



DATA PIPELINE

UVIT data reduction pipeline: A CCDLAB and UVIT tutorial

JOSEPH E. POSTMA*  and DENIS LEAHY

Department of Physics and Astronomy, University of Calgary, 2500 University Dr NW, Calgary, Alberta T2N 1N4, Canada.

*Corresponding Author. E-mail: joepostma@live.ca; jpostma@ucalgary.ca

MS received 26 October 2020; accepted 19 December 2020

Abstract. For the AstroSat 5-Year Special Issue, we present a tutorial on the usage of the CCDLAB Pipeline for UVIT data reduction from Level 1 raw data to completed science images. The tutorial informs us of the unique data-processing requirements for reducing UVIT data, including one unique development borne out of such reduction applicable to the field of astronomy in general in the form of a novel approach to solving world coordinate solutions.

Keywords. Instrumentation: detectors—methods: data analysis—techniques: image processing.

1. Introduction

The Ultra Violet Imaging Telescope (UVIT) (Kumar *et al.* 2012) is one of the main instruments onboard India's first major orbital observatory AstroSat (Agrawal 2006), launched by the Indian Space Research Organization (ISRO) on September 28 in the year of 2015. UVIT is composed to two co-pointing telescopes comprising three imaging detectors covering the far ultraviolet (FUV, 120 nm to 180 nm), near ultraviolet (NUV, 200 nm to 300 nm), and visible (VIS, 320 nm to 550 nm) wavelength channels. One telescope is dedicated to the FUV detector system while the other telescope utilizes a beam splitter to separate the NUV and VIS wavelengths to orthogonally-placed detectors. The design image resolution in the ultraviolet is approximately one arcsecond.

After a period of degassing with the telescope in a safe mode the system was opened for first light in December of 2015. While the aim of most familiar and common telescope hardware is to provide a stable and precise platform for integration imaging, requiring fine-guidance pointing and the like, AstroSat is commanded to oscillate its pointing on orthogonal

UVIT image axes at a rate of a few arcseconds per second with an amplitude of a few arcminutes. The purpose of this oscillation is to protect the detector components from bright objects, and while such a procedure performed on a typical integrating detector would ruin the image, for a two-dimensional photon counter such as UVIT the nominal image field at instrumental resolution may be recovered by de-shifting and coadding the count centroids as a function of the pointing oscillating. Within UVIT circles we call this oscillation “drift” and the time-sequence of the oscillation the “drift series”, and this series is measured by tracking the positions of point sources in the VIS channel images or alternatively by tracking sources within the FUV/NUV centroids themselves. While the UVIT detectors are photon counters scanning a 512×512 CMOS at 28.7 Hz over a 28×28 arcminute field, the VIS channel is run in a special integrating configuration mode at lower voltage where photon counts are integrated on the chip for 1 s, thus allowing identification of point sources and subsequent tabulation of the drift series at that spatio-temporal cadence.

Given the above consideration, viewing the first light image was not simply a matter of converting downloaded (from orbit) bits directly into some display format. Firstly, UVIT image data is composed of photon count centroids which must be stacked in order

This article is part of the Special Issue on “AstroSat: Five Years in Orbit”

to form an image, and secondly, the centroids must be corrected for the drift series as well as have other instrumental and calibration corrections applied to form science-ready images. Naturally the drift series correction requires precise timing knowledge between the FUV/NUV detector clocks and the VIS detector clock, and those relative to the spacecraft clock, and it also requires software to perform the necessary drift-series tracking and then shift-and-stack operations, etc. Along with other data corrections such as flat field and distortion mitigation, etc., a complete software data reduction package is called a “pipeline”. The first-light image reduced for UVIT is shown in Fig. 1, and a candid photo capturing the moment is shown in Fig. 2.

Initial results and performance evaluation of UVIT may be found in Tandon and Hutchings (2017), in-orbit calibration in Tandon and Subramaniam (2017), and additional calibration may be found in Tandon and Postma (2020).

2. Discussion

In general, a data reduction pipeline for some mission should not be run by end-users, that is, by the scientists who are subscribing for time and the resulting

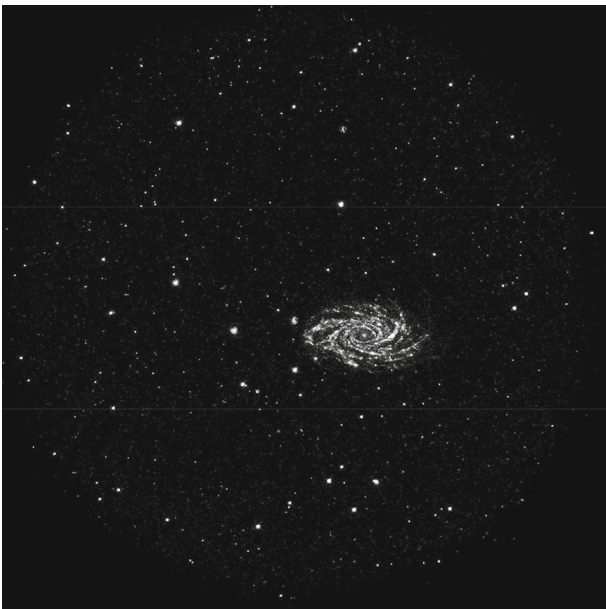


Figure 1. First-light image produced by the CCDLAB pipeline of UVIT in NUV of spiral galaxy NGC 2336. NGC 2336 is approximately 200,000 light-years in diameter, making it twice as large as the Milky Way galaxy, and is approximately 90 million light-years distant (December 18, 2015).

scientific data which they are interested in procuring. Only in the simplest of cases where the reduction is trivial is it reasonable to leave the processing of data to end users, for example if for image data the only corrections required are bias, dark field, background, and flat field. Otherwise, the peculiarity of instrumentation may often grow beyond the scale in ability, interest, and relevance of the end-user to be left to process themselves. That being said, peculiarities of instrumentation and the resulting data reduction may be made trivial by sufficiently well-written pipeline software, although for this to be achievable naturally requires well-behaved data as well increasing expenditure of software developer-hours. In the case where one is confronted with peculiar instrumentation, and complex methods required for the data reduction, and data which is not necessarily well-behaved, then data reduction for a mission should likely be left to the person or small group of people who are themselves responsible for the writing of the reduction pipeline software. Of course, this requires the software developers themselves to have some familiarity with and respect towards the handling of scientific data and its purposes, and it is helpful if such people also have some participation in the development of the instrumentation in the engineering phases. By “well-behaved data” we define the opposite of “poorly behaved data”: data which is prone to novel sources of variation which are unpredictable and which render existing solutions to reduction of data from the same instrument under the same configuration non-functional. Of course, we hope that there is a limit to such poor behavior. In the end, the question of data reduction complexity goes from the trivial case of requiring no user input or knowledge of the process whatsoever, to requiring constant user-input and full-time management of the reduction sequence from raw data input to science data output; UVIT has tended more towards the later of this range.

