Edge Detection as Part of the Demosaicing Process

Ted Cooper
Sony U.S. Research Labs
San Jose, California

Abstract

Traditional image processing workflows place edge enhancement at the end of the process, after demosaicing, white balancing, and chromatic enhancement steps have been applied. This paper investigates employing the edge enhancement simultaneously with the demosaicing process.

Introduction

Edge enhancement is used to produce a more "realistic" image for the viewer. Because the human visual system has sophisticated spatical differencing mechanisms (both chromatic and achromatic)[1,2] the visual response to an object's edge is stronger when the pixels or rosettes occurring in the region of an edge show a greater contrast difference than what is measured optically. Figures 1a and 1b illustrate a simple black to white edge transition going from white to black. Figure 1c illustrates how "unsharp masking" would change the luminance level to accentuate the transition Studies show HVS edge detection increases sharply with only moderate amounts of unsharp masking applied.[3,4] There are many mathematical forms of edge enhancement. Any over-shoot of the edge followed by an under-shoot correction typically provides the optimal class of edge enhancement as long as the transitions occur within 2 units of minimal spatial acuity.

In a traditional DSC, edge enhancement occurs at the end of the information processing chain. At this point, all of the high frequency chromatic information has been subsampled and lost. This means that the edge enhancement process – which attempts to provide "super-nyquist" information into the image – is based upon a low-pass filtered subset of the original information.

One of the most obvious artifacts that results from this "late stage" edge enhancement is the chromatic moire[5] that produces blue and red fringes in the region of a black and white edge. When Bayer-patterned10 CCDs encounter an achromatic edge, the process of reconstructing missing pixel signals distributes the edge transition over a much larger region. Figure 2 shows schematically for an RGB CCD array what happens across the transition. On line 1, the B value occurs on the transition edge. It correctly averages that transition that occurs within that pixel. But the R and G values need to be interpolated from neighboring pixels, which typically are 1.0 to 1.4 pixel units removed from the transition edge. Regardless of whether nearest-neighbor or bi-linear interpolation is used, the estimation of the missing pixel signals, R and G, on line 1 are low-pass filter estimates of the original data. Because of this, large errors in R are typical. The G error is typically half the R error because the G samples occur twice as often in a regular pattern such that a true G sample is never more than 1.0 pixel removed from any pixel location in the image.

Figure 1a. Image with a brightness Step

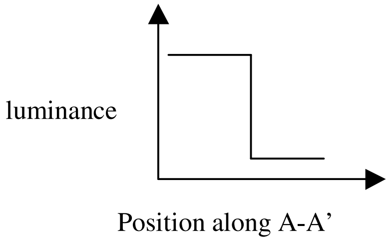

Figure 1b. Normal Edge Brightness

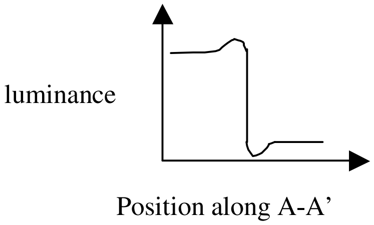

Figure 1c. Edge Enhancement Brightness

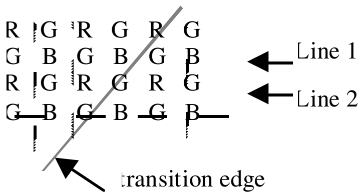

Figure 2. Color Filter Array with an Edge

On line 2, the R pixel is centered on the transition edge. Again the missing pixel signals B and G, for the R pixel cell, are spatially separated from transition region. In this case, the B pixel signal is strongly low-pass filtered. Because of the systematic organization of B and R pixels, an over-estimate of the missing R pixel on one line will typically be followed by an over-estimate of the B pixel on the next line (or some modulo of lines depending on the angle of the achromatic edge to the CCD pixel array). The same logic holds for under-estimates resulting from the pixel interpolation scheme. The net result is that the a-chromatic edge has red steps, followed by blue steps along the edge. Figure 3 gives an example (greatly magnified) for a black and white text label on MacBeth Color Chart. Note the blue and red pixel alternations along the number “2” in the image. This is chromatic moire, and is greatly exaggerated because the edge enhancement (unsharp masking) occurs late in the signal processing workflow.

