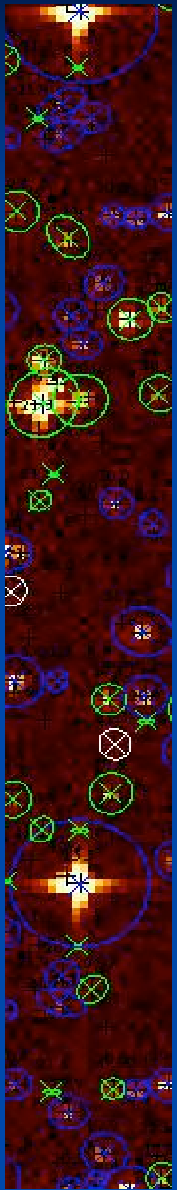
On-board detection, selection and confirmation algorithms

Observatoire de Paris-Meudon
S. Mignot, F. Arenou, C. Babusiaux

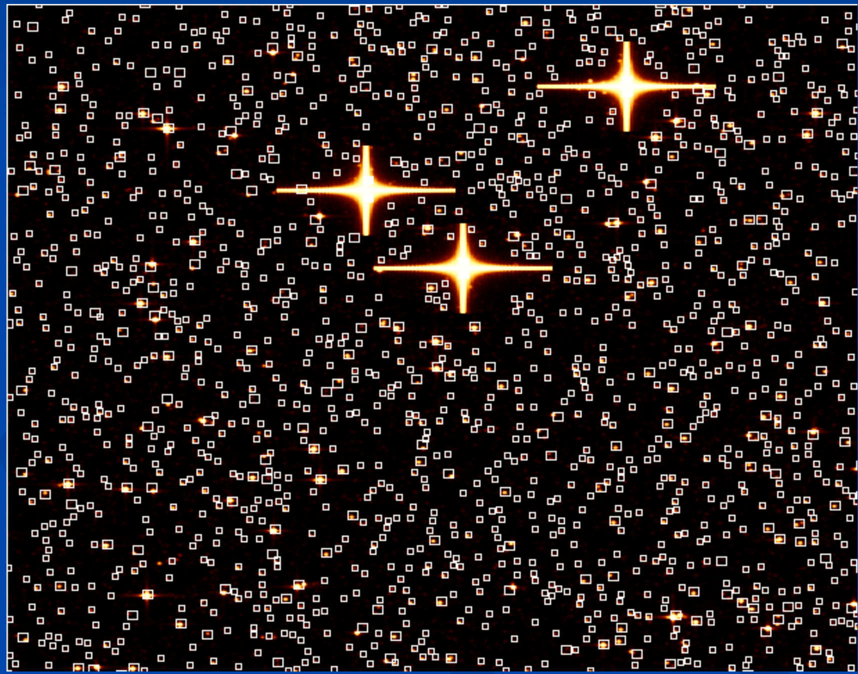
Overview

- Inputs to PDHE
- Architecture of the processing
  - Flowchart
  - Summary of outputs
- Algorithmic principles
  - Main drivers
  - Detection, selection and confirmation
- Development status
  - Features and tests
- Code stability and possible evolutions
  - Architecture
  - Methods
  - Implementation

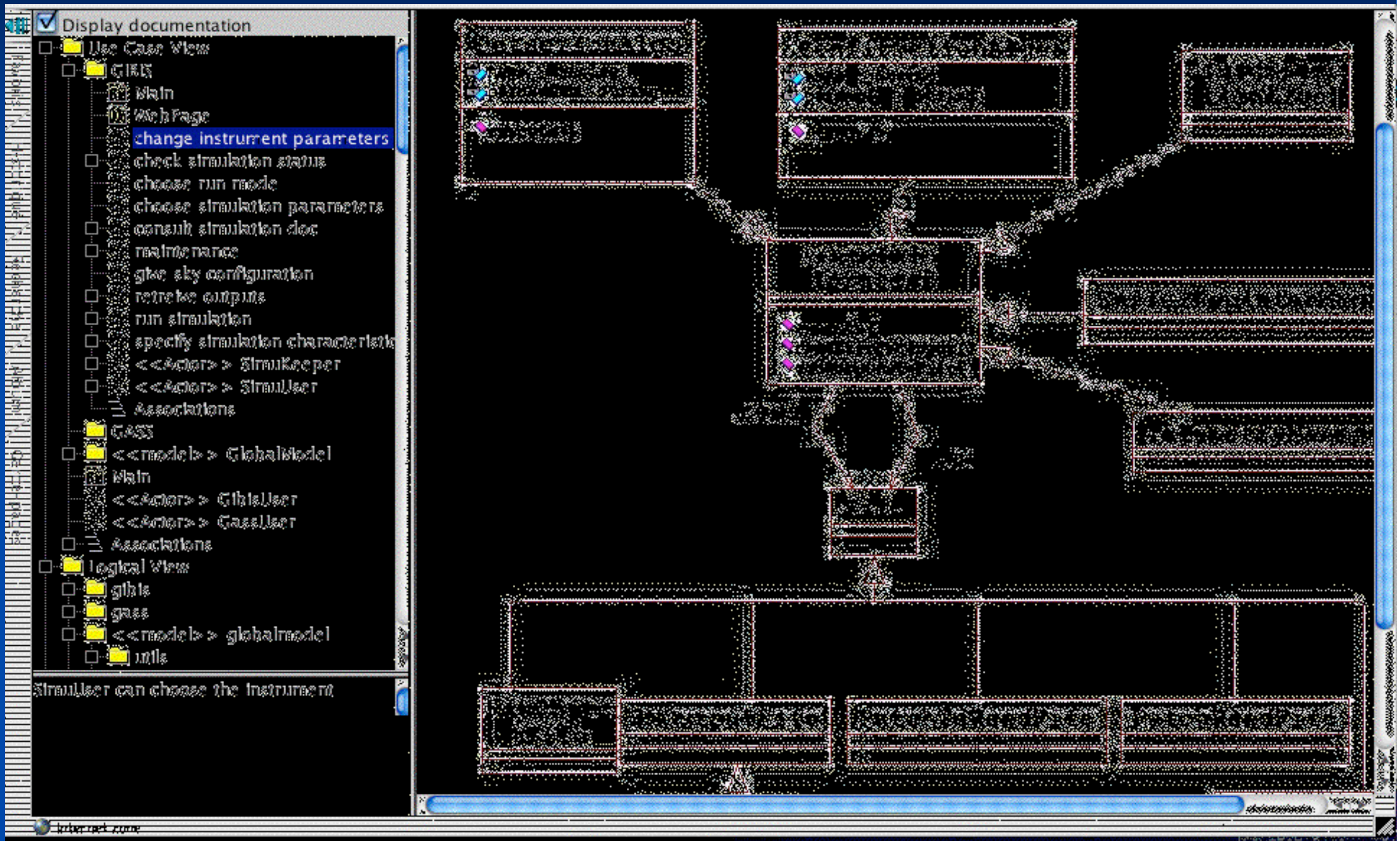
The input



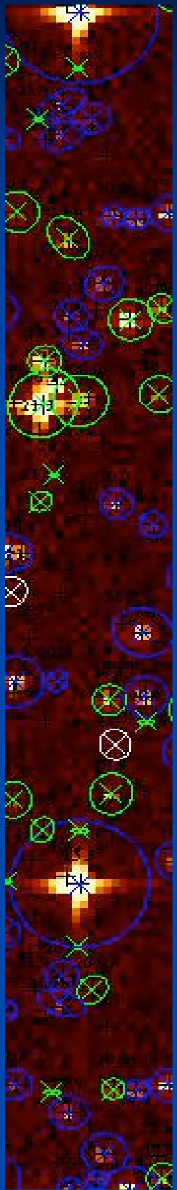
- Ressources
  - Detection (2 algorithms)
  - Preliminary selection/tracking of objects
  - Tests on representative data
  - Reading CCD, organising, packetting, compressing data is not included
- Several needs:
  - to take into account the Gaia2 design (e.g.saturation)
  - to test/implement new parts
  - to have an “object-oriented” software
    - Maintenance
    - to further use an UML model
- A large effort (~13manxmonth) for coding since april



