MTF Evaluation of White Pixel Sensors

Albrecht Lindner, Kalin Atanassov, Jiafu Luo, Sergio Goma

Qualcomm, Incorporated, 5775 Morehouse Drive, San Diego, CA, 92126

ABSTRACT

We present a methodology to compare image sensors with traditional Bayer RGB layouts to sensors with alternative layouts containing white pixels. We focused on the sensors' resolving powers, which we measured in the form of a modulation transfer function for variations in both luma and chroma channels. We present the design of the test chart, the acquisition of images, the image analysis, and an interpretation of results. We demonstrate the approach at the example of two sensors that only differ in their color filter arrays. We confirmed that the sensor with white pixels and the corresponding demosaicing result in a higher resolving power in the luma channel, but a lower resolving power in the chroma channels when compared to the traditional Bayer sensor.

Keywords: Color Filter Array, MTF, White pixel, WRGB sensor, Camera-phones, Image quality.

1. INTRODUCTION

The RGB Bayer layout is well established for color filter arrays (CFA) in color image sensors. However, there are ongoing efforts to explore alternative designs with different spectral sensitivity curves and more complex CFA patterns [1]. One possible alternative CFA design uses clear/white[1] pixels, i.e., pixels that collect photons across the entire visible spectrum (WRGB sensor) [2] [3] [4]. The benefit of white pixels is that they are more sensitive and thus have better performance in low-light conditions. The drawback is that white pixels are achromatic and thus do not provide any information related to color. In order to assess the benefits and drawbacks of WRGB versus RGB layouts, we propose a thorough evaluation method that measures the resolving power in the luma and chroma channels. We evaluated the overall resolving power of the entire imaging pipeline.

We constructed a test target that had four quadrants and was a slightly modified version of the one presented by Goma and Atanassov [5]. Two quadrants were achromatic, where the luma channel was modulated, and the other two quadrants were chromatic, where the chroma channel varied, but the luma channel was constant. Of the two achromatic (chromatic) quadrants, one was a Siemens star [6] and the other was a zone plate. We arranged the four quadrants into one image and placed markers in the corners for automatic detection. The target was then printed on non-glossy paper and mounted on a foam board.

We used two sensors that differ only in their CFA layout — one with a traditional Bayer layout and one with a WRGB layout. We took images with each sensor at 10 different lux levels and adjusted the sensor gain in order to have the same digital count on the sensor.

We automatically analyzed the target photos to estimate the sensors' MTF[2] curves for all lux levels. The analysis shows that the added white pixels improve the resolving power for the luma channel at the expense of resolving power in the chroma channels. In regards to the chroma channels, we found that at the spatial frequency where the MTF of the Bayer sensor falls off to half of its starting value, the MTF of the WRGB sensor is already at one fourth of its starting value.

The subsequent sections explain the study more exhaustively. Section 2 describes the test target and the image acquisition procedures. Section 3 details the evaluation procedure and measures of resolving power. Lastly, Section 4 concludes the article with a discussion of the results.

[1] *Clear* pixel is a more accurate term as there is no color filter, but we use *white* as it is more widely used.
[2] Strictly speaking we measure the spatial frequency response, but it is an indication of the MTF.

Image Quality and System Performance XII, edited by Mohamed-Chaker Larabi, Sophie Triantaphillidou,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 9396, 939608 · © 2015 SPIE-IS&T
CCC code: 0277-786X/15/$18 · doi: 10.1117/12.2080193

2. TEST TARGET

2.1 Target Design and Acquisition

The test target is reproduced in Figure 1 below. It consists of a zone plate (top) and a Siemens star (bottom), which are in gray levels (left) and with constant luma (right). If Z is a one channel image of a zone plate or Siemens star with values in the interval [0 255], the gray level side is generated by replicating Z over the three color channels R, G, and B. The constant luma side ($luma = 128$ in our case) is produced by:

$$R = Z \qquad B = G = \frac{luma - 0.299Z}{0.701} \tag{1}$$

The frequency in the zone plate is a linear function of the radius, i.e., distance to the center. The constant patches in the center represent the maximum and minimum values of the sinusoidal modulation. More details regarding the creation of the test target are included in the references [5]. To facilitate automatic processing of the acquired photos, we added control and tick marks around the target.

