MIT
Center for Space Research **Chandra X-Ray Center**

MEMORANDUM
September 3, 2002

To:	Martin Elvis, SDS Group Leader
From:	Glenn E. Allen, SDS
Subject:	acis_detect_hotpix
Revision:	1.0
URL:	http://space.mit.edu/CXC/docs/docs.html#hotpix
File:	/nfs/cxc/h2/gea/sds/docs/memos/memo_acis_detect_hotpix_1.0.tex

1 acis_detect_hotpix

1.1 Description

This tool uses the event data or bias data for a given observation to search for pixels and columns that may be temporarily or permanently "bad." The tool is designed to find pixels that are not included in the bad pixel file for the observation. Known bad pixels and columns (and masked regions) can be excluded from the search.

1.2 Input

1. A Level 1 event data file (acis*evt1.fits) or a Level 0 bias map (acis*bias0.fits)
2. A Level 1 bad pixel file (acis*bpix1.fits)
3. A Level 1 window mask file (acis*msk1.fits)

1.3 Output

The output file has the same format as a Level 1 bad pixel file. If an input bad pixel file is specified, the output is the union of this file and any newly identified bad pixels and columns.

1.4 Parameters

1. infile,s,a,"",,,"Name of input event file or bias files"
2. outfile,s,a,"",,,"Name of output bad pixel file"
3. badpixfile,s,a,"CALDB",,,"Name of input bad pixel file (CALDB — ⟨filename⟩ — none — NONE)"
4. maskfile,s,a,"default",,,"Name of input mask file (default — ⟨filename⟩ — none — NONE)"
5. logfile,s,h,"stdout",,,"Destination of output messages (stdout — STDOUT — ⟨filename⟩)"

6. probthresh,r,h,1.0e-3,1.0e−10,1.0e−1,“Minimum significance of potentially bad pixels in sigma”
7. regwidth,i,h,5,3,256,“Size of reference region for comparison (e.g. 5x5)”
8. clobber,b,h,no,,,“Overwrite output file if it already exists?”
9. verbose,i,h,0,0,5,“Amount of messages produced (0=none, 5=most)”
10. mode,s,h,“ql”,,,

1.5 Processing

1. Verify that the specified input files exist. If the parameter clobber = “no,” verify that the output file does not exist. Verify that the values specified for the parameters probthresh and regwidth are in the allowed ranges from 1.0e−10 to 1.0e−1 and from 3 to 256, respectively.
2. Create a 1024×1024 element array $A^i_{\rm pix}$ and a 1×1024 element array $A^i_{\rm col}$ for each active CCD i.
3. Read the input bad pixel file (if one is specified). If the parameter badpixfile = “CALDB,” then use the appropriate bad pixel ARD file. If a pixel or column is designated as “bad,” then set the corresponding elements of $A^i_{\rm pix}$ to be -4096. If a column is designated as “bad,” then set the corresponding element of $A^i_{\rm col}$ to be -4096. Do not set the elements of $A^i_{\rm pix}$ and $A^i_{\rm col} = -4096$ if the only STATUS bit(s) set equal to one in the bad pixel file are 5, 6, or 8–12. Since pixels and columns that have only these bits set are not truly bad, they should be included in the search.
4. Read the input mask file (if one is specified). If the parameter maskfile = “default,” then use the appropriate mask specified in the acis_build_mask document at http://space.mit.edu/CXC/docs/docs.html#msk. If a portion of a CCD is “masked out” (i.e. no events can be reported for the region), then set the values of the appropriate elements of $A^i_{\rm pix}$ and $A^i_{\rm col}$ to be -4096.
5. If the input file specified by the parameter infile is an event data file, then read the values of CCD_ID, CHIPX, and CHIPY for the first event and add one to the corresponding elements of $A^i_{\rm pix}$ and $A^i_{\rm col}$ (unless the value of $A^i_{\rm pix}$ or $A^i_{\rm col} = -4096$). Repeat this process for the rest of the events in the file. If the file(s) specified by infile is a bias map(s), then read the value of the first element of the first bias map, set the corresponding value of $A^i_{\rm pix}$ equal to the value in the bias map (unless the value of $A^i_{\rm pix} = -4096$), and add the value in the bias map to the appropriate element of $A^i_{\rm col}$ (unless the value of $A^i_{\rm col} = -4096$). If the value in the bias map is greater than or equal to 4094, then set the corresponding element of $A^i_{\rm pix} = -4096$. The values in the elements of the array $A^i_{\rm col}$ are sums over all of the pixels in the corresponding CCD column.
6. Compute the total number of pixels and columns to be searched. Set $N_{\rm pix}$ equal to the total number of pixels for which $A^i_{\rm pix} \neq -4096$. If six CCDs are active and the default window masks are used for the observation, $N_{\rm pix} \approx 6 \times 10^6$. Set $N_{\rm col}$ equal to the total number of columns for which $A^i_{\rm col} \neq -4096$. If six CCDs are active, $N_{\rm col} \approx 6 \times 10^3$. Write the values of $N_{\rm pix}$ and $N_{\rm col}$ to the logfile (for some value of verbose).
7. Compute the minimum significance of the pixels and columns identified as bad. This significance depends on the total number of pixels or columns searched (i.e. the number of trials) and the specified value of the parameter probthresh. The probability threshold for the search for bad pixels

