean" counts in the object frame remain constant. For situations where there is a gain difference between detectors, the mean flat for each detector will in effect be normalised by the ensemble average counts over all detectors thereby ensuring correct inter-detector gain normalisation.

To be more precise, as a function of position x, y and wavelength λ:

$$D_n(x,y) = \int_\lambda G_n(\lambda) g_n(x,y,\lambda) f(\lambda) S(x,y,\lambda) d\lambda \tag{1}$$

where $D_n(x,y)$ is the data as observed on detector n, G_n is the gain correction for chip n, g_n is the relative gain map for chip n (the flatfield), f is the remaining system response curve and S is the source function. The obvious caveat with defining differential gain corrections this way is that they are potentially a function of source colour due to inherent variation of the QE curves of the detectors (indeed this can even be the case at the individual pixel level).

Form initial confidence map: the initial confidence map for each exposure is formed from the mean current flatfield, and will therefore be the same for all images in a given flatfield sequence. Pixels outside a specified tolerance range are given a confidence of zero. The confidence for unrejected pixels is defined on a percentile scale such that the median confidence of unrejected pixels is 100%. Confidence maps are carried through further processing stages in conjunction with the "average" data frame background level noise variance.

Defringe: if the fringe spatial pattern is stable and if flatfields can be generated without fringing present, it is possible to decouple sky correction and fringe correction and apply a defringing method similar to the one we have developed for optical imaging.

Sky subtraction: alternatives. Standard NIR processing often subtracts sky first and then flatfields. We can see why this can be advantageous compared with dark-correcting, flatfielding and sky-correcting by considering the following encapsulation of the problem,

$$D(x,y) = ff(x,y)\ [S(x,y) + F(x,y) + O(x,y) + T(x,y)] + dk(x,y) \tag{2}$$

where $D(x,y)$ is observed, $ff(x,y)$ the flatfield function, $S(x,y)$ is the sky illumination, $F(x,y)$ is the fringe contribution, O(x,y) is the object contribution, T(x,y) is the thermal contribution, $dk(x,y)$ is the dark current and without loss of generality we have dropped any explicit wavelength- and time-dependence for clarity. Stacking with rejection produces an estimate of the terms

$$\hat{I}(x,y) = ff(x,y)\ [S(x,y) + F(x,y) + T(x,y)] + dk(x,y) \tag{3}$$

therefore,

$$D(x,y) - \hat{I}(x,y) = ff(x,y)\; O(x,y) \tag{4}$$

obviating the need for dark-correcting and fringe removal as both separate data gathering requirements and as separate data processing steps; and minimising the effect of systematic and random errors in the flatfield function by removing the largest potential error terms.

This of course glosses over the question of the spatial and temporal variability of the sky background. If the sky characteristics change significantly over an MSB then this method will still leave sky residual patterns exacerbating the problems, for example, of forming useful mosaic tiles.[4]

Reset anomaly: Many NIR detector systems have noticeable residual structure left in the background after the reset correction. For multi-sector reads, reset anomalies could leave a challenging background variation over the detector to deal with. This is analogous to the problem of dealing with rapid variations in sky background level/structure during stacking/mosaicing. We have had to develop ad hoc techniques to deal this problem for CIRSI data.

Image retention: ghosts of images/artefacts from preceding frames may be present. Strategies for dealing with this involve assessing the time decay characteristics and adjacency effects (*i.e.* ghost image spreading). Correcting for image persistance will either involve updating and maintaining a persistence mask, or accumulating a persistence map, running over MSBs if necessary, to subtract from the current image.

