HSCデータ解析パイプライン

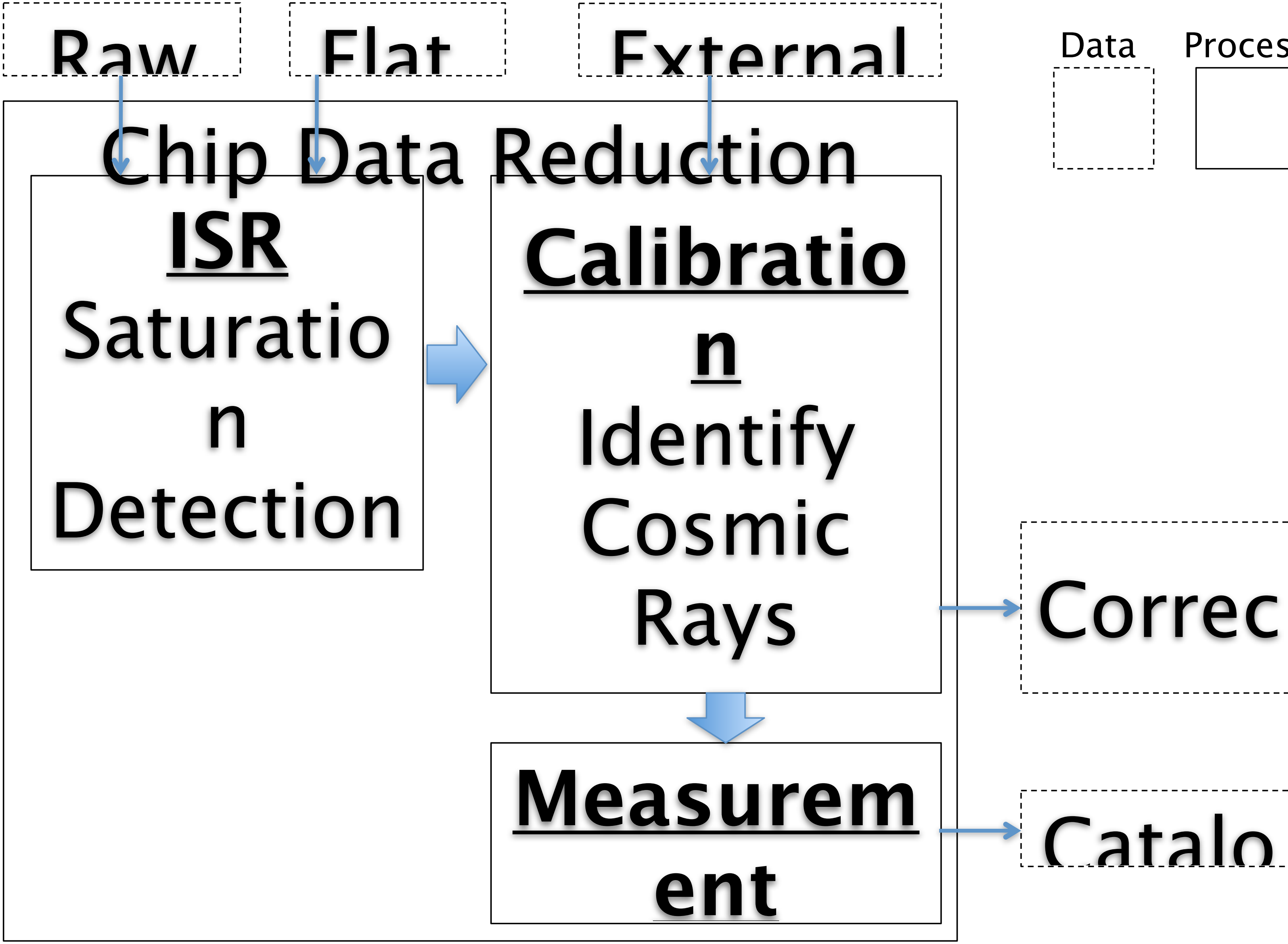
安田直樹(Kavli IPMU)

HSC Data Analysis Pipeline

- Being developed in the collaboration with Princeton University led by Robert Lupton who is the author of SDSS photometric pipeline.
  - PU: RHL, Craig Loomis, Paul Price, Jim Bosch, H. Miyatake
  - NAOJ: H. Furusawa, T. Takata, Y. Okura, H. Yamanoi, Y. Yamada, M. Koike
  - Kavli IPMU: S. Mineo, N. Katayama, T. Saito, Steve Bickerton, N. Yasuda
- Framework is based on LSST pipeline written in C++ and Python.
- HSC specific tools and functions are developed as add-ons.
  - Detailed information output to support online analysis and health check of instrument
  - Optical distortion, astrometry, flux calibration

Major Concerns

- Large amount of data
  - 10 times volume compared to Suprime-Cam
  - Processing
    - Multi-Process using multi-core CPUs and multi-node cluster system
  - Distribution
    - SDSS model
    - Individual FITS files and Catalog Database
- Wide field and large distortion
  - Inevitably suffered from scattered light
  - ~1000 pixel displacement due to distortion at the edge of FoV


Raw
Flat
External
Data
Process
Chip Data Reduction
ISR
Saturation Detection
Calibration
Identify Cosmic Rays
Correc
Measurement
Catalo

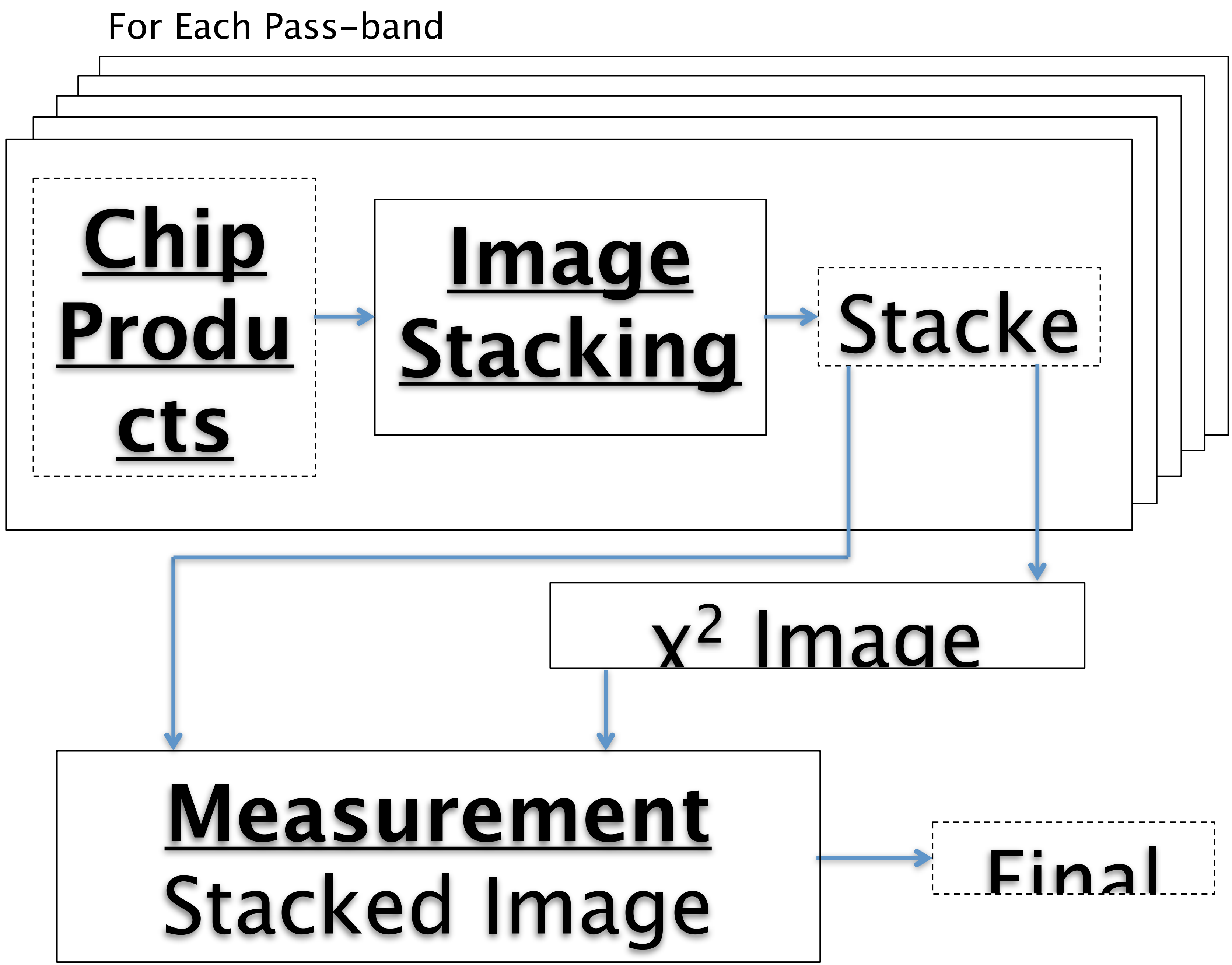
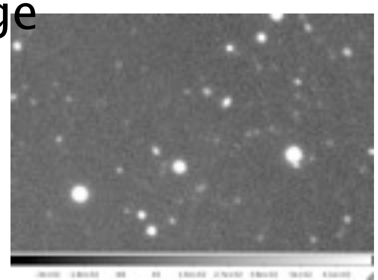