At this time, five years post-launch, the CCDLAB UVIT pipeline (Postma & Leahy 2017) has been developed to a point where a significant fraction of the data reduction sequence may be run as automated, although options remain in the software program settings to scale the automation back and be run with user-input for each phase instead: there are scenarios of observed image fields or observational peculiarities which sometimes require user intervention. We shall refer to steps in the data reduction sequence as “phases”. There remain several phases of the reduction procedure which unavoidably require user input given that these sequences are dependent upon the

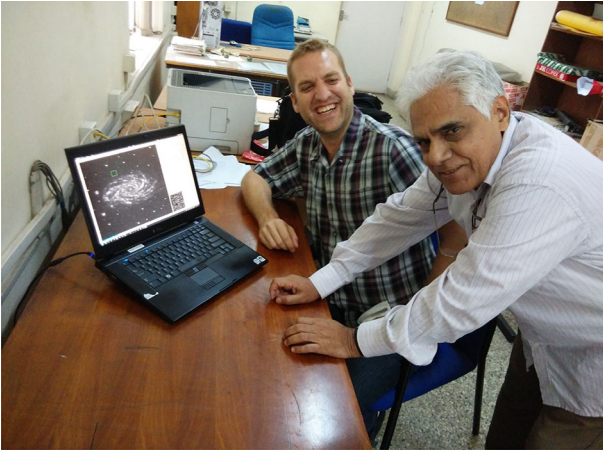


Figure 2. Shyam Tandon (Indian instrument P.I., *right*) and Joseph E. Postma (UVIT technical support, *left*) enjoying the first-light UVIT image (December 18, 2015; Picture location: Indian Institute of Astrophysics, Bangalore; Photographer: Koshy George).

observed image field itself, and there is simply so much variation in target fields that it is not practically possible to develop software to identify and process all scenarios...they are better left to the software of the human mind. These cases will be discussed in the text ahead. What follows is a tutorial on the usage of the CCDLAB UVIT pipeline as run in maximally-automated mode. Automated reduction scenarios which fail, and there are any number of failure scenarios, are best left to the “experts” to reduce with manual management of the sequence. We shall describe each phase of the reduction with discussion towards the instrumental peculiarities which necessitate each sequence.

For reference, the development and data processing machine which the CCDLAB UVIT pipeline runs upon uses an Intel 7th generation chip with 4 physical cores and hyperthreading, supplying 8 processor threads running at 4.4 GHz. It is equipped with 16 GB of DDR4 RAM, and a PCIe 3.0 NVMe 2GB solid state drive. The CCDLAB UVIT pipeline and FITS image processor has been developed in Microsoft Visual Studio C++ over several releases of that product.

2.1 Phase A: Extraction, digestion, and drift series

The general scenario for a given target field for some observation proposal is that it will typically be observed over multiple orbits. UVIT can only observe while in the shadow of Earth, limiting any particular

observation to a maximum of around 1800 s, with typical observations in the vicinity of 1ks, whereas proposals typically request 10 ks to 100 ks observation times. Thus, the Level 1 (L1) data product from ISRO for a given proposal comprises a zip archive containing multiple individual orbit-wise data sets as FITS binary tables. We select the archive file in the usual manner of an “open file dialog” initiated by double-clicking the “Extract L1 zip Archive” item in the CCDLAB UVIT menu as shown in Fig. 3.

A single-click of the zip Extraction menu will bring up the automation options for this phase, where each option is dependent upon the previous (higher) item having been completed. Thus, if the first option is deselected, then all following options are automatically deselected as well, and so on if intermediate items are deselected. Before we describe the meaning of each option, we must first understand the “digestion” processing of the data, the options of which are found in the next menu item “Digest L1 Fits File(s)”, as shown in Fig. 4.

In non-automated mode one would be required to double-click the menu item shown in Fig. 4, and use an “open file dialog” to open the FITS binary table files extracted out of the zip archive from the previous menu procedure. In both automated and non-automated modes, there are a series of options which apply corrections and perform data fidelity checks upon the extracted L1 data centroid files:

- *PC Mode Apply FPN Correction*: This option applies the “fixed pattern noise” correction to the centroid data. As discussed in the engineering development phases (Hutchings et al. 2007; Postma et al. 2011), the weighted-mean of an undersampled photon event or PSF results in a centroid which systematically tends toward the center pixel of the centroid kernel. Because we can centroid photon event PSFs to sub-pixel accuracy using a 3×3 pixel kernel, we gain resolution higher than the instrumental pixel scale itself, however, we must correct for the systematic bias in the centroid calculation of the undersampled PSFs. This effect was calibrated on-ground and the correction tables are part of the UVIT Calibration Database.
- *PC Mode Apply CPU Distortion Correction*: By CPU we refer to the “camera proximity unit” of UVIT, i.e., the main camera and all of its components. Of course, there are three such cameras comprising the visible (VIS), near ultraviolet (NUV), and far ultraviolet (FUV)

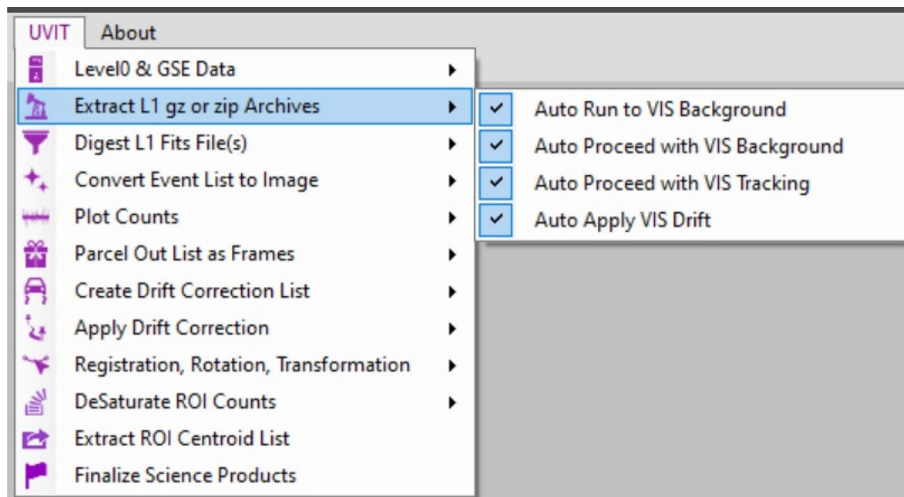


Figure 3. Level 1 data archive extraction.

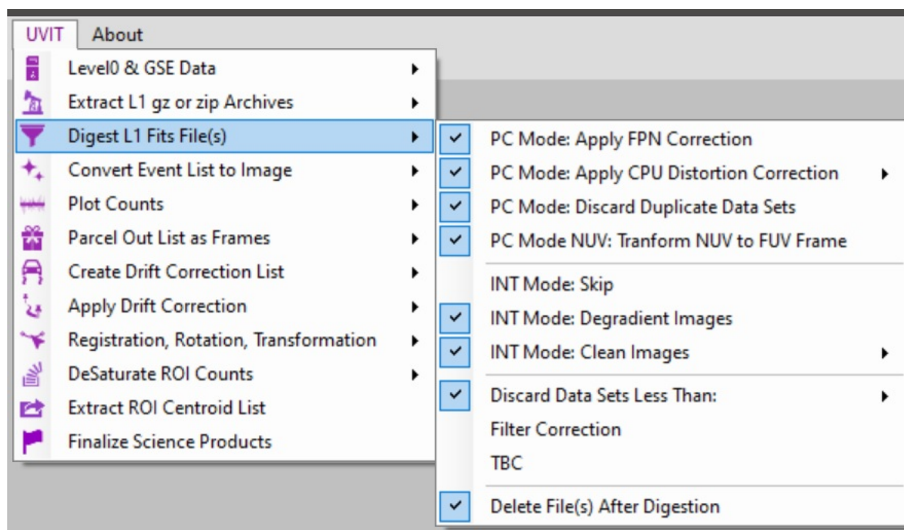


Figure 4. Level 1 FITS digestion.

channels. Systematic field distortion likely arises mainly from the fiberoptic taper channeling photon pulses from the phosphor screen down to the 512×512 CMOS array. Field distortion was calibrated on-ground with further improvements developed from in-orbit data, and field distortions maps are part of the UVIT Calibration Database. A single-click of this item will open options for the interpolation scheme within the map: either no interpolation (CMOS pixel scale only) or bilinear interpolation, with bilinear being the default-selected option.