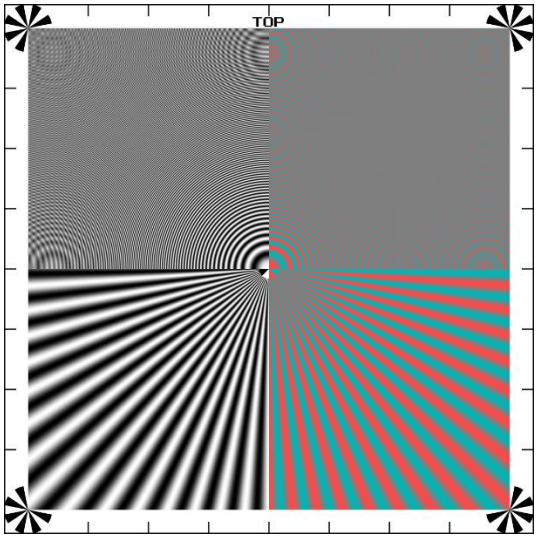


Figure 1. Test target with control and tick marks. The left side is achromatic; i.e., only the luma varies, and the right side has a constant luma value; i.e., the chroma values vary.

2.2 Scaling and Printing

The highest frequency on the Siemens star is obtained in the center at the border to the constant patches. For the zone plate, the highest frequency is obtained on the center-left and center-right borders and on the top. The corners contain higher frequencies, but these parts are not used in the current implementation of the evaluation.

We scaled the frequencies in the image so that the highest frequency in all sectors was half the sampling frequency if the image was acquired at a width of 1000 pixels. The width is measured for the inner part of the target without the cross and tick marks. After the frequencies were set, we generated a high resolution image of the test target with a resolution that was five times higher than the original. Therefore, the highest frequency was not two pixels per cycle, but 10.

The test target was printed on a non-glossy standard white paper and mounted on a foam board for rigidity. The width of the inner part with the sinusoidal target was 20 inches; the width of the entire image with the control and tick marks was 22 inches. We additionally left a white border totaling in a target size of 24 x 24 inches.

2.3 Image Acquisition

We acquired images of the test target with two similar sensors that have different color filter arrays. One is a traditional Bayer RGB and the other a WRGB layout. We used the following guidelines for the acquisition:

1. Place the target near the center (as perpendicular to module lens as possible), with uniform illumination. Place it in a distance so that the center image (zone plate and Siemens star) occupies approximately 1000 pixels.
2. To reduce the number of parameters, use 30fps and full frame exposure for all pictures.
3. For each sensor, maximum light level, A, is the light level where analog gain of 1x results in average luma of about 200DN in 10bit (after subtracting OB level) and minimum light level, B, would be the light level where analog gain of 16x results in the same average luma level.
4. Take 10 images for each sensor under different light levels, evenly from maximum level A to minimum level B.
5. Save the raw sensor images.

The acquired raw images are then demosaiced for the automatic evaluation described in the next section. An example of an acquired and demosaiced image is shown in Figure 2.