For Each Pass-band
Chip Products
Image Stacking
Stacke
χ^2 Image
Measurement
Stacked Image
Final

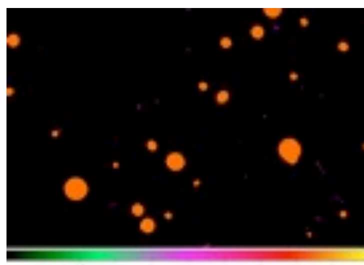
Image Format

HDU0: Header only
HDU1: Science Image



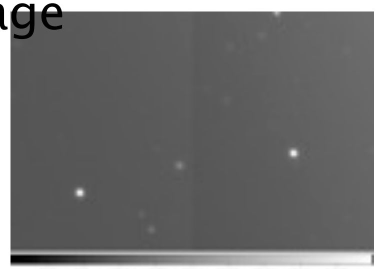
Calibrated astrometry and photometry

HDU2: Mask Image

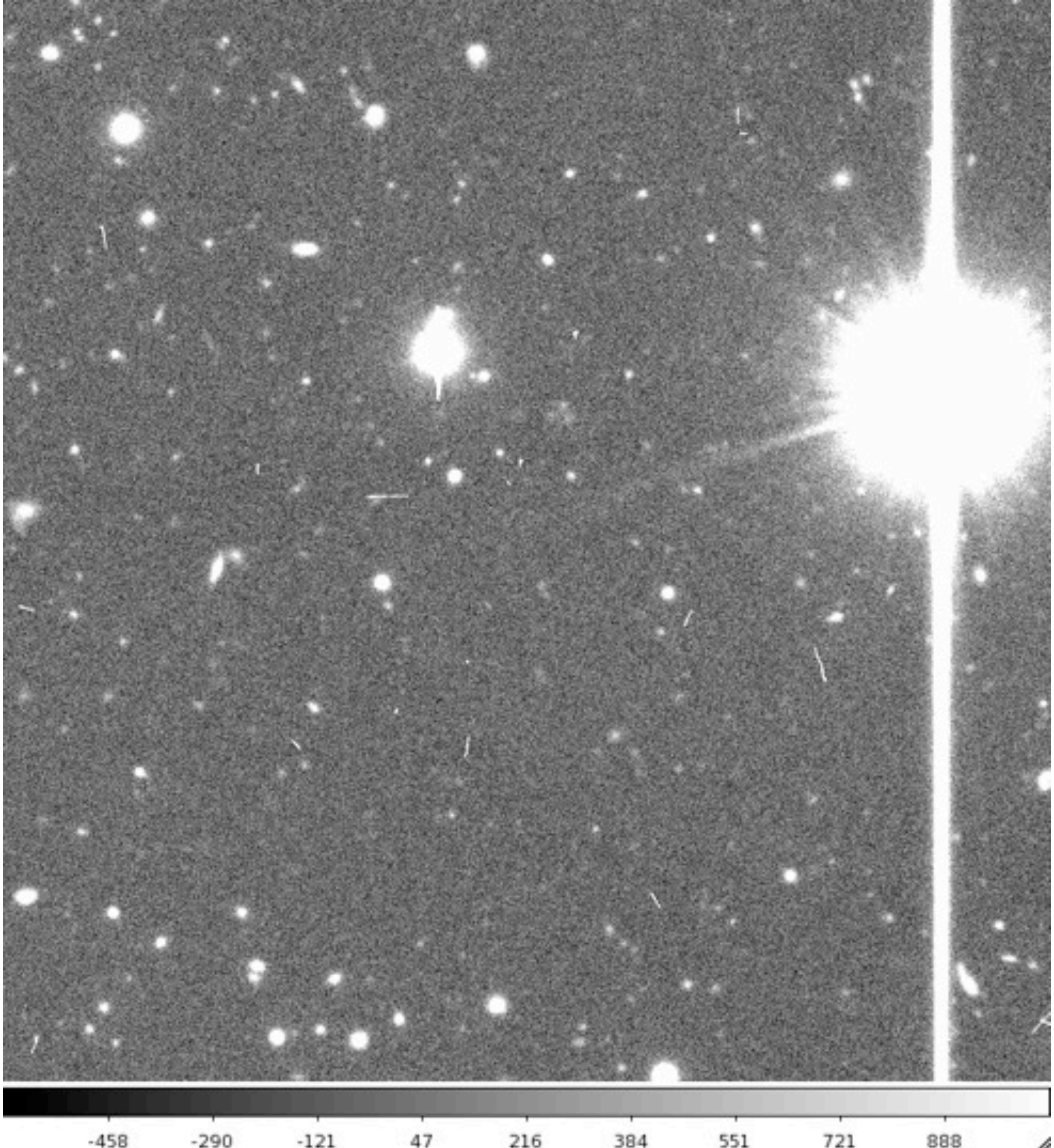


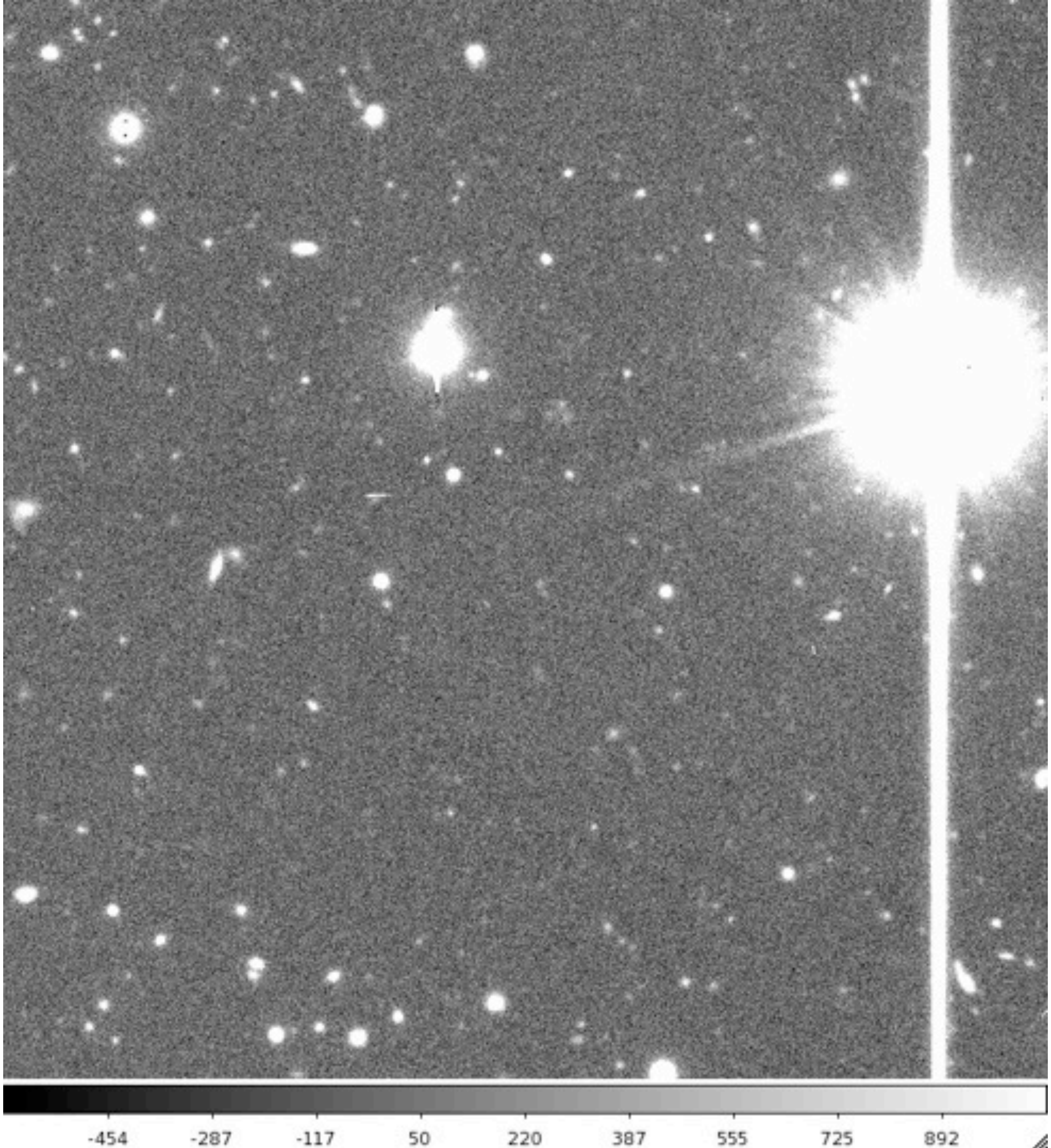
Plane 0 -> BAD
Plane 1 -> SAT
Plane 2 -> INTRP
Plane 3 -> CR
Plane 4 -> EDGE
Plane 5 -> DETECTED
Plane 6 -> DETECTED_NEGATIVE
Plane 7 -> CROSSTALK

