Data Processing For VISTA

Jim Lewis, Mike Irwin

CASU

VISTA Pipelines

- Summit Pipeline for rapid QC assessment
- ‘Garching’ pipeline for QC monitoring of the instrument and to provide calibration frames for summit pipeline.
- Both of the above are CPL pipelines with the same code base
- Science pipeline at CASU for creation of final data products.
  - Written in CASU’s own C/Perl/CFITSIO infrastructure

Data Organisation at CASU

Data Organisation at CASU

- Discs arrive from ESO roughly weekly

Data Organisation at CASU

- Discs arrive from ESO roughly weekly
- Plugged into a Linux box and copied to staging area

Data Organisation at CASU

- Discs arrive from ESO roughly weekly
- Plugged into a Linux box and copied to staging area
- Check for missing images

Data Organisation at CASU

- Discs arrive from ESO roughly weekly
- Plugged into a Linux box and copied to staging area
- Check for missing images
- Detect and correct missing header items

Data Organisation at CASU

- Discs arrive from ESO roughly weekly
- Plugged into a Linux box and copied to staging area
- Check for missing images
- Detect and correct missing header items
- Rename files (eg. v20091125_00001.fit)

Data Organisation at CASU

- Discs arrive from ESO roughly weekly
- Plugged into a Linux box and copied to staging area
- Check for missing images
- Detect and correct missing header items
- Rename files (eg. v20091125_00001.fit)
- Create index files (global & for each obs type)

Data Organisation at CASU

- Discs arrive from ESO roughly weekly
- Plugged into a Linux box and copied to staging area
- Check for missing images
- Detect and correct missing header items
- Rename files (eg. v20091125_00001.fit)
- Create index files (global & for each obs type)
- Ingest header info into raw data archive

Data Organisation at CASU

- Discs arrive from ESO roughly weekly
- Plugged into a Linux box and copied to staging area
- Check for missing images
- Detect and correct missing header items
- Rename files (eg. v20091125_00001.fit)
- Create index files (global & for each obs type)
- Ingest header info into raw data archive
- Shift data to final raw archive location

