CRDIR: Cosmic Ray Damaged Image Repair for DSLR Cameras

Primary Author: Kevin Moser
Committee: Don Figer Ph.D., Carl Salvaggio Ph.D., and Mark Fairchild Ph.D.

***Abstract*—High velocity particles in space known as cosmic rays can strike the electronics in an imaging sensor and create permanent damage. This cosmic ray damage can lead to slightly structured random high value impulse noise scattered throughout the scene which visually and analytically alters the image. In order to combat the cosmic ray damage, a new system needs to be developed to repair the damage effects in these images for future use. We proposed an algorithm to process the raw data NASA captures from Nikon cameras on the International Space Station in order to remove the effects of cosmic ray damage in a way that is compatible with NASA workflow. A statistical z-score method and a structural convolution method were evaluated against marked images to calculate false positives and false negatives results. The convolution method correctly detected 98 percent of damage in an image while avoiding identifying all stars and other objects of interest in the image and repaired the color filter array data with the local median. The algorithm saves the restored raw data format for use later in Adobe Suites to be finally printed or displayed for NASA outreach.**

I. Introduction

Consumer grade DSLR cameras, such as the Nikon D3s and D4, are used on the International Space Station (ISS) to obtain images of the station, Earth, stars, and the astronauts themselves. The photography and videos captured by the astronauts are some of NASA's greatest outreach to the general public. Also, imagery is one of the best forms of communication between ground control and the station when inspecting damage and solving problems onboard the spacecraft. Great care is taken by NASA analysts to ensure image quality and accuracy throughout the capture and processing workflow. Overall, 10,000 to 15,000 images can be captured by the astronauts within a week. However, once these DSLR cameras leave the protective cover of the Earth's atmosphere, the sensors are subject to damage from high velocity particles, such as protons or atomic nuclei, traveling through space. The damage at the sensor level spreads into the images making them less visually appealing and less useful for analysis. Particularly, any analysis looking for station damage or examining small stars could have error in the calculations from the cosmic ray effects. With normal image processing and interpolation, single and double high value pixels will spread across an image, clearly appearing in large displays and prints of the images. Once a pixel is damaged, the cosmic ray effects on that pixel are permanent and the longer the camera is in use the more total damage the sensor acquires. Most cameras on the ISS have a lifetime of six months to a year before the damage forces them to be retired. For these commercial cameras, very little can be done in the electronics or hardware to prevent cosmic ray damage to the pixels. Therefore, to improve the quality of these images and the lifetime of the cameras, a software approach is needed to remove the effects of cosmic ray damage from the captured data without affecting other important image features. The program must be developed in a system to integrate with NASA's current workflow of image collection and processing and be able to quickly handle the large number of images captured.