$$P_{\rm pix} \approx \frac{\rm probthresh}{N_{\rm pix}} \tag{1}$$

and the probability threshold for the search for bad columns

$$P_{\rm col} \approx \frac{\rm probthresh}{N_{\rm col}}. \tag{2}$$

If the input file specified by the parameter infile is an event data file, then the corresponding significance thresholds in terms of the parameter σ of the Gaussian function $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}$ are obtained by solving the equations

$$P_{\rm pix} = {\rm erfc}\left(\sigma_{\rm pix}\right) \tag{3}$$

and

$$P_{\rm col} = {\rm erfc}\left(\sigma_{\rm col}\right) \tag{4}$$

for $\sigma_{\rm pix}$ and $\sigma_{\rm col}$. For example, if $N_{\rm pix} = 6 \times 10^6$ and probthresh $= 10^{-3}$, then $P_{\rm pix} = 1.67 \times 10^{-10}$ and $\sigma_{\rm pix} = 6.39$. If $N_{\rm col} = 6 \times 10^3$ and probthresh $= 10^{-3}$, then $P_{\rm col} = 1.67 \times 10^{-7}$ and $\sigma_{\rm col} = 5.23$. Write the values of $P_{\rm pix}$, $P_{\rm col}$, $\sigma_{\rm pix}$, and $\sigma_{\rm col}$ to the log file (for some value of verbose).

8. Begin the search for bad pixels. Perform the search separately for each node of each active CCD because the event rate and bias can vary from node to node on the same CCD.

9. If the file specified by the input parameter infile is an event data file, then compute the *mean* value of the number of events per pixel

$$\mu^{ij}_{\rm pix} = \frac{\Sigma A^{ij}_{\rm pix}}{N^{ij}_{\rm pix}} \tag{5}$$

The sum is performed over all valid pixels of node j on CCD i. $N^{ij}_{\rm pix}$ is the total number of valid pixels on the node. Here, invalid pixels are those for which the values of $A^{i}_{\rm pix} = -4096$. Note that the mean is a real number, not an integer.

If the file(s) specified by the input parameter infile is a bias map(s), then compute the *median* value of the bias:

$$\mu^{ij}_{\rm pix} = {\rm median}\left(A^{ij}_{\rm pix}\right). \tag{6}$$

Only include the valid pixels of node j on CCD i. Here, invalid pixels are those for which the values of $A^{i}_{\rm pix} = -4096$.

Write the value of $\mu^{ij}_{\rm pix}$ to the logfile (for some value of verbose).

10. If the file(s) specified by the input parameter infile is a bias map(s), then subtract the value of $\mu^{ij}_{\rm pix}$ from all of the valid elements of $A^{i}_{\rm pix}$ (on the same node). Do not perform the subtraction if the value of an element $A^{ij}_{\rm pix} \leq -4095$.

11. If the file(s) specified by the parameter infile is a bias map(s), then compute the minimum significance of the pixels identified as bad:

$$\sigma_{\rm pix} \approx \left(\frac{\Sigma\left(A^{ij}_{\rm pix} - \left\langle A^{ij}_{\rm pix}\right\rangle\right)^2}{N^{ij}_{\rm pix}} \right)^{\frac{1}{2}} x_{\rm pix} \tag{7}$$

where $\left\langle A^{ij}_{\rm pix}\right\rangle$ is the mean (real-numbered) value of the valid elements of the pixel array and $x_{\rm pix}$ is the solution of $P_{\rm pix} = {\rm erfc}(x_{\rm pix})$. The sum is performed only over the valid elements of the array. $N^{ij}_{\rm pix}$ is the total number of valid pixels on the node. Here, the invalid pixels are those for which the values of $A^{i}_{\rm pix} \leq -4095$ and those for which $\left|A^{ij}_{\rm pix} - \left\langle A^{ij}_{\rm pix}\right\rangle\right| > 3$.