Interleave: the on-sky detector pixel scale for WFCAM will undersample typical seeing conditions on Mauna Kea. To recover some of this lost resolution, an observation at a particular pointing optionally consists of several microstepped exposures. Microstepping is done by shifting the telescope by a very precise non-integral pixel distance. In the main, WFCAM will probably use a 2x2 microstep sequence and the shift for this sequence will be an integer number of pixels plus a half. Interleaving is done for this sequence by creating an output image that is a regular interwoven pattern of **all** the input pixels. Therefore each output image pixel samples the sky on a finer grid that the input image pixels and helps recover some of the lost resolution for undersampled images. However, interleaving does nothing about removing bad pixels. In situations where the MSB includes a dither pattern, this latter can be used to remove bad pixels and cosmic rays. Note that interleaving and/or dithering will result in a new confidence map for the output image.

Rough World Coordinate System (WCS): basic information in the FITS header as well as knowledge of the instrument will allow us to give each processed frame a WCS that is good to of order several arcseconds. The basic information needed is the RA and Dec of the pointing, a (stable) reference point on the detector grid for those coordinates (this is usually the optical axis of the instrument), the projected pixel size, the rotation of the camera, the relative orientation of each detector and the geometrical

[4] For the sparse field situation with mainly point sources and small extended objects, any object extraction program with a background following algorithm will do an excellent job at subtracting out the background flux for each object at the extraction stage. Note that the background flux will be a combination of the sky emission, scattered light and diffuse astronomical emission (*i.e.* halo of a bright star or the disc of a galaxy). So whether you subtract out the sky component explicitly or not, background following will still be necessary for accurate extraction.

distortion of the telescope + camera optics, which defines the astrometric projection to use. Although most of this can be gleaned off-line, deriving accurate values through on-sky observations will be an important task at commissioning time.

First Pass WCS fit: once we have a rough WCS for the processed frames, we will fit a more accurate WCS using secondary faint astrometric standards (*e.g.* USNO V2.0). If we ultimately wish to mosaic (tile) several of these frames together then this WCS will be used to map the input frames onto the output grid. The procedure is standard and straightforward: extract a (shallow) catalogue from the frame; extract a section of the astrometric catalogue for this region and use the rough WCS to help match the celestial positions with the (x, y) positions on the detector; do iterative clipped least-squares solution; update WCS.

Tiling: the pros and cons of stitching interleaved/dithered images together to form a contiguous mosaic tile are:
pros: making better use of fuzzy "dither" edges round each of the 16 "detectors" in a contiguous tile – a large fraction of the total contiguous area is on more than one "detector stack"; generating "smooth" mosaic diagnostics and internal calibration statistics from combining many input detector stacks; reduces by factor $\times 16$ later complexities of overlap cross-calibration; user expections for contiguous tiled mosaics
cons: involves non-linear resampling and possible use of sub-pixel interpolation schemes;[5] sky variation may make it difficult to achieve; not necessarily using individual integrations optimally.

An important aspect of tiling involves the output WCS. The input projection geometry will have no meaning in a tiled image as there is no unique centre of reference for the projection. The tiling algorithm is therefore free to assign a new projection for the output image and to map the input pixels directly to it. We suggest that output mosaic tiles should all have the FITS TAN projection with a reference tangent point defined to be the centre of the output image. Another possiblity is to keep the same projection as used to describe the telescope+camera, Arc projection with Zenithal polynomial distortion. The disadvantage of the latter is the much higher radial distortion term $\times 200$ and the uncertainty over support for the projection in image analysis/display packages.

Tile catalogue generation: generate object catalogues for the tile using the associated confidence map. Whether or not to use tile-generated catalogues as the primary science product is still an open question, so both sets of catalogues will be produced.

Second Pass WCS: If tiling is done, it is a good idea to refit the WCS using all the objects on the tile and using the new WCS geometry. This step is done in exactly the same way as the first pass, except that the input file is now a tile catalogue.