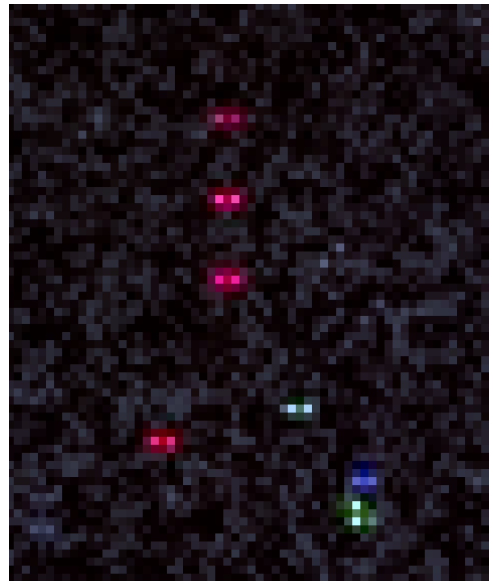

Fig. 1. Cosmic ray damage effects on color image

II. Background

A. Cosmic Ray Damage

Cosmic ray damage has been prevalent in cameras since they first went to space. The high velocity particles from deep space are typically absorbed by the Earths atmosphere, shielding us from the effects. However, once satellites and telescopes started to orbit the Earth, people have had to design way to handle damage from the cosmic rays. The Hubble Wide Field and Planetary Camera 2 experienced a large number of cosmic ray hits and has been the study of many past papers. For this camera, the effects of the damage can be handled by capturing dark current frames or by accurately measuring the damage across multiple images [4]. However for our design, the program needed to be designed to work on individual images with no prior knowledge of the camera, such that a random image can be chosen and processed at will. Because the astronauts are interested in quickly capturing interesting scenes from a photography standpoint, most images in the NASA archives have no dark currents frames associated with the cameras. Also, many of the images will be processed individually for later printing or display without the set of frames from the single camera. Some research has been done on processing the damaged Hubble images individually. These algorithms depends on the local area histogram [6], the sharpness of the edges in the object [7], and convolution kernels to identify damage [5]. All these past techniques were examined and developed into our new methods. Difficulty arises between applying the techniques that worked on the Hubble WFPC2 sensor, a carefully designed CCD for astronomy imaging which is created to handle the radiation of space, onto the CMOS sensor design for that of the normal consumer camera. On the DSLR cameras, cosmic ray damage presents itself by generating a strong increase in dark current on individual or pairs of pixels along the same color channel. The pairs of pixels are a strange artifact with no known current explanation. The prevalent pattern of two pixels could be from damage to the readout of the sensor or leaks between pixels, but until a camera is returned to Earth, the sensor damage can not be properly characterized. The high dark current associated with cosmic ray damage leads to much brighter pixels in the image as seen in Fig. 1. While these values are clearly near the max value, there is some variability in the damage effects in different images captured with the same camera, dependent on exposure time, ISO, and scene brightness. Therefore, the standard salt and pepper noise reduction approaches do not apply to cosmic ray damage. To best detect and repair the effects of cosmic ray damage, the images had to be handled in their most raw format captured by looking at the individual pixel detectors on the imaging sensor plane.

B. Bayer Pattern and Raw Images

The standard digital camera sensor measures one irradiance value of light at each individual pixel, which is merely enough to create a grayscale image. Therefore, to gain color information in the image, a color filter array (CFA) is applied to the sensor to sample specific colors at a pixel. Based on the cones in the human visual system, the important information of a scene is best captured with red, green, and blue filters. Bryce Bayer invented the most common CFA design, simply called the Bayer pattern in his namesake [1]. Half of the pixels are sampled green, which visually plays a key role in conveying luminance, and the remaining half are split between the reds and blues. Each color channel is subsampled in a checkerboard type pattern, which can be modeled as a sinusoidal function based on the pixel location as demonstrated in figure 2.

To generate the full color image, the missing pixels are filled in with calculated values in a process called demosaicking, which could be as simple as the average of the two nearest pixels of the same color channel or extremely complicated to preserve the sharpness of edges and purity of the colors. However, any high values on the CFA will lead to similarly high values being spread to the nearby pixels in the demosaicking process. Particularly, the effects of cosmic ray damage spreads in color images, creating what can appear to be colored stars or bright edges, when the damage is realistically only one or two pixels long in the CFA. For most DSLR cameras, the color filter array image is saved in a raw format so demosaicking can be handled on a computer with much higher processing power. The raw format allows us to repair the effects of the cosmic ray damage on the CFA and save it back to the raw format to be demosaicked through NASA's choice of raw image processing programs.