Rose (UML) Model



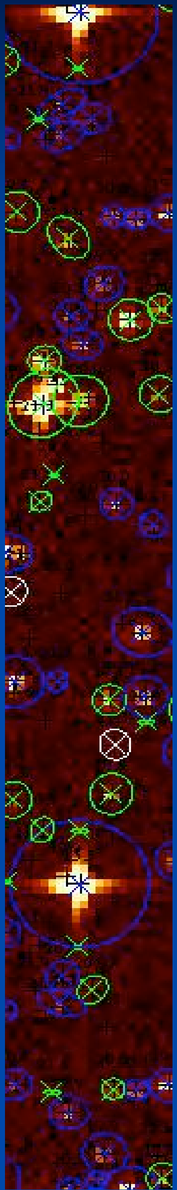
Documents



- Requirements
  - Input to PDHE contract, 14-DEC-02
  - Gaia_Detect documentation, 29-AUG-02
  - SWA users guide
  - Flowcharts of the SWA and GD detection algorithms, 29-OCT-02
  - A document summarizing scientific requirements to be done soon
- Proposals
  - UML model
  - First implementation of the selection algorithm, 25-NOV-02
  - Technical characteristics and design issues for the GAIA astrometric focal plane

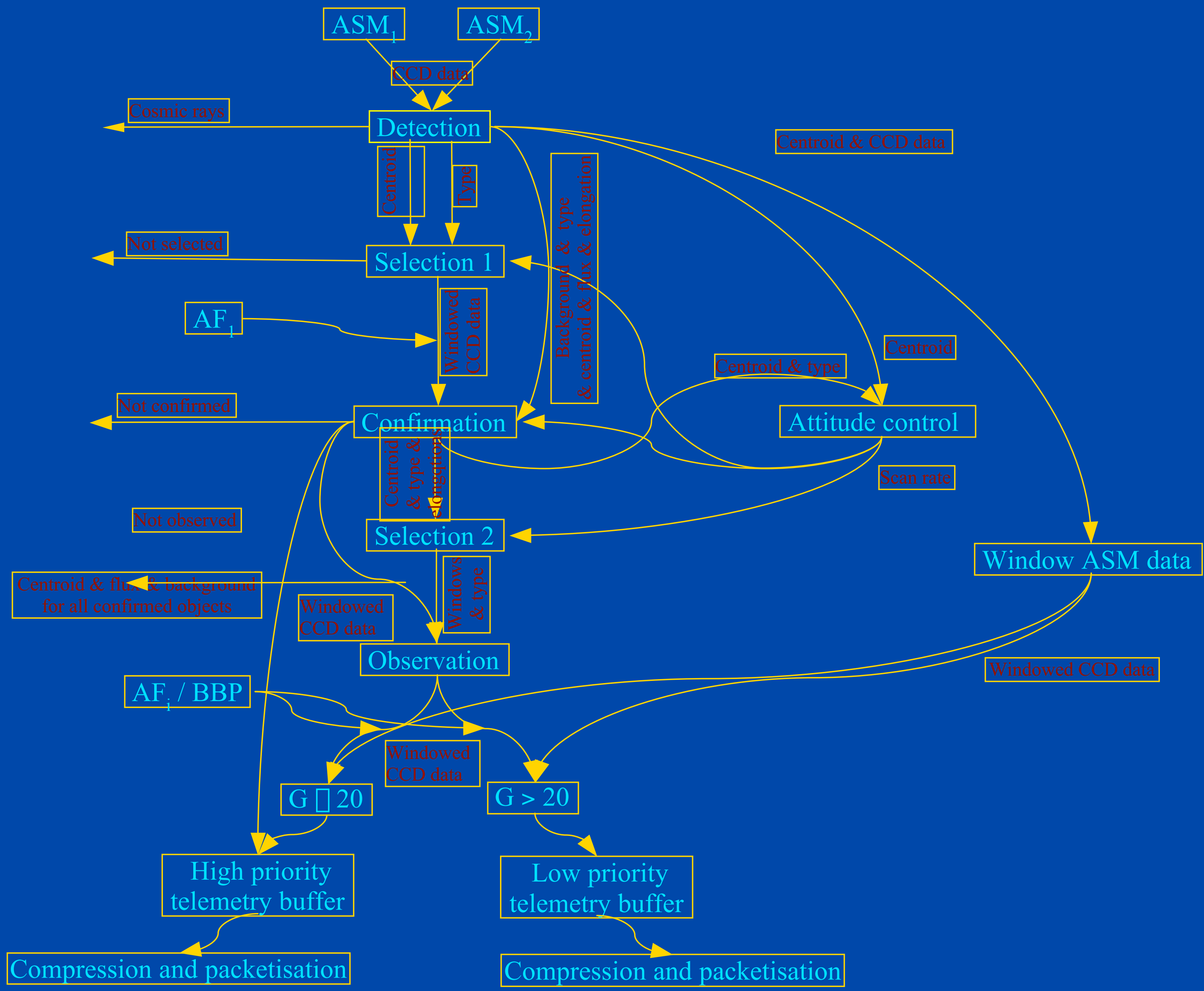
- OBD-CoCo-03
- OBD-FC-01
- SWA_guide.ps
- OBD_SM_02
- GAIA-CUO-117

- http://www.ast.cam.ac.uk/~carine/GAIA/SIMU/UML_MODEL/model.html
- OBD-HC-01
- UB_GDAAS_TN_001_V30.doc