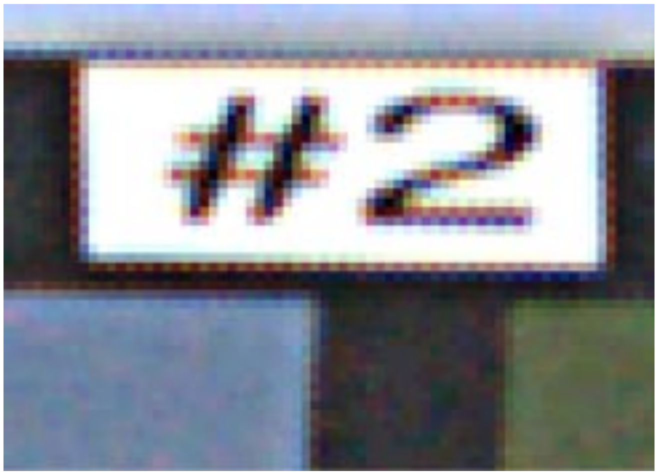

Figure 3. Traditional Unsharp Masking late in the Workflow

Algorithm Overview

By employing edge enhancement at the same time as the demosaicing process, the edge enhancement algorithm has access to the full high-frequency data set digitized from the Bayer-patterned CCD.

The first stage of the algorithm is to create a “gray” (or brightness) map of the image using the raw RGB data. Any type of demosaicing method may be employed to extract a first order approximation for the R, G, and B pixel values to create the gray map.[6,7,8,9] Attempts using a simplified luminance approximation[4] for the gray map found in Equation 1 failed noticably whenever chromatic edges were present. Instead, a luminance based on L*a*b* is much more robust. However, it is not required to use the full CIE L* calculation.

$$Lumiance = 0.177*R + 0.813*G + 0.011*B[failed] \qquad (1)$$

Once the “gray” map is created, a threshold criteria is applied to each pixel to determine if it is involved in a significant brightness change compared to its neighbors. Typical values for the threshold range from 20 – 35 units based on 8-bit color primary imagery. Once a central pixel is determined to have passed the threshold criteria, then the algorithm investigates the 5x5 pixel neighborhood around the central pixel. Each of the twenty five pixels in the neighborhood are evaluated to determine if they also belong to the group that “passed” the threshold criteria or “failed” it. For each neighborhood pixel that “passed”, that pixel is entered into the average value computation as well as a search to find the maximum and minimum. Likewise, each neighborhood pixel that failed the criteria is also entered into a separate average value computation and a separate search for the maximum and minimum brightness for the “failed” criteria group.

At this point, there are two average brightness values. One for the neighborhood pixels that were undergoing significant brightness transitions, and another for pixels that had relatively constant brightness (and/or color).

Next a test is applied to compare that central pixel brightness (which must have “passed” the threshold criteria to even be considered) with the two average values. If the central pixel’s brightness is closer to the “passed” group, then a further test is made to compare the central pixel brightness to the maximum and minimum brightness of that group. Whichever is closer is declared the winner. The RGB values for the winner are stored in the RGB values of central pixel in question. Similarly, if the central pixel’s brightness was closer to the “failed” group’s average brightness, then that group’s maximum and minimum are compared to the central pixel and a winner is declared. Again, the RGB values of the winner from the “failed” group are stored into the RGB value for the central pixel. At any other pixel in the image whose central pixel is not part of a strong gradient, the original demosaiced RGB values are used.

The foregoing steps have created a “mask” image. The mask pixels that are different from the original image are located in strong brightness gradient areas. The final “color” of each changed mask pixel is either a maximum or minimum from the pixels that surround it. The mask highlights the sharp edges of the image and only has colors that are found in the original image (to the extent they exist in the demosaiced approximation). The changed mask pixels has slowly varying colors and looks somewhat block-like in appearance.

Since the "mask" image has strong visual steps in its contour, it needs to be smoothed by adding a fractional part of the original demosaiced image to it.

$$Output = Mask * (1-\alpha) + Original * \alpha \qquad (2)$$

A blending function displayed in equation 2 results in the final RGB output. This yields the original demosaiced RGB values everywhere except where strong gradients have been found.