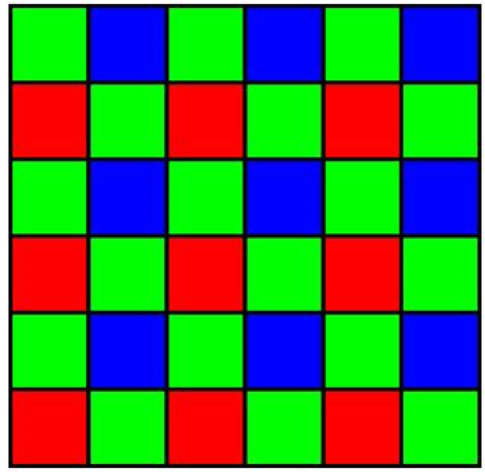

Fig. 2. Bayer Pattern color filter array applied to the camera sensor

For Nikon DSLRs, the raw CFA pattern is saved into the Nikon Electronic Format (NEF). The NEF file is organized in a Tagged Image File Format (TIFF) with tags for all the image metadata and sub-files for the raw CFA data and a low resolution RGB image for thumbnails and quick display of the file. Each camera company has its own raw image format for saving all this data, and each format may undergo some form of lossless compression or reorganization of tags and sub-files. To easily read and write the CFA of the raw image, we converted all our raw format images to uncompressed 16 bit Adobe Digital Negatives (DNG), which still preserves all metadata and sub-files, but allows us unconstrained access to the CFA data. The DNG format is a consistent structure and can be created from any camera raw file so if NASA decides to start using Cannon or other product, our workflow remains untouched.

Once the CFA image is read, the red, blue, and green channels can be isolated by knowing the pattern of the Bayer filters. Across each color channel, 3 by 3 pixel local neighborhoods were analyzed to create full median color planes to replace detected cosmic ray damage. By simply replacing every pixel in the image with the median on a surrounding small neighborhood, all high value pixels will be removed, removing the worst of the damage effects. However, in astronomy imaging, detail amongst the stars and small bright points of light is crucial to preserve. The median filter alone will remove all bright details leaving only a dark sky. So an adaptive method

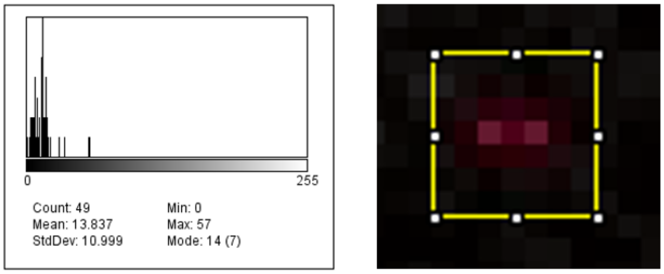

Fig. 3. Histogram of cosmic ray damage and background where the damage corresponds to the outliers of the distribution

-1	-1	-1	-1	-1
-1	-1	-1	-1	-1
1	-1	1	-1	-1
-1	-1	-1	-1	-1
-1	-1	-1	-1	-1

-1	-1	-1	-1	-1
-1	-1	-1	-1	-1
-1	-1	1	-1	1
-1	-1	-1	-1	-1
-1	-1	-1	-1	-1

Fig. 4. Design of the two horizontal kernels used to identify cosmic ray damage, similar vertical and diagonal iterations are also used

is needed to determine if a pixel should be replaced or left unchanged in the image [3]. From this point, we developed two methods for detecting the cosmic ray damage to be replaced with the median, one algorithm that looks at the local statistics in 3 by 3 pixel regions on the individual color channels and a second technique that analyzes the structure of the relative brightness of the image with convolution kernels.

III. Methods

A. Modified Z-Score

An important identifying factor of the effects from damaged pixels in a single image is that they have much higher values than their neighboring pixels. The local histogram of the regions without damage follows a Gaussian distribution, while the damage itself clearly displays as outliers beyond the distribution. For a Gaussian distribution, 99.7 percent of the data should be within 3 standard deviations of the central mean value and anything beyond can be considered an outlier. The number of standard deviations a value is away from the mean can be calculated and is called a z-score. For our cosmic damage calculations, the z-score equation needed to be slightly modified. First, in order to avoid a false mean from the high damage values and to increase runtime, the pre-calculated 3x3 median values were used as the central point of the distribution. Second, the standard deviation is calculated in the 3x3 regions while ignoring the center pixels to remove that error in calculation if it is a damaged pixel. So our modified z-score equation is

$$ZScore = \frac{S - M_3[S]}{\sigma_3'} \tag{1}$$

where S is the value of the central pixel, $M_3[S]$ is the median of the local 3 by 3 pixel neighborhood, and σ_3' is the standard deviation of the local neighborhood without the central pixel. Z-scores are calculated for every pixel along each color channel and a mask of potential damage points is then created by marking z-score values greater 2.7. A secondary minimum brightness over median mask is created to help remove falsely identified damage from areas with extremely low standard deviations.

$$Offset = S - M_3[S] \tag{2}$$

A threshold is then applied to the offset at a low value that was determined through experimentation on visually marked images to remove false positives. Once the two masks are combined, the damaged pixels are replaced with the median at that location, and the repaired color channels are placed back into a Bayer pattern and saved to the DNG format. The image can then be opened through Photoshop to observe the algorithms effect on the output RGB image.

