Evaluation Methodology for Bayer Demosaic Algorithms in Camera Phones

Sergio Goma, Kalin Atanassov,

Qualcomm Inc., 5775 Morehouse Drv., San Diego, CA, 92126

ABSTRACT

The current approach used for demosaic algorithm evaluation is mostly empirical and does not offer a meaningful quantitative metric - this disconnects the theoretical results from the results seen in practice. In camera phones, the difference is even bigger due to the low signal to noise ratios and also due to the overlapping of the color filters. This implies that a demosaic algorithm has to be designed to allow for graceful degradation in presence of noise. Also, the demosaic algorithm has to be tolerant to high color correlations. In this paper we propose a special class of images and a methodology that can be used to produce a metric indicative of a real case demosaic algorithm performance. The test image that we propose is formed by using a dual chirp signal that is a function of the distance from the center.

Keywords: Demosaic Algorithms, Cameraphones, Image quality.

1. INTRODUCTION

Most of today camera phones are Bayer color pattern based. As a consequence, one of the main blocks of an image processing pipe is the demosaic block, which needs to interpolate a sub-sampled color pattern (Bayer pattern). The demosaic problem although well understood [2], does not have a well defined evaluation procedure. As today camera phone sensors are striving to deliver image quality close to the image quality delivered by DSCs, the demosaic algorithm plays a very important role in camera phone image processing pipe. Development of demosaic algorithms usually is one of the points, even though there are available solutions with reasonable quality[6][8], every year new demosaic algorithms are proposed, making the evaluation of those algorithm a complex problem, as this is mostly done through visual inspection. The common approach used in demosaic algorithm evaluation[1] is:

- A reference image is sub-sampled according to the Bayer pattern (eg. Lighthouse, Sail, etc.)
- The proposed algorithm is used to demosaic the sub-sampled image
- The result is visually compared with the reference image – sometimes MSE is used to quantify the difference.

As reference image, the usual choice is the famous 'Lighthouse'(actually the picture of the Portland Head Light Station in Cape Elizabeth – Maine) picture used from the Kodak picture set. The picture has several features that are making it a very good candidate for demosaic algorithms with the most notable features the fence and the siding of the first building that are offering a very good indication of the artifacts introduced by a demosaic algorithm. In [1] the authors use the power spectral density of this picture to derive a novel demosaic algorithm. As good as this picture is in benchmarking the performance of a demosaic algorithm it's missing some of the features that are typically challenging in a camera phones. In this context, the main assumptions that are not true in camera phones are:

- The reference image used for demosaic benchmarking is already white-balanced and the colors are decorrelated – the reference image has a good saturation. In a real camera phones this is not the case, the image to be demosaiced is not always white-balanced, and usually the colors have high crosstalk. The color crosstalk is usually higher than in a DSC/DSLR due to the fact that the pixels are very small and to

Digital Photography VI, edited by Francisco Imai, Nitin Sampat, Feng Xiao, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 7537, 753708 · © 2010 SPIE-IS&T · CCC code: 0277-786X/10/$18 · doi: 10.1117/12.839103

compensate this, a higher overlap of color filters needs to be used – not to mention additional electrical and optical crosstalk. For this reason, a demosaic algorithm that makes use only of the luma component and disregards the chroma component would offer a good performance when benchmarked over a reference image as the lighthouse, but in reality would offer an unexpected lower performance.

- A real camera phone image will have noise due to the low SNR of the sensor used and also due to the fact that a major usecase for camera phones is low-light indoor scenes. This is usually a concern for demosaic algorithms especially ones that use H/V gradients (a classical example is [2]), as it can create artificial patterns in the final image due to the fact that in the presence of noise, the algorithm will switch between horizontal and vertical gradients. It follows that in this case, a demosaic algorithm that performs well on a clean reference image might offer in reality a lower performance by introducing unexpected artifacts.

Noting that one of the most important qualities of a demosaic algorithm is its frequency response, but also the fact that a higher frequency response of the optics in a camera can lead also to undesired artifacts, a demosaic algorithm should offer a good frequency response while ideally not introducing any artifacts above the sampling frequency of the Bayer pattern – which is one of the feature that is checked visually in the lighthouse picture. The artifacts introduced by demosaic algorithms above the sampling frequency are noticeable as false colors (zipper artifact). The artifacts are hard to quantify using simple MSE approach so perceptual studies are conducted[3] to evaluate performance of different demosaic algorithms. The latter are however very subjective and usually non-repeatable.

