The HST Treasury Program on Eta Carinae

Technical memo number 8 -- January 2005

IDENTIFYING AND PATCHING OVEREXPOSED PIXELS IN STIS CCD DATA

John C. Martin, University of Minnesota

1. Introduction

This report outlines the procedure which was used to identify and patch overexposed pixels in the Eta Carinae Treasury Project database. Overexposed pixels are pixels which have been saturated with photons/electrons during an exposure and have been pushed past the regime where the CCD has a linear flux response. They can also affect the pixels surrounding them by bleeding charge or changing the effective well depth of neighboring pixels (especially when they occur in large clusters). In our program data, they are primarily found in the very bright Balmer Hα and Hβ emission features. A table of the exposures which have overexposed pixels is included in Appendix A to this document.

2. Identification of the Overexposed Pixels

Overexposed pixels are flagged in the DQ array of the FITS files. Unfortunately, the Goddard version of CALSTIS (which our reduction pipeline was built on) does not save an audit trail for each pixel but instead saves only the "worst" thing that happened to that pixel in the reduction process. As a result the fact that a pixel is saturated often ends up replaced with some "worse" flag.

Overexposed pixels are easily identified in the *raw* CCD data as those pixels with enough counts to exceed the maximum useable well depth. It is then a far less trivial task to map the locations of these pixels into the reduced data; especially considering that the reduced pixels have been interpolated so that their values are not independent from the values in the pixels surrounding them (Davidson, 2004). One must take a more sophisticated approach to correctly identify the overexposed pixels in the processed data.

Fortunately, nearly all of the observations with overexposed pixels also have a matching shorter exposure at the same epoch. In those cases, we can identify the affected pixels by comparing the long (overexposed) and short exposures to find the pixels which differ significantly between them. Normally these exposures are very short and only separated by the read out time of the CCD. Therefore, it is unlikely that the pointing of the HST drifted significantly between exposures. Note that charge bleeding and overexposed "trails" are both caught by this method since a pixel affected by either of these should differ between the long and short exposures. The pixels identified in this manner were used as seeds which were grown into patches that encompassed adjoining pixels. This was an important step since the pixel interpolation methods used by CALSTIS to "rectify" the two-dimensional spectra effectively blur the influence of the overexposed pixels. In the case of the interpolation method which the Eta Carinae Treasury Project uses, the influence of one overexposed pixel was blurred over three original pixels on both the dispersion and cross-dispersion axes (Davidson, 2004). The values of the pixels identified in this manner were changed to NaN (in accordance with standards IEEE 754/IEEE 854) in the FLUX array of the FITS file.

The revision name covering the identification of overexposed pixels for the Treasury Project is "Jo March." Therefore, FITS files in the database containing the words "Jo March" in their header have had their bad pixels replaced by the value NaN. (A quick way to check if this revision has been applied is to "grep" the header.) In that case, the following HISTORY entry has been added to both the primary and FLUX headers:

```
Jo March was here. Pixels affected by saturation flagged.
```

Some of the FITS files in the database which have overexposed pixels do not have a companion short exposure that can be used to positively identify the resulting bad pixels. Those FITS files contain the following HISTORY entry in their primary and FLUX headers:

```
*WARNING* The raw data for this file contained saturated pixels.  These
Pixels have not been reset to NaN in the flux data and may not be proper
ly identified in the DQ array.
```

3. Combined & Patched Exposures

We have created another data product for the database by replacing bad pixels in the long (overexposed) exposures with the data from the short exposures. This can be a bit more complicated than a simple replacement of the bad pixels in the long exposure with the corresponding pixel values from the short exposure because different exposure lengths yield different noise characteristics. The longer exposure always enjoys a higher signal to noise ratio since the integrated flux is significantly less over most of the short exposure. Therefore, in order to create a single dataset from two separate exposures we must have a region that smoothes the transition from one dataset to the other around the bad pixels.

For this operation we used a transition region which is five reduced pixels wide (five pixels in the reduced data = 2.5 original pixels) where the pixels were combined as an average of the pixel values in the short and long exposures with the weight of the short exposure pixel values decreasing outward from the region of bad pixels. Following this procedure ensured that the smooth function of flux was preserved across the pixels without any significant discontinuities. The elements from the respective ERR arrays are combined in a similar manner. The masks used to weight and combine the long and short images are included as the MASK array following the ERR array in the FITS files of the patched and combined data.

The FITS data products of this operation will eventually be available as part of the Treasury Product database. Due to their processed nature, they will be used to compliment rather than replace the exposures that were combined to produce them.

References:

Davidson, K. D. 2004, "Our Adopted Scheme for Subpixel Modeling", Eta Carinae Treasury Project Technical Report #1, http://etacar.umn.edu/treasury/publications/pdf/tmemo001.pdf

Appendix A

The following is a complete list of observations which are overexposed in part of the observed spectrum. The list was made by checking the raw HST data for pixels with values over 30,000 counts. Saturated pixels which were caused by cosmic ray hits or hot pixels were ignored in this analysis.