B. Structure Convolution

As the algorithm process developed, a realization was made that processing across the individual color channels can lead to some artifacts when looking at the finalized color image which will be discussed further in the results section. In order to abate this newly created damage, the process needed to be redesigned to work across the full raw CFA data. In particular, the Adobe suites of products contains software that automatically handles many singular hot pixels while ignoring all of the more problematic damage from the patterns of connected pairs of pixels on a single color channel. These damaged pixels can be found in horizontal, vertical, or diagonal iterations throughout the image. The pair of damaged pixels should be the only relatively high values to in a local region which can be analyzed with a convolution kernel. A convolution kernel calculates the sum of the local region with each pixel weighted by a specific value. Common convolution kernels are used for averaging, blurring, sharpening, and edge detection throughout an image. For analyzing the structure of the pairs of damage on the color filter array, the convolution kernel weights for the horizontal damaged pixel pair appear as seen in Fig. 4. The convolution kernel is then run on the color filter array minus the median values reformatted to the Bayer pattern, leading to a structure equation,

$$Structure = (CFA - M_3[CFA]) * K_5 \tag{3}$$

where CFA is the raw color filter array image, M_3 performs the median across each color channel but maintains the CFA format, and $*$ performs the convolution with one of the 5 by 5 structure kernels. A total of eight structure maps are created for each direction of the kernel. A threshold at zero is then applied to the each output structure map to generate masks. The number of negative weighted pixels will generally lead to a negative result. Especially near a star or other event that is brighter than the median across multiple color channels, the output from the convolution will be negative and not marked as damage. However, the effects of the damage will be detected as the large sum of the two bright pixels will offset the low negative weighted values nearby, adding up to a positive result. Applying the threshold to the convolution image will create a mask for that particular pattern, so convolution kernels must be run for the horizontal, vertical, and diagonal iterations on the CFA and the masks must be added together to create a total structural mask. Similar to the z-score method, a minimum brightness offset over the median threshold mask is multiplied to the total structure mask to reduce the spread of false positives from convolution kernels that glanced on the edges of a star or damaged pixels. The marked pixels in the final mask are then replaced in with the medians and the repaired CFA is placed in the DNG image.

IV. Results

Many of the raw images from the International Space Station were processed through the z-score and structure algorithms and the output color images were visually compared to their unprocessed counterparts. By looking at Fig. 5, one can see that while the z-score method detected many of the damaged pixel artifacts in the image and replaced the high value with the local median, the algorithm also created new artifacts within the bright star. This issue brings us back to the difficulty with processing each color channel separately. Due to the Bayer pattern of sampling, the star which is bright in all color planes, will only appear in half as many pixels on the red and blue channels as compared to the green channel. By processing the red, blue, and green pixels separately, there is only enough data points on the green channel to determine the point is a star. The red and blue pixels are consequently flagged, leading to their replacement with the median. The originally bright white star now is processed to appear only green in the center. Also, as the algorithm moves away from the low signal areas of space to the high level of variation along the station and solar panels, the distribution of the small region becomes much less of a Gaussian distribution. The solar panel region in particular has a lot of thin edges and variability of pixel values leading to larger standard deviation. The high standard deviation compared to normal means low z-score values and damaged pixels being ignored by the algorithm.

Now looking at Fig. 6, the structure convolution algorithm preserves all the details within the star while replacing all potential damage artifacts in the image with the local median. The convolution structure clearly avoids these small bright points of interest as long as the values are high compared to the median across multiple color channels. Even on difficult areas with brighter backgrounds and lots of variation between the pixels like on the solar panels of the station, the convolution method detects most of the damaging effects without disrupting important features in the final image. The few locations that the effects from the damage was missed appear within sharp edges and in the center of star streaks from long exposures. In these situations, important image features should be preserved even at the risk of leaving some cosmic ray damage artifacts behind. Overall, the convolution method for detecting damage from raw images handled the damage smoothly and produced clean outputs for display.