Considering the above, we propose a methodology that tries to enhance the artifacts introduced by a demosaic algorithm and also tries to stress the areas outlined above, which are missing in a traditional benchmark against reference pictures like the lighthouse. Our methodology relies on a test chart that includes the following features:

- For frequency response estimation we have both a chirp pattern (zone plate) area and a segment of a Siemens star;
- For luminance tolerance we developed a constant luminance pattern in which the luminance plane is a flat gray and visible details exist only in the chrominance plane;
- For noise tolerance we used combinations of the above in presence of noise.

The structure of our paper is as follows: we present further the methodology we propose to be used in demosaic benchmarking: the chart and feature details, data measurements and a benchmark of several demosaic algorithms using the proposed methodology.

2. METHODOLOGY

2.1 Problem description

In a typical imager system, image capturing process involves sampling on a three-color grid. During front end processing, the three different color planes are up-sampled (demosaic-ed) to artificially bring each color plane resolution to the original resolution[7]. The demosaic process is critical to image quality due to its three-color processing and frontal position in a typical pipeline. Artifacts introduced by the demosaic process are often very objectionable and generally are hard to remove in subsequent processing. Evaluating demosaic process performance is a critical task when designing image pipelines[5]. Moreover producing a metric that can be used to represent demosaic-ing quality is useful tool when absolute or relative performance of different algorithms need to be assesses. In this paper we propose a special class of images and a methodology that can be used to produce a metric or higher dimensional curve that represents demosaic algorithm performance.

2.2 Test Chart

The test image that we propose is formed from using a dual chirp signal that is a function of the distance from the center (zone pate) and a Siemens star. Additionally there is a grey patch that is used to measure noise characteristics and color white pattern and constant luma pattern and has resolution of 1024x1024. The zone plate image (single channel) is generated as:

$$\mathrm{d} = \left(\mathrm{h}^2 + \mathrm{w}^2\right) \cdot \mathrm{sc}$$
$$\mathrm{dist(i,\,j)} = \sqrt{(\mathrm{i} - \mathrm{center_h})^2 + (\mathrm{j} - \mathrm{center_w})^2}$$
$$\mathrm{Z(i,\,j)} = 256 \cdot \left(\frac{\cos\left(\mathrm{dist(i,\,j)}^2 / \mathrm{d}\right) + 1}{2} \right)^{\mathrm{R}}$$

where the parameters are:

- ***h*** and ***w*** - image height and width in pixels (for that particular chart is ***h = w = 512***);
- ***center_h*** and ***center_w*** - vertical and horizontal zone plate center offsets from the top left corner (for that particular chart ***center_h = center_w = 256***);
- ***sc*** – control the zone plate frequency (chosen to be ***0.0006***);
- ***R*** – sinusoid duty ratio (chosen to be 0.8 for the B&W chart and 1 for constant luma chart);
- ***i = 1:h*** and ***j = 1:w***

The black and white zone plate is generated as

$$R = G = B = Z$$

The constant luma pattern is generated as (cyan – magenta)

$$R = Z,$$
$$B = G = \frac{luma - 0.299 \cdot x}{0.701}$$

Where ***Z*** is the zone plate image generated as described above and ***luma*** is the desired luminance level and can be from 77 to 178 (is chosen to be 160 for this particular example).

The Siemens star is generated as follows

$$V(i,j) = \left(i - h/2\right) + \sqrt{-1} \cdot \left(j - h/2\right)$$
$$\alpha(i,j) = \mathrm{angle}\left(V(i,j)\right)$$
$$\mathrm{d(i,\,j)} = \left|\mathrm{V(i,\,j)}\right|;$$
$$\mathrm{S(i,\,j)} = \begin{cases} \mathrm{round}\left(128\left(\cos\left(2\alpha\right)+1\right)/2\right) * 255 & \text{if } \mathrm{d(i,\,j)} < \mathrm{p} * \mathrm{h} \\ 128 * \left(\cos\left(\alpha * \mathrm{freq}\right)+1\right) & \textit{else} \end{cases}$$

Where the parameters are:

- ***h*** star height and width in pixels (for this particular chart ***h = 512***);
- ***freq*** – angular frequency in terms of number of rays (for that particular chart ***freq = 120***);
- ***p*** – fractions of the radius used for the center shape frequency (chosen to be ***0.05*** for the test chart)
- ***i = 1:h*** and ***j = 1:w.***

Black and white and constant luma figures are generated the same way as for the zone plate image.