12. For each valid pixel of a node, determine the value of the corresponding element of the pixel array (S, i.e. the "source" value), the value to use as a reference (B, i.e. the "background" value), and the number of background pixels used to compute the reference value (n).

$$S = \text{the value of the appropriate element of } A^{ij}_{\rm pix}. \tag{8}$$

$$n = \text{the number of valid pixels to use to compute } B. \tag{9}$$

Typically, $n = \text{regwidth} \times \text{regwidth} - 1$, but the value of n is smaller if one or more of the pixels in the reference region are invalid. The reference region is a square "centered" on the source. For example, if regwidth = 5 and the pixel being examined is CHIPX = 372 and CHIPY = 948, then the reference region is the region defined by $370 \leq \text{CHIPX} \leq 374$ and $946 \leq \text{CHIPY} \leq 950$. If regwidth = 6, then the reference region for the same pixel is $370 \leq \text{CHIPX} \leq 375$ and $946 \leq \text{CHIPY} \leq 951$. The pixel being examined is excluded from the computation of n and B. Any other pixels in the reference region for which $A^{ij}_{\text{pix}} \leq -4095$ are also excluded from the computation of n and B. The reference region may not extend off of the node. For example, if the pixel being considered is CHIPX = 512 and CHIPY = 2, then the reference region is $510 \leq \text{CHIPX} \leq 512$ and $1 \leq \text{CHIPY} \leq 4$ (if regwidth = 5). If there are no valid pixels in the reference region or if regwidth = 256, then n is equal to the total number of valid pixels on the entire node.

$$B = \Sigma A^{ij}_{\text{pix}}, \tag{10}$$

where the sum is performed over the n valid elements of the reference region.

13. If the file specified by the parameter infile is an event data file, then the statistical significance of S for a pixel given the values of B and n for the pixel (i.e. the likelihood that the value of A^{ij}_{pix} for the pixel is significantly larger or smaller than the values of A^{ij}_{pix} for the reference pixels) is

$$\sigma = 2^{\frac{1}{2}} \left(S \ln \left(\frac{(n+1)\, S}{S+B} \right) + B \ln \left(\frac{n+1}{n} \frac{B}{S+B} \right) \right)^{\frac{1}{2}} \tag{11}$$

(Li and Ma (1983)).

If the file specified by the parameter infile is a bias map, then the statistical significance of S for a pixel given the values of B and n for the pixel is

$$\sigma = S - \frac{B}{n} \tag{12}$$

14. If the absolute value of the significance $|\sigma| \geq \sigma_{\text{pix}}$, then flag the pixel as potentially bad. Set the value of $A^{ij}_{\text{pix}} = -4095$.

15. Repeat steps 9–14 for each valid pixel of each node of each active CCD.

16. Begin the search for bad columns. The procedure used to search for bad columns is very similar to the procedure used to search for bad pixels (steps 9–15) with appropriate substitutions of A^{i}_{col}, A^{ij}_{col}, μ^{ij}_{col}, N_{col}, and x_{col}, for A^{i}_{pix}, A^{ij}_{pix}, μ^{ij}_{pix}, N_{pix}, and x_{pix}, respectively. In steps 10, 11, 12, and 14, substitute the value -4094 for the value -4095.

17. Write the output to the file specified by the parameter outfile unless a file by that name already exists and the parameter clobber = "no," in which case write an error message to the logfile and exit. The format of the output file is the same as the format of an ACIS Level 1 bad pixel file. If an input bad pixel file is specified by the parameter badpixfile, then copy the data in the input file to the output file. Newly identified bad pixels and columns are added to the output file. Bits 15 and 14 (0f 0-31) are set to one for the newly identified bad pixel and columns, respectively. For each new bad pixel or column, write the values of CCD_ID, CHIPX, CHIPY, S, B, n, and σ to the logfile. If no new bad pixels or columns are identified, do not create an output file and write an appropriate message to the logfile.

1.6 References

Grant, C. 2002, private communication
Isobe, T. 2002, private communication
Li, T.-P., and Ma, Y.-Q. 1983, ApJ, 272, 317
Tibbetts, K. 2002, private communication