- *PC Mode Discard Duplicate Data Sets*: Early in the mission it was found that L1 data sets were being provided with metadata which made the

data sets appear to be originating from unique observations; however, many data sets were simply duplicates of previous observations. The reason for this had to do with the onboard data buffer only refreshing itself after a certain amount of data collection, whereas the entire buffer is repeatedly downloaded to ground for processing into L1 format. Recent L1 data no longer suffers from this duplication issue as software mitigation has been applied at the L1 level, but early L1 data may still suffer from it. It is a relatively trivial fidelity check which consumes little computation time, and so it is good to leave this option selected.

- *PC Mode NUV Transform NUV to FUV Frame*: The NUV camera shares a telescope with the

VIS channel, where the wavelengths are directed to orthogonally-placed cameras by means of a beam splitter differentiating the visible from near ultra violet wavelengths. This inverts the NUV field relative to the FUV and VIS fields, and additionally due to mounting orientation the NUV field is rotated by approximately 32 degrees relative to those channels. Thus, this option transforms the NUV centroids through an inversion and rotation matrix such that the NUV centroids nominally share the same field orientation as the FUV and VIS fields.

- *INT Mode Skip*: This option would typically only be used when a user is investigating problems with PC mode data. The VIS images require the most time for extraction given that typically $\sim 10^4$ images are available to be processed for a given target campaign, comprising many Gigabytes of data, whereas the FUV and NUV centroid lists are typically less than 100 MB each.
- *INT Mode Degradient Images*: The VIS drift-tracking data are full-frame reads of the CMOS array with an integration time of 1s, and the line-scans generate a relatively uniform horizontal gradient of several hundred ADU's on top of a bias of approximately 1400 ADU's, and so the gradient is significant. Sufficient uniformity in the horizontal gradient allows for its correction by the subtraction of the median of each column of the CMOS array from each column as such. Correcting this gradient assists in the identification and selection of sources in the images to use for drift tracking, and assists in the reliability of the subsequent source-tracking routine.
- *INT Mode Clean Images*: The VIS channel detector system frequently develops artefacts due to the various effects of in-orbit space radiation upon the detector system hardware. These artefacts manifest as multiple well-separated bright lines running horizontally across half of the image field, although the artefacts are inconsistent in placement, and degree, for any particular appearance of them. This menu item has options for thresholds to detect these artefacts with default values supplied. The algorithm is such that, as the VIS images are extracted out of the FITS binary tables into individual images, they are scanned row-by-row for a given number of bright pixels above the given threshold; if a bright line artefact is

detected as such, then the offending pixels in the row are replaced by the average of the pixels above and below the given line. These artefacts can render VIS image data unusable for their purpose of drift-tracking, given that sources can drift over these line artefacts and the artefacts are typically much brighter than the sources. The solution described here mitigates this problem.

- *Discard Data Sets Less Than*: This item opens a drop-down list to select a value specifying the number of minutes under which an exposure should be discarded from the reduction. There are frequent 30 s observations which are performed simply for brightness-safety checks for the detectors before a full exposure is commanded, but such short exposures have too-low of SN to be accurately combined in registration with other science exposures. The default value is 2 min.
- *Filter Correction*: Early in the mission there was difficulty in aligning the time-position of the filter wheels with the time of the centroid image data, because these two systems use different electronics hardware. The detector system electronics unit has its own internal clocks, whereas the filter wheel system is a completely separate piece of hardware with its own clocks. So-called "housekeeping" metadata files are supplied in the L1 archives which serve to correctly align which filter was being used for a particular observation. This option generally is not required any longer at this point as the correction now occurs at the L1 creation level, but for older data it is still sometimes required.
- *TBC*: This refers to "Time Bit Correction" which mitigates a stuck 20th bit in the clock on one of the UVIT channels. The mitigation for this problem is now corrected at the L1 level, and so is not required.
- *Delete Files After Digestion*: This option deletes the FITS binary table files after the FUV/NUV centroid and VIS image data has been extracted and digested. It is somewhat redundant because CCDLAB will ask the user if they wish to delete all intermediate processing files at the very last step later when the science images are finalized, although it helps to reduce disk space usage at intermediate processing steps given that the intermediate-processing data files can grow to order

10^2 of Gigabytes. The original zip archive L1 file will never be deleted by CCDLAB.

When CCDLAB is installed all optimal default options will be preselected. We may now return to the “Extract L1 zip Archive” menu item of Fig. 3 which will initiate the extraction, digestion, and much of the processing of the data upon double-click in automated mode when all sub-menu items are selected. We shall describe the effect of each item:

- *Auto Run to VIS Background*: This option will have the CCDLAB pipeline run through the extraction of the zip file, collating all extracted orbit-wise data into individual directories and sub-directories parented by the directory of the selected zip file, perform all data-fidelity checks and instrumental corrections as explained under the “Digest L1 Fits File(s)” menu item description, and will automatically determine the VIS-channel background image to use for subtraction from the VIS image drift-tracking data. Typically this sequence of the phase requires several minutes.
- *Auto Proceed with VIS Background*: This option will have the pipeline automatically proceed with the subtraction of the VIS background from all VIS image data. Typically there are order 10^4 of VIS images which require background correction, and the process requires several minutes.
- *Auto Proceed with VIS Tracking*: This option will have the pipeline process each orbit-wise directory of VIS tracking images, automatically determining sources in the initial image and then tracking their centroids through the sequence of images following for each orbit. This is the first non-trivial area where it is possible for the automated mode to fail: failure would be found in the VIS image fields simply not having any good sources to use for automated tracking, in which case the user would be required to manually select very faint sources and observe when and where failure in tracking occurs and then remember to not use those sources in following attempts; this may occur in a few percent of observations. This phase displays the tracking sequence graphically on the CCDLAB image window, showing the paths for each source being tracked as in Figure 5. The automated drift-tracking algorithm will attempt to track as many sources as it can, and it will discard poor tracks and subsequently-untrackable sources “on the fly” as it runs

through the orbit-wise image sequences. The drift series are determined as differentials from the initial position for each source, and thus the multiple source tracks can be merged at the end of the process as a mean.

- *Auto Apply VIS Drift*: This option will apply the drift series from all orbits determined in the previous step to all orbits of FUV/NUV data. A drift series typically takes on the forms for the x - and y -axes as shown in the plots of Fig. 6. In Fig. 6 there occurs an apparent delta-function in the drift series just after the half-way mark on the x -axis; this originates from one of the several individual sources contributing to the series being skewed by some noise at that instant, and such variations are mitigated by taking a “robust mean” of the multiple drift series from the multiple sources tracked. By “robust mean” we mean an iterative average where the values of a sequence which exceed 3 standard deviations of the sequence relative to the sequence’s mean are replaced by the median of the sequence, until no values of the sequence exceed 3 sigma from the new mean.

The exposure map for the image of the centroid list is created at the stage of the application of the drift series for the given centroid list. The exposure map is created by following-along the pointing of the telescope as measured by the drift series, where the active field-of-view of the detector, measured on ground from flat-field calibration, is moved about the padded field of view as unity pixel values summed into the exposure map for each frame read of the observation.