Results

Figure 4 shows the unmodified demosaiced image whose raw data also created Figure 3. Significant color fringing is present in the image. Figure 5 displays the "mask" image generated from Figure 4. Note that the mask image has fewer colors and more smoothly varying colors than the original. The visibility of the textual edges is more pronounced, but also quite block structured. Perhaps the "mask" image's most striking feature is the lack of chromatic moire around the slanted edges of the text. Figure 6 shows the final output with $\alpha = 0.5$.

Figure 4. Unmodified Demosaic Image using Bi-linear Interpolation. (8 times magnification)

Figure 5. Mask image (8 times magnification)

Figure 6. Demosic plus Mask with α = .5 (no magnification)

An explanation for the removal of "color fringing" or chromatic moire is as follows. Most objects in an image have a relatively constant color hue with changing luminance levels depending on the curvature of the object and/or the geometry of the scene. This translates into an object having a relatively constant R/G and B/G ratio across its surface and out to its edges.[11] When the CCD captures an object's edge, the interpolation errors in determining the missing R and B values cause the red and blue fringing discussed earlier. The G values have a much smaller error because they have twice the sampling frequency.

When a pixel is part of an object's edge, it will typically have a high brightness gradient depending on the brightness of the object next to it. By searching for the maximum brightness pixel in the 5x5 cell neighborhood, one chooses the brightest pixel in the neighborhood that could have been involved in the edge. Since the G pixel typically has the highest component of a pixel's brightness and the G pixel is sampled the most often, the brightness measure of an edge will be the most accurate component compared to its two chrominance values. Since R/G and B/G ratios should remain relatively constant along an object's edge, and the G value should be the most accurate, then by summing or averaging the R and B values in the adjacent pixels, one arrives at the best "average" color at the edge. When one finds the average brightness for all pixels that were above the gradient threshold, one has averaged at least two sets of adjacent R and B pixels. This average brightness found from the summation is a very good measure of brightness at the central pixel.

A similar argument holds for those pixels who "failed" the threshold average in the 5x5 cell neighborhood. The average brightness value is the best measure of the brightness in a constant, non-changing region of the image.

Now to explain why the maximum or minimum brightness reading in the 5x5 neighborhood is chosen. Once the average value has selected the "passed" or "failed" class, the final choice is whether the central pixel brightness is closer to the selected class's maximum or minimum. Since the R/G and B/G ratio should be constant along an object's edge, then the brightest pixel in the neighborhood should have the largest R, G, and B values. The large R, G, and B values provide the added boost needed for edge enhancement like the unsharp masking provides. It has the added feature that larger primary values typically yield lower chromaticity, and this in turn reduces the red and blue fringing at edge boundaries.

Similar logic applies to the minimum brightness value. It has low values for R, G, and B. The smallest values of R, G, and B approaches the gray axis and yield a low chromaticity. The smaller primary values provide the undershoot of brightness needed for the unsharp masking effect.

Figures 7, 8, and 9 show the final combination of the mask image with the original demosaiced image for values of α in equation 2 of 0.5, 0.7, and 0.8, respectively. The optimal value depends upon what is the desired use for the image. If soft smooth edges are important – like in portrait and art images, then larger values of α are needed. For images with small text or fine jewelry, smaller values are needed. Images of bright colorful outdoor scenes where helped least by the edge enhancement since high contrasts required lower α values to decrease block structures. It should be noted that none of the images in this paper are White Balanced since that step follows the demosaicing process.