Basic catalogue generation: The basic catalogue generation software makes direct use of the confidence maps and produces the following global information:

- **DQC information:** includes mean sky brightness, sky noise, image detection threshold, average FWHM seeing, average stellar ellipticity and average saturation level;
- **Basic object descriptors:** fluxes, positions and shape information, described more fully in an Appendix;

[5] Note that in general stacking dithers could involve interpolation if accurate pixel registration is not used

- **Overlay file:** in addition to the object catalogue, the generation software produces a 'regions' file suitable for use with DS9 to overlay detection ellipses on an image;
- **Astrometric calibration:** as discussed previously, derived from all-sky astrometric catalogues;
- **Photometric calibration:** the generated catalogues will be matched automatically to whatever photometric standard fields are available. More details of this are this in the UKIDSS technical photometric calibration document.

Miscellaneous: these are all potential problems that we have encountered and had to cope with in processing other mosaic camera data:

- undersampling and intra-pixel sensitivity variation problems;
- how to deal with bright stars, ghostimages, trails ? One possiblity it to use input predictions for positions of bright stars, galaxies, globular clusters, solar system objects and so on;
- cross-talk between arrays and even between readout "sectors". If present usually well-defined and dealt with by measuring and applying a cross-talk matrix to remove effect.

5.5 Hardware Design

Summit processing will use a fast Linux PC per detector, processing data from all four detectors in parallel and use commodity disk RAID arrays for local storage. If image tiling is done at the summit then a 5th PC will be used to combine the individual detector output.

Precisely which system to use for data transfer between JAC and Cambridge is still under discussion but a strong possibility is to ship using pluggable IDE disks to minimise production and ingestion overheads (*e.g.* foreseeable next generation IDE disk will be 320 Gbytes and should hold ≈3 nights of data each).

In Cambridge, in order to provide a backup for the main on-line science archive at ROE and also to facilitate several of the advanced processing/reprocessing stages needed, we propose to store on-line the raw science frames and master calibration data of WFCAM in Cambridge, essentially an extension of the rest of the UKIRT raw data archive. Costs of suitable scale commodity RAID arrays (≈10 Tbytes capacity for each full year of operation) have already fallen sufficiently (currently £5k per Tbyte but likely to be half that price by 2004 when the first tranch of bulk storage will need to be in place) to make this a cost effective and viable (scalable) solution.[6]

With on-line storage in place at two sites we do not foresee the need for further (expensive) near-line "tape" storage systems providing that a high density transfer medium is used from Hawaii, whereby the ≈100 Gbytes of data expected per night is stored on one "tape" or shipped disk. If recovery from the raw data archive is required this would be feasible, even for on-the-shelf storage.

[6]The annual growth rate of the science storage requirements is estimated as follows: a typical WFCAM science frame will be 4 devices of 2048×2048 pixels, 2×2 interlaced, $\times 4$ bytes per pixel ≈ 250 Mbyte. Assuming 75% UKIRT time is devoted to WFCAM with 70% fill factor (weather, faults, other science programmes) we will have to deal annually with ≈ 200 nights of science data. Assuming ≈200–500 frames per night (an average of one every 1-3 mins over a 10hr night) then the pixel archive grows by ≈10–20 Tbyte/yr. All catalogue data (derived from pixels or elsewhere) will be $< 10\%$ the size of the pixel archive.

Transfer of processed data between Cambridge and Edinburgh will be via the internet. CASU currently share a 1 Gbit/s link into the LAN. Assuming WFAU have a similar level of access by 2003–4, 10% of that bandwidth would suffice to transfer each days processed data in a few hours.

We have sufficient computing power in place as part of the Cambridge Data Processing Centre activities to handle the development issues of the pipeline for WFCAM, including trialling different processing methods and benchmarking simulated WFCAM data processing. However, sometime during 2003 we will need to purchase and setup a modest multi-node PC cluster to deal with the real data flow expected to start toward the end of that year and a modest high performance PC-based system to begin serious investigation of advanced processing options and equally heavy duty simulations to assess catalogue completeness issues and the reliability of parameter error estimates.

A WFCAM Pipeline Science Requirements

This list combines all requirements from PDR, with additional input from JAC and UKIDSS. It describes the expected pipeline inputs (by defining the standard observing strategies which the pipeline will have to accommodate), the specific summit feedback parameters which will be used to guide observing in pseudo-real time, and the details of the final output from the Cambridge pipeline.