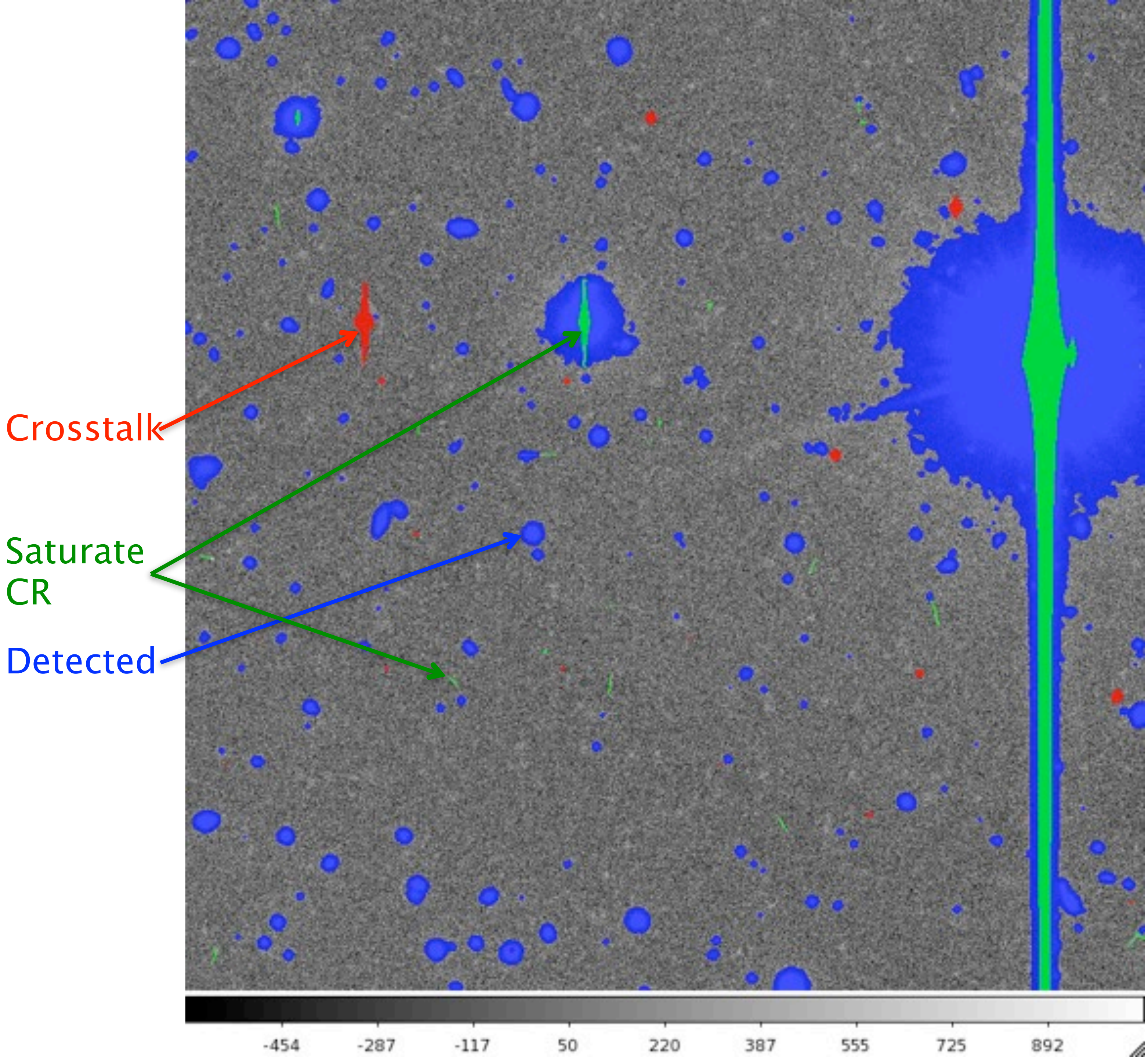
HDU3: Variance Image



ADU count / Gain


-458
-290
-121
47
216
384
551
721
888


-454
-287
-117
50
220
387
555
725
892


Crosstalk
Saturate
CR
Detected
-454
-287
-117
50
220
387
555
725
892

Catalog Parameters

Band Observed band
MJD Mean time during exposure

* Position *
xc, yc center of objects in pixel
ra, dec Equatorial coordinates
gLong, gLat, eLong, eLat, … ?

* Flag *
flag object flag

* Flux *
psfMag PSF fitted magnitude
apMag Aperture magnitude
kronMag Kron magnitude
modelMag Model fitted magnitude
extinction Galactic extinction

* Shape *
Stokes Q and U 2nd order moment
Other estimator for weak lensing

* Profile *
n Sersic index
reff Half light radius
AB a/b axis ratio
Phi Position Angle
BT Bulge total ratio

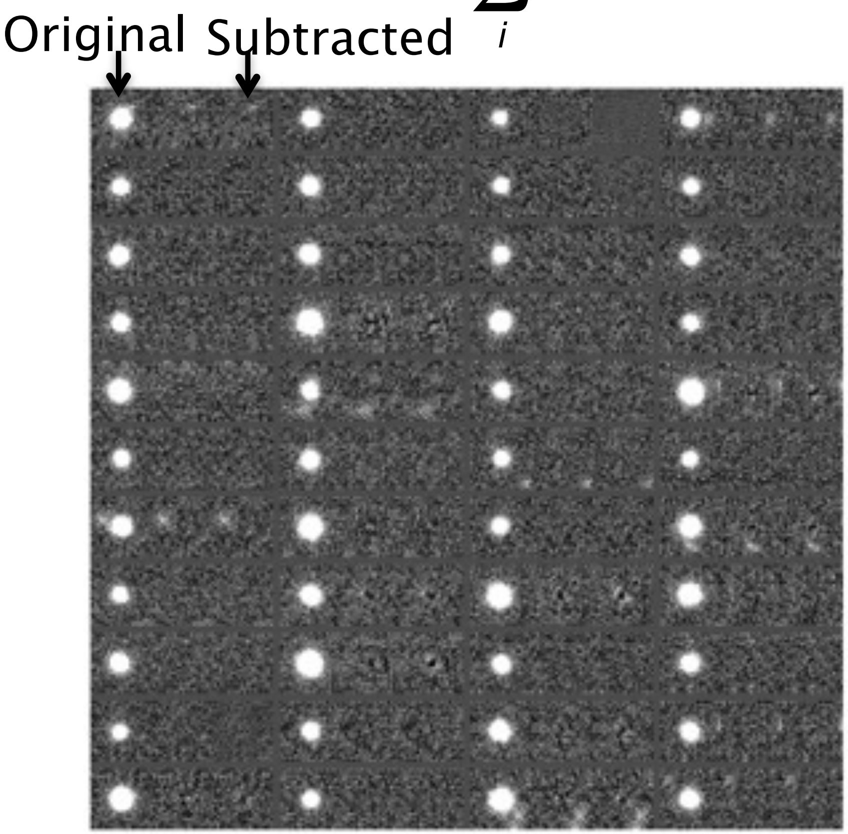
* Size *
Some dimension indicator

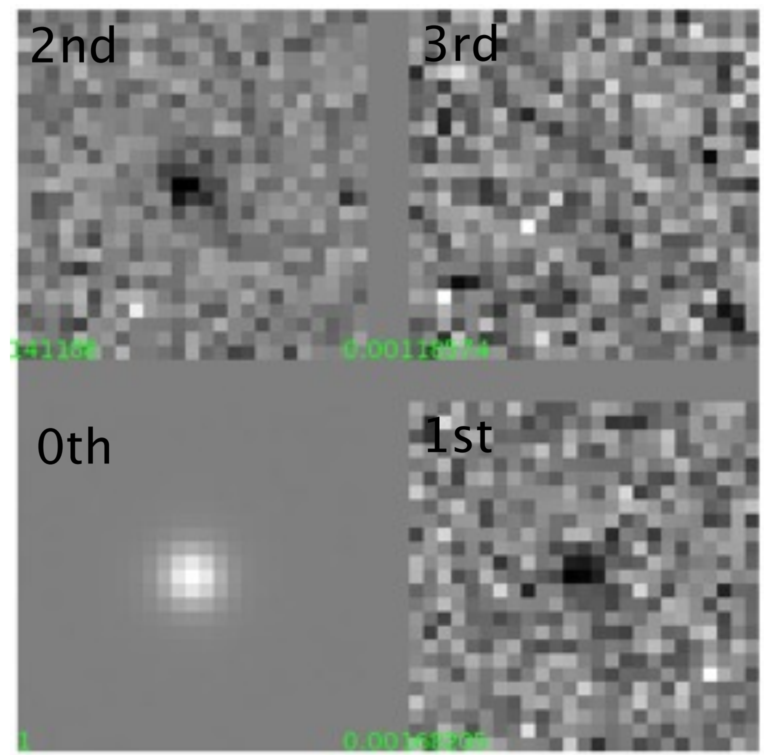
* Misc *
type Shape based star-galaxy
photo-z?