At the completion of the application of the drift series we have reached the limit of full-automation reasonable to develop for the pipeline. At this intermediate stage there will be drift-corrected images along with their exposure maps and other intermediate data files in each orbit-wise subdirectory for each channel and filter, and the images will be displayed as an image-set to the user in the main CCDLAB window. If only a single orbit was ever observed for a target, then only a single image will exist and the user can skip ahead to Phase D. Otherwise, multiple orbits must be registered together to align the fields. A typical proposal and its L1 data would have required ten minutes to reach this point.

Given that images have been created at the end of this sequence, we should mention the image-

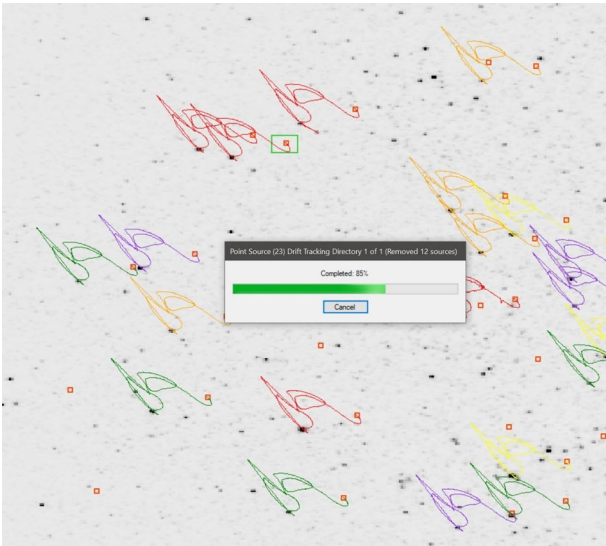


Figure 5. CCDLAB displays the paths as each source is tracked in the drift series.

creation options available in the menu item “Convert Event List to Image”, as shown in Fig. 7.

- **Filter Cosmic Ray Frame:** The option to filter cosmic rays will remove frames from the centroid list which exceed a certain specified number of counts in the frame, given that a nominal frame should only contain one or two dozen of photon events from real sources, whereas a cosmic ray event in the frame will generate many tens to thousands of events as a “splash”. This menu item presents options to set a cosmic ray frame detection threshold either by total count within frame, or by the number of standard deviations above which the number of counts in a frame should qualify a frame as

likely containing a cosmic ray; the latter is likely the safer option to use since this accounts for variation in background. This cosmic ray filtering option is never used and the frequency of cosmic ray “splashes” in the centroid list simply forms a part of the nominal background. However, the option could be used for improving the signal to noise ratio for faint sources; for example, the background in the M87 region reduces by approximately 15% (from 2×10^{-4} c/pix/s) in NUV when using a 4-sigma threshold, and likewise and significantly for FUV the background reduces by 65% (from 2×10^{-5} c/pix/s). If this option is used and cosmic ray “splash” frames are removed from the centroid list, the final integration time for the image is appropriately adjusted given the number of frames removed.

- **Apply Max-Min Threshold:** Each centroid comes with a “diagnostic” of the maximum-corner pixel minus the minimum-corner pixel of the 5×5 pixel kernel centered on the peak pixel of the photon event. By limiting this range as a threshold this option will ignore centroids which have potential “contamination” from coincident events. This option is never used for science purposes.
- **Centroid Image Padding:** This option pads the nominal field of view so that drift correction can move centroids into regions which may have exceeded the nominal instrumental field-of-view. A default option of 44 pixels around the 512×512 CMOS array is selected. The exposure map uses the same padding.

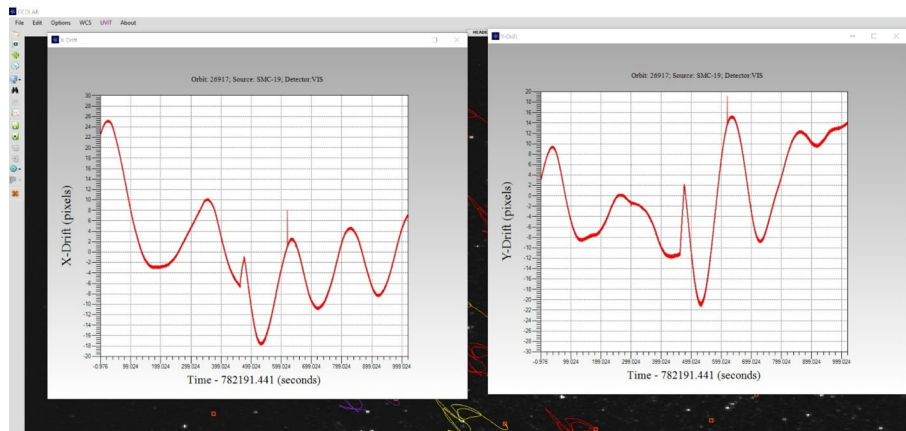


Figure 6. An example of the drift series in x - and y -axes. The drift series for each source are tabulated as a differential from their initial positions, and the plots here represent several source drift series plotted together in overlay.

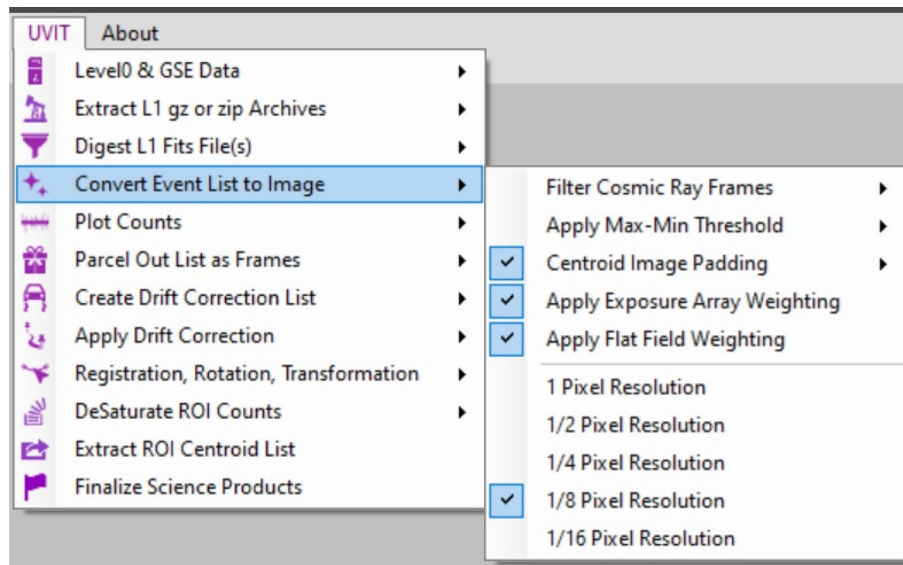


Figure 7. Menu for converting the centroid lists to images.

- *Apply Exposure Array Weighting:* This option scales each centroid by its location within the exposure map such as to normalize the image to a uniform exposure time. The exposure map is applied here much like a flat-field is applied, although nominally most of the field is uniformly observed and only the periphery of the field where the drift moved the sky in and out of the field-of-view receives a non-unitary scaling correction. The exposure map is included with the final science image and so the correction may be removed (by multiplying-in the exposure map) in order to get observed total counts, etc. The exposure map originates with the active field of view of the detector and thus it also captures any “bad pixels” as determined in ground calibration.
- *Apply Flat Field Weighting:* This option provides each centroid a weight, nominally of unity value, but with small variations given the flat field for each detector and filter and the original location of the centroid on the image.
- *Pixel Resolution:* Science images are finalized at 1/8th CMOS pixel resolution, although the intermediate phase of orbit-wise image registration typically uses 1/4th pixel resolution which is set in the registration menu item to be discussed in Phase B.