A.1 Top-level requirements

T1	The Summit and Cambridge pipelines must be functionally identical, produce identical output given the same input and the same processing steps, be based on the ORAC-DR recipe/primitive philosophy (and therefore be configurable at both the primitive and recipe level).
T2	The pipeline should be able to function in the absence, temporary or semi-permanent, of (data from) one or more WFCAM detector arrays.
T3	The summit pipeline in particular must be able to make sky-conditions/DQC information available to the OMP Query Tool. [PDR requirement 6.1]
T4	The DR system must be able to complete any processing required fast enough to keep up with data taking. [PDR req. 6.1]
T5	The DR system must be able to provide quick feedback on the quality of the raw data to a user of WFCAM. [PDR req. 6.2].
T6	The DR system must be capable of providing the QCPs (Quality Control Parameters) required to automatically monitor the quality of the data (using WFCAM survey recipes). [PDR req. 6.3 – note that OCDD no longer lists QCP parameters]
T7	The DR system must be capable of comparing the values of QCPs with defined limits, and raising alarms if the values are outside of those limits. [PDR req. 6.5]
T8	The DR system must capable of standalone operation, i.e. of running without supervision. The data reduction recipe to be used must be supplied in the raw data FITS headers. The observing sequence will ensure that suitable calibration observations are available. [PDR req. 6.4]
T9	The DR system must provide logging information to disk. [PDR req. 6.6]
T10	A recipe for automatically determining best focus from a focus run must be provided. [PDR req. 6.7]
T11	Pipeline version number must be recorded in FITS headers of reduced data
T12	Pipeline must provide recipes to reduce MSBs undertaken in nonphotometric conditions, to a point where the reduced frames can be merged with the rest of the survey concerned in the science archive.

A.2 Observing Sequences

The pipeline should be able to fully reduce data from the standard observing sequences noted below. With the choice of 32-channel readout, the minimum exposure time can be reduced; this may open the possibility of dithering to remove bad pixels, so “microstepped sequences” below should be taken as possibly meaning “jittered series of microstepped sequences” pending further discussion (of feasibility and data rate).

A.2.1 Individual Microstepped survey frames

O1	1×1 microstep
O2	2×2 microstep sequence
O3	3×3 microstep sequence

A.2.2 Complete Microstepped survey tiles

O4	Complete tile of four 1×1 microsteps
O5	Complete tile of four 2×2 microsteps
O6	Complete tile of four 3×3 microsteps

A.2.3 Variants

O7	All of modes 1–6 may be repeated immediately in a series of filters. It is not anticipated that filter changes will be made within the space of one microstep sequence, nor that any Observation will consist of Integrations in multiple colours.
O8	All of modes 1–6 may be carried out in narrow-band filters, in which case "on-line" and "continuum" frames will require subtraction

A.2.4 Standard fields

O9	2×2 microstep or microstep/mesostep sequence on a standard field (mesostep is an offset sequence large enough to allow determination of the vignetting function).

In all of modes 1–6, the survey tiles will be supported by observations of suitable standard fields, at appropriate intervals. Standard stars will always be observed in all broadband filters, unless the observing for the night is known to be constrained to a smaller number of filters.

A.2.5 Extended object

O10	Large extended field (complete tile interspersed with jitter frames on a remote sky patch)
O11	Crowded field: 2×2 or 3×3 microstep sequence with equal number of interspersed sky observations

A.3 Summit feedback

The summit pipeline must reliably generate a number of sky-related parameters which will be used to select from the available survey MSBs. These are to be filed as calibrations or otherwise made available to the Query Tool.