The goal for our algorithm was to detect and replace any of the visible effects of cosmic ray damage when displayed in a RGB color image after demosaicking. To this end, two result measurement techniques were created. First on the original images, all visible damage was marked by observers and mapped back to pixel positions on the color filter array. The images were then run through the damage detection algorithms and the output mask where damage occurred was compared to the visibly marked mask. False positives were calculated where the algorithm marked pixels that did not visibly display the effects of damage. False negatives were calculated where the algorithm failed to detect visibly apparent damage artifacts. Secondly, a low percentage of random value impulse noise was generated at random locations and adjusted to form pairs of pixels on the same color channel. This noise was then applied to fake color filter arrays that were generated by sub-sampling RGB images in the Bayer pattern. The images used were a mixture of very low damaged captures and artistic renditions of the aurora and nebula effects as seen in many of these International Space Station images. The generated damaged images were processed through the structure convolution method to look at the consistency of the algorithm to detect cosmic ray damage across different images. False positives and negatives were again calculated by comparing the simulated noise mask to the structure convolution mask.

Looking first at the visually marked image, the results of z-score and structure convolution are measured in Table 1. Clearly, in a situation with matching low levels of false positives between the algorithms, the z-score method fails to detect a large percentage of the damage. As discussed above, the method works well on the space regions but many of the false negatives can be contributed to the high standard deviation of the station or other bright objects of interest. The statistical outliers method with z-score is designed for the smooth background of space as seen in many of the normal astronomy images and demonstrates flaws in these widely diverse images captured on the International Space Station. The other issue that arises is the false positives are visibly clustered on the small bright regions such as the star that are important to preserve. On the other hand, the structure convolution method has a false negative percentage of only 2 percent and a similarly low false positive percentage when compared to z-score. Because of the structure convolution methods design, the false positives can be assured to not occur on the bright points across the color channels which make up most important features in the image. To confirm these results, the structure convolution method was also examined on the simulated cosmic ray damage data as seen in Table 2.

Across the set of images, the false negative and false positive rates remained fairly consistent. The best image had the darkest background which affords the most contrast for detecting damage and was the largest image observed. Similarly, the image with the highest false positive and false negative rate was the most oversaturated image and was the smallest image analyzed. The subsampling to create the color filter array

