CORONAL DIAGNOSTIC SPECTROMETER
SoHO

CDS SOFTWARE NOTE No. 46

Version 2 19 August 1998

Missing Pixels and Cosmic Rays

William Thompson
NASA Goddard Space Flight Center
Code 682.1
Greenbelt, MD 20771, USA

William.T.Thompson.1@gsfc.nasa.gov

S. V. H. Haugan
Institute of Theoretical Astrophysics
University of Oslo

S.V.H.Haugan@astro.uio.no

Peter Young
DAMTP, Cambridge University
Silver Street
Cambridge CB3 9EW
UK

P.R.Young@damtp.cam.ac.uk

1 Overview

Several routines that handle missing pixels and cosmic rays in CDS data are available. Some are generic, while others are specific to CDS. The two topics are related because routines which handle cosmic rays first convert them to the missing pixel flag value. Thus, we will first examine those routines which handle missing pixels.

2 Missing pixels

In the CDS data structure[**?**], missing pixels are given a special missing pixel flag value, which is stored together with the data.[1] This missing pixel flag value is then used in various places in the CDS analysis software, for example in the calibration of the data.[**?**]

One place where missing pixels are used is in the SERTS image display software[**?**]. In these routines, the keyword MISSING is used to pass in the missing pixel flag value. Optionally, one can use the routines SETFLAG and UNSETFLAG to control the flag value. Using these routines, one can set the missing pixel flag value once, rather than passing it separately into each routine.

There are also some general purpose routines which also take a missing pixel flag value, either through the MISSING keyword, or explicitly as a parameter in the call. They do *not* use the value set by the SETFLAG routine, because they can be used for other things besides image processing.

AVERAGE: Averages an array over one or all of its dimensions.

BASELINE: This function estimates the baseline value of the input array, i.e. the average of all pixels which do not contain any signal.

FILL_MISSING: Uses bilinear interpolation to fill in missing pixels in a data array.

FMEDIAN: Performs median filtering. Differs from MEDIAN in that the median filter extends smoothly to the edge of the array, and in that different widths can be set for the X and Y directions.

SIG_ARRAY: Serves as a complement to the AVERAGE routine by calculating the standard deviation value of an array, or over one dimension of an array as a function of all the other dimensions.

3 Cosmic rays

A number of different routines attempt to remove cosmic rays from data, using a variety of approaches:

[1] In CDS FITS files[**?**], missing pixels are stored in one of two ways. Integer data are stored with missing pixels set to the appropriate flag value, and the flag value is given in the header. However, floating point data are stored with missing pixels set to the special IEEE value NaN (not-a-number). This is a FITS standard. The NaN values are then converted to a selected missing pixel flag value when the data is read.

3.1 AVG_WO_CR

The routine AVG_WO_CR works in a similar fashion to the AVERAGE routine. However, it attempts to remove cosmic rays before doing the average. Cosmic rays are recognized as being more than 3 sigma above the rest of the pixels making up the average. Prior to doing the averaging, those pixels are set to the missing pixel flag value. If no flag value is set, then the program will determine its own value.

A critical assumption is made in this routine that all the pixels along the dimension being averaged should have the same value. Thus, the routine is best suited for averaging together data at the same spatial location. When there is significant spatial or temporal variation in the multiple exposures, then AVG_WO_CR will only be able to reject those cosmic rays with values large compared to the spatial variation.

Because this routine is designed specifically to work with exposures from the detector, it was decided that the routine should be made sensitive to the missing pixel flag value set by the routine SETFLAG. Thus, if no value representing missing pixels is passed to the routine, then it will then look to see if a value has been set with the SETFLAG routine.

3.2 CDS_CLEAN_IMAGE

This routine removes cosmic rays from CDS/NIS images built up by rastering with the mirror mechanism. In this procedure, cosmic rays are recognized as pixels which are very high compared to neighboring exposures. Thus, it can only be used on images which are built up from rastering the slit across the sun. The first dimension must represent the raster dimension, and the second dimension must represent points along the slit. Spectra are not supported.