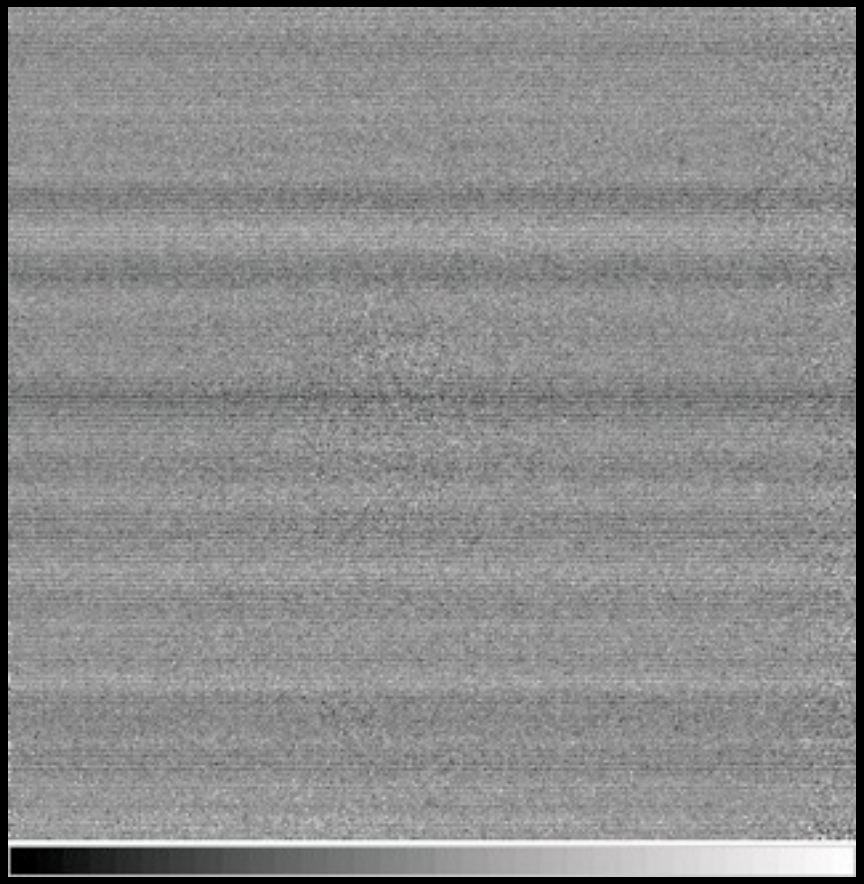
Data Organisation at CASU

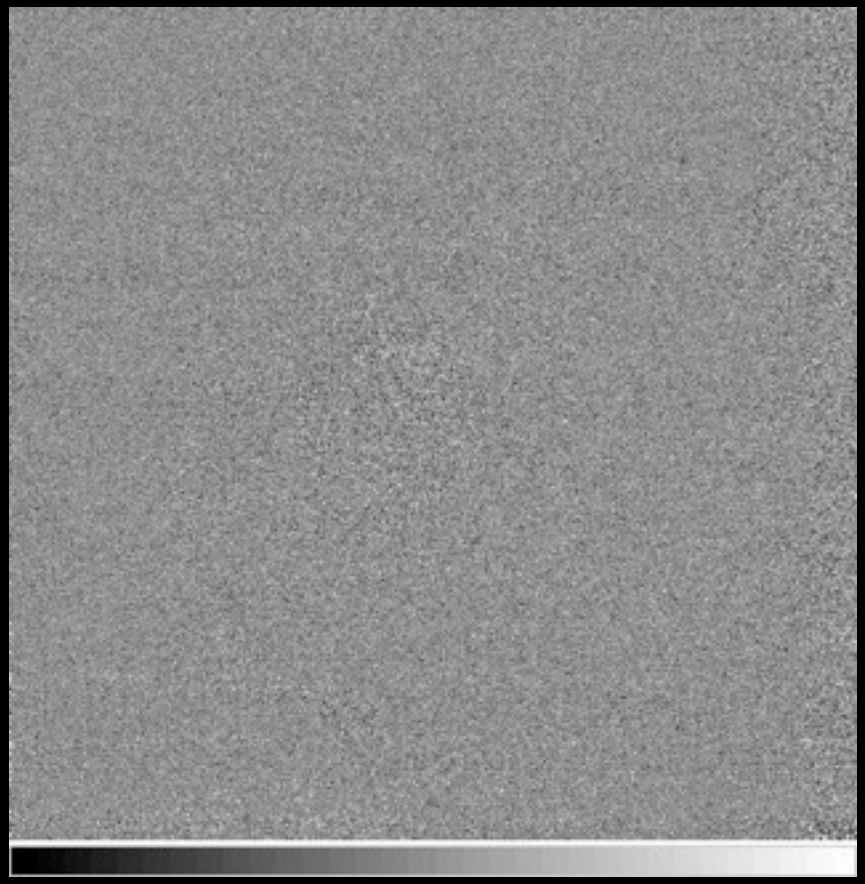
- Discs arrive from ESO roughly weekly
- Plugged into a Linux box and copied to staging area
- Check for missing images
- Detect and correct missing header items
- Rename files (eg. v20091125_00001.fit)
- Create index files (global & for each obs type)
- Ingest header info into raw data archive
- Shift data to final raw archive location
- Process typically takes about 3-4 hours per disc.

Processing Steps

- Reset correction (debias)
- Dark correction
- Linearity correction
- Flat field correction
- Sky background correction ******
- Destripe

VIRCAM Stripes





Not Processing Steps

- Detector crosstalk removal
- Persistence
- Defringing

‘Group’ Processing Steps

- Interleaving (?)
- Dithering/Jittering
- Catalogue generation
- Astrometric calibration
- Photometric zeropoint calibration
- Tiling