F1	Sky brightness and sky noise measurement (per filter)
F2	Mean stellar image diameter (determined by fitting within the processed frames, filed as a calibration after conversion to K-band equivalent).
F3	Instrumental zero point determined from 2MASS stars in the survey processed frames (or the difference between this and the expected value). Stored as a calibration and made available to the QT.
F4	"Locally photometric" flag (determined from variations in detected image properties in the frames in the microstep sequence).
F5	"Globally photometric" flag (determined from all previous observations of standard fields on the current night).
F6	Limiting magnitude (5s) in the processed frame.
F7	Mean axial ratio of suitable stellar images.
F8	Nightly photometric zero point from observations of standard fields

A.4 Summit QC parameters

The summit pipeline must produce the following parameters relating to quality control of the instrument and telescope. Note that these supplement QC monitoring done in the controller and Linux PC. The latter will be in the raw data FITS headers and it is assumed that they will be propagated. Daytime instrument checks will be carried out using the engineering pipeline provided separately by the WFCAM project.

Q1	Limiting magnitude (5s) in the superframe (possibly normalised to one second exposure time), recorded in an index file. A final check for subtle noise problems, used to trigger an alarm message by comparison to a predefined range.
Q2	Mean axial ratio of bright sources (probably those in F4) – a check for astigmatism and trailing, used to trigger an alarm message by comparison to a predefined range.
Q3	Image diameter (determined by fitting within the reduced superframes, filed as a calibration after conversion to K-band equivalent). The calibration value must either be the average of that from the four individual frame pipelines or (more likely given T2) that from one pre-specified array.
Q4	Sky brightness (one value per filter) – determined from mode or median of pixel values. Recorded in FITS headers and used to trigger an alarm by comparison to a predefined range.
Q5	Image periodic noise (looking for, e.g., 60Hz pickup) - to be done within each reduced frame.
Q6	Running log (filed as calibration) of last available photometric zero-point from the last-observed standard field in each filter. Anticipating possible operation with different filters in the four positions in a paddle, all available zeropoints should be included, not just the one appropriate to the current filter.
Q7	Number of bad pixels, determined from last-determined flatfield as the number of pixels lying outside a pre-specified range.
Q8	Number of hot pixels measured from dark frame
Q9	Dark count per second

A.5 Other summit output

D1	Display of intermediate frames should be configurable via the ORAC disp.dat system
D2	Diagnostic messages should appear on the oracdr
D3	Display of complete tiles in the case of O4-O6. Options selected through disp.dat.

A.6 Cambridge and summit pipeline functions

Note: nonlinearity will be dealt with in the acquisition PC and will not be a function required of the DR pipeline.

C1	Removes instrumental signatures: bias, dark, flat field (to precision limited by photon statistics and array systematics), bad and hot pixels, channel variations, scattered light, vignetting correction, fringing correction where possible. If necessary, applies cross-talk matrix to remove interference between channels and arrays. Maintains and uses running calibrations (superflat, bad pixel mask, master dark frame etc.)
C2	Removes (as far as possible) atmospheric OH emission and sky background, other 2-D background variations arising in the detector, and cosmic rays
C3	Removes persistent images from previous exposures
C4	Interleaves microstep frames to produce one superframe per array
C5	Produces and propagates "confidence map" for the superframe, and propagates this into tiles if produced
C6	Combines superframes into a tile if part of a tile observing pattern, and develops a consistent internal calibration scheme to put observations on an approximately uniform system.
C7	Generates source catalogue (after deblending in crowded fields) containing astrometric, photometric, shape and DQC parameters
C8	Flux-calibrates to 2% (limited by Poisson noise), using all-night or running photometric information, superflats etc. as appropriate.
C9	Astrometry on catalogues and reduced pixel images, the latter to use an agreed and appropriate WCS in all FITS headers. External accuracy limited by input catalogue; internal accuracy limited by s/n, detector stability etc. Goals are 250 and 100mas respectively, with an ultimate goal of 20mas for internal systematics.
C10	Measures running atmospheric extinction coefficient (summit pipeline), records in an index file
C11	Determines nightly extinction coefficient (Cambridge pipeline), records in FITS headers
C12	Provides a recipe which resamples and combines a complete tiled image, without generating catalogue (useful for large extended fields, as in O10)
C13	Provides a recipe to determine residual systematics and scattered light functions from a mesostep sequence on a standard field
C14	Provides a recipe to enable determination of best focus from a series of frames taken at different focus positions (focus value will be included in FITS headers of raw data).