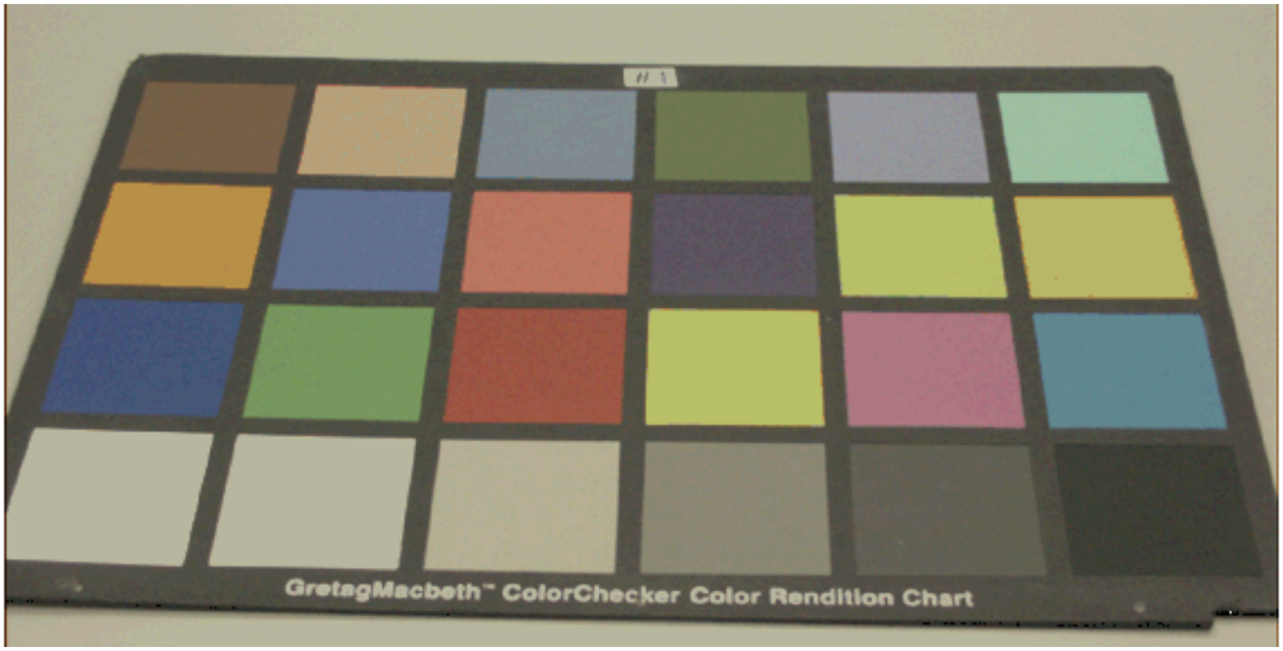

Figure 7. Demosaic plus Mask with α *=0.5 (1/2 magnification)*

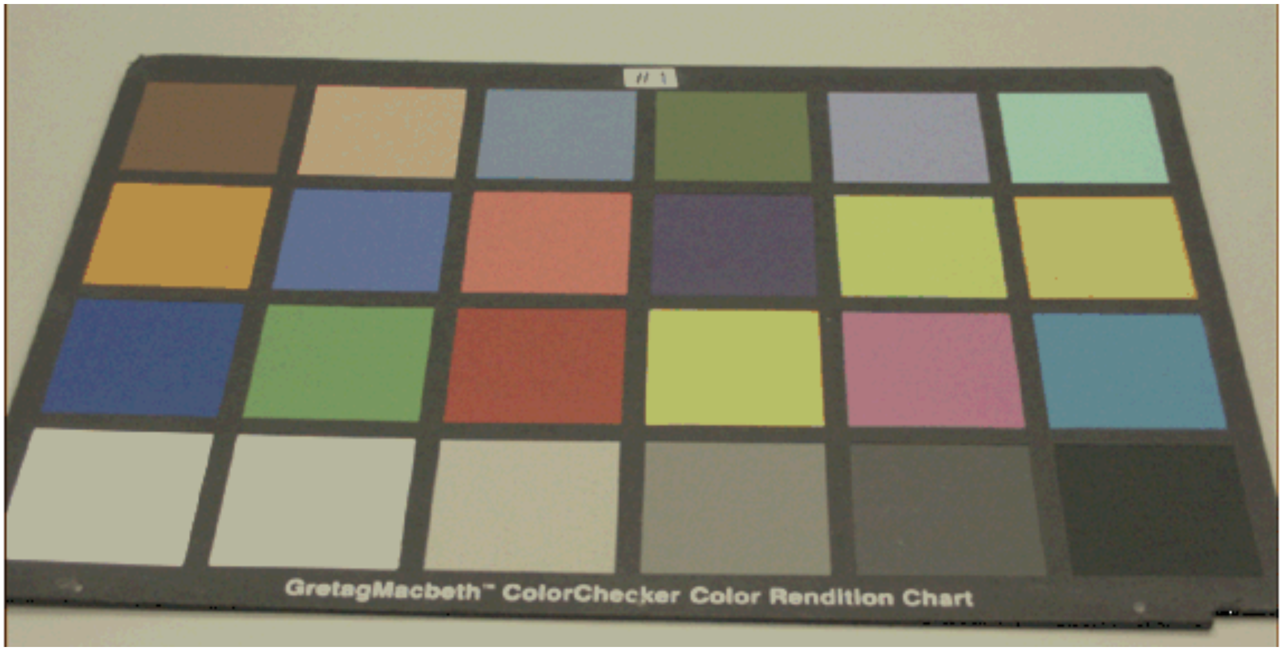

Figure 8. Demosaic plus Mask with α *=0.7 (1/2 magnification)*

Figure 9. Demosaic plus Mask with α =0.8 (1/2 magnification)

Conclusions

The current method provides good results for text and fine impulse details in images. The method is very fast because calculations for the 5x5 neighborhood only need to be performed when the central pixel contains a strong brightness gradient. Most images have only a few percent of the pixels that have strong brightness gradients.

Another reason for the success in employing edge enhancement as part of the demosaicing process is that color fringe reduction can occur at a very early stage of the process. When it occurs late in the workflow, the edge enhancement accentuates the red and blue fringing because the gamma correction process will boost the red and blue pixels that occur on white or neutral backgrounds.

Further extensions to this research would include using segmentation techniques to find edges of objects larger than the 5x5 cell size. If large edges could be localized to sub-pixel accuracy in an image, then the selection of which pixel values to include in the "passed" and "failed" groups could be limited to only those pixels that were changing in the direction of the edge. This would further refine the "mask" image and allow proper shading of sub-pixel components occurring in the edge. It would also permit better estimation of the true edge color rather than just using the maximum and minimum brightness pixels found in the neighborhood. While the edge enhancement will still need to exaggerate the brightness and darkness across the edge, a better image will result when a precise chromaticity value is known for both sides of the edge.