PSF measurement

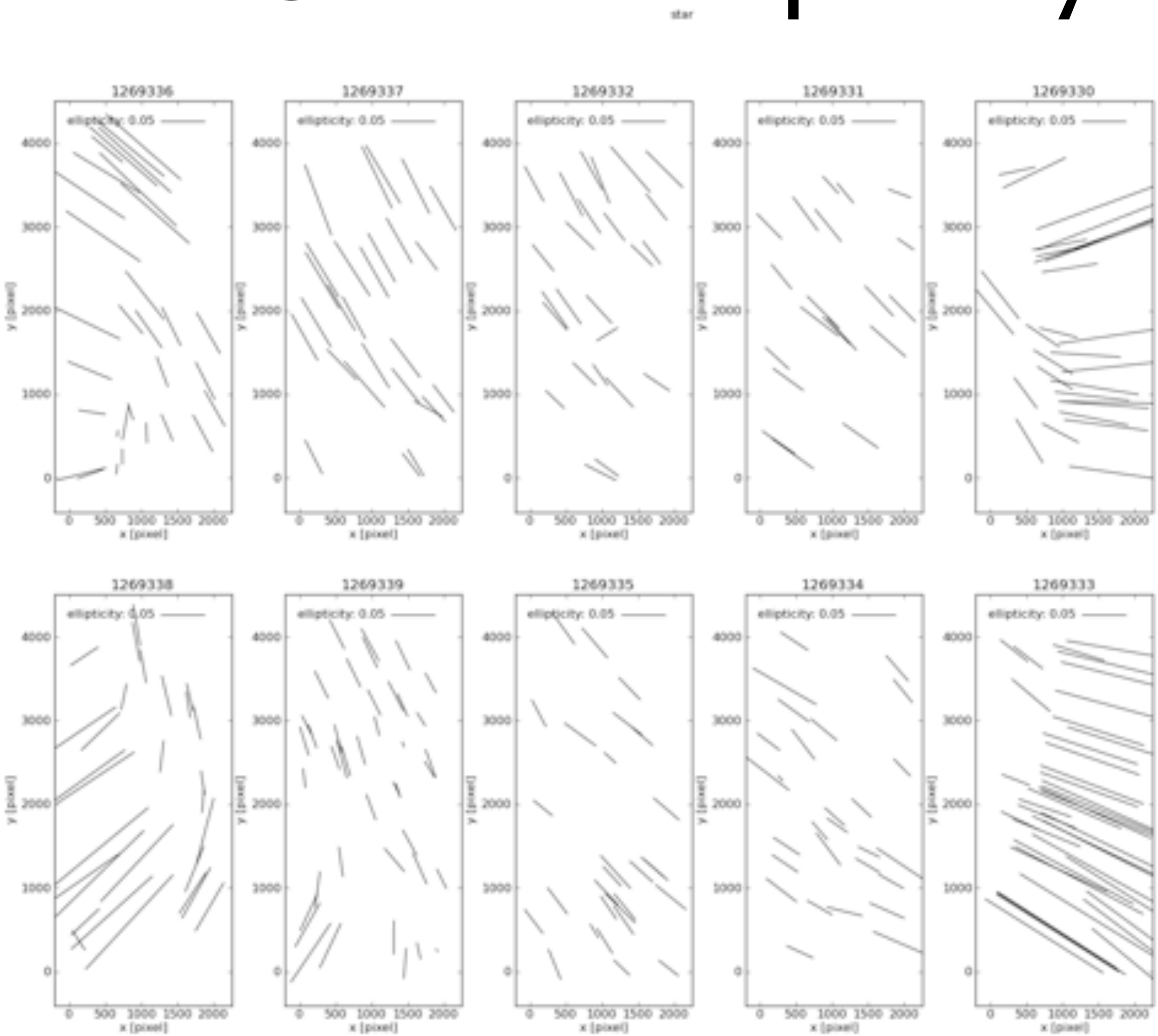
- KL expansion of the PSF

$$P(u,v) = \sum_i a_i(x,y) B_i(u,v) \qquad a_i(x,y) = \sum_{l=m=0}^{l+m\le N} b_i^{lm} x^l y^m$$