A.7 Cambridge pipeline data products

Note: there are a number of possible ways to catalogue in the case of overlapping frames and/or complete tiles. It is assumed that the summit and basic pipeline's jobs are to reduce the data to hand (microstepped superframe or complete tile) without reference to catalogues from frames in overlapping fields taken previously.

Note: anticipated content of FITS files coming from CASU to WFAU is up to three files per array per pointing: one of the superframe pixel data, one of the associated confidence map, one of the associated source catalogue. In the case of tiled observing modes: one of the mosaiced superframe, one of of the associated confidence map, one of the associated catalogues.

P1	FITS image and assoicated confidence map of each processed frame (one per array, N(arrays) per pointing; this applies in observing modes O1 to O3).
P2	FITS line-emission processed frame (in observing mode O8).
P3	Photometric calibration information and confidence maps for each processed frame
P4	Catalogue of objects detected in the processed frames (one per array; this applies in observing modes O1 to O3). Catalogued image properties are detailed elsewhere.
P5	For tiled observing modes, both the products detailed in P1–P4 as appropriate, plus catalogue and confidence maps for the completed tile will be supplied.

A.8 Catalogued Image Properties

I1	This table is either the Sloan list or the Cambridge list, or a merged/truncated version of both. For the summit pipeline the APM set would be adequate; users requiring Sloan-style parameters can wait for the data to be processed in the UK.
I2	Realistic error estimates should accompany all parameters for which this is possible

B Further Appendices

B.1 Expected sky count rates

At the moment the 'standard' integration time for WFCAM will be 10 seconds (in reality this may be the combination of several short sub-integrations using *e.g.* non-destructive reads). Table B.1 below compares expected count rates in the various filter passbands as a demonstration of the amount of time needed to achieve a 0.3% *rms* flatfielding accuracy.

Band	UFTI (0.091 ″ /pix) (photons/pix/s)	WFCAM (0.4 ″ /pix) (photons/pix/s)	10k photons/pix (time s)	0.3% photon noise (time s)
Y	???	???	???	???
J	49	950	11	110
H	280	5400	2	20
K	210	4100	2.5	25

Assuming similar efficiencies, the gain of UFTI is ≈7, so the final column gives the total integration time to get a 0.3% nominal self-calibrating pixel-to-pixel flatfielding accuracy. The current UKIDSS strategy for the shallowest surveys is for an integration time of 10 seconds, t_{total} = 40 seconds. (In the UKIDSS proposal, this is designated as t_{stare} = 10s, t_{pos} = 40s.) For the purposes of flatfielding, each detector pixel will see sky for 40s per pointing (assuming microstepping includes an integer pixel shift to move bright stars an amount larger than the PSF) ×4 pointings per tile - which is a total of 160 seconds. This is generally enough to flatfield J, H and K bands to 0.3%, if we wish to use the night sky observations for the flatfield. It is also possible to interlink different targets for shallow surveys within an MSB to meet self-calibration requirements. We have insufficient data to make the same calculation for Y.

B.2 Basic Processing Steps Revisited

- instrumental signature removal – bias, non-linearity, dark, flat, fringe
- sky background tracking and removal during stacking – possible need to also remove other 2D background variations from imperfect multi-sector operation of detectors
- define and produce a strategy for dealing with image retention from preceding exposures
- combine frames if part of an observed dither sequence or tile pattern
- consistent internal calibration to put observations on an approximately uniform system