The generated in that manner image has the following properties:

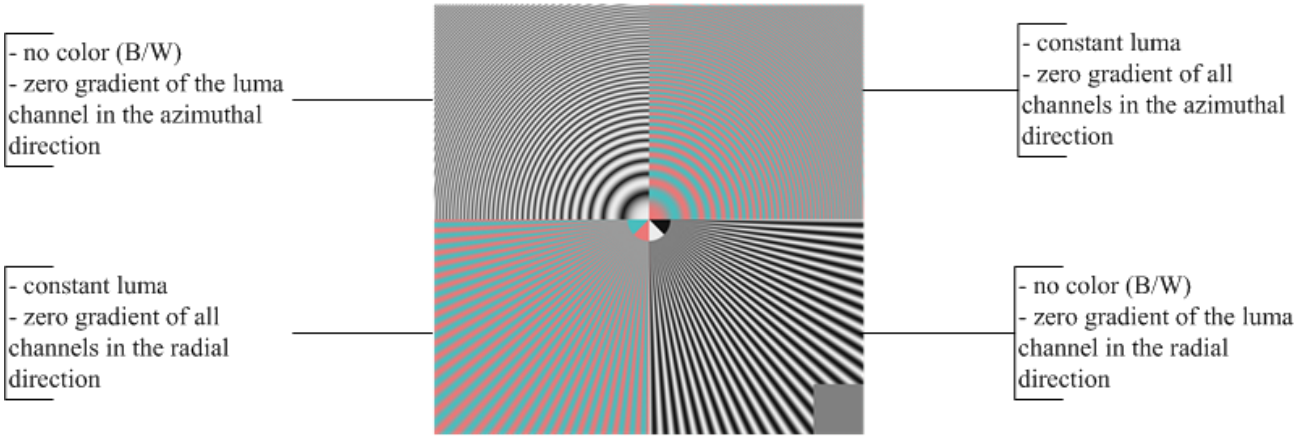


Figure 1 Test chart element - main features (frequency and constant luma)

Additionally, using the given above structure, the proposed chart is composed by using the test chart element shown in Figure 1 in the following conditions:

- Top-left: the test element white balanced and without any noise;
- Top-right: the test element without being whitebalanced and without any noise;
- Bottom-left: the test element without being whitebalanced and with noise;
- Bottom-right: the test element white balanced and with noise;

. The final test image is shown below:

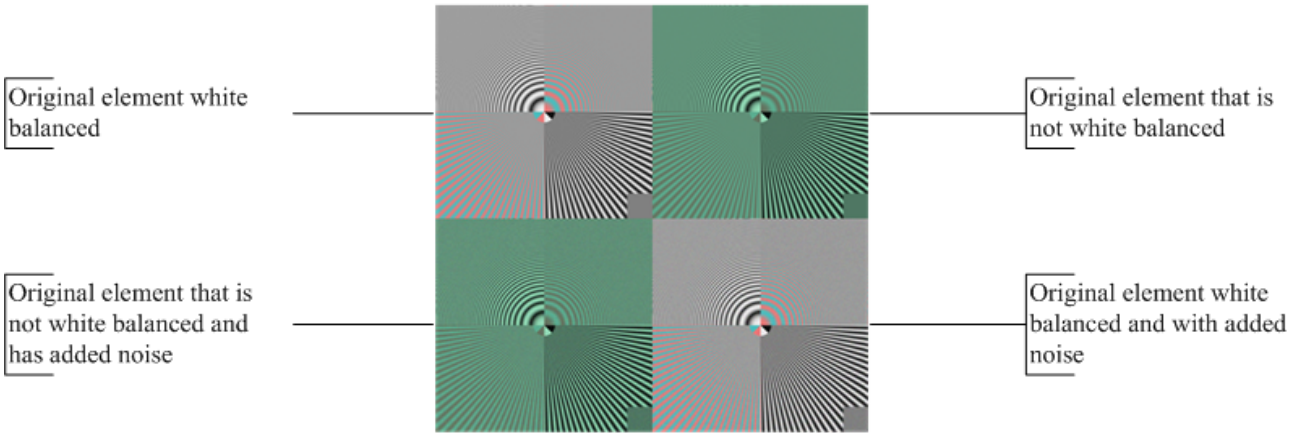


Figure 2 Final chart based on main features plus noise and white balance combinations