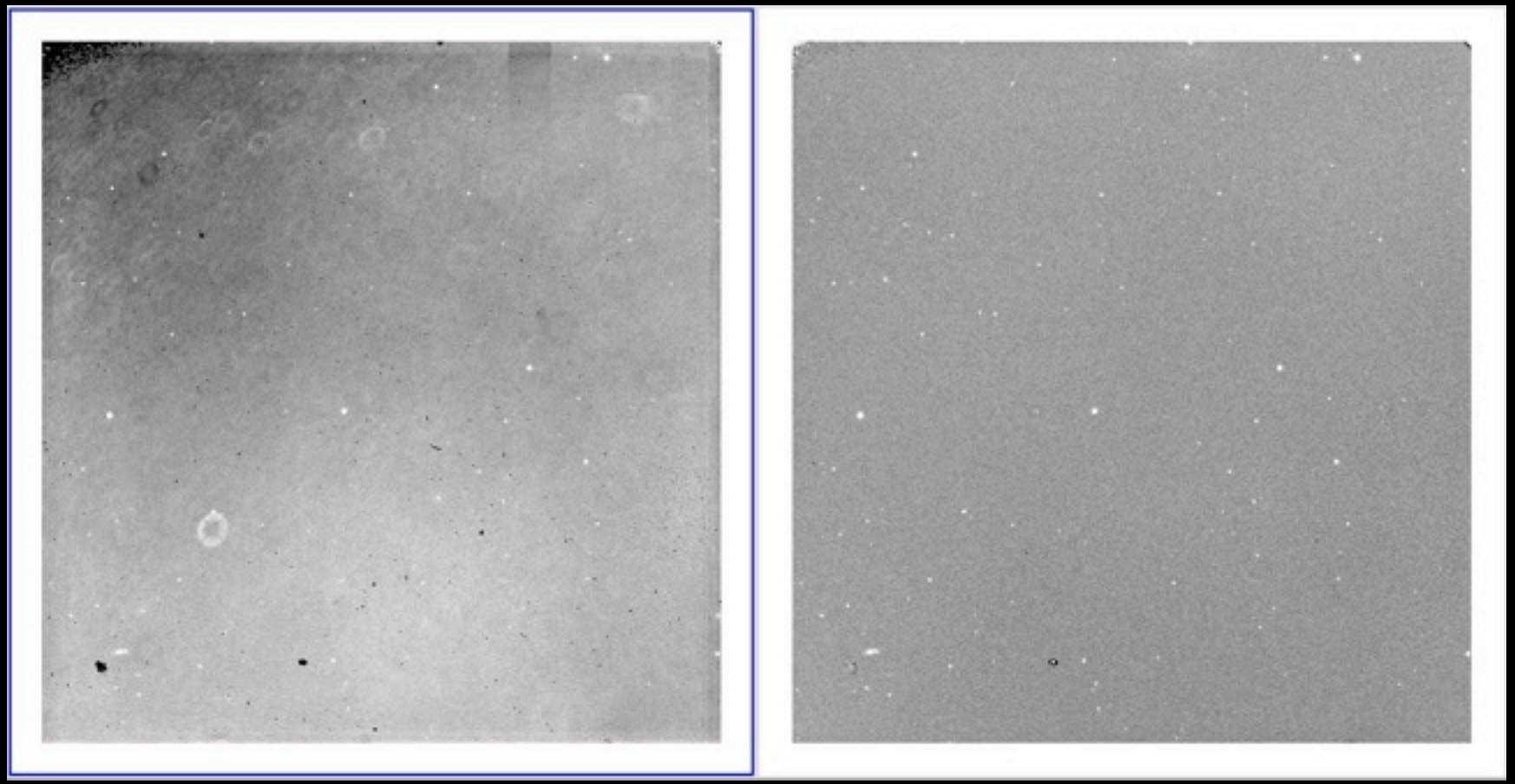
Data Products

- Products consist of:
  - Calibrated single exposures
  - Interleaved superframes + confidence maps
  - Shift/averaged stack frames (pawprints) + confidence maps
  - Filled area tiles + confidence maps
  - Object catalogues for pawprints and tiles
- All products are in MEF files
- All QC parameters are stored in MEF headers
- All image products are floating point and will be eventually compressed.