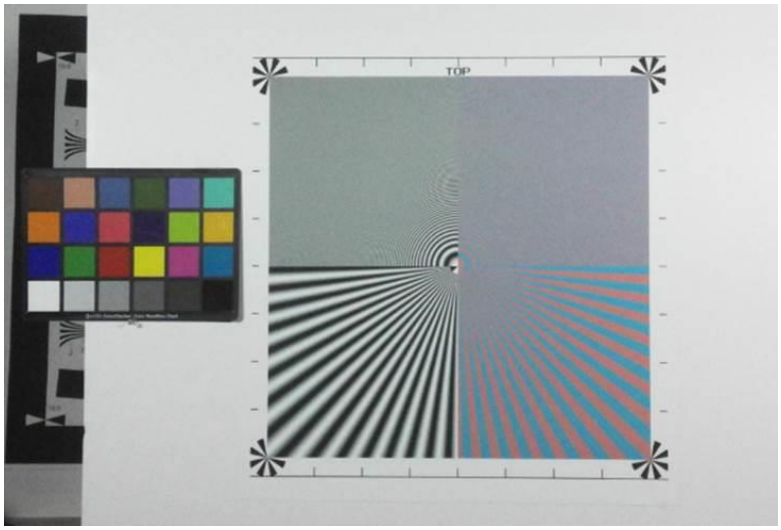


Figure 2. Acquired image of the test target. The target is approximately 1000px across so that the highest spatial frequency on the target is two pixels per cycle.

3. TARGET EVALUATION

The evaluation of the target image uses a demosaiced image as input. The evaluation methodology described in this Section is independent of the specific demosaicing algorithm and can thus be used to compare different combinations of sensor layout and demosaicing algorithm. Figure 3 shows the evaluation workflow in a brief overview; details for each block are given in the following subsections.

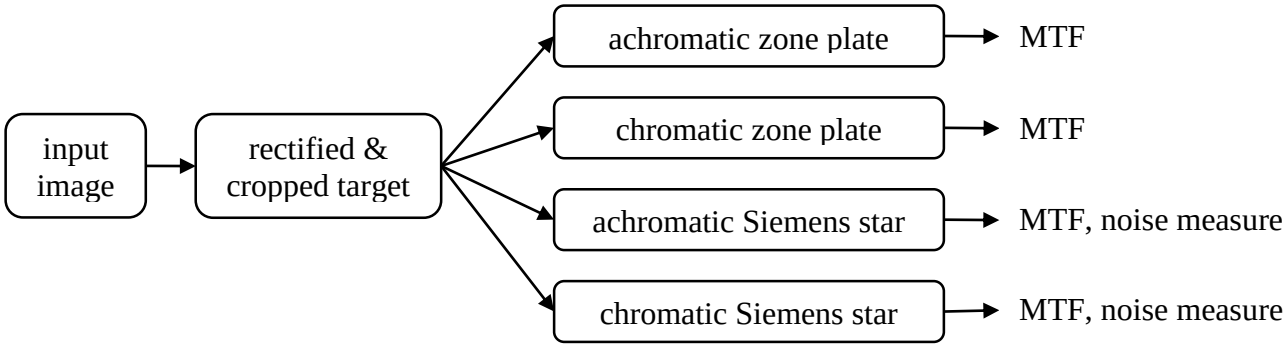


Figure 3. Block diagram of the evaluation workflow. The steps are detailed in the following subsections.

3.1 Corner Detection and Rectification

In order to find the four control marks in the corners of the test target, we first split the entire image horizontally and vertically in four equally sized sub-images. If the test target is approximately centered, each sub-image will contain one of the corners. We correlated a template of the control mark over the image and found the maximum position of the correlation for all four corners.

We calculated the four distances between the four corners at the top, bottom, left and right borders. According to the acquisition instructions, these distances should be 1000 pixels, but can vary slightly due to incorrect scaling or perspective distortions. We used an affine transformation to map the inside of the four control marks to a perfect square where the edge length is given by the average of the four previously measured distances. We used a linear interpolation for this transformation and applied it once per channel. Due to a very careful setup of the scene the corrections were minimal, e.g. the 1000 pixel distance was off by less than 10 pixels for all images. The impact of the interpolation on the final result is thus small. The image from Figure 2 is shown in Figure 4 after being rectified and cropped using this method.

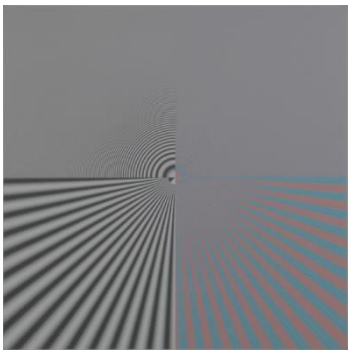


Figure 4. Rectified and cropped image from Figure 2

3.2 Analysis of the Siemens Star

The center of the Siemens star shows two constant patches that indicate the maximum amplitude. We estimated the amplitude as half the difference between the 1% and 99% quantiles in that region of the image as indicated by the green outline in Figure 5.