The white balance (and color correlation) is selected based on typical actual values and the corresponding matrix is

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 4.6497 & -1.3365 & -0.8135 \\ -0.1138 & 1.3318 & -0.2180 \\ -0.1441 & -0.8084 & 2.3811 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

The intention of introducing wrong white balance is find out if the demosaic algorithm is affected by an assumption of input image having a correct white balance. The purpose of adding noise to the image is to observe how the performance would degrade in the presence of noise.

2.3 Important Data Extraction

Because of the structure of the test elements all the processing is done after conversion to polar coordinates. Both structures used, zone plate and Siemens star, have zero derivative and maximum derivative in the radial or angular dimensions and that has been exploited by polar coordinates conversion. There are several statistics of interest that are used to describe a demosaic algorithm performance:

- saturation in the black and white blocks – specified additional color that is introduced solely by the demosaic algorithm;
- hue variations in the constant luma blocks in the polar dimension (radial or angular) which is flat by design;
- luma variations in the constant luma blocks;
- noise level and cross talk in the grey patch.

3. RESULTS

To illustrate the method, we choose several interpolation algorithms compared with the original as a reference. We used the following name coding: camera phone demosaic algorithm one – CAM_SG, Matlab included demosaic algorithm – MATLAB, original - ORIG, AHD demosaic[2] – AHD, camera phone demosaic algorithm two – CAM_CX, bilinear demosaic – BILN. Camera phone demosaic algorithm one and two are two demosaic algorithms used in commercial phones and shown here only for comparison purpose.

Three types of planes are extracted from the test chart.

Color saturation plane computed as

$$P_{SAT} = 1 - \frac{3}{R+G+B}\min(R,G,B)$$

Hue plane is computed as

$$\mathrm{R} = \mathrm{R}/255$$
$$\mathrm{G} = \mathrm{G}/255$$
$$\mathrm{B} = \mathrm{B}/255$$
$$\mathrm{z} = \frac{2\mathrm{R} - \mathrm{G} - \mathrm{B}}{2\sqrt{(\mathrm{R} - \mathrm{G})^2 + (\mathrm{R} - \mathrm{B})(\mathrm{G} - \mathrm{B})}}$$

$$\theta = \cos^{-1}(\mathrm{z})$$
$$\mathrm{P_{HUE}} = \begin{cases} \theta & if\ B \le G \\ 2\pi - \theta & if\ B > G \end{cases}$$

The luma plane is computed as

$$P_{LUMA} = 0.299 \cdot \mathrm{R} + 0.587 \cdot \mathrm{G} + 0.114 \cdot \mathrm{B}$$

For clearer illustration of the artifacts and for subsequent processing the results are converted from Cartesian to polar coordinates as illustrated below:

$$P(r,\alpha) = C(r \cdot \sin(\alpha) + x_{OFFSET}, r \cdot \cos(\alpha) + y_{OFFSET})$$

where ***C(x,y)*** is the Cartesian coordinate image and ***P(r,α)*** is the polar coordinate image. It has to be noted that all the four components (clean, noisy, wrong WB, and noisy with wrong WB) are treated independently and only comprise ¼ of the range of ***α***. Moreover, to simplify the processing all components are rotated so the center of the figure is in the bottom right corner.