Final images are in total counts, and the final exposure time in seconds is given by the header keyword “RDCDTIME”.

2.2 Phase B: Registration of orbit-wise images

Generally, we will have a series of nominally-drift-corrected orbit-wise images, and due to drift these images are almost always at some unique translational offset at the scale of a few arcminutes relative to each other. Additionally, rotation of the field may enter between (but not within) orbits, thus requiring a rotational transformation as well as the translational one in order to align the fields to a common frame. Naturally a transformation matrix must be applied to the centroids in order to thus align their image fields, and therefore requires a precise determination of the translation and rotation parameters for each field. This task is accomplished with user interaction and although it is perhaps theoretically possible to code an automated routine here, the complexity of such an algorithm has seemed to the developer to exceed the simplicity of the interactive approach.

Registration is accessed through the “Registration, Rotation, Transformation” menu item under the main UVIT menu, and is initiated by double-clicking the “General Registration” submenu item as shown in Fig. 8. The option to “Masterize Singles” is a simple housekeeping option that will format filenames and structure subdirectories as “master files” when there is a single orbit-observation for a given channel and filter combination. The option to “Folder Browse Scan for Most Recent XYInts Lists” will provide a folder browser dialog to select the channel and filter directories which one wishes to register, otherwise the user

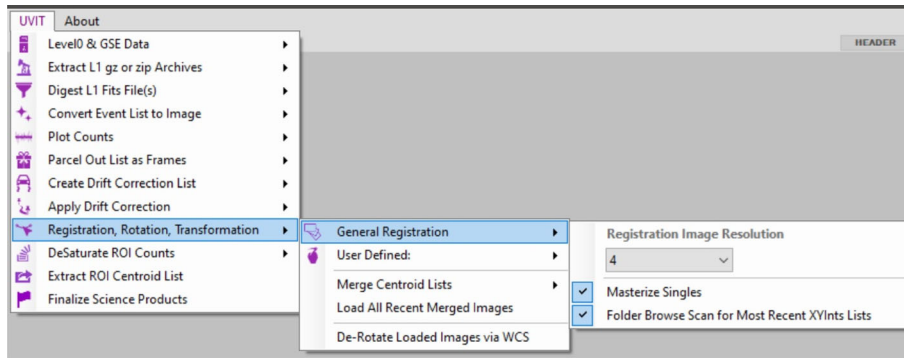


Figure 8. Menu for orbit-wise registration of images.

must use an open file dialog to then manually search for the specific centroid files which they wish to register. The numeric dropdown item is nominally set to 4, indicating $\frac{1}{4}$ pixel resolution at which to register the images, and this can be set to 2 ($\frac{1}{2}$ pixel) or 1 (full pixel) in scenarios where the increased signal from coarser binning is required in order to identify sources.

We must note that at this point the data files are structured in subdirectories such that the original location of the L1 zip archive is the parent directory, and under this parent are FUV, NUV, and VIS subdirectories. The VIS directory contains subdirectories of all orbit-wise image sets, whereas the FUV and NUV directories contain subdirectories first for a given filter, and then within each filter directory are orbit-wise subdirectories containing the centroid list data sets and their drift-corrected images. Thus, there are several options for the manner in which the registration of all orbit-wise images may be approached.

All available images will currently be loaded into CCDLAB as an image set and the user may blink through them making note of corresponding sources across the image set (either mentally, or with CCDLAB by marking the sources via a right-click on the image window), where each image will have some translational and possible rotational offset relative to the others. The results of this visual scan of the images are the following possibilities:

- (i) There are obvious corresponding sources across all orbit-wise channel-filter images. In this case, the parent directory for all files can be selected with the aforementioned folder browser dialog, and all images can be registered as a common set in this way.
- (ii) There are *not* obvious corresponding sources across all orbit-wise channel-filter image combinations, due to astrophysical and instrumental

differentiation of the brightness of sources detected. In this case, one must typically select and process only the FUV and NUV directories separately as there always is enough correspondence for a given channel across all of its filters to be able to identify common sources for registration. In this case a final registration to a common frame for all channel-filters will occur later with the orbit-merged centroid images in Phase C.

When the registration procedure is started, the user will receive simple instructions from CCDLAB to use the cursor to select sources in the initial image field. Registration is an iterative process and the procedure may be repeated as needed. For example, if there is no field rotation between the orbits then it is sufficient to select only a single source to track the translational shifts between orbits. If a user selects two sources then the registration will compute both translational and rotational shifts, and if three or more sources are selected then the rotation will compute a full 2D transformation matrix in order to effect the rotation and translation between orbits. It should be noted here that while a rotation and translation transformation should be sufficient for a given channel-filter, small residuals in the accuracy of the distortion maps as well filters which generate unique distortions and scale offsets (particularly for NUVB15 which has unique residual distortions relative to the other channel-filters on the scale of ~ 2 arcseconds) benefit in the application of a full 2D transformation in order to align all channel-filters to the same scale and field rotation.

If the first registration iteration uses only a single point, and upon the translation transformation it is then witnessed in the CCDLAB image window when blinking through the images that some rotation exists, then the user may iterate the process again and select more points for transformation. The first point that the

user selects will become the “anchor point”, and then the secondary points may be “grabbed” by the cursor in order to rotate them about the anchor into the position aligned with their sources in subsequent images after the first image has been used for the point selection.

If all images across all channel-filter combinations could have been registered in this way then registration will be finished at this point, however, if the orbit-wise NUV and FUV centroid images could only be registered separately at this point then registration will be iterated a final time after the merging of orbit-wise fields in Phase C.

2.3 Phase C: Merging channel-filter orbit-wise data

At this phase, we assume that any given filter-channel orbit-wise data will be aligned via registration. Thus, we simply merge the orbit-wise data into a “master” data file where all centroid data of contributing orbits are merged into a master list, from which a master image may then be produced. The exposure maps must also be correctly merged in this process so that the total exposure time may be correctly evaluated for regions of the periphery of the final images, and this is handled internally. Merging is initiated by double-clicking the “Merge Centroid Lists” menu item as shown in Figure 9, under the “Registration, Rotation, Transformation” menu, and the option to “Delete Contributing Directories” cleans up the previous intermediate orbit-wise data subfolders from each channel-filter directory during the merge, and the “Folder Browse Scan for Most Recent XYInts Lists” option allows the user to have to only select the parent directory of the FUV and NUV folders instead of searching for the data files manually in order to begin the merge.

After the merge one will have a “master” data set for each channel-filter observed for the proposal. If all fields were able to be registered previously then the registration will remain and the user may proceed to Phase D. Otherwise, if the FUV and NUV fields were not able to be registered previously due to large differentials in source detection, then with the increased signal-to-noise of the merged files, having typically ten-times the total exposure compared to a single orbit-observation, one will now be able to identify common sources such as to effect the final registration for all master channel-filter fields, and such registration may again be performed iteratively if required.

2.4 Phase D: Optimizing the PSF

Due to the drift-series tracking occurring at a sampling rate of 1 Hz, and the occasional occurrence of the movement of the Scanning Sky Monitor (SSM) camera on the spacecraft inducing high-frequency differentials in the drift track, it is not possible for the VIS drift series to be a perfect representation of the telescope’s pointing oscillations. Additionally, we have found a thermal stick-slip inducing a small differential in the pointing of the FUV telescope relative to that of the VIS telescope at the CMOS pixel scale which is much larger than the instrumental resolution of 1/3rd pixel; the stick-slip is shown in Fig. 10. The effect upon the point sources when this phenomenon manifests is to extend a point source into a small “streak”.