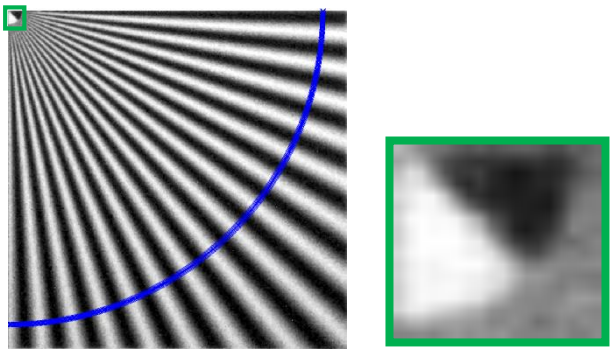


Figure 5. Siemens star with constant patches in the center (highlighted in green). We estimate the maximum amplitude as half the distance between the 1% and 99% quantiles in that region. The blue line indicates a circular line of constant spatial frequency. The circular blue line is used in Figure 6

We use bilinear interpolation to compute image values along the circular line shown in blue in Figure 5. Figure 6 shows the interpolated values next to its Fourier domain. Again, the DC component is removed for better scaling.

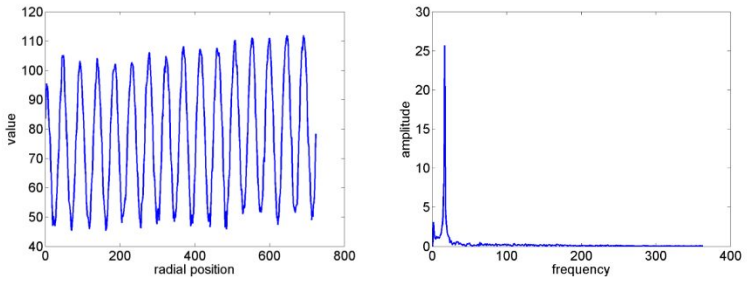
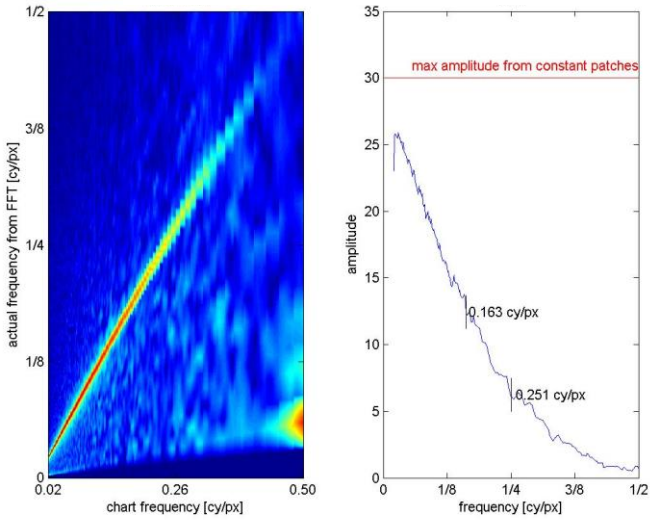


Figure 6. Left: Image values along circular line (see Figure 5). Right: Signal in Fourier domain

If the above procedure is computed for multiple circular lines, i.e., chart frequencies, we get multiple FFT curves as in Figure 6. We can concatenate all FFT curves for different chart frequencies to a color coded image as shown in the left portion of Figure 7. The prominent ridge along the diagonal direction indicates the spatial frequency response of the camera system. The values off the diagonal indicate noise in the frequency spectrum and ideally should be zero. For the further analysis we ignore the values not on the diagonal. The right portion of Figure 7 shows the values along the diagonal line. In addition to the spatial frequency response curve, we see the maximum amplitude estimated from the constant patches (see Figure 5) indicated by the red line. The two tick marks indicate where the curve falls off to half and one quarter of its initial value, respectively.