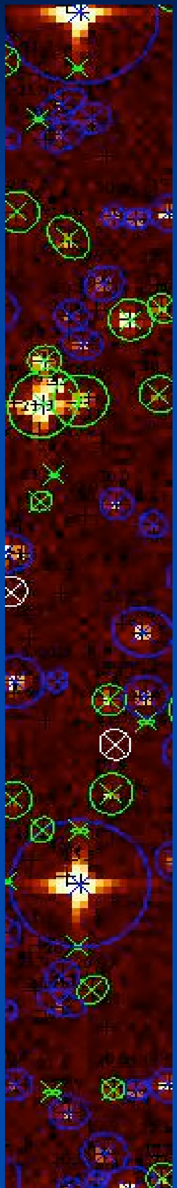
Architecture of the processing

Flowchart

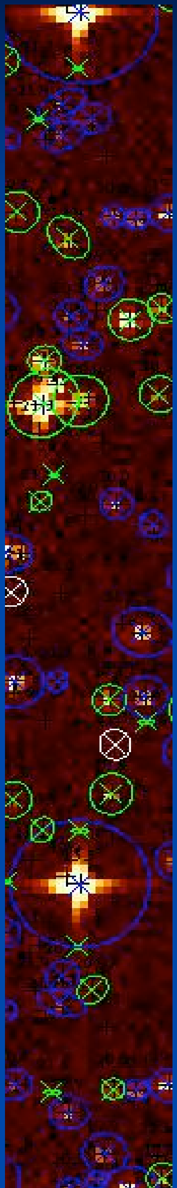


ASM_1
ASM_2
CCD data
Cosmic rays
Detection
Centroid & CCD data
Centroid
Type
Background & type & centroid & flux & elongation
Not selected
Selection 1
AF_1
Windowed CCD data
Centroid
Centroid & type
Not confirmed
Confirmation
Attitude control
Centroid & type & elongation
Scan rate
Not observed
Selection 2
Window ASM data
Centroid & flux & background for all confirmed objects
Windowed CCD data
Windows & type
Observation
Windowed CCD data
AF_i / BBP
Windowed CCD data
G 20
G > 20
High priority telemetry buffer
Low priority telemetry buffer
Compression and packetisation
Compression and packetisation

Output summary

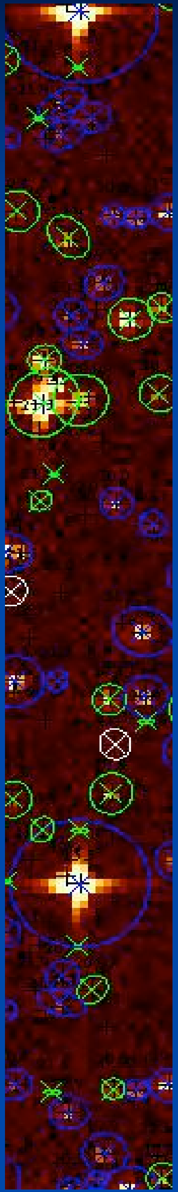


Inputs	Task	Outputs
ASM1 and ASM2 samples (all CCD)	Detection	(c1) centroids (f1) flux (t1) types (e1) elongation backgrounds cosmic rays (rejected)
(c1) centroids (ASM1 & ASM2) (t1) types scan rate	Selection 1	(w1) windows (position & type, AF1) list of not selected
(w1) windowed samples backgrounds (c1) centroids (t1) types (f1) flux scan rate	Confirmation	(f2) flux (c2) centroids (e2) elongation (t2) types list of not confirmed
(c1) centroids (c2) centroids (e1) elongation (e2) elongation (t2) types scan rate list of not confirmed	Selection 2	(w2) windows (position & type, AF2-11 & BBPs) list of not observed
(c1) centroids (c2) centroids (t2) types	Scan rate computation	scan rate
ASM1 and ASM2 samples (all CCD) (c2) centroids (w2) AF1 windowed samples	ASM & AF1 windowing	(w3) windowed data (ASMs) (w4) windowed data (AF1)
(w1) windowed samples (G<20) (w4) windowed samples (G<20) (w3) windowed samples (G<20) (c2) centroids (f1) flux backgrounds	Packetisation Compression (high priority)	high priority packets
(w1) windowed samples (G _ 20) (w4) windowed samples (G _ 20) (w3) windowed samples (G _ 20)	Packetisation Compression (low priority)	low priority packets



Algorithmic principles

Algorithmic drivers



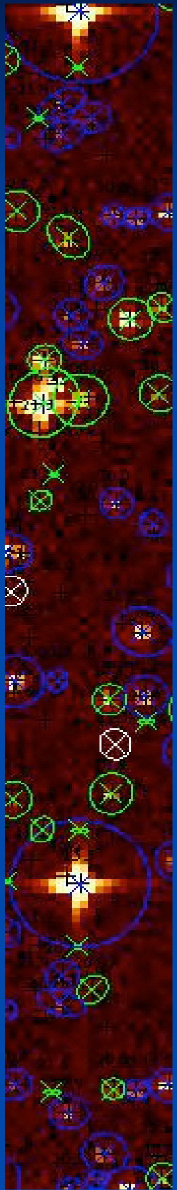
- **Simplicity:**
  - Strong real-time constraints (read-out each 0.735 ms)
  - Limited computing power
- **Efficiency:**
  - Telemetry & computing resources (false detections)
  - Statistical study of the Galaxy (completeness & homogeneity)
  - Completeness of telemetered information (ground reduction)
- **Robustness:**
  - Multiplicity of instrument modes
  - Instrument deterioration
  - Diversity of observed objects & unpredicted ones

Detection: overview

- 1. Buffering
- 2. Denoising (filtering)
- 3. Source detection
- 4. Data collection
- 5. Measurements
- 6. Classification

Background estimation (→ 3. Source detection or 4. Data collection)

Buffering



SWA	GD

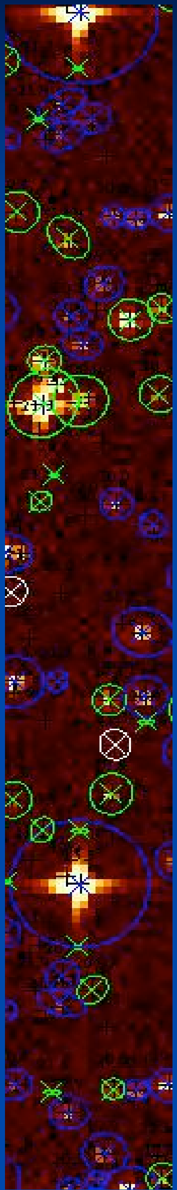
SWA

- Local approach
- Per column basis:
  - Denoising -> 3
  - 4-connectivity -> 3
  - Background estimate ->21

GD

- Region approach
- Per column "pack" basis:
  - Background estimate -> 32

Denoising (filtering)



SWA	GD

- Noise
  - Photon noise
  - Read-out noise

→ Centered gaussian model

convolution filtering with kernel:

1	2	1
2	4	2
1	2	1

Source detection

SWA

- Maximum-based:

  1 max ⇔ 1 object

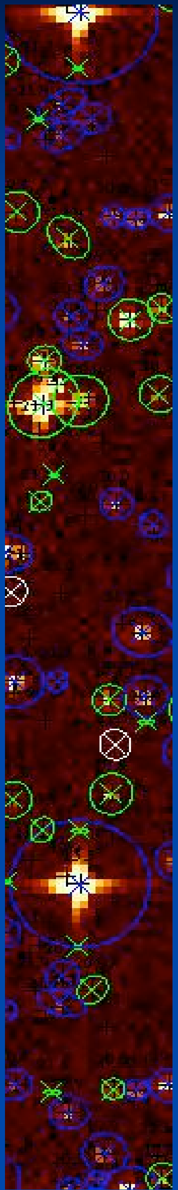
  - 8-connectivity
  - χ^2 test for relevance

GD

Background estimation
- Model: smooth
- Trimmed median
  - On a region-basis (32x32 samples)
- 2d linear interpolation