Because the routine compares against the neighboring exposures, rather than the average of all the exposures, it's less sensitive to spatial variation. Therefore, it can more readily distinguish a cosmic ray that would be masked by the procedure used in AVG_WO_CR.

When a cosmic ray is detected in a pixel, then that pixel is replaced by the interpolated value from the two adjacent exposures. Depending on how the routine is called, it can also interpolate into missing pixels.

Like AVG_WO_CR, this routine is sensitive to the value set by the SETFLAG routine.

3.3 CDS_CLEAN

This routine uses a similar methodology to that of CDS_CLEAN_IMAGE. The main difference is that the argument to CDS_CLEAN is a quicklook data structure[**?**], rather than a two-dimensional array. The routine can handle any kind of NIS data, so long as there are multiple exposures.

3.4 CLEAN_EXPOSURE

This routine acts on single NIS exposures. It compares pixels with the local median inside a rectangular box to find cosmic rays. Most naked-eye cosmic rays are identified, with only a few false alarms. The most usual false alarms occur at the very edge of the slit for strong lines.

Given values `LIMIT`, `MAX_VAR_LOW` and `MAX_FACTOR_HI`, it flags pixels with value `P` and local median `M` as a cosmic ray if

```
    P LT LIMIT  AND  (P - M) GT MAX_VAR_LOW
OR  P GE LIMIT  AND  (P / M) GT MAX_FACTOR_HI
```

The algorithm sometimes leaves minimally affected pixels undetected on the border of cosmic ray hits. These pixels are not way off, but they do not have the "correct" value and may thus affect e.g., background estimates. To counter this, neighbouring pixels (directly above/below or left/right) are flagged as cosmic rays as well. This behaviour can be turned off by setting the keyword /NO_NEIGHBOUR.

Sensible default values for the algorithm parameters (including the sizes of the local median box) are provided for *debiased, non-calibrated* NIS data.

Although designed primarily for NIS exposures, this routine could be useful for data from any slit spectrograph given suitable settings of the keyword parameters.

If a MISSING value is supplied, cosmic rays are flagged as missing by default. If no MISSING value is supplied, or if the keyword /FILL is set, cosmic rays are filled with the value of the local median.

3.5 CDS_CLEAN_EXP

This routine applies CLEAN_EXPOSURE to all the exposures in a quicklook data structure containing NIS data. The data should *not* be calibrated, but may be debiased. Debiasing will be performed if not already done.

Cosmic rays will be filled with the local median *unless* the keyword /NOFILL is set.

Apart from the /FILL vs /NOFILL distinction, and the fact that the MISSING value is taken from the quicklook data structure directly, all keywords for CLEAN_EXPOSURE are propagated through CDS_CLEAN_EXP

3.6 CLEAN_SPIKE

CLEAN_SPIKE applies the principle that cosmic rays are confined to single exposures and so look like spikes when neighbouring exposures are compared. On the input of a 3-D array with dimensions (X,Y,Z), a specific X-Z plane is taken to be an exposure, and CLEAN_SPIKE looks along the Y direction for any pixels that are greater than their immediate Y-neighbours by an amount specified by CUTOFF. If the value of the pixel is P and the average of the immediate Y-neighbours is AV, then

```
(P - AV_Y) / AV GT CUTOFF    (1)
```

flags the pixel as a potential cosmic ray. This criterion will often be satisfied by many pixels in regions of low counts (i.e., noisy data) and so it is necessary to also have a lower limit below which the "cosmic ray" is simply classed as noise and so not flagged. The continuum seen in CDS data varies considerably and so rather than have a specific quantity by which the strength of a pixel value is judged, it is necessary to use the FMEDIAN routine to give an estimate of the local background around the pixel of interest.[2] If we call this median value M, we then require the pixel value, P, to also satisfy

```
(P - M) GT BCKGRND        (2).
```

The above method works for spikes found between different exposures. However, a feature of CDS data, particularly that of transition region lines, is *real* spikes corresponding to short lived brightenings. Such spikes differ from cosmic rays in being continuous in the X and Z directions, and so an additional criterion for a pixel to be identified is that *either*

```
(P - AV_X) / AV_X GT CUTOFF   (3)
```

or

```
(P - AV_Z) / AV_Z GT CUTOFF   (4)
```

be satisfied, where AV_X is the average of the two neighbouring pixels in the X direction, and AV_Z is the average in the Z direction.