85

Figure 7. Left: Color coded Fourier transforms of sinusoidal signal along different circular lines. Right: Values along the diagonal ridge of the figure on the left indicate the MTF. The red line indicates the maximum amplitude estimated from the constant patches shown in Figure 5. The two tick marks indicate where the curve falls off to half or one quarter of the initial value.

3.3 Analysis of the Zone Plate

The zone plate is designed so that the frequency increases linearly with the distance to the center. The decrease of the amplitude along a radial line thus is a direct measure of the MTF. This is indicated with the red envelope in Figure 8.

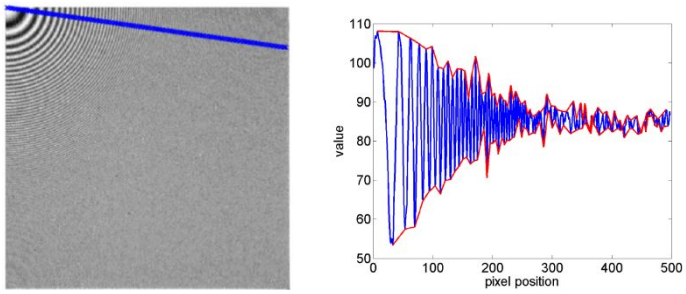


Figure 8. Left: the frequency increases linearly along a radial line on the zone plate. Right: Image values along this line show the dampening of the signal amplitude with increasing spatial frequency. The red envelope is a direct estimate of the MTF.

The zone plate also can be used for noise analysis as shown in Figure 9. Ideally the values along a circular line (see left portion of Figure 9) are constant and thus the Fourier domain (see right portion of Figure 9) only contains a DC component. The average power in the FFT domain is thus an indicator for the presence of noise.

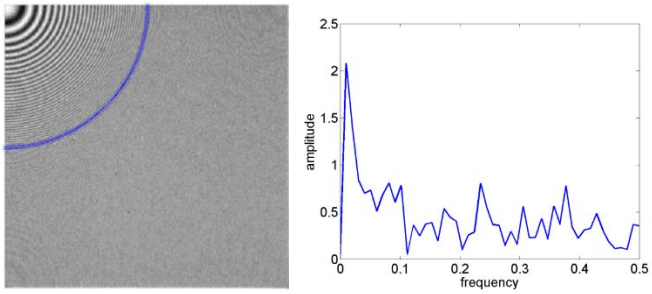


Figure 9. Fourier transform (right) of values along a circular line (left). Ideally, the values in the Fourier domain all vanish except at frequency 0. The value at frequency 0 has been set to 0 for better scaling

Figure 10 shows a summary of the above steps for an entire zone plate. The left portion shows a color coded concatenation of MTF curves (derived from the envelope in Figure 8) for different directions. The center shows an averaged MTF curve for all direction from horizontal to vertical and their variance. Finally, the right portion shows the average noise as a function of spatial frequency, i.e., the radius of the circular line as shown in Figure 9.