Source detection
- SNR threshold-based
  - Background substraction

Data collection



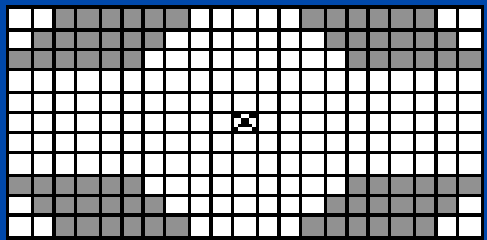
SWA

Background estimation

- Trimmed median between PSF's spikes

Data collection

- Template-based:

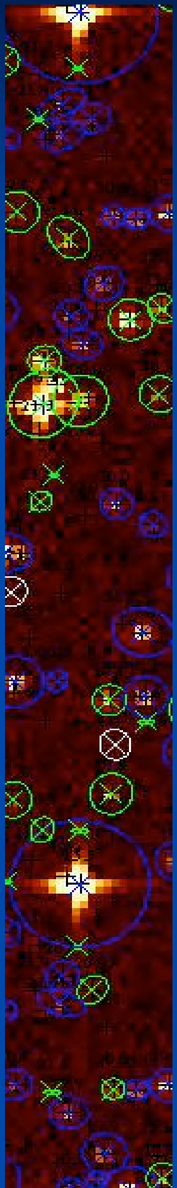


- Object SNR filtering

GD

- Connected-component based:
  - 4-connectivity
- Object “SNR” filtering
- Border effects handling
  - Restore split objects

Measurements



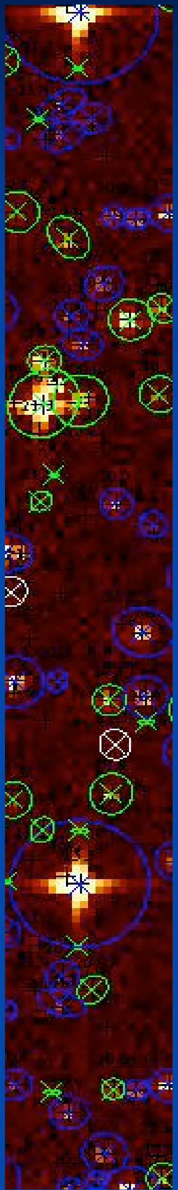
SWA

- Centroid
  - Barycentre on flux-adaptive sized neighbourhood (3x3 or 5x5)
  - Refined with an offset for PSF assymetry
- Flux
  - Sum of sample values on 3x3 neighborhood + correction
  - Refined to account for PSF extension beyond the neighborhood
- Background
  - Previous estimate

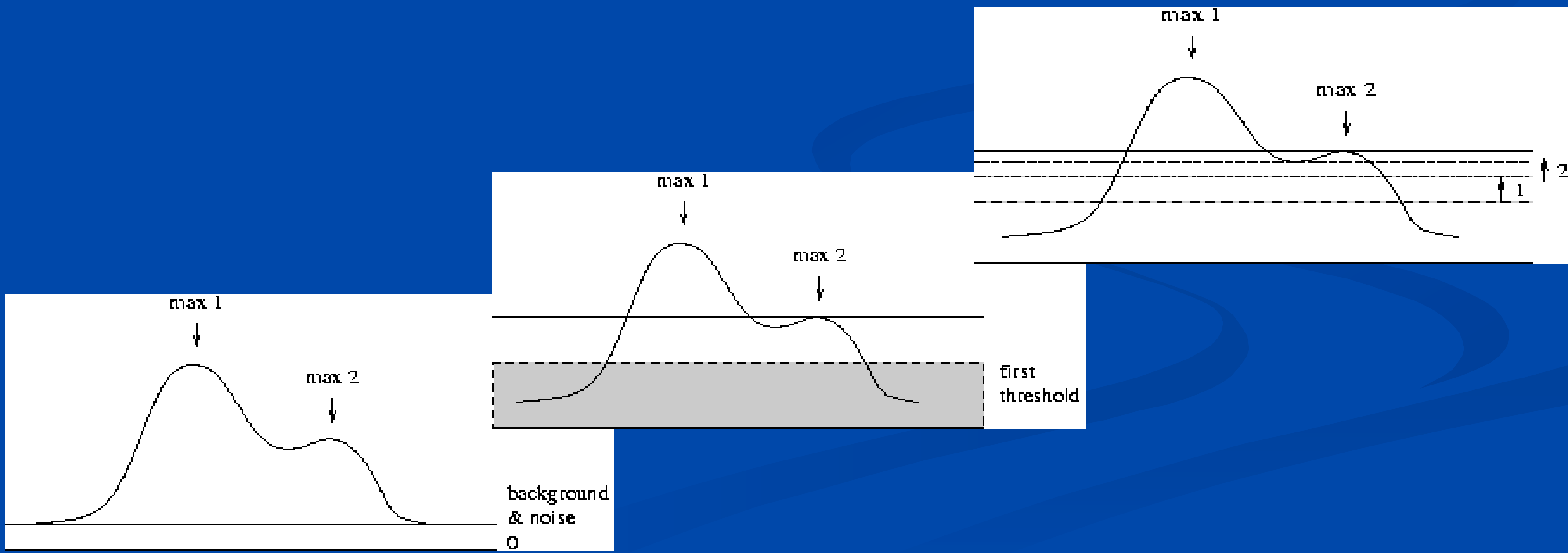
GD

- Centroid
  - Barycentre of connected-component's samples
  - Refined with LSF-fitting (AL)
- Flux
  - Sum of samples
  - Refined using a spline-interpolated error look-up table
- Background
  - Mean value of connected-component samples' previously interpolated background estimates
- Type
  - First-order moments for shape parameters (orientation & excentricity)
  - Use flux, background, χ^2 test

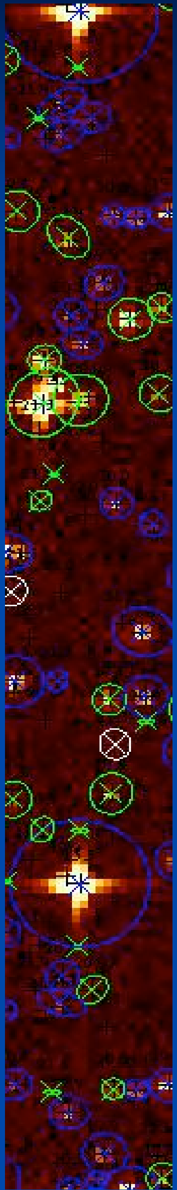
GD: mutliple objects



- Maxima-based compound objects resolution
  - Separate sub-objects 2 by 2 recursively
  - Optimal threshold search

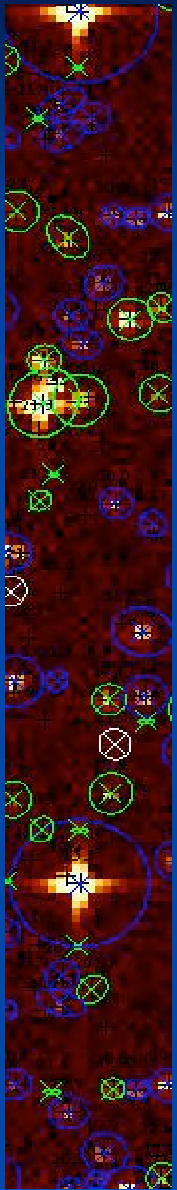


Selection/tracking: why and how



- After detection:
  - Which objects to observe
  - What to download
- Constraints on resources
  - How (centering, size, overlapping, …) in each CCD (tracking=observing strategy)
  - Limited number of samples per row
  - No sample overlapping
  - Storage (possible downlink failure)
  - Telemetry (some great circles with a large number of objects)
- For each object
  - Some priority to bright objects
  - Not implemented for the moment