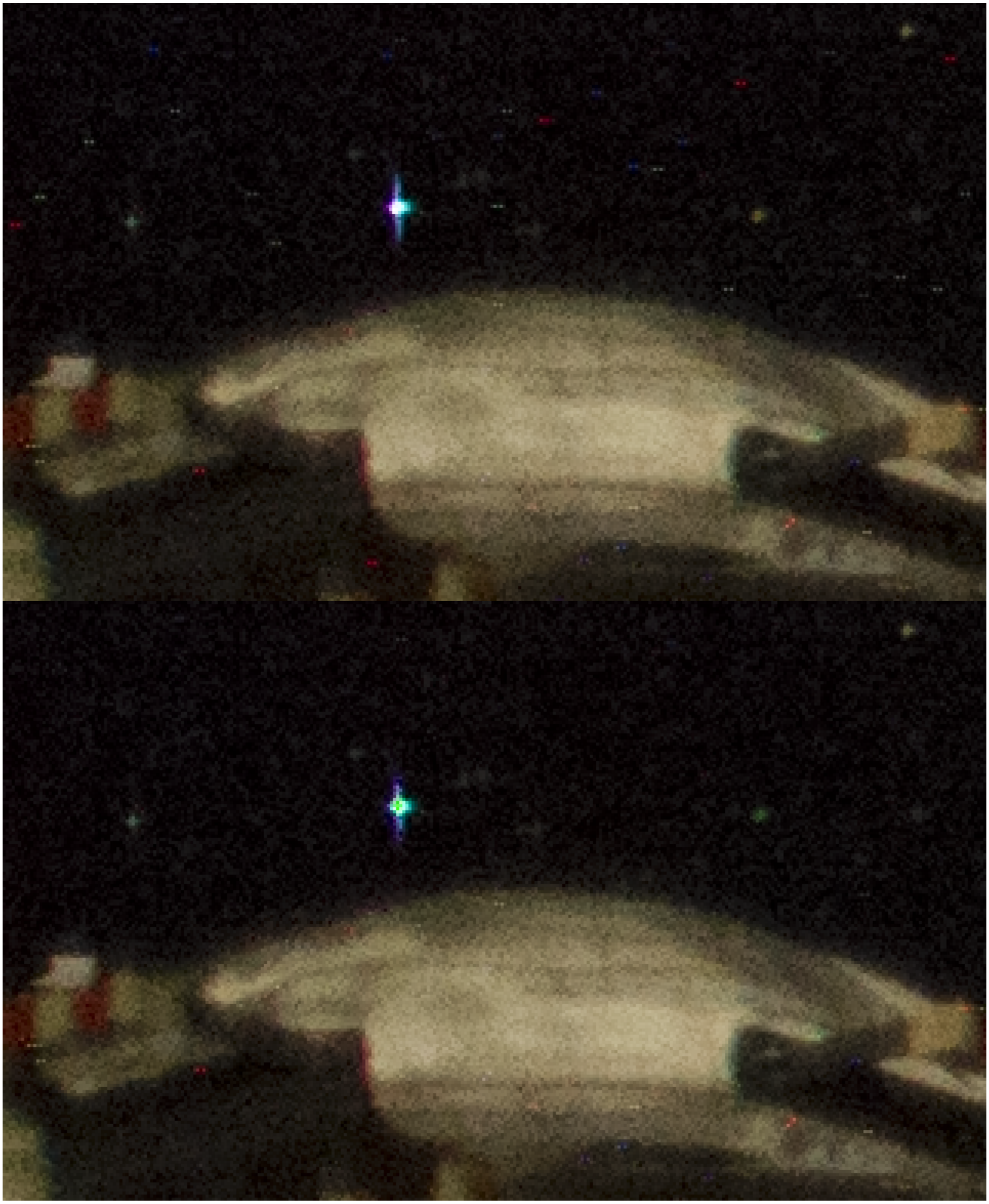

Fig. 5. Before and after z-score damage removal algorithm, new artifacts created in center of stars and close edges from process and damage is missed on the station structure