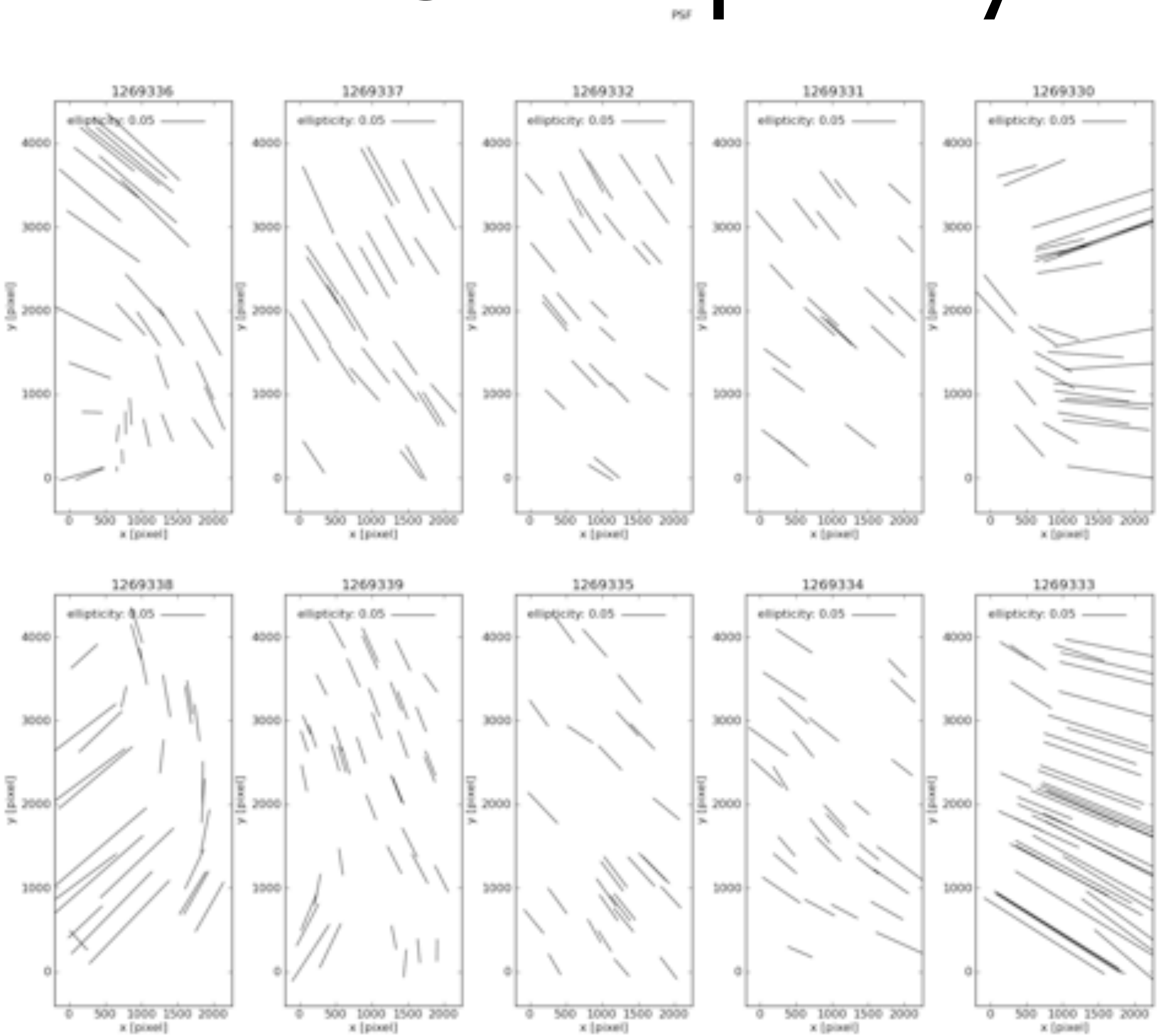




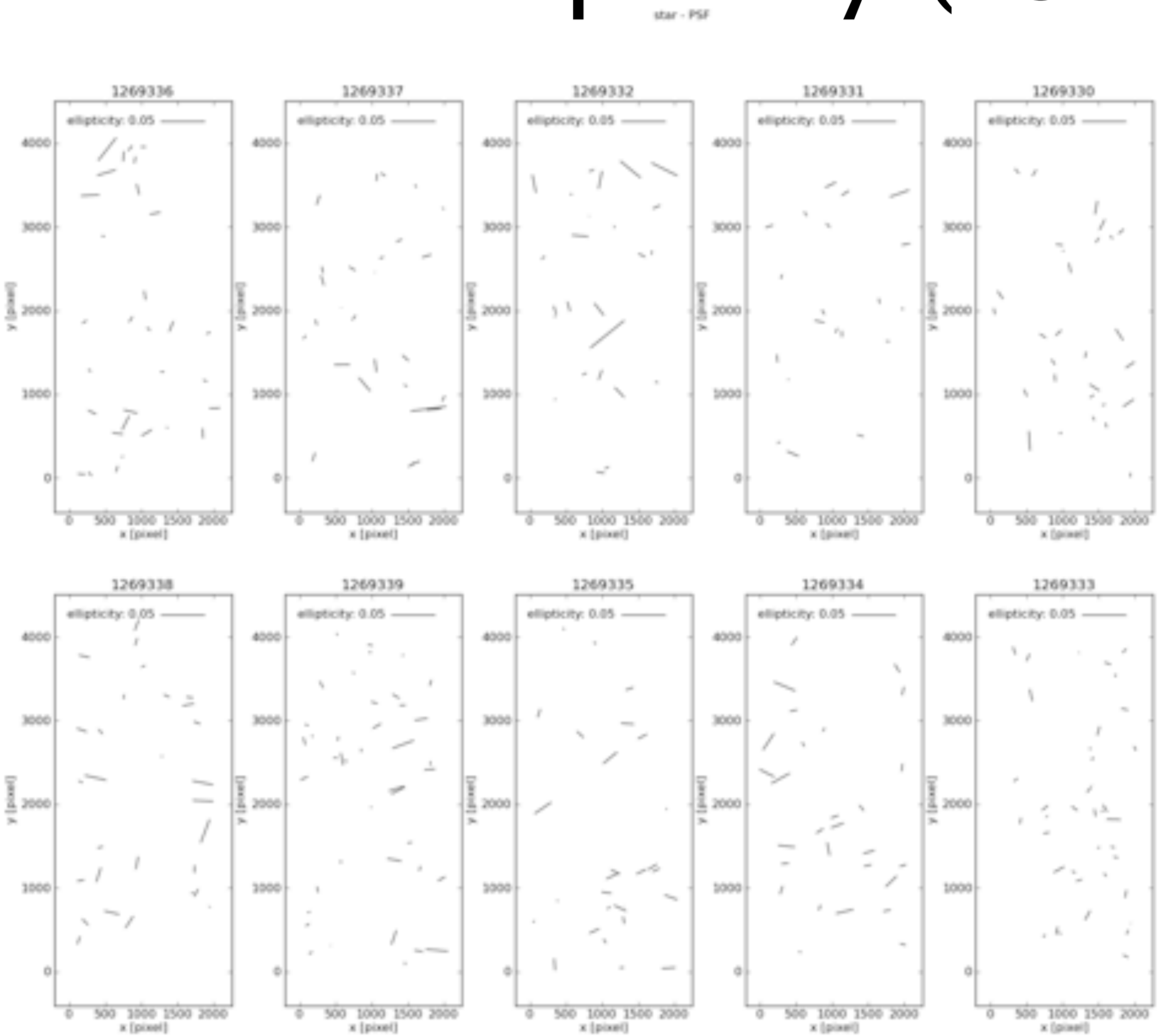
Stellar Ellipticity



PSF Ellipticity



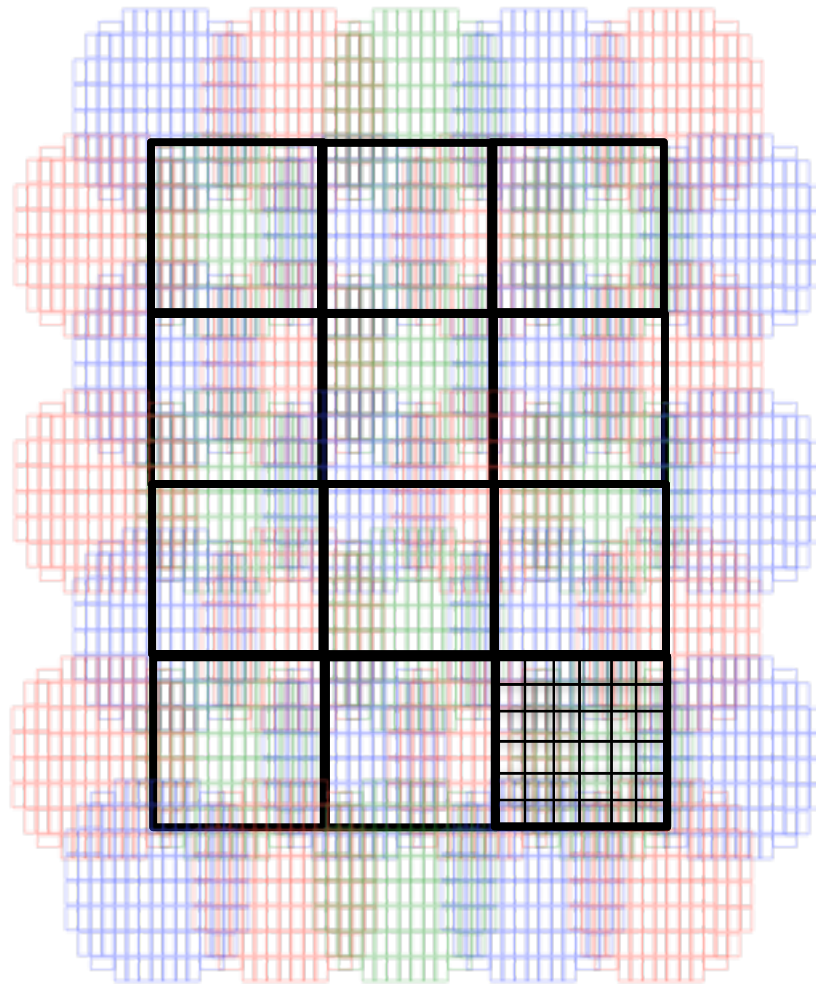
Residual Ellipticity (PSF–Star)



Mosaicking / Stacking

- Wide layer : pre-define skyTiles with the size of ~1.5 deg
- Mosaicking / stacking based on skyTiles
- Each skyTiles will have each WCS
- Every exposures will be warped to this skyTile and stacked
- Each skyTile will be divide into smaller sub-images sharing the same WCS

- Deep/Ultra-Deep layers : mosaicking / stacking based



Solving Distortion and

Pixel coordinates

$$\begin{pmatrix} u^{s,e} \\ v^{s,e} \end{pmatrix} = \begin{pmatrix} \cos\boxed{\theta_c} & -\sin\theta_c \\ \sin\theta_c & \cos\theta_c \end{pmatrix} \left(\boxed{\begin{matrix} x^{s,e} \\ y^{s,e} \end{matrix}}\right) + \left(\boxed{\begin{matrix} X_c \\ Y_c \end{matrix}}\right) = \begin{pmatrix} x^s\cos\theta_c - y^s\sin\theta_c + X_c \\ x^s\sin\theta_c + y^s\cos\theta_c + Y_c \end{pmatrix}$$

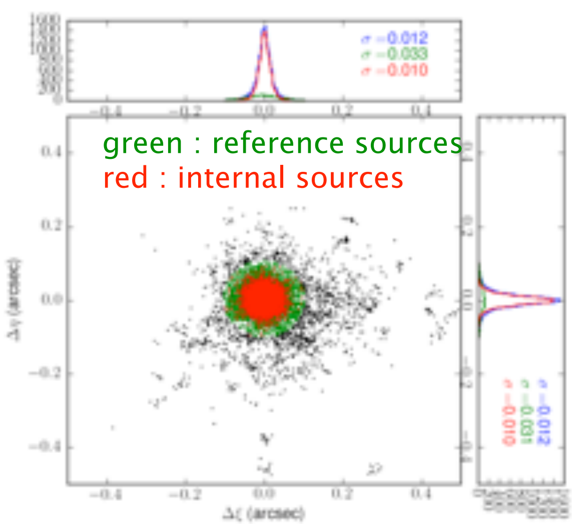
Intermediate (Projected) coordinates

$$\begin{pmatrix} \xi^{s,e} \\ \eta^{s,e} \end{pmatrix} = \begin{pmatrix} \xi(\alpha^s, \delta^s, A^e, D^e) \\ \eta(\alpha^s, \delta^s, A^e, D^e) \end{pmatrix}$$

□ : parameters to be determined
□ : observed values

Minimize the difference using multiple objects on multiple exposures

$$\chi^2 = \sum_e \sum_s \left\{ \xi^{s,e} - \sum_k \boxed{a_k} (u^{s,e})^{i(k)} (v^{s,e})^{j(k)} \right\}^2 + \sum \left\{ \eta^{s,e} - \sum_k \boxed{b_k} (u^{s,e})^{i(k)} (v^{s,e})^{j(k)} \right\}^2$$



From the processing of Suprime-Cam data
we can achieve ~10mas (relative) and ~30mas
(absolute)
accuracy in astrometry
We can expand this to entire survey region
--> ubercalibration of astrometry

Flux Calibration – model for “stellar flat”

$$\boxed{m_0^{star}} = \boxed{m_{obs}^{star,exp,chip}} + \boxed{dm^{exp}} + \boxed{dm^{chip}} + dm(u^{star}, v^{star})$$