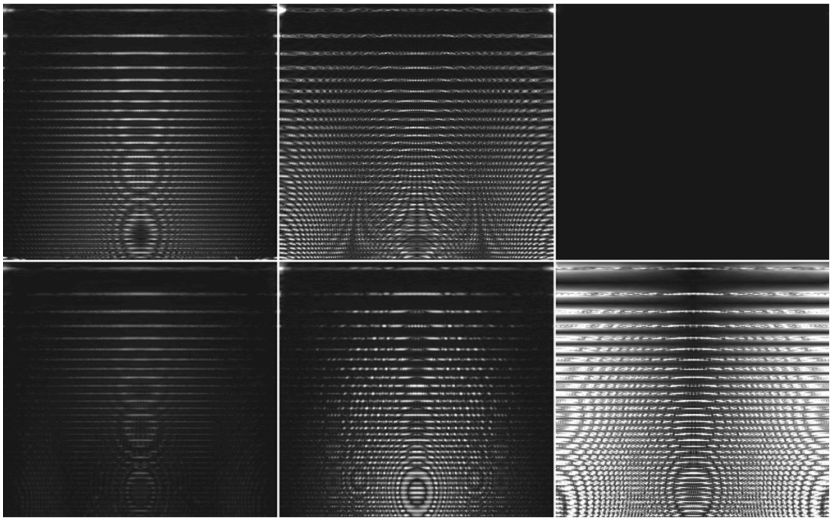


Figure 3 Saturation Zone Plate B&W (polar coordinates) left to right top: CAM_SG, MATLAB, ORIG bottom: AHD, CM_CX, BILN

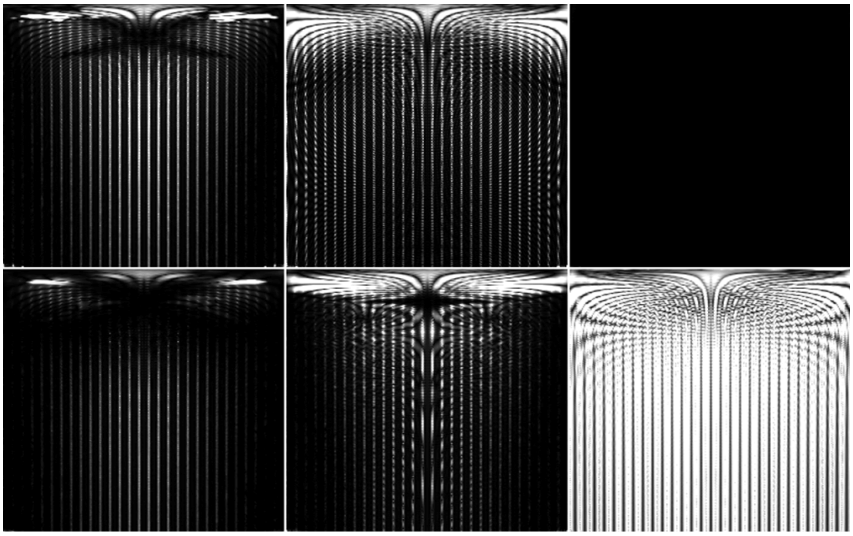


Figure 4 Saturation Siemens star B&W (polar coordinates) left to right top: CAM_SG, MATLAB, ORIG bottom: AHD, CM_CX, BILN

Computation of the hue metric aims to produce a number that correlates well with the perceived deviation from the structure of the hue pattern. That is best illustrated in polar coordinates where original image is comprised of alternating horizontal lines with increasing toward the bottom frequency. The artifacts manifest themselves as a low band perturbation that changes the constancy of the horizontal component of the image. The artifacts can be best estimated in polar coordinates as follow:

- Let H(i,j) be the hue component of the image. Lowpass filter the hue component as

$$H_{LP} = H \otimes h_{AV}$$

where $\boldsymbol{h_{AV}}$ is a square "brick wall" filter with size of $\boldsymbol{n}$ (used value of $\boldsymbol{n}$ is 7).

- Subsample the resulting low pass signal by $\boldsymbol{m}$ ($\boldsymbol{m<n}$, used value for $\boldsymbol{m}$ is 4)

$$H_{SS} = H_{AV} \downarrow_{4}$$

- Filter with horizontal only high pass filter (that is for zone plate constant luma chart, for the Siemens star constant luma chart vertical only high pass filter is used), take the square, and average

$$H_{HP} = H_{SS} \otimes h_{HP} \quad \rightarrow \quad h_{HP} = [-1 \quad -1 \quad 4 \quad -1 \quad -1]$$
$$m_{HUE} = \sum \left(H_{HP}(3 : N_{HP_x} - 2, nN_{HP_y} : N_{HP_y} \right)^2 \Big/ 100$$