To correct this residual in the drift series, the user may optimize the PSF of the images via the menu item as shown in Fig. 11. This procedure may be run for images which do not have a noticeable problem with the PSF and such images will also benefit in significant improvement of the profiles. At this point the merged data set images will all be loaded into

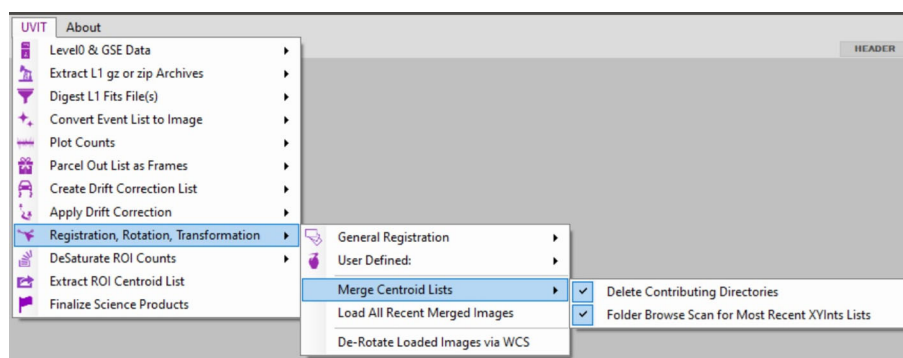


Figure 9. Merging orbit-wise centroid lists.

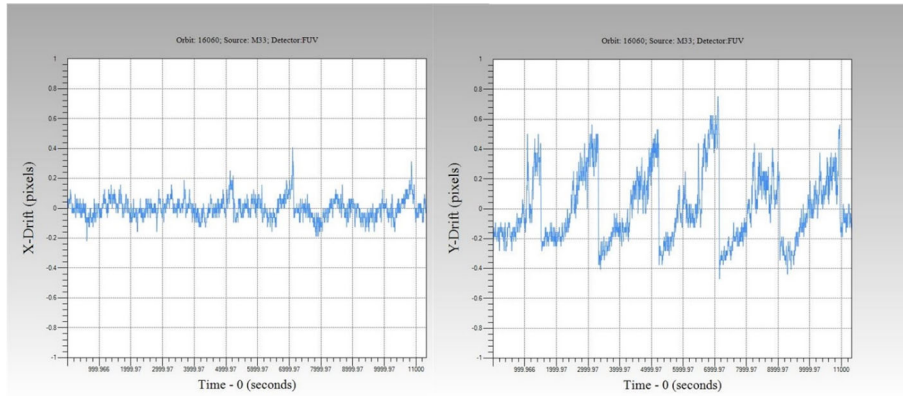


Figure 10. Plot of thermal stick-slip which causes a differential in the pointing between the VIS and FIV telescopes.

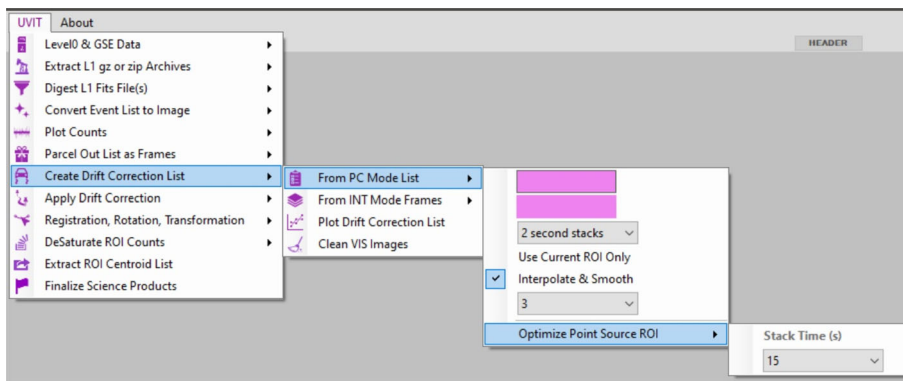


Figure 11. Optimizing the point source spread function.

CCDLAB for viewing, and with the cursor the user may move the region-of-interest sub window over the brightest sources in the image and right-click the mouse to mark these source coordinates. If multiple sources are selected, then double-clicking the “Optimize Point Source ROI” button as shown in Fig. 11 will have CCDLAB automatically determine the best solution to optimize the PSF for all sources. If only a single source is selected, then the user must specify the “Stack Time” via the drop-down submenu, and then initiate the procedure; the stack time specifies how many seconds to use for sampling the source centroids for averaging their position within the sub-window. The region of interest subwindow size should be set such that it just contains the PSF of the brightest source...typically 11×11 science-image pixels.

It is important to make a note here about the optimization of PSFs. If the metric for optimal PSF is either the narrowest FWHM, or a maximized peak value, then the best result is simply found in correcting all centroids from a given source PSF to fall within a single 1/8th pixel bin, i.e., in correcting the

PSF spread of centroids for a given source into a delta function. But then what is the effect of doing this to all other centroids and sources to which those corrections would be applied across the image? The instrumental PSF is produced by ostensibly random effects, and thus, correcting a single source into a delta-function induces a convolution of its PSF into all other sources, thus degrading the resolution for the rest of the image. That is, a single source would be artificially narrowed into a delta-function to produce its “best PSF”, but the effect of the corrections required to perform this operation on the single source applied to the rest of the centroids of the image would be to convolve the PSF of the single source into all other centroids, thus degrading the resolution for the rest of the image. This is why we cannot automatically apply the corrections for the “best PSF” of a single source across the entire centroid list, and why multiple sources should be used to determine systematic effects of residual drift across those sources. A stack-time for a single source of “20” (seconds) would thus be a good choice for cases where only a single point source is available.

2.5 Phase E: World coordinate solution and de-rotation to sky coordinates

Given that there are inter-orbit variations in the field rotation aspect of the telescope, likewise the final merged master images will be rotated at some random value relative to sky coordinates. We also wish to have a world coordinate solution (WCS) in any case so that locations in the image may be mapped to catalogue and sky coordinates. Given that we are handling centroid data, we may derotate the image field via the photon event centroid list in order to align the image axes to sky coordinates as necessary once we have a determination of the field rotation from the WCS. CCDLAB implements the trigonometric algorithm as described in Postma and Leahy (2020), and this is accessed through the WCS menu item on the main CCDLAB menu bar as shown in Fig. 12. The trigonometric algorithm is an entirely novel solution to the problem of solving World Coordinate Systems, and is borne specifically out of the problem of solving WCS for UVIT images in the FUV; the algorithm

however is generally applicable to any image in astronomy and can determine solutions instantaneously in most cases.

CCDLAB has also implemented “astroquery” (Ginsburg *et al.*, 2019) via Python script kindly supplied by Dr. Eric Rosolowski of the University of Alberta, and this menu item is shown in Figure 13. The user may simply double-click the “AstroQuery” menu item in order to download the GaiaDR2 catalogue file relevant to the region of the image. The catalogue region to download for UVIT images is given by the header keywords “RA_PNT” and “DEC_PNT”, the field radius is specified as 17 arcminutes, and the user may specify either a circular or square region for the catalogue query (as shown in Fig. 13); these specifications may be modified within the menu item for other images with different header keywords and field sizes, etc. The RA and Dec key values may be in either numeric degree format or as textual sexagesimal format. The trigonometric algorithm is not an absolute blind solver (at this point) and so it does require a catalogue region specification of coordinates within $\sim 1/2$ of the field center with respect to the field width, although more accurate specifications improve performance of the trigonometric algorithm.