Implementation



- In general, implementation of CUO-113 proposal
- Saturated stars
  - It is looked into the 6 spike patches and if there is a companion inside a patch: priority to bright stars is assumed, the bright star patch is kept; otherwise the patch of the first to arrive is kept (optional).
- Double stars
  - A window centered on geocentre or two windows
- Optimization of the patch selection
  - Allowing up to 1/3 patch shift to cope with DMS/crowding
  - Allowing patch (but not sample) overlapping
- Based on a configuration file

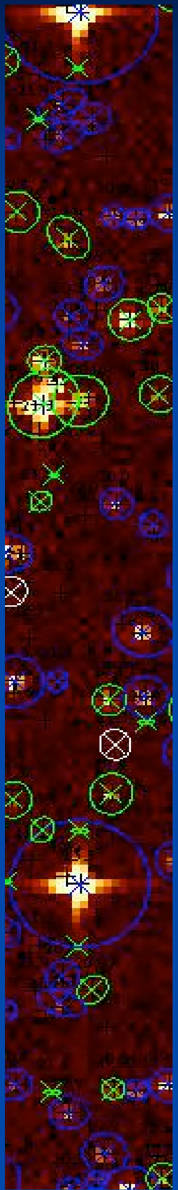
Patches for double stars (TBC)

- Patch sizes

	single	double
G < 12	(3+3) 16x2	-
12<G<17	16x12	16x12
G>17	6x10	16x12

- Selection
  - 12<G<16: if along-scan and across-scan separation are <= 12 pixels, one patch is taken, centered on the geocentre else two patches are taken, with the large single star patch size
  - G > 16: if along-scan and across-scan separation are <= 8 pixels, one patch centered on the geocentre else two small patches are taken

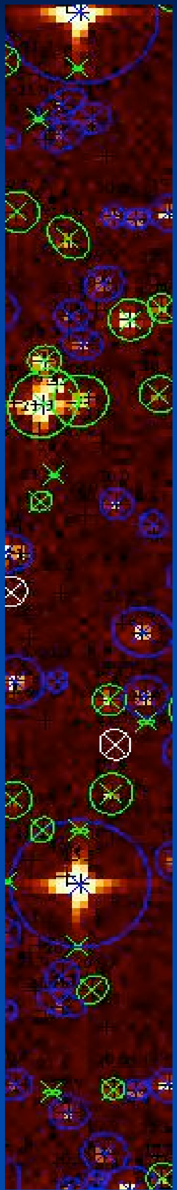
Confirmation: overview

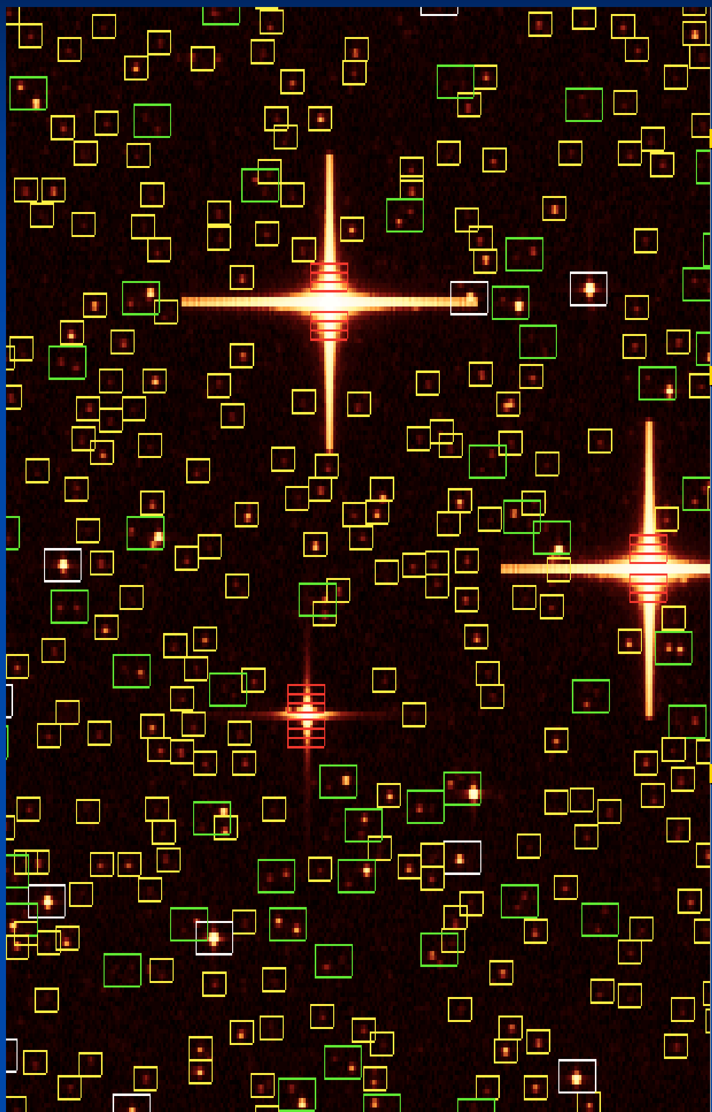


- "GD-based" simplified detection
  - Predict sample-wise background in AF1 from ASM data (different exposures)
  - Source detection: SNR threshold (unfiltered data)
  - Data collection: connected-component based (4 connectivity)
  - Measurements: centroid, flux (no bright star refinement) and type
- Cross-matching
  - 1 <-> 1: threshold on distance between centroids (NEO-oriented threshold)
  - ? <-> 0: not confirmed
  - > 1 detections during confirmation: cross-match each with the detection(s)
    - No additional detections
    - Thresholds on centroids' distance and flux difference

Development status

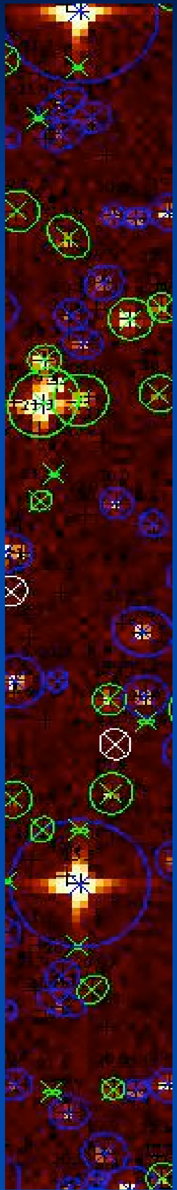
Demonstration





- Data observed by HST WFPC2
  - Real Baade window data
  - Gives ~4 million stars/square degrees
- GIBIS simulation
  - Gaia 2 Astrium design in ASM/AF1
  - 32" corresponds to ~0.5 s scan time
  - Image duplicated across-scan to get 1966 px
  - Sampling 2x2, exposure time 1.915 s, RON=8.7e
  - Constant background 22.5 mag.arcsec-2
- Input to PDHE
  - A program is given to duplicate the image along-scan
  - Or it can be done on the fly

