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

Chip Data Reduction

ISR
Saturatin
Detection

Calibration
Identify Cosmic Rays

Measuremt

External

Dat Process

Correc

Catalo
For Each Pass-band

- Chip Products
- Image Stacking
- Stacked Image
- $\chi^2$ Image
- Measurement
- Stacked Image
- Final
Image Format

HDU0: Header only
HDU1: Science Image
HDU2: Mask Image
HDU3: Variance Image

Calibrated astrometry and photometry

- 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

ADU count / Gain
Saturate
CR
Crosstalk
Detected
# Catalog Parameters

<table>
<thead>
<tr>
<th>Band</th>
<th>Observed band</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJD</td>
<td>Mean time during exposure</td>
</tr>
</tbody>
</table>

* **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 2\(^{nd}\) 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

\[ P(u, v) = \sum_i a_i(x, y) B_i(u, v) \]

\[ a_i(x, y) = \sum_{l=m=0}^{l+m\leq N} b_i^m x^l y^m \]
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

\[
\begin{pmatrix}
  u_{s,e} \\
  v_{s,e}
\end{pmatrix} = \begin{pmatrix}
  \cos \theta_c & -\sin \theta_c \\
  \sin \theta_c & \cos \theta_c
\end{pmatrix} \begin{pmatrix}
  x_{s,e} \\
  y_{s,e}
\end{pmatrix} + \begin{pmatrix}
  X_c \\
  Y_c
\end{pmatrix} = \begin{pmatrix}
  x_s \cos \theta_c - y_s \sin \theta_c + X_c \\
  x_s \sin \theta_c + y_s \cos \theta_c + Y_c
\end{pmatrix}
\]

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

\[
\chi^2 = \sum_{e} \sum_{s} \left\{ \xi_{s,e} - \sum_{k} a_k u_{s,e}^{i(k)} (v_{s,e}^{i(k)}) b_k \right\}^2 + \sum_{e} \sum_{s} \left\{ \eta_{s,e} - \sum_{k} b_k u_{s,e}^{i(k)} (v_{s,e}^{i(k)}) \right\}^2
\]

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”

\[ m_0^{\text{star}} = m_{\text{obs}}^{\text{star,exp,chip}} + dm^{\text{exp}} + dm^{\text{chip}} + dm^{(u^{\text{star}}, v^{\text{star}})} \]

- Correction over focal plane
  \[ dm(u, v) = \sum_{i+j=n} f_{i,j} \times u^i \times v^j \]
  May need to include the effect of rotation

- 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 \( m_0^{\text{star}}, dm^{\text{exp}}, dm^{\text{chip}}, f_{i,j} \) by requiring the same star has the same

Overall zeropoint will be determined by the comparison with reference catalog
Flux Calibration – An Experiment

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

External $\sigma \approx 0.03$ mag

Internal $\sigma \approx 0.01$ mag

RMS of relative comparison over focal plane
PSF mag. vs Aperture Mag.

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

Galaxies

Stars $0.001 \pm 0.023$ mag (3σ clipping)
Color–Color plot for stars
Deblender

- Port of SDSS deblender
- There are still some fails but works reasonably well
$\chi^2$ Image Detection

- Optimal way to use all color information available (Szalay et al. 1999)
- Define parameter $\chi$ as
  $$g_i = \frac{f_i - \mu_i}{\sigma_i} \quad i = 1, \ldots, N_{\text{band}}$$
  $$y = \sum_{i=1}^{N_{\text{band}}} g_i^2$$
- Probability distribution of $y$ is $\chi^2_{N_{\text{band}}}$ with degree of freedom
- We can judge which pixel is drawn from
Image Detection

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

Objects with odd colors can be detected in \( \chi^2 \) image.

Szalay et al. 1999

\[
R = \sqrt{y} \quad B \quad V \quad I
\]
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
  - Monthly
- Stacked Images per Unit Sky Tile
  - Every 6 months
- Multi-band Images
- Color Images (gri, riz, izy?)
  - PSF information as separate files
  - QA information (Completeness limit)

Catalog Data
- Single CCD Catalogs
- Multi-band Catalogs
- XMatch for each object
- Multi-band Catalogs
- XMatch
- External Catalog?

- SDSS like SQL-based DB access

Direct http or rsync 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