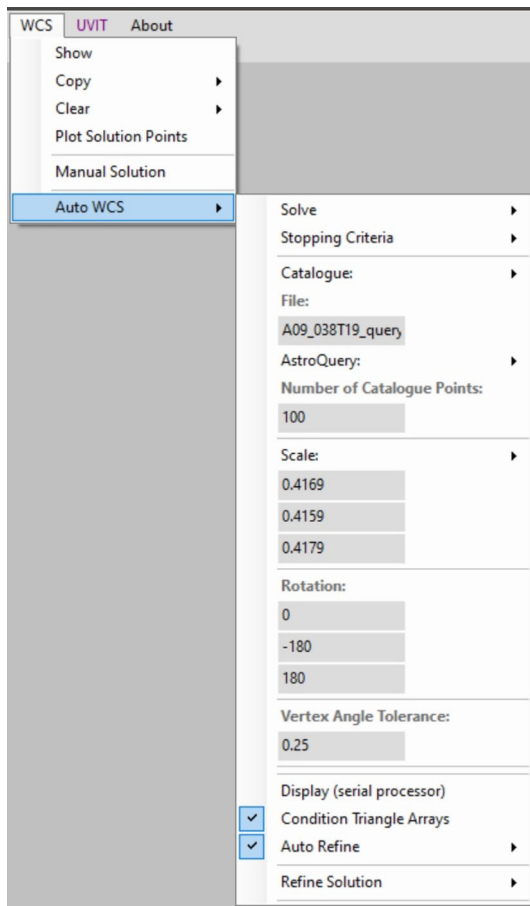


Figure 12. CCDLAB implements an automatic world coordinate solver.

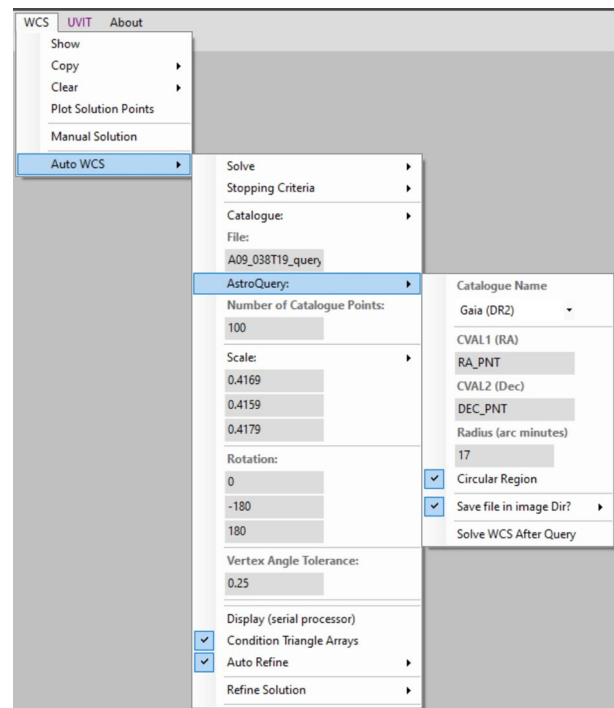


Figure 13. Astroquery is used to download the GaiaDR2 catalogue information for an image region.

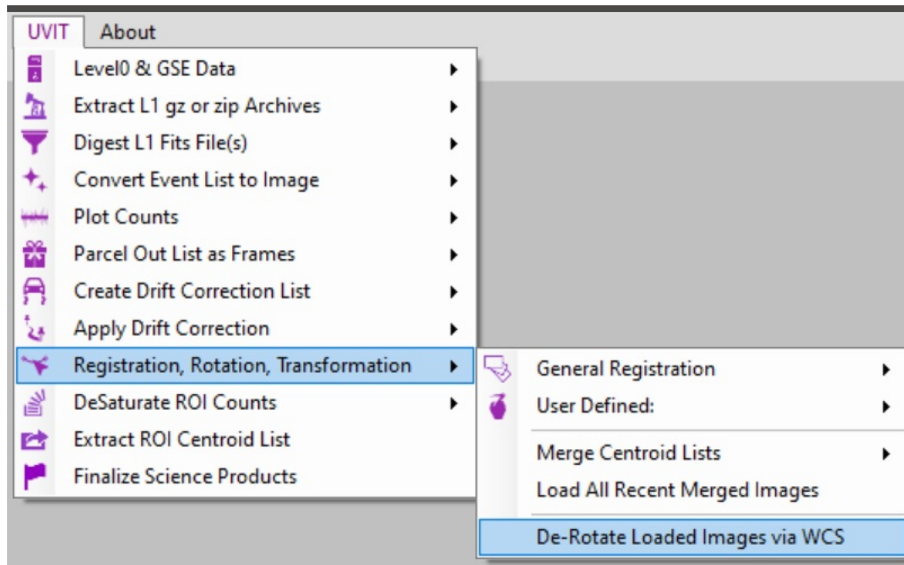


Figure 14. The images may be de-rotated to align the axes with sky coordinates.

Once astroquery is finished, the user may then click the “Solve” menu item and the WCS will be determined by the algorithm. Thus, at this phase, CCDLAB should have all channel-filter master images loaded for viewing and they should all be aligned from registration, and the user may then use the WCS menu to solve the WCS solution for the set of aligned fields. Please note that the WCS solution will be computed for the current image being viewed only, and given that we are querying the GaiaDR2, the image which the user should choose to determine the solution for should be the filter which is closest to visible wavelengths.

As a final correction to the data, we may derotate the image field based on the initial solution of the WCS. This task is initiated by selecting the menu item as shown in Fig. 14, “De-Rotate Loaded Images vis WCS”. This procedure will use the WCS solved for the currently-viewed image, and use the solved field rotation therein to de-rotate all of the fields such as to align the axes to sky coordinates, i.e., vertical is increasing declination and leftward is increasing right ascension. A new WCS for the derotated field will automatically be computed and this WCS will be copied into all of the other headers of the final images.

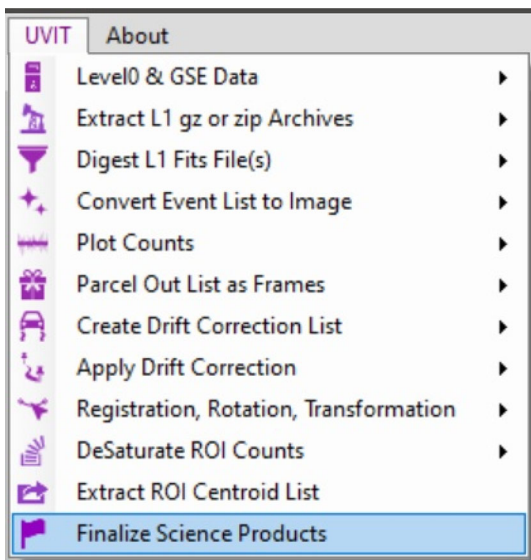


Figure 15. Menu item to finalize science image products.

2.6 Phase F: Finalize science products

The last step is to package the final science image files for distribution to end-users, and to clean up all intermediate processing folders and files from the computer system. If the user clicks the “Finalize Science Products” menu item as shown in Fig. 15, then CCDLAB will create a zip file in the parent directory containing the final science images along with their exposure maps, with appropriate naming of the files such as to distinguish them. The units of the science image will be in total counts, its exposure time is given by header keyword RDCDTIME, and the exposure map is normalized to RDCDTIME so that most of the exposure map is unit value. The user will be asked if all intermediate processing files should be deleted, and if so, the original L1 zip archive will still

be allowed to remain so that the source data is not lost and could be re-processed without having to download it again in the future if the need arises. The “intermediate files” are the centroid lists which include lists of the centroids, their solar-system barycentric Julian dates, flat-field and exposure weighting, and their detector frame numbers and frame times. These may be of interest if one wishes to examine temporal variations within the span of a single orbit, but otherwise the orbit-wise images of approximate ten-to-twenty-minute integrations and their mean time of observation may be taken as data points for light curves, etc.