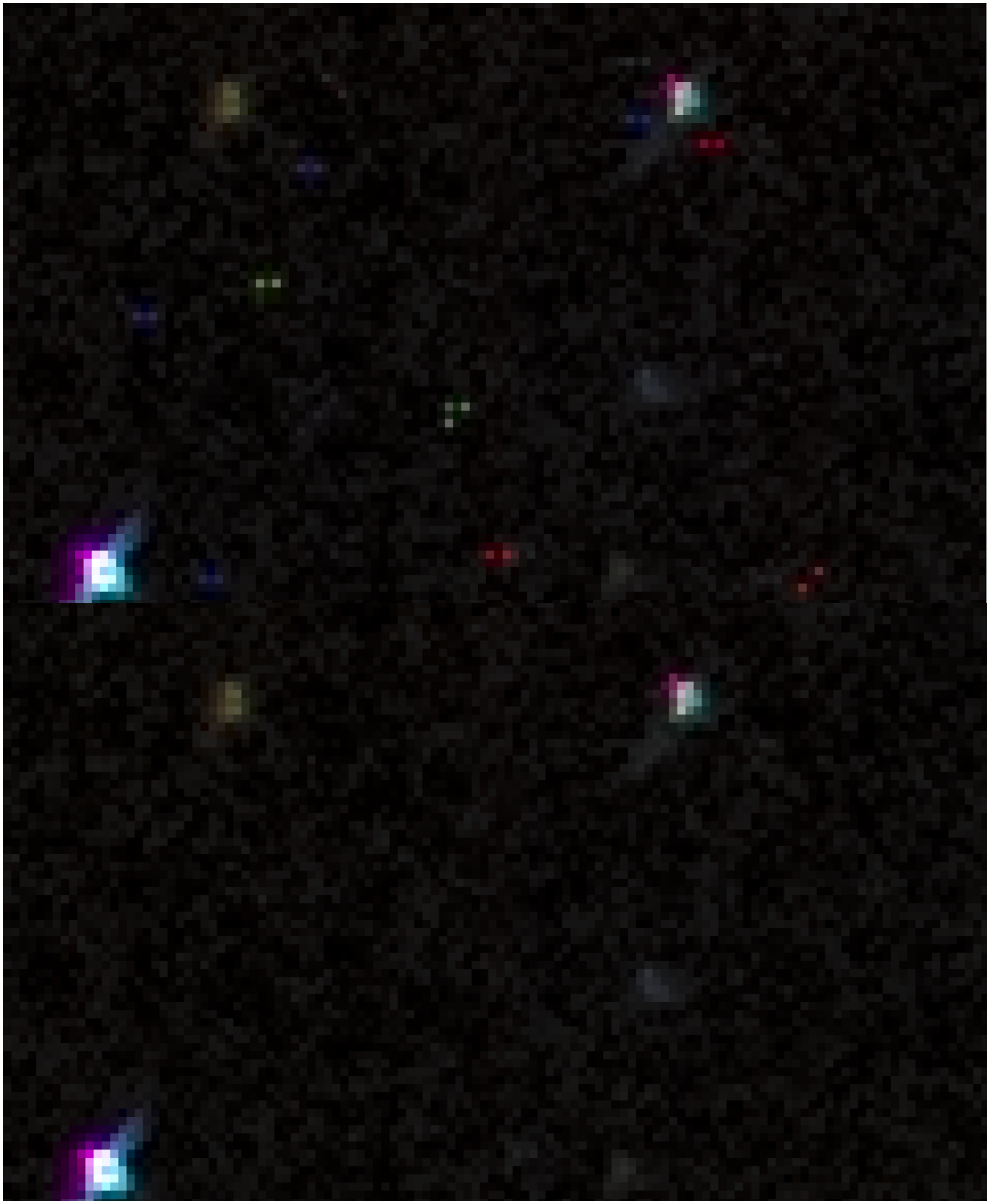

Fig. 6. Before and after structure convolution damage removal algorithm, all damage is removed and stars are left completely unchanged

Method	False Negatives	True Positives	False Negative Rate	False Positives	True Negatives	False Positive Rate
Z-score	386	2185	17.7%	1529	1913015	0.08%
Structure	44	2185	2.0%	2268	1913015	0.12%

Fig. 7. Table 1 depicts the z-score and structure algorithms false positive and negatives for detecting visible damage in output images

Image	False Negatives	True Positives	False Negative Rate	False Positives	True Negatives	False Positive Rate
Aurora 1	265	16693	1.59%	10486	8277707	0.127%
Aurora 2	394	23380	1.69%	5898	12029612	0.049%
Aurora 3	122	8216	1.48%	2902	4087784	0.071%
Aurora 4	36	1023	3.52%	1643	588801	0.279%

Fig. 8. Table 2 depicts the structure algorithm false positives and negatives for a number of simulated scenes

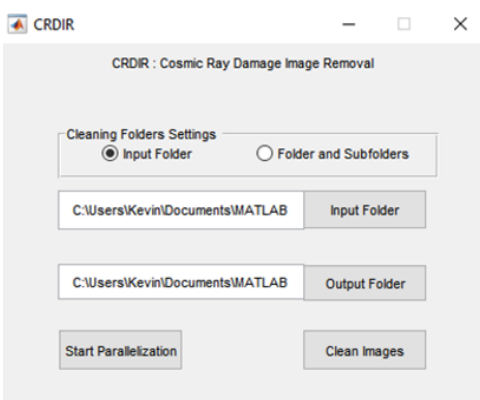

Fig. 9. Matlab GUI designed for processing folders of images

can also remove the number of samples across the small objects in the image leading to the increase in the false positive rate. Overall, the structure convolution method clearly produced the best visual results of the algorithms considered and consistently removed near 98 percent of the cosmic ray damage with minimal effects to other features in the image.