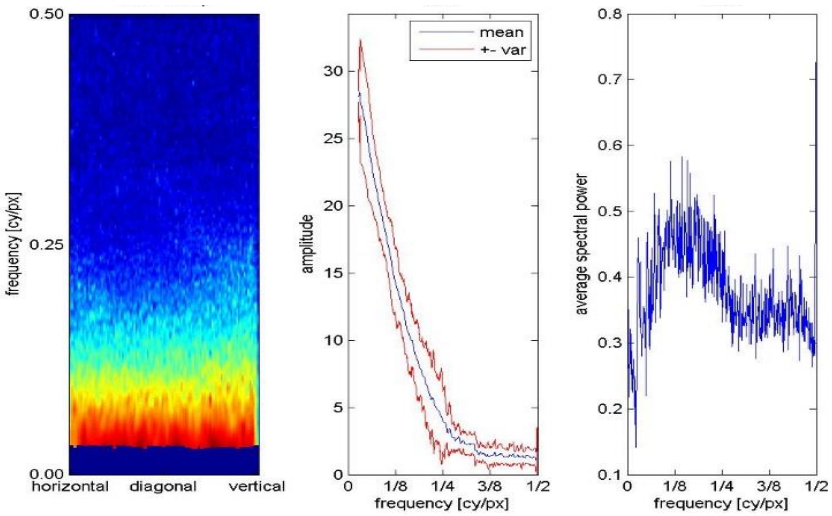


Figure 10. Summary of results extracted from a zone plate. Left: Color coded MTF curves for different directions from horizontal to vertical. Middle: average and variance of all MTF curves. Right: Average noise as a function of spatial frequency, i.e., radius of circular line on zone plate.

4. RESULTS AND FUTURE WORK

We worked with two sets of raw images: the first set was shot with an WRGB image sensor and the second with a comparable camera, but with a conventional Bayer RGB sensor. The raw RGB images were processed with different demosaicing algorithms, two advanced algorithms (denoted *algo1* and *algo2*) and a naïve bilinear approach. The WRGB images were processed with two different algorithms (denoted *algo3* and *algo4*).

Figure 11 shows results for images processed with *algo3*. It shows for increasing gains, i.e., decreasing light intensity, the frequency at which the MTF curve drops to one half and one quarter of its initial value (see also Figure 7). The figure clearly shows that the WRGB sensor has increased resolving power in the lightness channel at the expense of the chroma channel when compared to the conventional RGB sensor. For example the MTF of the WRGB sensor's Y channel drops to one quarter at 0.25 cycles per pixel, whereas this decay already happens at 0.21 cycles per pixel in the case of the traditional Bayer RGB.

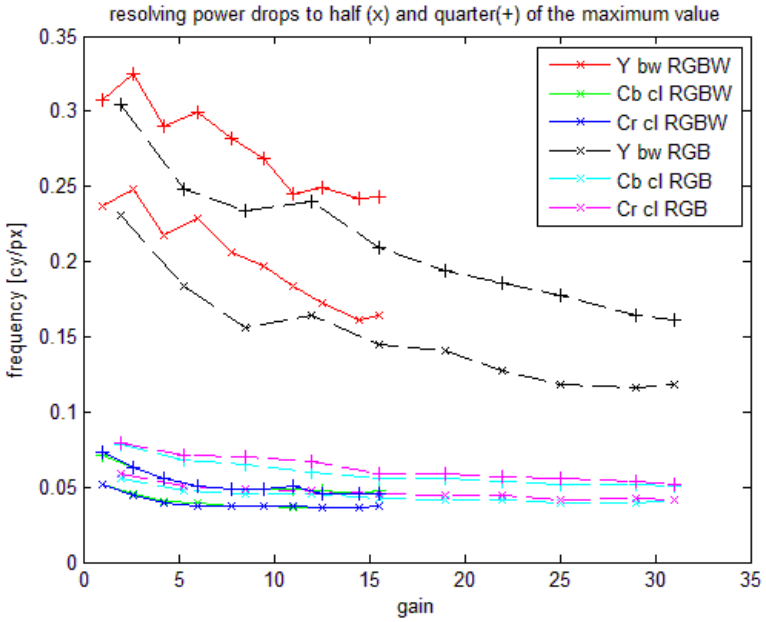


Figure 11. Frequencies where the MTF curve drops to one half (x) and one quarter (+) of its initial value for varying gain levels. The WRGB sensor increases the resolving power for the lightness channel at the expense of the chroma channels.

Figure 12 shows a comparison of all five demosaiced image sets on the Cb channel. The results of the Cr channel are qualitatively the same as for the Cb channel and are not reproduced. The results show that bilinear and *algo2*'s demosaicing of the RGB images have a high resolving power because they do not have noise suppression enabled (see bottom noise plot). We further observe that *algo1* on the RGB images and both algorithms on the WRGB images have equal amounts of noise suppression. However, the WRGB images have a poorer resolving power in the chroma channel due to the added white pixel.