References

1. Semir Zeki, A Vision of the Brain, Blackwell Scientific Publications, Oxford, 1993, pp. 122-130.
2. Brain Wandell, Foundations of Vision, Sinauer Associates, Sunderland, Mass.,1995, pp. 135-138 and 393-402.
3. Mark Fairchild, Color Appearance Models, Addison-Wesley, Reading, Mass., 1998, pp.30-33.
4. W.K. Pratt, Digital Image Processing, John Wiley & Sons, New York, 1978, pg. 737 and pp. 322-326.
5. Ozawa and Takahashi, IEEE Trans. on Electron Devices, Vol. 38, No. 5, May 1991, pp. 1217-1225.
6. J.E. Adams, Design of Practical Color Filter Array Interpolation Algorithms for Digital Cameras, Part 2, 1998 International Conference on Image Process, vol. 1, pp. 488-492.
7. Kuo-Tang Kuo and Sau-Gee Chen, Fast Integrated Algorithm and Implementations for the Interpolation and Color Correction of CCD-Sensed Color Signals, IEEE transactions, 0-7803-4455-3, 1998, vol. 4, pp. 225-228
8. T. Sakamoto, C. Nakaishi, and T. Hase, Software Pixel Interpolation for Digital Still Cameras suitable for a 32-bit MCU, IEEE Transactions on Consumer Electronics, Vol. 44, No. 4, Nov. 1998, pp. 1342-1352
9. B.Tao, I. Tastl, T.Cooper, M. Blasgen, and E. Edwards, Demosaicing using Human Visual Properties and Wavelet Interpolation Filtering, 7th Color Imaging Conf. IS&T, Nov. 1999 pp.252-256.
10. Bayer Color Filter Array pattern, U.S. Patent 3 971 065, 1976.
11. Ron Kimmel, Demosaicing: Image Reconstruction from Color CCD Samples, IEEE Trans. on Image Processing, Vol. 8, No. 9, Sept 1999, pp.1221-1228.