- basic catalogue generation including astrometric, photometric, shape and DQC information
- final astrometric calibration from the catalogue with an appropriate and agreed World Coordinate System (WCS) in all FITS headers
- basic photometric calibration from catalogue using suitable pre-selected standard areas covering entire field-of-view to monitor and control systematics
- each frame and catalogue supplied with provisional calibration information and overall morphological classification embedded in FITS files
- propagation of error arrays and use of confidence maps
- realistic errors on selected derived parameters
- nightly extinction measurements in relevant passbands (also possibly JCMT information on water vapour content since can alter apparent extinction/ effective bandpass more than airmass)
- pipeline software version control – version used recorded in FITS header

B.3 Glossary of Terms

channel – data channel from each 2kx2k detector through to the backend of the pipeline

pointing – a telescope slew to position with new guide star acquisition. Note that microstep sequences and dither sequences are done within one pointing.

read – a read is the act of physically reading a detector. This generally means the act of digitising the data, and can be either destructive (*i.e.* followed by a reset) or non–destructive

exposure – an exposure is a sequence of one or more reads, that are used to produce a single output image

integration – an integration is the result of a sequence of one or more exposures. Whereas the exposure is the result of a sequence of reads, the integration is a summation (or similar) of a sequence of exposures

darks – reference exposures of the same duration as the images in an exposure sequence used for correcting detector artefacts

MSB – a Minimum Schedulable Block is a sequence of integrations that makes up a largely self-calibrating and useful scientific observation. The component parts of an MSB are defined below and in turn are constrained by the scientific requirements on flatfielding accuracy, astrometric and photometric calibration and sky removal.

mesostep sequence – a type of MSB being a sequence of exposures designed to completely sample across the face of a detector to monitor residual systematics in the photometry

focus run – a type of MSB with a series of exposures for use in providing an estimate of best focus

microstep sequence – a sequence of 1x1, 2x2 or 3x3 integrations taken with accurate sub–pixel (including small integral) shifts combined by interleaving the pixels

offset – a macrostep to another region of sky within an MSB sequence which may be either another target region or an offset region used for sky estimation observation – an observation is a sequence of integrations that makes up a usable scientific observation. An observation is the level at which calibration frames (flats, skies, darks *etc.*) are collected

reduced frame – result of reducing an integration

superframe – the result from combining (interleaving) the reduced frames from a microstep sequence. The superframe can be 2kx2k, 4kx4k or 6kx6k depending on microstep mode

loops – a contiguous sequence of multiple exposures in one place to increase integration time

dithers – sequence of integrations offset in "spiral" pattern for use in bad pixel rejection, deep stacking *etc.*

tile – the four pointings and observing sequences necessary such that a combination of them covers a contiguous area of sky involving all four detectors

coadd – image resulting from combining exposures taken sequentially during an integration

stack – image resulting from combining frames from a single detector, including combining superframes

mosaic tile – complete contiguous image combined from using exposures from all detectors

confidence map – a normalised inverse variance weight map defining the "confidence" associated with the intensity value in each pixel, this also encodes for hot pixels, bad pixels, dead pixels and so on. The raw confidence maps for each frame output by the Data Acquisition System (DAS) will be derived from regular (frequency TBD) analysis of the calibration flats and darks. All output processed frames (stacked, tiled, mosaiced *etc.*) will also have an associated derived confidence map.

B.4 Data Quality Monitoring

WFCAM data will come with the following DQC measures that will greatly enhance the usefulness of data products for the end user:

- Night–time observing conditions/weather
- Instrument status values: detectors/filters/controllers

Individual frames will have the above information written into their headers; the quick–look summit pipeline will then derive further DQC measures and insert into headers:

- sky brightness – saturated? Too bright? No signal?
- sky noise – detectors OK? Pattern noise?
- average stellar ellipticity – trailed images?
- average seeing measure – for each detector, in focus?
- stellar aperture correction – weird PSF?
- no. of spurious images – corrupted data?
- astrometric errors – wrong RA, Dec?
- first pass photometric calibration – system throughput?

All processing done by the UK–side pipeline will record progress in image FITS headers, possibly derive further DQC measures, and version and time stamp the processing stages.

