Crowded Field Processing

Colin Slater
Rubin Observatory/University of Washington
Milky Way science is integral to LSST

• “Mapping the Milky Way” is one of the survey’s four science pillars.
• All parts of the MW are important
  • E.g., there’s greater interested today in interactions between the MW disk and the stellar halo; one benefit of the uniform coverage of surveys like PS1 and Gaia. Lots of value in measuring the MW uniformly.
• “Exploring the Transient Sky” is another science pillar
  • Galactic transients occur where the stars are, and stars in the Milky Way are predominantly in the disk. Image differencing is strongly complemented by deep, well-measured direct photometry.
• We’ve done testing of our current deblender, see DMTN-077.lsst.io
  • The first-order effect is that deblending drops in completeness with increasing source density, as closely-neighboring fail to get detected and passed to the deblender.
  • This affects ~8% of wide-fast-deep area at single-visit depth, 15% at 10-year depth. Affects substantially all of the Galactic plane area.
• But crowded fields have very different needs from high-latitude, and might not be well suited to deblender assumptions

• The ideal solution would be one deblender that smoothly transitions to heavily-point-source biased solutions at low Galactic latitude

• That is a serious challenge, and there’s already lots of work that needs to be done on the deblender for high latitude cases.

• There are easier, well-known solutions for crowded fields. Risk-reduction.
ISR, Astrometric & Photometric Calibration, ...

Coadd Construction

Cell-based Coadds

Deblending-based Pipeline
Including Galaxy Measurement
Runs over the entire survey area

“Normal” Catalogs

“Crowded Field” Pipeline
Simultaneous PSF Fitting
Only in relevant areas

CF Catalogs

See DMTN-129.lsst.io
SIMULTANEOUS PSF FITTING

• The “Crowded Field Branch” would only fit PSFs to sources, no galaxy model fitting.

• The core of the photometry algorithm is DAOPHOT-like. Because we know the PSF model, we can compute every star’s contribution to the counts in a given pixel.

• Uses a sparse matrix with source fluxes and pixel values as parameters, least-squares problem to solve. Early prototype running inside the LSST stack.
Coadd measurements for deep detection

• Need to use coadds for detecting sources deeper than single-epoch imaging.
• We’ve already heard about the benefits of cell-based coadds
• All the same benefits apply to crowded fields as well
Challenges at LSST-scale

- LSST is not just a static-sky survey. There’s lots of Milky Way science in stellar motions and variability.
- Want to also use the extra depth from many exposures, without having resort to full multifit.
- Short-period coadds (think 6-12 months) give us depth, without washing out motion.
- For light curves, fit fluxes on individual images, but with coadd-defined positions.
“Crowded Field” Pipeline Branch

- Cell-based Coadds

Including “Short Period” (e.g. annual) coadds for proper motion

- Source Detection

“Detect-fit-subtract” cycle has lots of important details; requires scientific judgement

- Simultaneous PSF Fitting to Coadd, Source Subtraction

- Object Positions & Motion

- Simultaneous PSF fitting to each epoch

“ForcedSource” per epoch Fluxes
Iterative Source Detection

• Because we assert that we know the shape of all sources, subtracting identified sources to find faint neighbors is possible
• This cycle happens on the coadd, not data-intensive
• Will also want to iterate on measuring the PSF; removing neighbors to get clean PSF wings
• Many heuristics and tuning parameters in this process. Requires careful testing.
Positions, motions, and fluxes

- Fits on individual short-period coadds give us positions
- Matching objects between different coadds gives us motions
- Position + motion lets us do force photometry on individual epochs
• Mean fluxes would come from the mean of individual epoch fluxes, not from the coadd measurement.
• Possible extensions to also measure centroids or spatial derivatives to fit parallax; untested and unproven.
• Our requirements say we must measure PM down to the single epoch depth, going deeper is a bonus we want to achieve.
• Lots of little details matter.
  • PSF accuracy is important, and it’s hardest to measure in crowded fields.
  • Building a source list is full of heuristics. Will require experimentation and testing
  • Iterative detection and subtraction is sensitive to diffuse light
  • Background subtraction also plays a subtle role; may be tie-ins from here to the deblending pipeline
Conclusions

• We have a plan for a crowded-field data product, see DMTN-129.lsst.io.
• Basic algorithmic elements are well-known.
• The challenge is in assembling the complete system, to work robustly at scale, and addressing the time-domain