Features (1)



- ## Nominal case:

  Supported

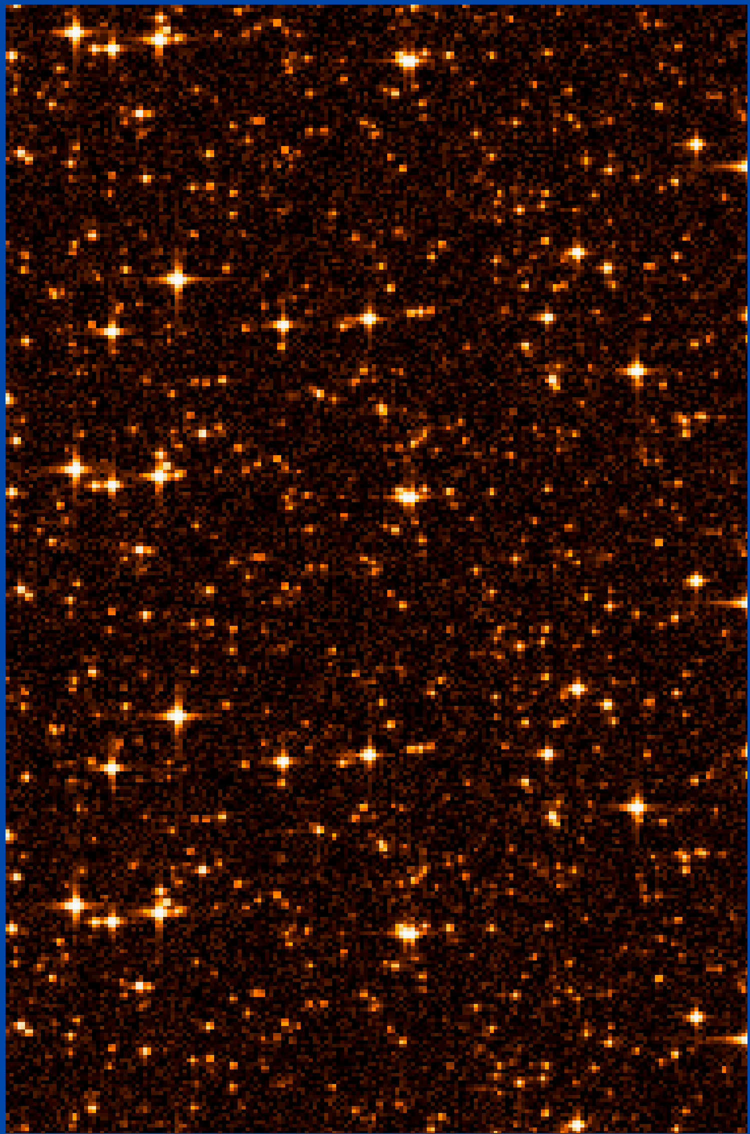
  - Stellar content: most populated magnitude classes (*G* = 15...23 mag)
  - Expected noise budget: photon noise and read-out noise (expecting pre-processed CCD data)
  - Targeted towards the high densities stringent requirements

  Missing

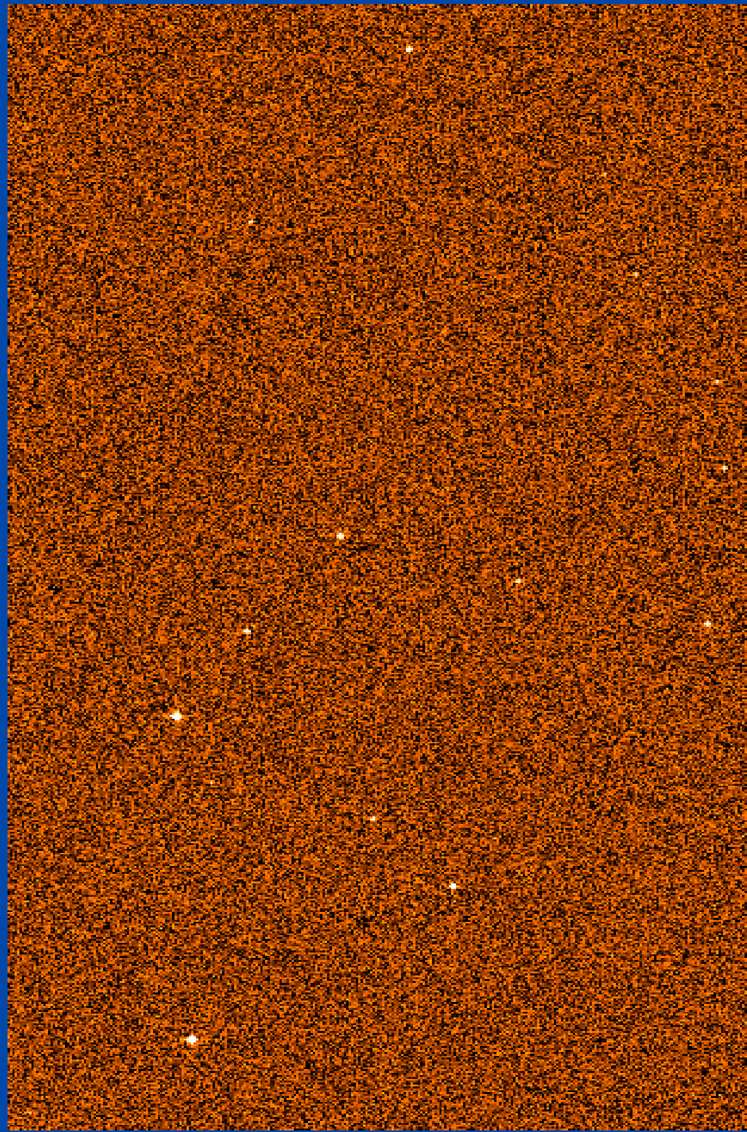
  - Objects:extended objects (galaxies, globular cluster etc...),
    saturated objects (*G* = 2..15 mag),
    fast displacement (NEOs etc...)