The column listing the number of saturated pixels is included to give a hint as to the degree that observation is over exposed. The more saturated pixels, the more overexposed the spectrum is. An star (*) next to the Treasury Project (TP) Label denotes that long exposure has no corresponding short exposure. Therefore, the "bad" pixels in those data remain unidentified.

Raw HST Dataset	TP Label	λ_0	# of Pix Saturated
o4j802080	c800_008	Hα	69
o4j8010q0	c821_022	Hα	115
o4j8010s0	c821_024	Hα	90
o4j8010t0	c821_025	Hα	305
o4j8011d0	c821_042 *	Hα	10
o55601hnq	c890_002	Hα	185
o55601040	c890_004	Hβ	20
o556010b0	c890_011	Hα	135
o556020m0	c914_022	Hα	183
o556020n0	c914_023	Hα	4508
o556020r0	c914_034	Hβ	35
o55602150	c914_041	Hα	181
o55602180	c914_044	Hα	77
o5f102030	cA20_003	Hα	477
o5f102070	cA20_007	7283	117
o5f1020e0	cA20_014 *	9336	16
o5f103010			35
o5f103020			137
o5kz01cqq	cA22_001	Hα	12232
o5kz010l0	cA22_021	Hα	159
o5kz010o0	cA22_024 *	4451	24

Raw HST Dataset	TP Label	λ_0	# of Pix Saturated
o5kz010u0	cA22_031	Hβ	199
o5kz02020	cA22_037	Hα	268
o5kz02030	cA22_038	Hα	13922
o5kz02040	cA22_039 *	Hα	757
o5kz02050	cA22_040 *	Hα	240
o5kz020s0	cA22_063	Hα	14848
o5kz020t0	cA22_064 *	Hα	626
o5kz020u0	cA22_065 *	Hα	57
o5kz020v0	cA22_066 *	Hα	3
o5kz021b0	cA22_082 *	Hα	9
o5kz021f0	cA22_086 *	Hα	4
o5kz021i0	cA22_089	Hβ	1366
o5kz021j0	cA22_090 *	Hβ	46
o5kz022c0	cA22_119	Hβ	1369
o5kz022d0	cA22_120 *	Hβ	11
o5kz022e0	cA22_121 *	Hβ	6
o5kz02360	cA22_149	Hα	278
o5kz023a0	cA22_153	Hα	282
o62r01010	cB29_001	Hα	5038
o62r010k0	cB29_020	Hα	174
o62r010v0	cB29_031	Hβ	206

Raw HST Dataset	TP Label	λ_0	# of Pix Saturated
o6ex03010	cB75_001	Hα	641
o6ex03030	cB75_003 *	9336	926
o6ex03080	cB75_008	7283	7718
o6ex03090	cB75_009 *	7283	14
o6ex030a0	cB75_010 *	4451	53
o6ex030c0	cB75_012 *	5216	129
o6ex030d0	cB75_013 *	2375	43
o6ex030e0	cB75_014	Hβ	328
o6en01010	cB75_016	Hα	
o62r02010	cB90_001	Hα	628
o62r020d0	cB90_013	Hα	54
o6ex02010	cC05_001	Hα	622
o6ex020k0	cC05_020	Hα	603
o6ex020v0	cC05_031	Hβ	240
o6mo02010	cC51_001	Hα	592
o6mo02030	cC51_003 *	Hα	27
o6mo020v0	cC51_031 *	7283	6
o6mo021m0	cC51_058	Hβ	246
o8gm01010	cC96_001	Hα	574
o8gm01060	cC96_006	Hα	17
o8gm12010	cD12_001	Hα	627
o8gm120v0	cD12_031	Hβ	208
o8gm21010	cD24_001	Hα	580
o8gm41010	cD34_001	Hα	483
o8gm330m0	cD37_022	Hβ	179
o8gm32010	cD37_029	Hα	480
o8gm32030	cD37_031 *	Hα	8
o8gm52010	cD41_001	Hα	497
o8gm52030	cD41_003 *	Hα	2
o8gm521k0	cD41_056	Hβ	202

Raw HST Dataset	TP Label	λ_0	# of Pix Saturated
o8gm630c0	cD47_012	Hβ	201
o8gm630h0	cD47-017 *	6252	188
o8gm62010	cD47_020	Hα	462
o8gm620v0	cD47_050 *	7283	4
o8ma72010	cD51_001	Hα	475
o8ma82010	cD58_001	Hα	395
o8ma821o0	cD58_059	Hβ	180
o8ma92010	cD72_001	Hα	605
o8ma920w0	cD72_032	Hβ	347
o8ma83010	cD88_001	Hα	695
o8ma940e0	cE18_014 *	2697	9
o8ma940f0	cE18_015 *	4451	17
o8ma940p0	cE18_025	Hβ	468
o8ma940r0	cE18_027 *	6252	26
o8ma940t0	cE18_029	Hα	887
o8ma940v0	cE18_031	Hα	88
o8ma940w0	cE18_032 *	Hα	104