The FITS image headers list “BJD0” and “MEANBJD” as the Barycentric time of the start of the image and the mean Barycentric time of the image, in Julian Days. This is done by converting the detector clock times of the centroids to UTC via Level1 “housekeeping” files provided inside the L1 archives, and then converting UTC to geocentric Julian Day and then correcting to Barycentric time. The algorithm is listed in Appendix A.

3. Conclusion

We have presented a tutorial for the processing of Level 1 data into final science image products. A video-tutorial may also be watched at this YouTube address: https://www.youtube.com/watch?v=4_48yRcN3nc.

Also, CCDLAB may be installed on Windows by downloading from this address: <https://github.com/wer29A/CCDLAB/releases>.

The UVIT Calibration Database may likewise be downloaded from: https://drive.google.com/file/d/1dD4R7qvsW7Eny93AqgD0IE1weWTAHj0_/view?usp=sharing.

If any user desires more information, source code, or assistance with CCDLAB, please contact the authors.

Acknowledgements

This publication uses the data from the AstroSat mission of the Indian Space Research Organisation (ISRO), archived at the Indian Space Science Data Centre (ISSDC). This publication uses UVIT data processed by the payload operations centre at IIA. The UVIT is built in collaboration between IIA, IUCAA, TIFR, ISRO and CSA. We thank the referee(s) for their review.

Appendix A: MATLAB algorithm for determining barycentric Julian Date

```
function result = HJDC(JD,lat,long,ra,dec,varargin)
% Usage: [nx3]answer = HJDC(JD,lat,long,ra,dec);
% Result is returned in arbitrary “answer” variable
% as a nx3 row-major
% matrix, where n is number of JD input values. 1st
% column is Barycentric
% Julian Day corrections, 2nd column is Barycentric
% radial velocity
% corrections, 3rd column is airmass.

% Input:
% JD = full Geocentric Julian Day day.day
% lat = latitude of observation deg.deg
% long = west longitude of observation deg.deg
% RA = right ascension of target hours.hr
% dec = declination of target deg.deg

%DEFINITIONS
vrot_eq = 465.1;
% earth equatorial rotational linear velocity in m/s,
% based on spherical
% earth using quadratic-mean (polar-equatorial)
% radius; can be improved to
% take into account non-sphericity and geographical
% elevations, but these
% are 2nd order corrections at best.
au = 1.49597870e11;
% astronomical unit (m)
cs = 173.14463348;
% speed of light (au/d)

% BEGINNING OF COMPUTATIONS
% west longitude of observatory in hours:
L = long/15;
% Greenwich Mean Sidereal Time at JD:
GMST = rem(18.697374558 + 24.06570982441
908*(JD - 2451545.0),24);
% Local Sidereal Time at JD and longitude:
LST = GMST - L;
%local hour angle of target:
ha = LST - ra;
ha = ha*pi/12;
%latitude in radians:
lat = lat*pi/180;
%right ascension in radians:
ra = ra*pi/12;
%declination in radians:
dec = dec*pi/180;
%altitude:
```

```

alt = asin(sin(lat).*sin(dec)+cos(lat).*cos(dec).*
cos(ha));
% true zenith angle; don't care about stuff below
horizon:
zt = (pi/2 - alt); zt(zt > pi/2-pi/2/50) = NaN;
A = ( 1.002432*cos(zt).^2 + 0.148386*cos(zt) +
0.0096467 ) ./ (cos(zt).^3 + ...
0.149864*cos(zt).^2 + 0.0102963*cos(zt) +
0.000303978);
% Airmass of target: Young, A. T. 1994. Air mass
and refraction. Applied
% Optics. 33:1108–1110.

% exact decimal day number from J2000.0 UT 12hr:
n = JD-2451545.0;
% mean anomaly, in radians, at day number n:
g = rem((357.528 + .9856003.*n).*pi/180,2.*pi);
% mean longitude, in radians, at n:
L = rem((280.46 + .9856474.*n).*pi/180,2.*pi);
% ecliptic longitude, in radians, at n:
lam = L + 1.915.*pi/180.*sin(g) + .020.*pi/
180.*sin(2.*g);
% ecliptic obliquity, in radians, at n:
eps = 23.439.*pi/180 - .0000004.*pi/180.*n;
% distance of earth from sun in au's at JD:
R = 1.00014 - 0.01671.*cos(g) - 0.00014.*cos(2.*g);
% rectangular coordinates of earth wrt solar system
barycenter referred to
% equinox and equator of J2000.0, in au's:
X = -R.*cos(lam);
Y = -R.*cos(eps).*sin(lam);
Z = -R.*sin(eps).*sin(lam);
% first deriv's of XYZ above, wrt time in days (au/d).
Note: deriv's of
% XYZ w d/dt eps ~0:
Xdot = .0172.*sin(lam);
Ydot = -.0158.*cos(lam);
Zdot = -.0068.*cos(lam);
% rv in direction of target due to earth's rotational
motion, +ve away, m/s
rv_rot = cos(lat).*cos(dec).*sin(ha).*vrot_eq;
% rv due to earth's orbital motion, +ve away au/d:
rv_orb = -Xdot.*cos(ra).*cos(dec) - Ydot.*sin(ra).*
cos(dec) - ...
Zdot.*sin(dec);

```

```

rv_orb = rv_orb.*au/86400; %convert to m/s
RVC = rv_rot + rv_orb;
RVC = round(RVC.*100)/100; %rounded to 2 dec-
imal place = 0.01 m/s.
% Radial Velocity due to earth's rotation and orbital
motion referred to
% barycenter, in direction of target, +ve away.
BJDC= 1/cs.*(X.*cos(ra).*cos(dec) + Y.*sin(ra).*
cos(dec) + Z.*sin(dec));
% Barycentric Julian Day Correction. ADD this to
geocentric input JD
% Otherwise known as Heliocentric JD (HJD) but
there is ambiguity here.
% Difference between HJD and BJD << BJDC
(~ 1s/~ 5min) in almost all cases,
% so only important for very-high precision timing
if ~isempty(varargin)
    if varargin{1} == 'X'
        result = A(:);
    end
    if varargin{1} == 'H'
        result = BJDC(:);
    end
else
    result = [BJDC(:) RVC(:) alt(:)];
end

```

References

Agrawal P. C. 2006, *Adv Space Res* 38, 2989
Ginsburg A. *et al.* 2019, *AJ* 157, 98G
Hutchings J. B., Postma J., Asquin D., Leahy D. 2007, *PASP* 119, 1152
Kumar A., Ghosh S. K., Hutchings J. *et al.* 2012, *Proc SPIE* 8443, 84431N
Postma J., Hutchings J. B., Leahy D. 2011, *PASP* 123, 833
Postma J. E., Leahy D. 2017, *PASP* 129, 115002
Postma J. E., Leahy D. 2020, *PASP* 132, 054503
Tandon S. N., Hutchings J. B. *et al.* 2017, *J. Astrophys. Astr.*, 38, 28
Tandon S. N., Postma J. *et al.* 2020, *AJ* 159, 158
Tandon S. N., Subramaniam A. *et al.* 2017, *AJ* 154, 128