Where $\boldsymbol{N_{HP_y}}$ is the width and $\boldsymbol{N_{HP_x}}$ is the height. The first and last 2 horizontal entries are omitted due to filter response tail. The vertical computation starts at $\boldsymbol{n}$ $\boldsymbol{N_{HP_y}}$ ($\boldsymbol{n = 3/5}$ for that particular example) to skip the very low frequencies which do not cause visible artifacts. The normalization by 100 is done to limit the range of the metric. Similar computation is done for the Siemens star (in this case $\boldsymbol{n = 1/3}$).

$$m_{HUE} = \sum \left(H_{HP}(3 : nN_{HP_x}, : \right)^2 \Big/ 100$$

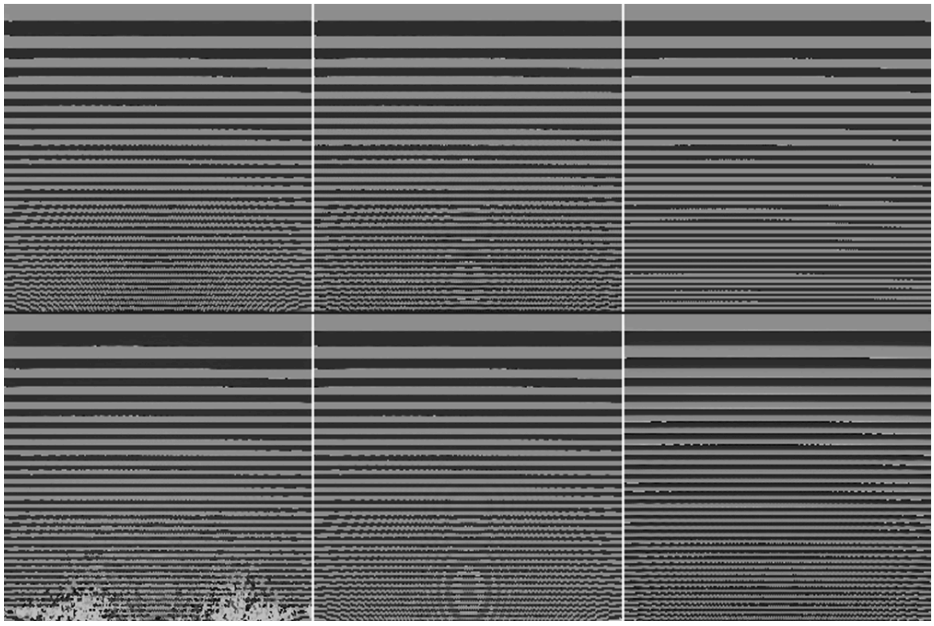


Figure 5 Hue for zone plane for constant luma (polar coordinates) left to right top: CAM_SG, MATLAB, ORIG bottom: AHD, CM_CX, BILN

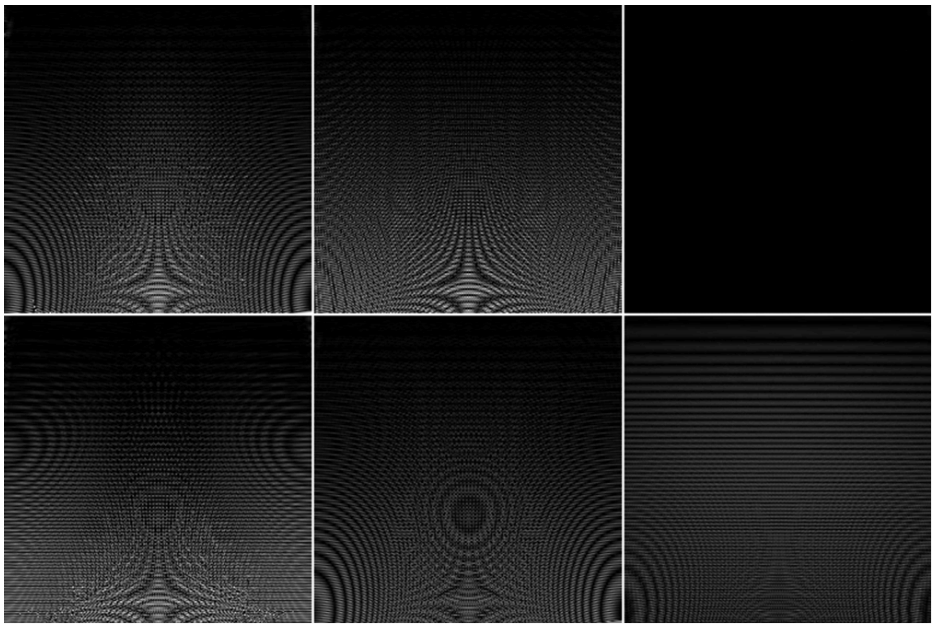


Figure 6 Absolute luma deviation – zone plate (polar coordinates) left to right top: CAM_SG, MATLAB, ORIG bottom: AHD, CM_CX, BILN