Correction over focal plane

$$dm(u,v) = \sum_{i+j \le n} \boxed{f_{i,j}} \times u^i \times v^j$$

May need to include the effect of rot

Offset between chips
Basically 0.0 if correct flat is used

Offset between exposures
Mainly due to atmospheric extinction

Observed instrumental magnitudes

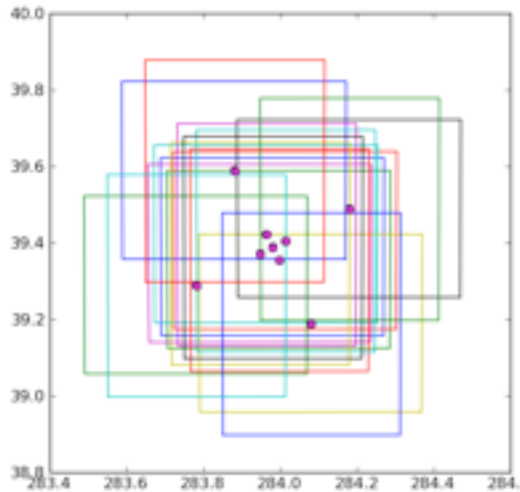
True magnitudes of stars

Determine m_0^{star}, dm^{exp}, dm^{chip}, $f_{i,j}$ by requiring the same star has the same m_0^{star}

Overall zeropoint will be determined by the comparison with reference catalog

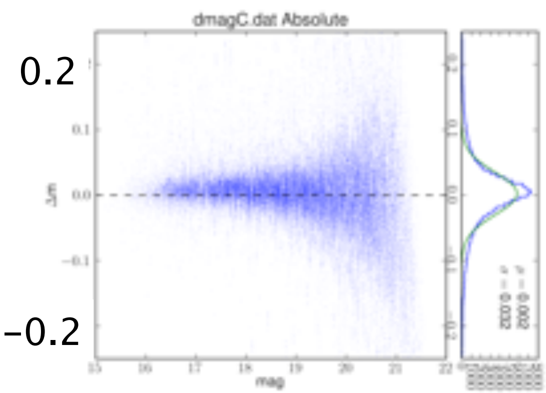
Flux Calibration – An Experiment

30sec x 18 exposures
at stellar dense field
(~600stars / chip)

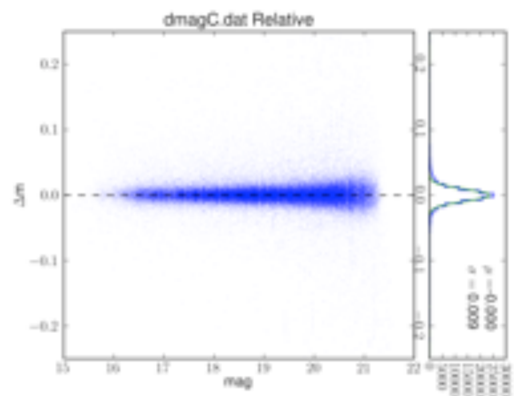


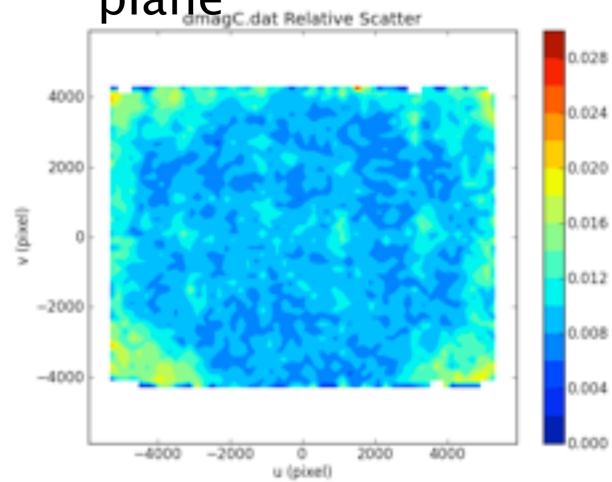
RMS of relative comparison over focal plane

External σ~0.03mag

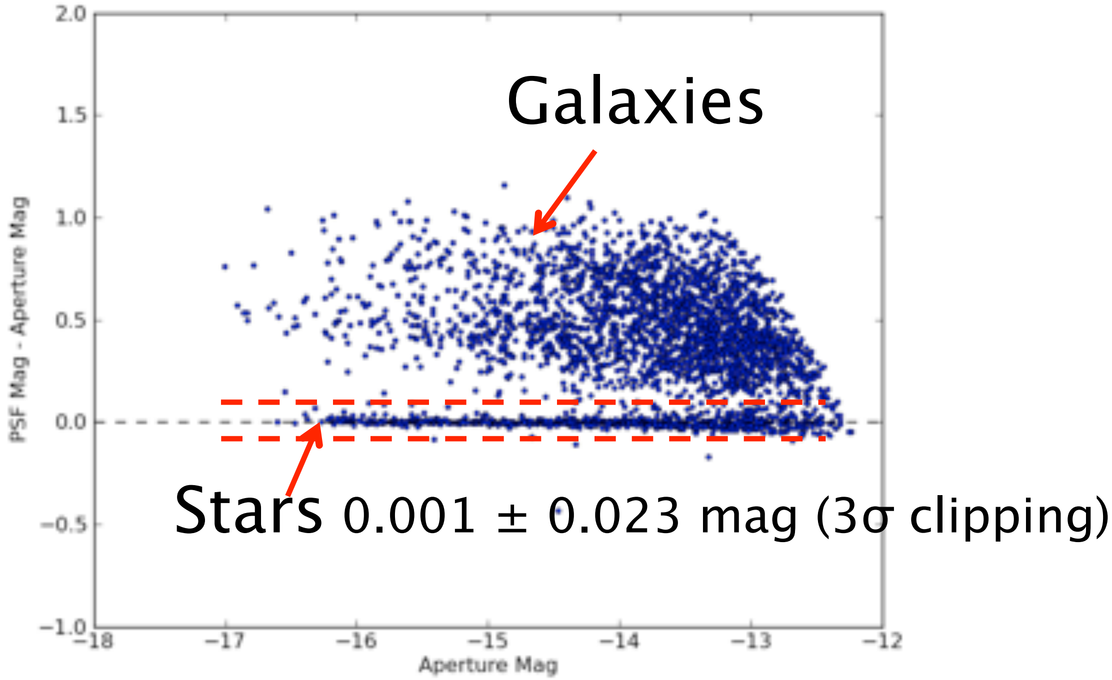


Internal σ~0.01mag



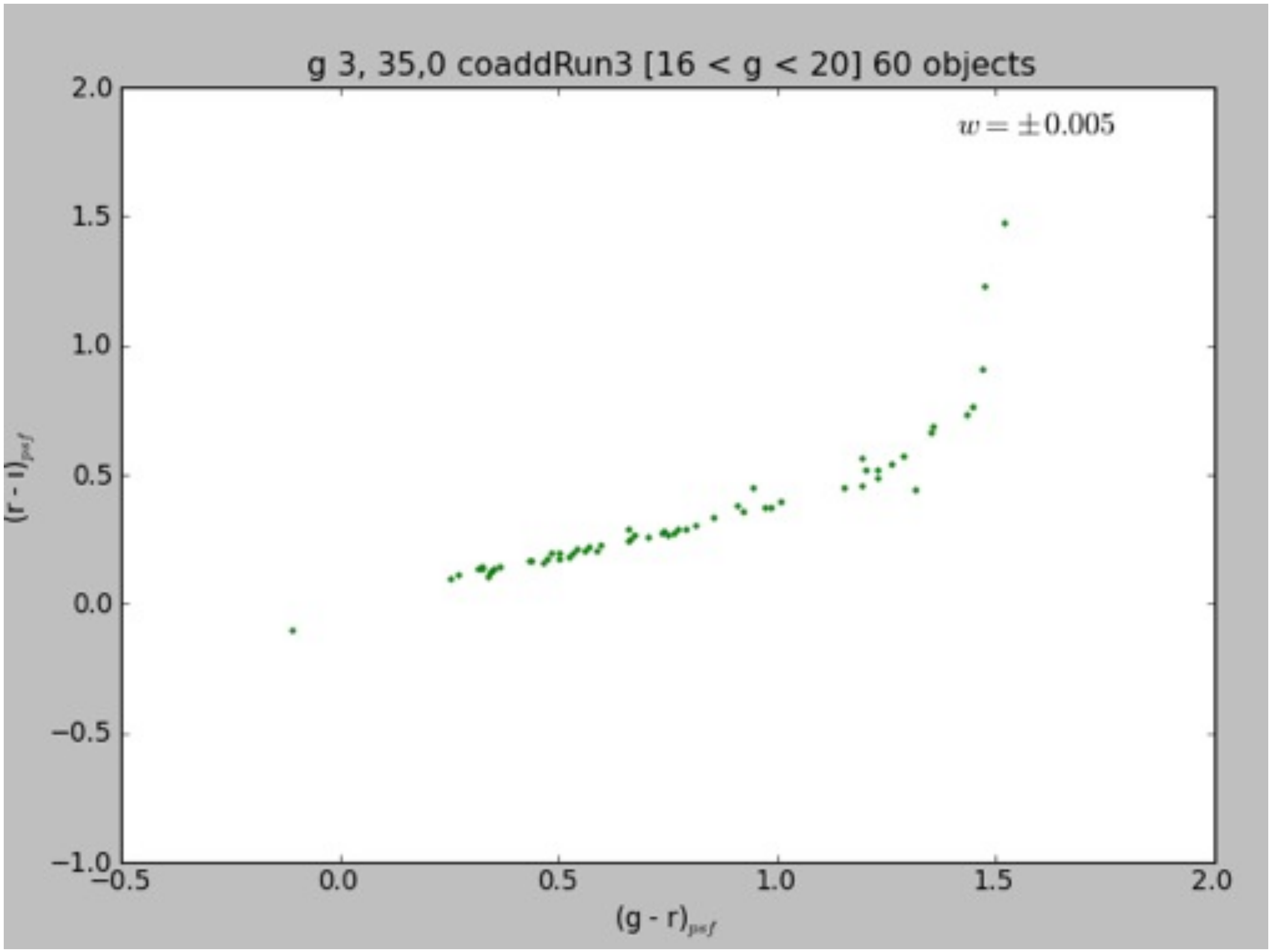


PSF mag. vs Aperture Mag.