IR Data Reduction Worries

- IR detectors are currently inherently more unstable than optical CCDs.
  - Some odd electronic effects
  - Not as good cosmetically as CCDs
- Sky emission > 100x brighter than most objects of interest
  - And it's variable both spatially and temporally!
- Exposure times are short, so nightly data rates are very high.
  - ~250 Gb/night
  - Rice tile compression can save factors of 3-4 in 32 bit integer data. It's lossless and very fast.
  - "Bookkeeping" can be a NIGHTMARE!
  - Experience with WFCAM on UKIRT

Before And After Background Correction



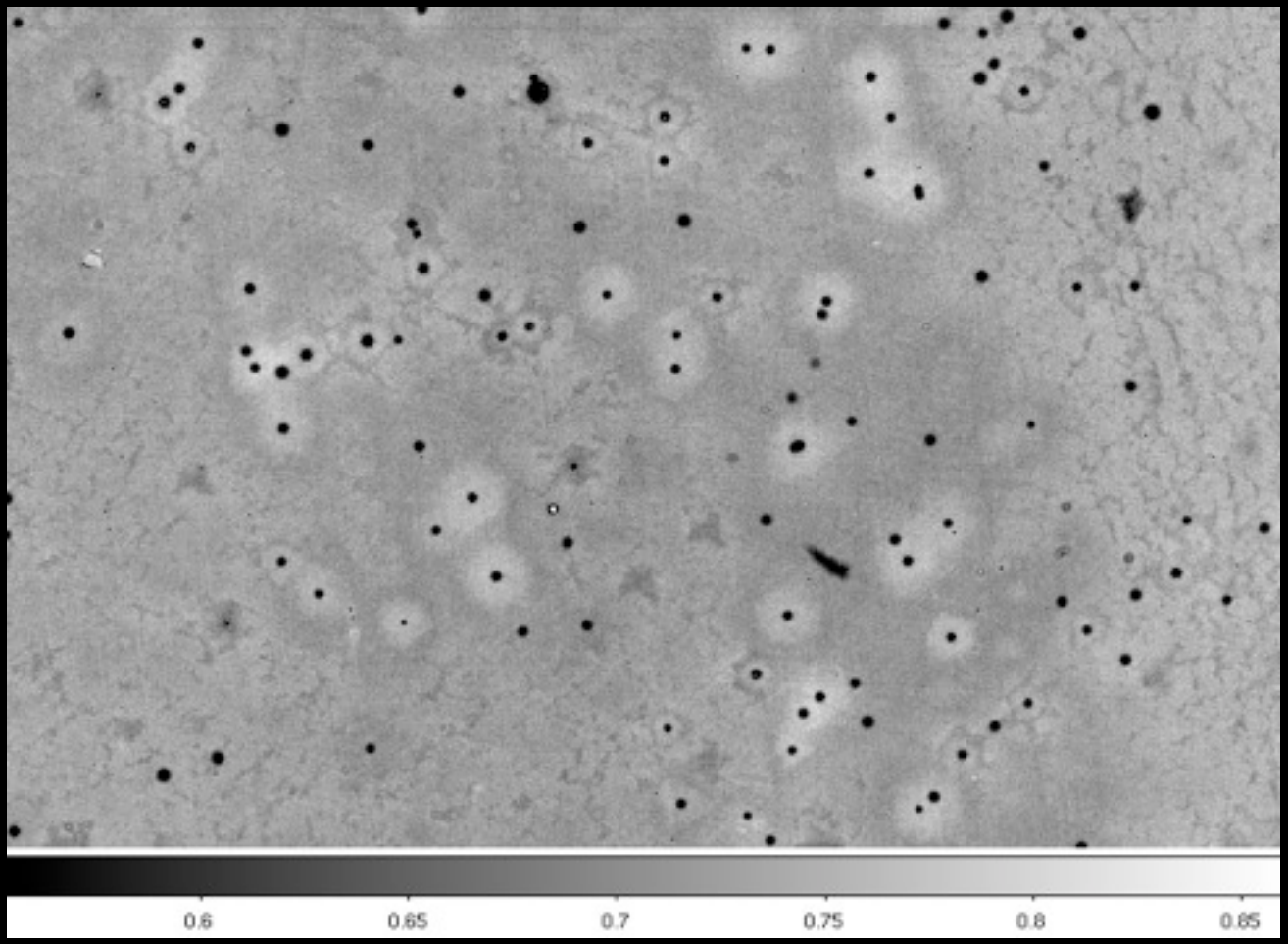
Jittering Schemes

- A standard filled tile consists of 6 ‘pawprints’
- PJ jitter – all the jittered exposures for a given pawprint are completed before moving to the next pawprint.
- JP jitter – an exposure is done in each pawprint before moving to the next jitter position

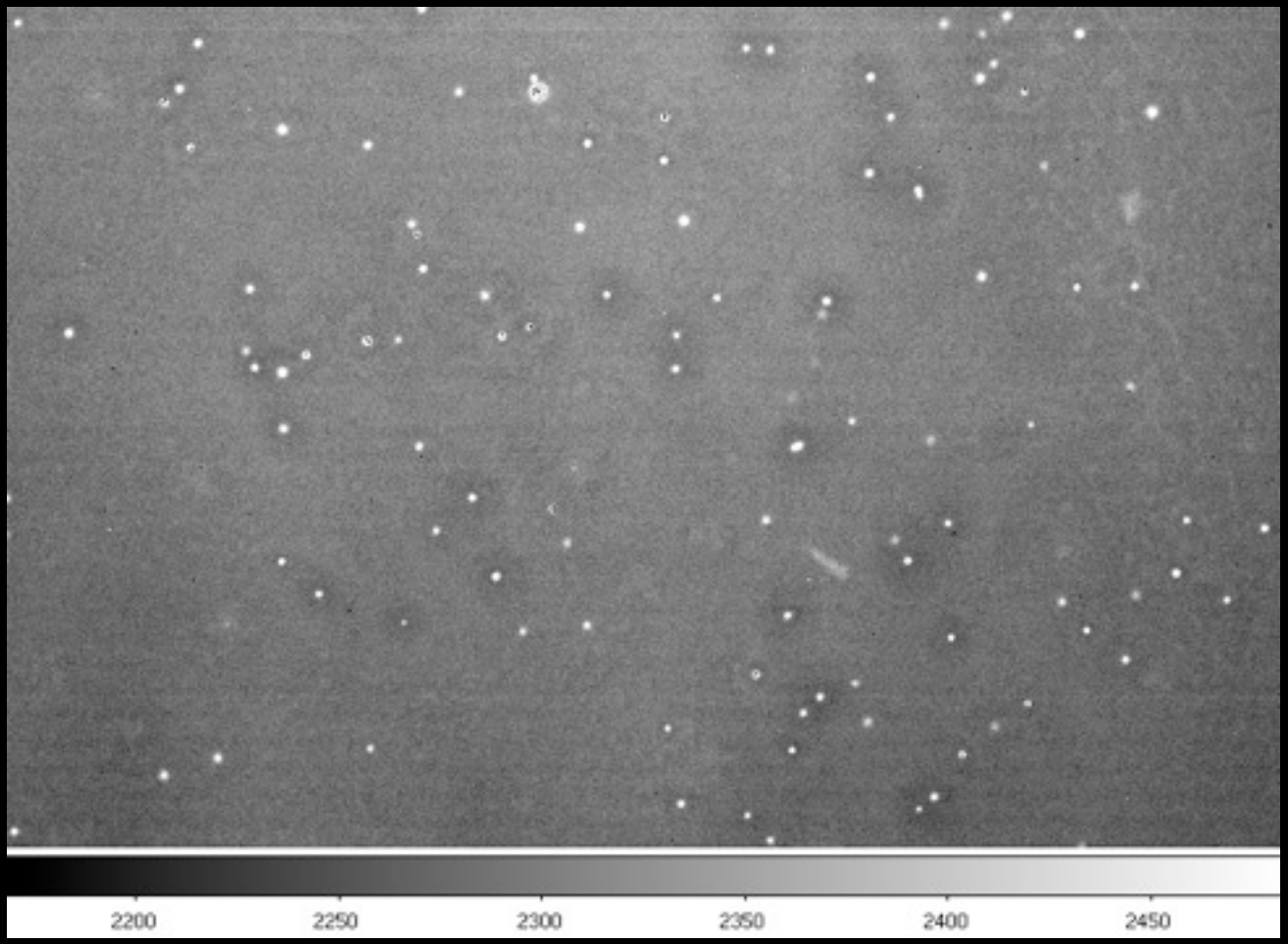
Available Background Subtraction Algorithms