Many cosmic rays appear as a single large peak, but some have two neighbouring pixels with large values. If such a cosmic ray appears within a line profile then there is a chance that neither conditions (3) or (4) will be triggered. To account for such cases it is neccessary to look at the pixels *two* away from the pixel of interest in the Z direction and apply condition (4) to these.

3.7 CDS_CLEAN_SPIKE

This routine applies CLEAN_SPIKE to the data in a quicklook data structure containing NIS data. The data should *not* be calibrated, but may be debiased. Debiasing will be performed if not already done. For most of the NIS spectral data, extracting the window data using GT_WINDATA will leave an array with dimensions (lambda,solar_X,solar_Y) and this is sent to CLEAN_SPIKE for cleaning. For wide slit data, however, the array returned is 4-dimensional and of the form (solar_X,1,solar_Y,time). In such a case the (solar_X,solar_Y,time) array is extracted and REARRANGE'd to the form (solar_X,time,solar_Y) for cleaning.

Cosmic rays will be filled with the local average *unless* the keyword /NOFILL is set.

A keyword /INFO can be set to give information on the number of pixels removed in each window.

Note for some particularly large cosmic rays it may be neccessary to run CDS_CLEAN_SPIKE twice to remove all of the badly affected pixels.

[2]The median of the 5×5 (Y,Z) area around the pixel of interest is used.

3.8 NEW_SPIKE

This is a revised version of CLEAN_SPIKE and, although the method is almost identical, the code has been completely overhauled to make it run quicker hence the routine has been given a new name.

NEW_SPIKE works on 3D data arrays just as CLEAN_SPIKE does, and a key change is that pixels that neighbour flagged cosmic rays can also be flagged, in a similar manner to CLEAN_EXPOSURE. This is done by setting the keyword /NEIGHBOURS. The number of affected pixels can be displayed by setting the /INFO keyword, while the /FILL keyword replaces cosmic ray pixels with the median of the surrounding 5×5 pixel area.

3.9 CDS_NEW_SPIKE

CDS_NEW_SPIKE is analogous to CDS_CLEAN_SPIKE and works on CDS quick-look data structures, applying NEW_SPIKE to each of the NIS data windows. By default, the routine flags the neighbours of cosmic rays and fills them with the median of the surrounding 5×5 pixel area. To switch off these defaults, the keywords /NO_NEIGHBOURS and /NOFILL can be used.

3.10 XCDS_COSMIC

This routine is an interface to execute CDS_CLEAN, CDS_CLEAN_EXP or CDS_CLEAN_SPIKE, allowing the user to view and possibly correct any errors manually at the pixel level. This routine is also available from the DSP_MENU program.

To enable visualization, manual correction and saving/restoring information about the cosmic ray pixels, the automatic cleaning routines are called from XCDS_COSMIC with the /NOFILL keyword set. You can, however, also call the CDS_FILL_MISSING routine to fill the missing values.

3.11 CDS_FILL_MISSING

This routine may be used to fill in all pixels flagged as MISSING inside a quicklook data structure. It uses a median filter on each NIS exposure when possible by default, but uses FILL_MISSING in situations where this fails. If the keyword /INTERPOLATE is set, FILL_MISSING is always used.

3.12 Saving and restoring information about missing pixels

If you use e.g., XCDS_COSMIC to scrutinize the data to ensure that all cosmic rays have been detected and that no false alarms have occurred, or in any other way manipulate the data in a quicklook data structure to flag pixels as missing, you may want to save your work for use in a later IDL session.

Short of saving the whole quicklook data structure with SAVE_QLDS, you may save a list of the missing pixels only, by using CDS_SAVE_MISSING. In a later session you could read in the data structure again, and then use CDS_RESTORE_MISSING to re-flag the pixels as missing instead of doing the job manually over again.