B.5 Table of Proposed Catalogue Parameters

APM/SuperCOSMOS/INT WFC/CIRSI analysis produces 32 4-byte parameters per detected object. As an example the INT Wide Field Camera survey pipeline produces the following set (which has evolved from the original APM set):

No.	Name	Description
1	Seq. no.	Running number for ease of reference, in strict order of image detections
2	Isophotal flux	Standard definition of summed flux within detection isophote, apart from detection filter is used to define pixel connectivity and hence which pixels to include. This helps to reduce edge effects for all isophotally derived parameters.
3	Total flux	An attempt at generating automatically a total flux using a curve–of–growth technique. The closest description as to how it works is given in Hall & Mackay (1984). This is better than the isophotal flux in terms of systematic error, but still not perfect and definitely worse for random error.
4	Core flux	Best used if a single number is required to represent the flux for ALL objects. Basically aperture integration with radius rcore (in the FITS header) but modified to simultaneously fit 'cores' in case of overlapping images.
5	X co-ord	Intensity–weighted centre–of–gravity in X
6	Y co-ord	Intensity–weighted centre–of–gravity in Y
7	Gaussian sigma	These are derived from the intensity–weighted second moments. The
8	Ellipticity	equivalence between them and a generalised elliptical Gaussian distribution
9	Position angle	is used to derive Gaussian sigma = $(\sigma_a^2 + \sigma_b^2)^{1/2}$
		Ellipticity = $1.0 - \sigma_a/\sigma_b$
		Position angle = angle of ellipse major axis wrt x axis
10	Peak height	in counts relative to local value of sky
11	Areal profile 1	Number of pixels above a series of threshold levels relative to local sky.
12	Areal profile 2	Levels are set at T, 2T, 4T, 8T . . . 128T where T is the threshold. These
13	Areal profile 3	can be thought of as a sort of poor man's radial profile. Note that for
14	Areal profile 4	deblended, *i.e.* overlapping images, only the first areal profile is computed
15	Areal profile 5	and the rest are set to –1 flagging the difficulty of computing accurate
16	Areal profile 6	profiles.
17	Areal profile 7	
18	Areal profile 8	
19	Core 1 flux	A series of different radii core/aperture measures similar to parameter 4.
20	Core 2 flux	Together with parameter 10 these give a simple curve–of–growth analysis from
21	Core 3 flux	peak pixel, $1/2\times$ rcore, $\sqrt{2}\times$ rcore, $2\times$ rcore, $2\sqrt{2}\times$ rcore
22	Core 4 flux	
23	RA	RA and Dec explicitly put in columns for overlay programs that cannot, in
24	Dec	general, understand astrometric solution coefficients. Derived exactly from WCS in header and XY in parameters 5 & 6
25	Classification	Flag indicating what image is most probably: –1 stellar, +1 non–stellar, 0 noise
26	Statistic	An equivalent N(0,1) measure of how stellar–like an image is, used in deriving parameter 25 in a 'necessary but not sufficient' sense
27	Core 5 flux	$4\times$ rcore, extra to ensure $\sim 99\%$ of PSF flux
28	Sky level	Local interpolated sky level from background tracker
29	Sky variance	Local estimate of variation in sky level around image
30	Child/parent	Flag for parent or part of deblended deconstruct

31 Blank
32 Blank

This will be enhanced to include error estimates for a TBD subset of the parameters.

As an alternative the SDSS archive has a much larger set of parameters per object (see for example `http://www.sdss.jhu.edu/`) – many are error estimates on parameters analogous to the above. Others include different flux measures – eg. Petrosian magnitudes, and list–driven flux measures expressed in asinh magnitudes whereby flux estimates at the same position in all other wavebands are obtained given a significant detection in any one passband.

It is proposed that the parameters derived from WFCAM images in a single passband be similar to the above INT–WFC set with extra parameters (*e.g.* Petrosian fluxes, profile fitting) added where feasible, possibly in further processing stages, after liaison with the WFCAM user community.