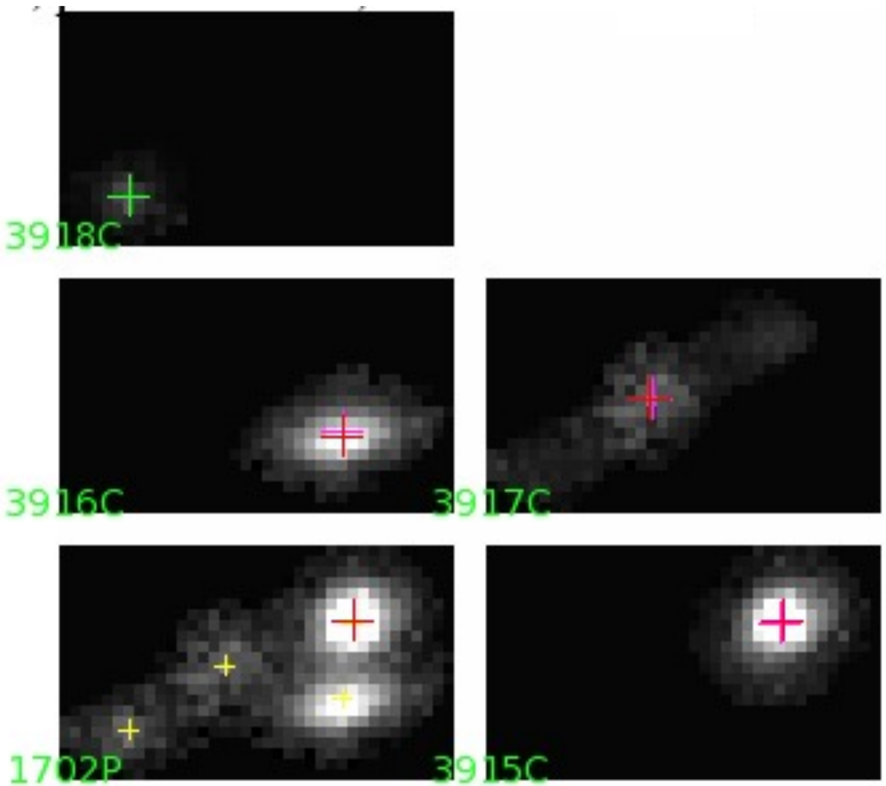
Photometry: Aperture, PSF, Kron, Model, etc...