- Tilesky – 2 pass combination of all observations in a tile
- Pawsky – Single pass combination of all observations in a pawprint with object masking done ‘dynamically’
- Pawsky with standard mask – as above, but the mask is given beforehand
- Offset sky – Use a sky taken elsewhere

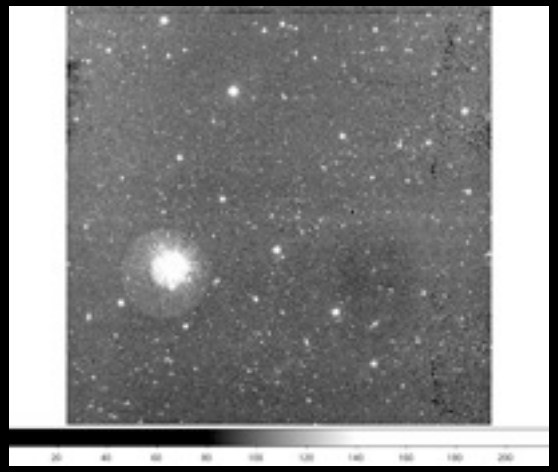
Flat Field Holes

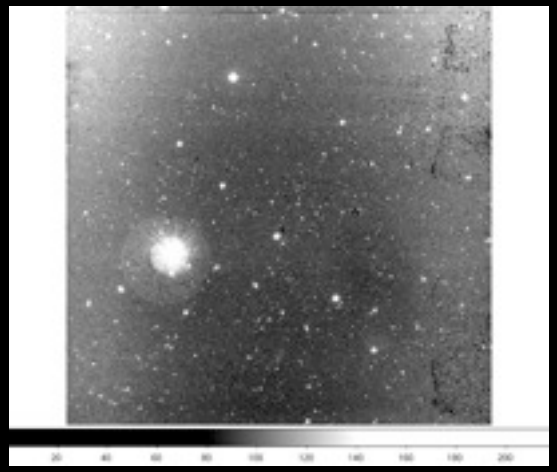


Sky Frame Dots (Stars? No!)



Time Variable Sky





General Detector Properties

General Detector Properties

- The VIRCAM detectors are not linear.

General Detector Properties

- The VIRCAM detectors are not linear.
- The VIRCAM detectors are not 16 bit.

General Detector Properties

- The VIRCAM detectors are not linear.
- The VIRCAM detectors are not 16 bit.

Detector	Linearity (%)	Saturation (ADU)
1	2.33	33000
2	3.32	32000
3	3.79	33000
4	3.50	32000
5	1.98	24000
6	2.98	28000
7	1.99	35000
8	3.38	33000
9	3.31	35000
10	4.44	35000
11	4.64	37000
12	2.55	34000
13	9.99	33000
14	2.72	35000
15	1.74	34000
16	3.28	34000

DIT times

- The sky saturates very quickly in H in particular
- 15s H exposures => background ~ 15-20K very often
- Leaves very small dynamic range for objects of interest
- Break up long exposures
  - e.g. 45s total should be observed as (DIT=9, NDIT=5) or (DIT=5, NDIT=9) and not DIT=15, NDIT=3

Offset Sky Exposures (how)

- For each object pawprint there should be an offset sky pawprint – maybe.
- DIT, NDIT, NJITTER, Filter should match object pawprint
- Unguided
- Large jitter offsets (~ half a detector width)

Offset Sky Exposures (where)

- Depends on what you're observing.
- Observers must look at offset sky regions in some visualisation tool to make sure this doesn't happen.
- Large jitter offsets can save us.