V. User Interface

Finally, the structure convolution algorithm needed to be developed into a system to integrate with NASA's current workflow of image collection and processing and need to be able to quickly handle the large number of images captured. The incoming NEF images can be easily batch processed with the Adobe DNG Converter to uncompress the raw data and create DNG images for reading into the program. Once a folder of all images is established, the algorithm must run through each image and detect the damage, replace with the local medians, and save to a new output file in a specified folder location. For this system, the CRDIR structure convolution code was developed in Matlab, a graphical user interface (GUI) was created, and an executable was generated that could be easily installed on any machine. Within this process, options were developed to allow for parallelization of the repair process so large numbers of images could be run at once on multiple cores. On our machine with the parallelized code running, a runtime was achieved of 32 images in 40 seconds. Our tests were conducted on a PC with a 2.5 GHz Intel Xenon E5-2680 v3 processor with 12 cores for parallel processing. The GUI design as seen in Fig. 9 first allows a user to select an input folder or higher level directory with multiple folders of images for repairing. Next, the user is prompted to select an output location for the repaired images and given a button option to start parallelization if possible. Once all the settings are selected, the clean images button can be selected and a waitbar will appear to indicate the progress through the images.

VI. Conclusion

In conclusion, the cosmic rays traveling outside the Earth's atmosphere can create permanent effects in NASA's DSLR camera images from the International Space Station. While this damage will appear problematic for display and measurement purposes, the effects in the images can be removed by identifying the damaged pixels in the raw image and replacing their values with the local median of that region. The structural convolution method across the color filter array proved the best algorithm for handling the damage effects, quickly detecting nearly 98 percent of the damage effects and with minimal effects to other objects in the scene. The algorithm was evaluated on a mixture of real and simulated data and was developed into a system to deliver to NASA for future use. The progress of this research will assist in the processing of the historical library of data, the latest images from the International Space Station, and any future projects. As spacecrafts venture further from the Earth, the amount of cosmic rays hitting the sensors will greatly increase, leading to progressively shorter lifetimes for these DSLR cameras. The successful removal of damage effects in images means more images can be captured and less equipment is necessary to send on these missions. Across all potential uses, the structure convolution algorithm and workflow developed here will greatly aid in the processing and display of these astounding images.

ACKNOWLEDGMENT

The author would like to thank Don Figer for his role as advisor, Carl Salvaggio and Mark Fairchild as the committee members, Peter Blacksberg as the project sponsor, and David Messinger and Jeff Pelz for their help and support on the outreach of this project.

REFERENCES

[1] Bayer, Bryce E. *Color Imaging Array*, US Patent 3 971 065, 1976.

[2] *Digital Negative (DNG) Specifications*, Adobe Systems Incorporated, San Jose, CA, USA: June 2012

[3] Dong, Yiqiu, and Shufang Xu. *A New Directional Weighted Median Filter for Removal of Random-Valued Impulse Noise*, IEEE Signal Processing Letters, USA: March 2007, 14.3: 193-196.

[4] Freudling, Wolfram. *An Image Restoration Technique for the Removal of Cosmic-Ray Hits from Dithered Images*, Publications of the Astronomical Society of the Pacific, USA: PASP, January 1995, 107: 85-89

[5] Rhoads, James E. *CosmicRay Rejection by Linear Filtering of Single Images*, Publications of the Astronomical Society of the Pacific, USA: PASP, May 2000, 112: 703-710

[6] Pych, Wojtek. *A Fast Algorithm for Cosmic-Ray Removal from Single Images*, Publications of the Astronomical Society of the Pacific, USA: PASP, February 2004, 116:148153

[7] van Dokkum, Pieter G. CosmicRay Rejection by Laplacian Edge Detection, Publications of the Astronomical Society of the Pacific, USA: PASP, November 2001, 113:14201427