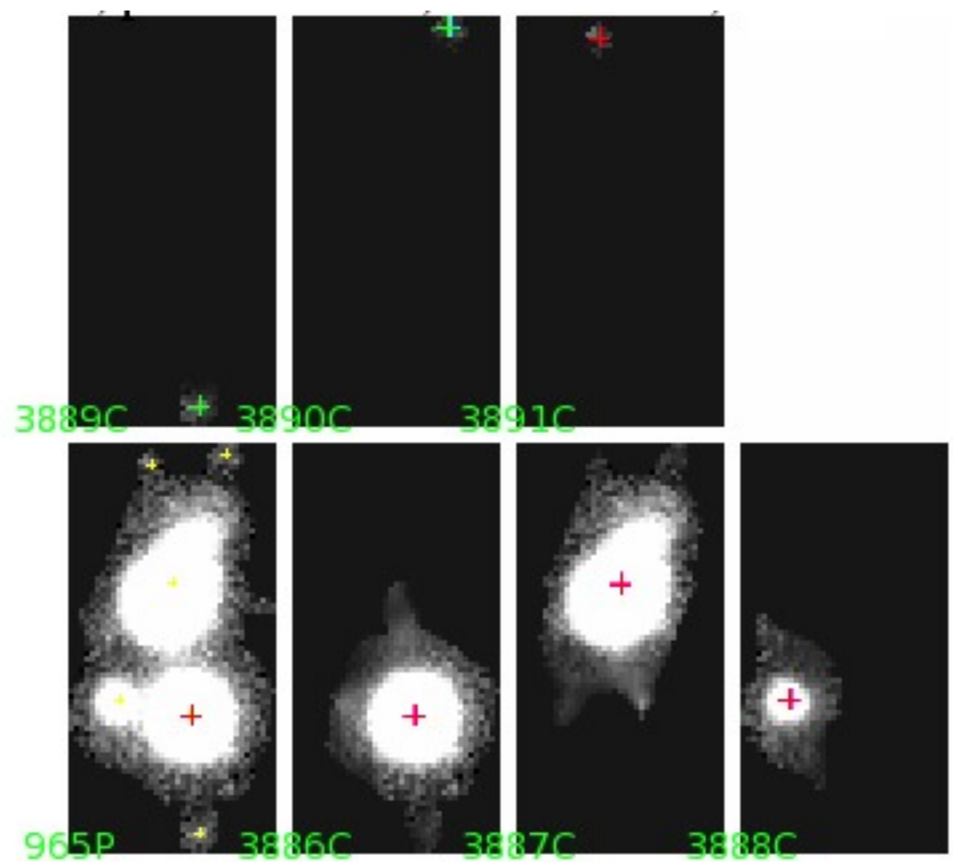
Color-Color plot for stars

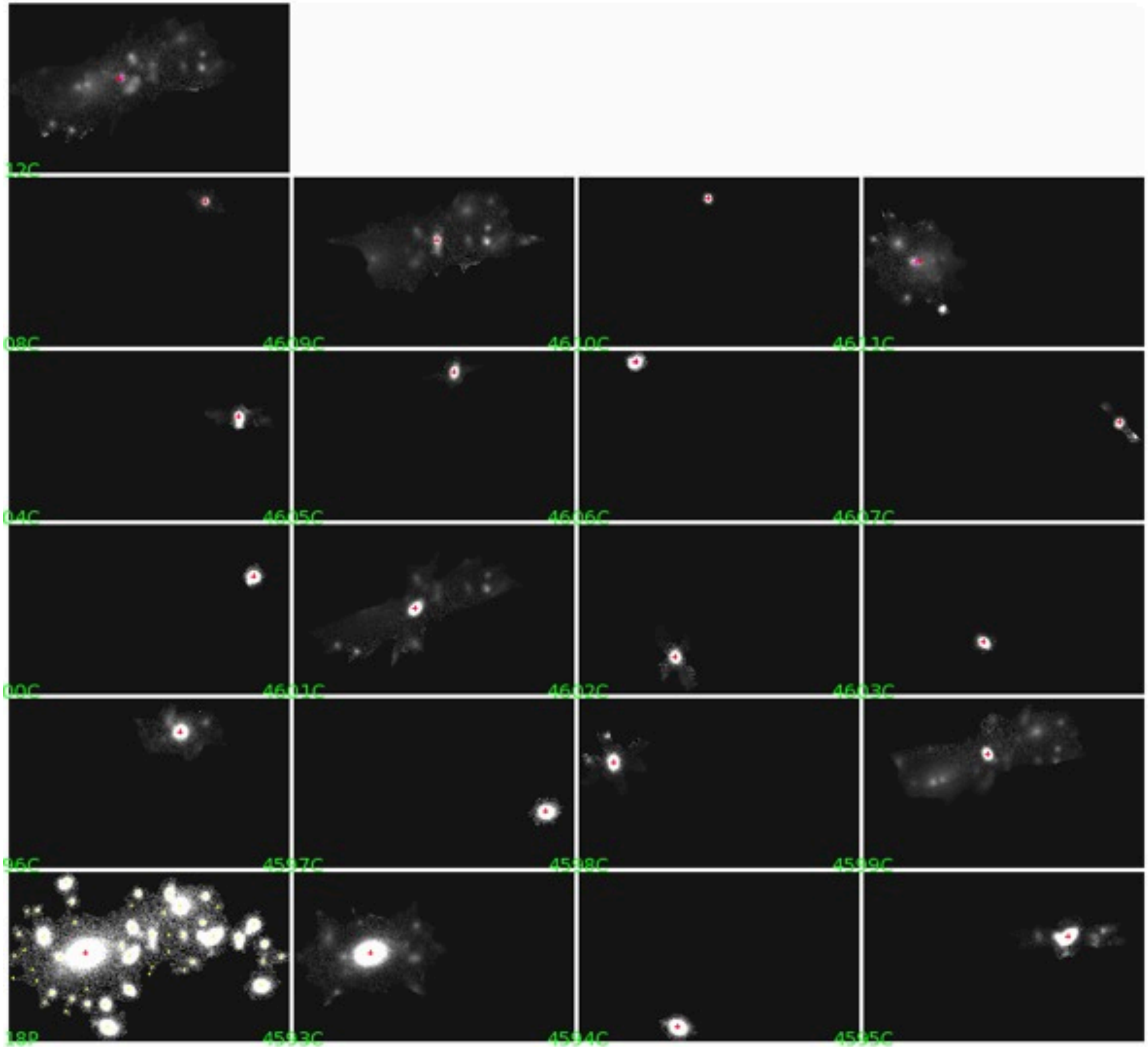


Deblender

- Port of SDSS deblender
- There are still some fails but works reasonably well






12C
08C
4609C
4610C
4611C
04C
4605C
4606C
4607C
00C
4601C
4602C
4603C
96C
4597C
4598C
4599C
18P
4593C
4594C
4595C

χ^2 Image Detection

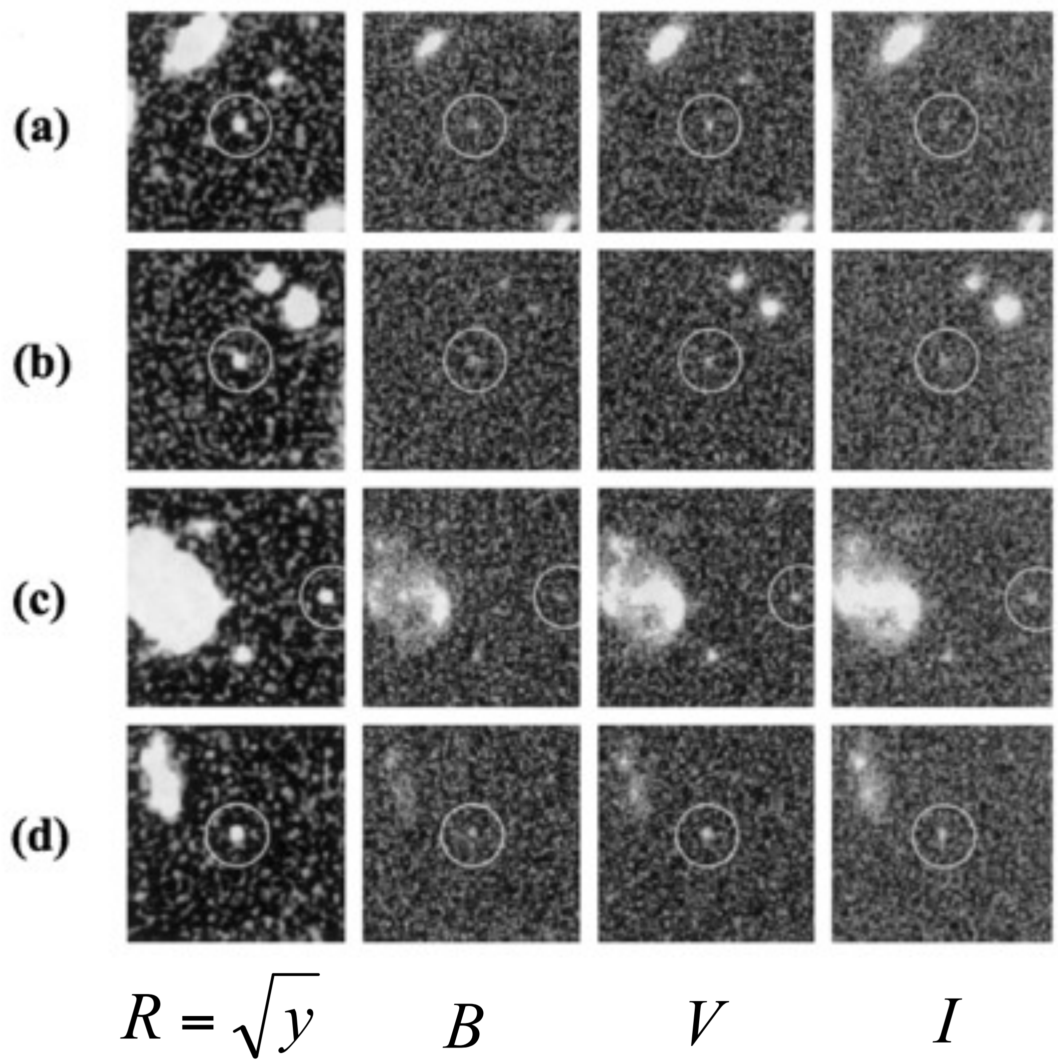
- Optimal way to use all color information available (Szalay et al. 1999)
- Define parameter y as

$$g_i = \frac{f_i - \mu_i}{\sigma_i} \quad i = 1,\ldots,N_{\mathrm{band}}$$

$$y = \sum_{i=1}^{N_{\mathrm{band}}} g_i^2$$

- Probability distribution of y is χ^2 with N_{band} degree of freedom
- We can judge which pixel is drawn from

χ^2 Image Detection



Bright in V-band but weak in other two. Obvious in the R image.

Objects with odd colors can be detected in χ^2 image.

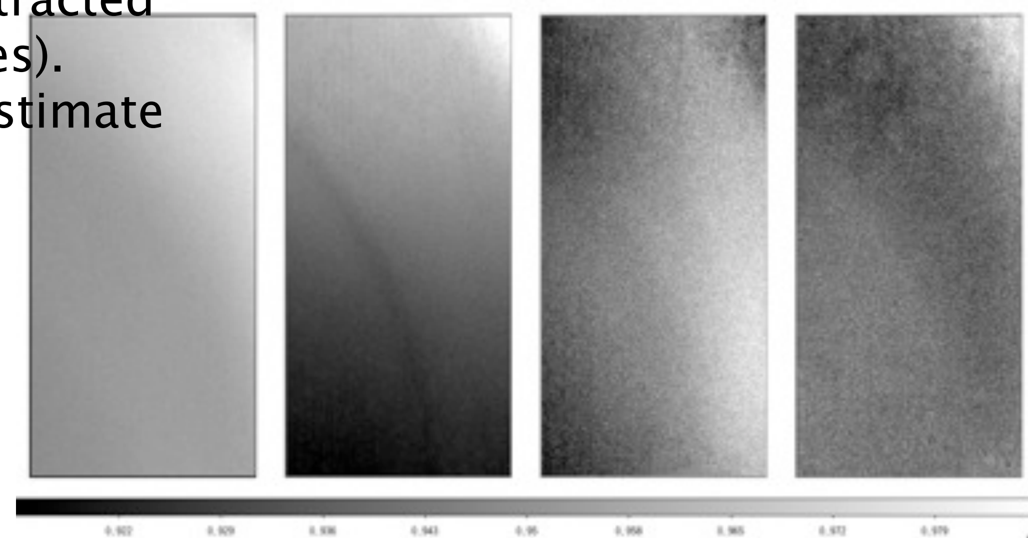
Szalay et al. 1999

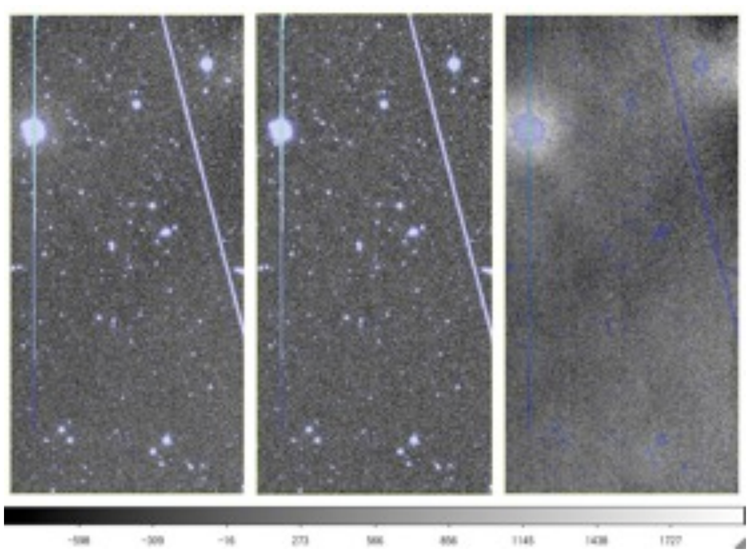
Sky Subtraction with PCA

Bright/large objects will be over-subtracted in a conventional method (sub-images).
Multiple exposures may be used to estimate the sky background.

$$B^e(x,y) = \sum_i k_i^e E_i(x,y)$$

Can be applied to very big galaxies' photometry





Light around bright star is surviving.
Not sure about large scale fluctuation.

How many order do I need to represent background?
Will coefficients correlate with other parameters?

- Elevation, rotator angle, moon position, …

Data Products

Image Data

Single CCD Images → Single CCD Catalogs

Monthly

Stacked Images per Unit Sky Tile

Every 6 months

Multi-band Images → Multi-band Catalogs

↓

Color Images (gri, riz, izy?)

PSF information as separate files

QA information (Completeness limit)

Direct `http` or `rsync` access

Catalog Data

Single CCD Catalogs

↕ XMatch for each object

Multi-band Catalogs

↕ XMatch

External Catalog?

SDSS like SQL-based DB access

Summary

- HSC data analysis pipeline is being developed as a collaboration effort
- Pipeline is tested using Suprime-Cam data
- Waiting for HSC commissioning data for final tuned up