As future work we propose to use the proposed methodology to perform a more exhaustive evaluation of image sensors and demosaicing algorithms, because both aspects go hand in hand to achieve an optimal resolving power. The observed trend that white pixels increase the resolving power in the lightness channel, but decrease it in the chroma channels, can then be further assessed.

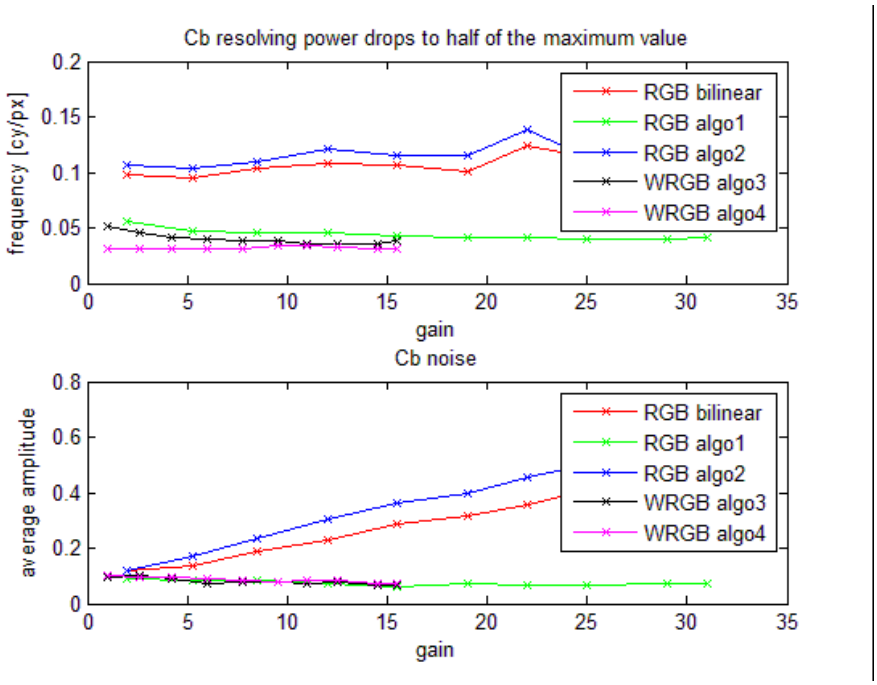


Figure 12. Comparison for all five demosacied images for Cb channel. The top figures show the frequency at which the MTF curve drops to half of its value. The bottom figures plot the corresponding noise levels

REFERENCES

[1] Z. Sadeghipoor Kermani, L. Yue and S. Süsstrunk, "Optimum Spectral Sensitivity Functions for Single Sensor Color Imaging," in *Proceedings of IS&T/SPIE EI: Digital Photography VIII*, 2012.

[2] H. Honda, Y. Iida and Y. Egawa, "A Color CMOS Imager With 4 × 4 White-RGB Color Filter Array for Increased Low-Illumination Signal-to-Noise Ratio," *IEEE Transactions on electron devices*, vol. 56, no. 11, pp. 2398-2402, 2009.

[3] A. Getman, J. K. Kim and T.-C. Kim, "Imaging system having White-RGB color filter array," in *IEEE International Conference on Image Processing*, 2010.

[4] C.-s. Wang and J.-w. Chong, "An Improved White-RGB Color Filter Array Based CMOS Imaging System for Cell Phones in Low-Light Environments," *IEICE TRANSACTIONS on Information and Systems*, vol. 97, no. D, pp. 1386-1389, 2014.

[5] S. Goma and K. Atanassov, "Evaluation methodology for Bayer demosaic algorithms in camera phones," in *SPIE 7537, Digital Photography VI*, 2010.

[6] C. Loebich, D. Wüller and B. Klingen, "Digital Camera Resolution Measurement Using Sinusoidal Siemens Stars," in *SPIE-IS&T Electronic Imaging Digital Photography III*, 2007.