- **Data set:** Astro: Baade's window (3. 10^6 */deg^2) + Spectro: l=74

  Astro+Spectro: average case (25 000 */deg^2)

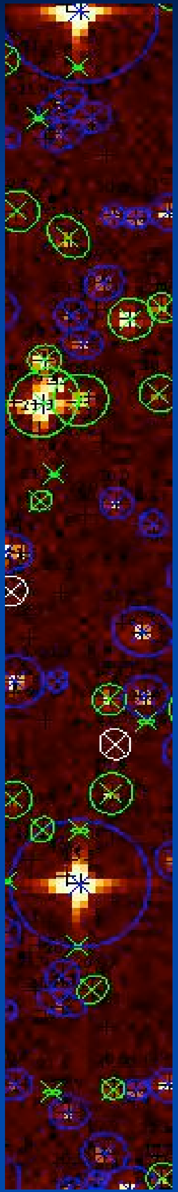


Baade's window



Average case

Features (2)

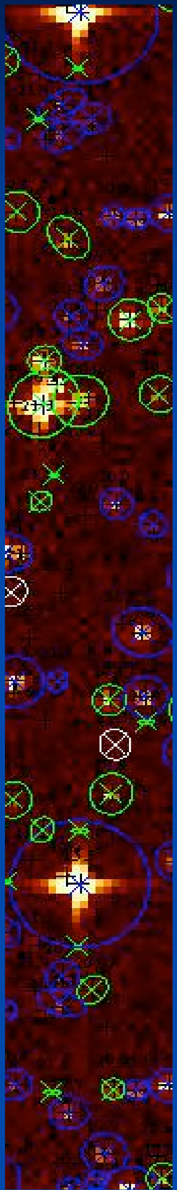


- **Instrument behavior:**

  Missing

  - Modes: beginning of mission, recovery from attitude disruption, lagging behind, redundancy, instrument degradation, low density attitude control
  - Combination of FOV (implemented but not tested)
  - CCD real output data (cosmic rays,hot pixels,CTE,etc)
  - Calibration (on-board PSF determination or uploaded from ground)

Tests



Data flow variability:

- Expected cosmic rays effects:

   Possibly an issue

- Undetected stars & false detections (without cosmic rays)
  - Completeness (ASM):

Baade:	GD:	100%	up to	G = 18.4	(84.9% for G < 20)
	SWA:	100%		G = 18.3	(86.9% for G < 20)
average:	GD:	100%		G = 20	
	SWA:	100%		G = 20	

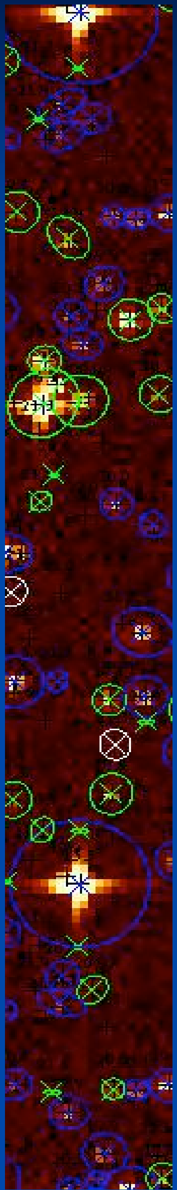
  - Completeness (spectro):

average:	GD:	90.4%	G = 20
	SWA:	95.8%	G = 20

  - GD/SWA comparison: SWA: complete at low magnitudes
    GD: improved detection rates at high magnitudes

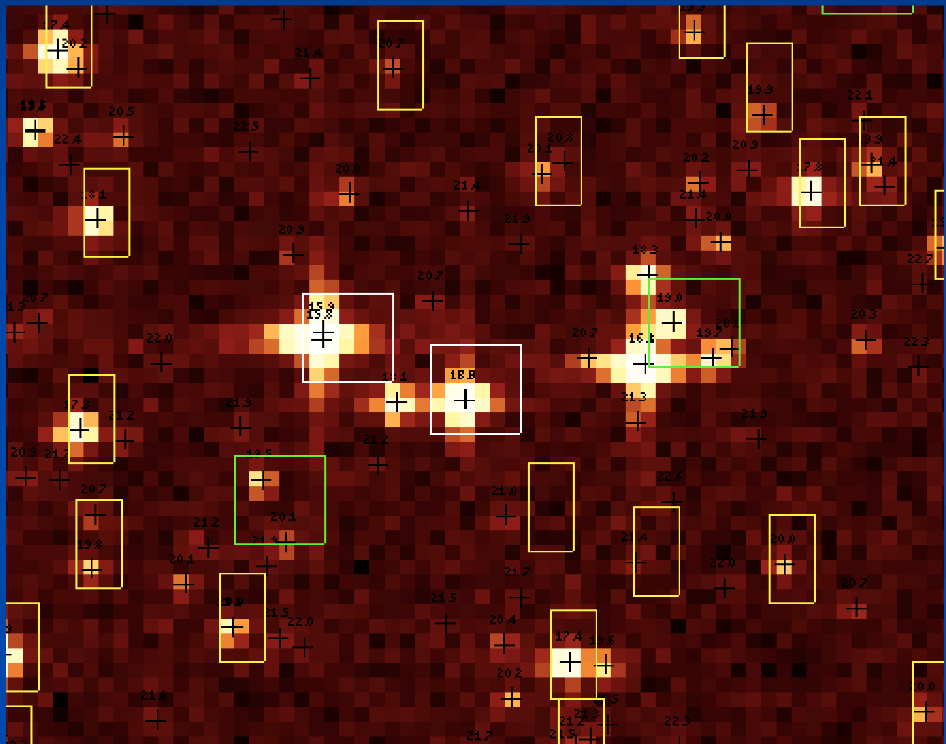
Code stability & possible evolutions

Methods



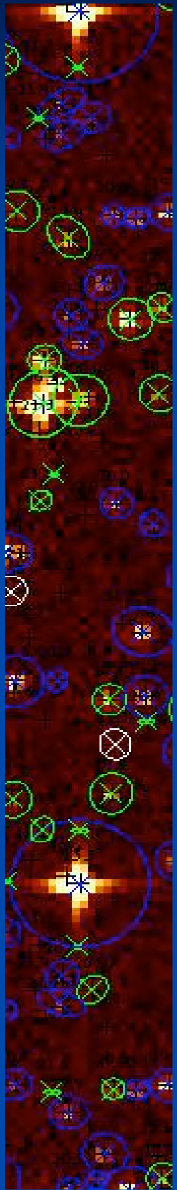
- ## Short term:
  - *On-going tests: learning from GD & SWA to build a synthesis*
  - Stable:
    - Maximum search with additional handling of extended maxima: improves completeness
    - Region-growing algorithms: evolve towards a one-pass object detection and data collection
    - Requires efficient traversing of the CCD data set for simple element-wise computations
  - Planned:
    - Improvements for double stars
    - Predictive processing: slowly varying image parameters
    - Paralleled centroid determination (relying on scalar products)
- ## Prospective
    - confirmation strategy: cross-matching, MBP
    - Rewriting of the selection function: selection criteria (priority to magnitude), windowing (position for bright stars to be adapted to design, patches for extended objects, shifted position of windows, short windows in AF11/BBP, odd number of patches in MBP)
    - Real false detection-oriented selection: explicit decision tree
    - Cosmic-ray rejection at the detection stage
    - Specific objects' handling (galaxies, saturated etc...)

Example: double stars



- A large fraction of stars are double stars
- The sampling has began to be studied
- The strategy is not completely fixed
- What is implemented is one patch centered between components
- Alternative strategy is to take supplementary patches
  - Increase telemetry rate
  - More complicated reduction
  - May be better for astrometry

Architecture

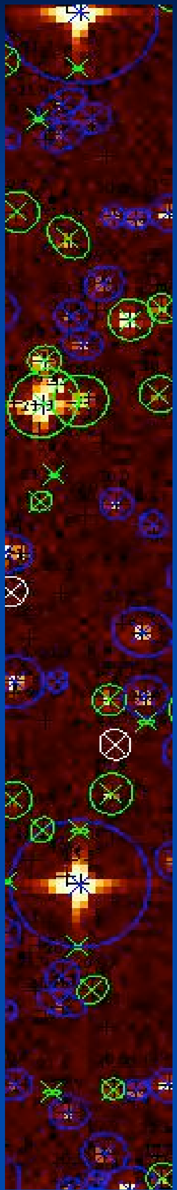


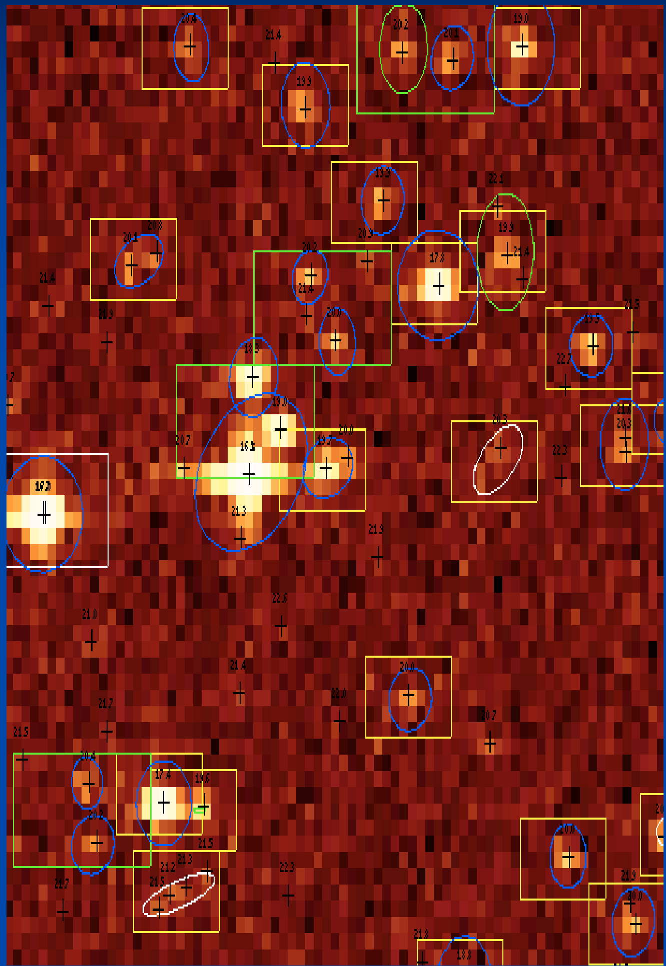
- Instrument modes & optical inhomogeneity in the focal planes
- Real-time and on-board algorithmic constraints conformance
- Validation of star-based attitude control in low stellar densities
- Overall distribution of observations in the two FOV during the mission
- Post-processing: numerical binning, RVS CCD summation

Implementation adjustments

- Exact CCD output data type
- Handling of CCD rows independence
- Combination of the FOV and effects of AC velocities
- Bug fixes (bug reports very welcome !)

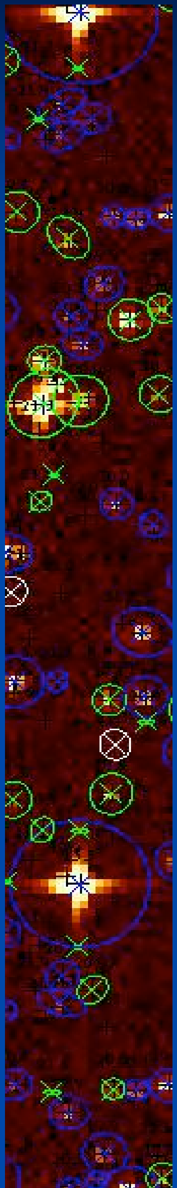
Identified Bugs

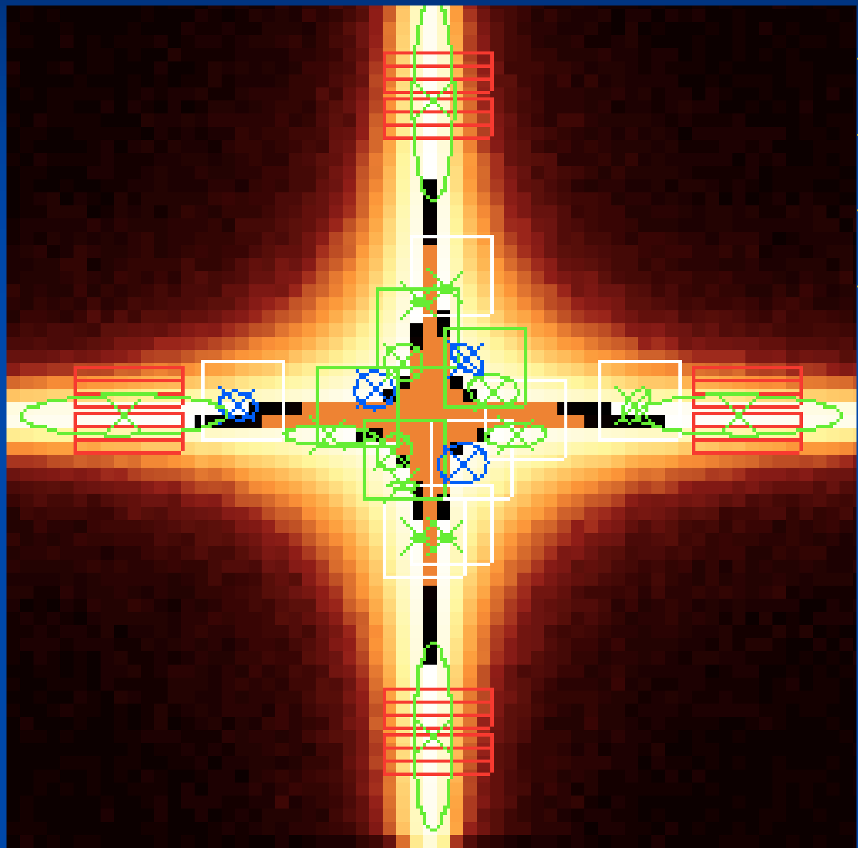




- Patches for extended objects
  - same as stars assumed
- Patches for bright objects
  - based on the detected magnitude, but with limits based on the Gaia-1 design
- Bad sample position
- Bad confirmation rate in RVS
- Memory leaks in all algorithms
- Star leaks in selection…

Example : bright-star bug





- No ASMO in the new design
- Saturation, ghosts, CTE complicate
- 1% of stars but the best astrometry
- Improving the detection/selection depends on realistic simulations

Conclusion

- Several missing features
- Pending questions and issues
  - Adaptive processing (object type, attitude according to stellar density)?
  - Telemetry issue
  - On-board reduction for high priority objects not downloaded, not observed ?

Yet : main ideas / building blocks gradually falling into place
representative of main requirements

Hope for efficient interaction & feedback from the industrial feasibility study.

People in relation to this activity

- People
  - F. Arenou
  - C. Babusiaux (simulation/detection)
  - E. Hoeg (sampling)
  - D. Katz (spectro instrument)
  - S. Mignot (detection/selection)
  - J. Portell (focal plane)

- Contact
  - F. Arenou:

Email:	Frederic.Arenou@obspm.fr
Phone:	(33/0) 1 45 07 78 49
Fax:	(33/0) 1 45 07 78 78