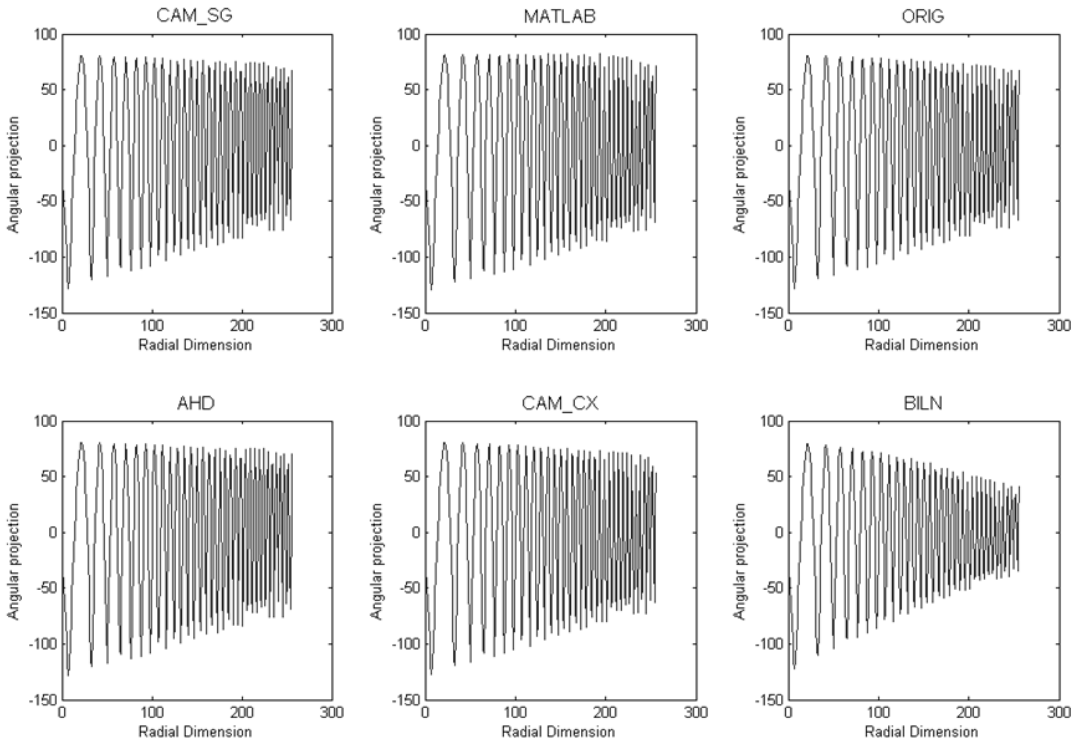


Figure 7 Luma High Frequency – zone plate left to right top: CAM_SG, MATLAB, ORIG bottom: AHD, CM_CX, BILN

The above plots are generated by using the luma plane for the B&W zone plate figure. The luma plane is converted to polar coordinate and the integration (averaging) is performed over the angular dimension as:

$$S(r) = \frac{\sum P_{LUMA}(r,\alpha)}{N_\alpha}$$

Where $P_{LUMA}(r,\alpha)$ is the polar conversion of the luma plane for the B&W zone plate figure and N_α is the number of points in the α dimension. The resolution capabilities of a particular demosaic method are computed as the maximum of the squared absolute value of the Fourier transform of the last ***m*** samples of the angular projection with removed mean.

$$S_F(R) = \Im\left(S(N_R - m : N_R) - \sum S(N_R - m : N_R) \Big/ m+1 \right)$$

$$HF = \max\left|S_F(R)\right|^2$$

Where N_R is the number of points in the radial dimension and ***m = 10***

The visual indication is consistent among figures 5-7, and to further characterize, we used the following metrics:

- ***Saturation MS (zone plate - B&W) -*** Average over the squared saturation for zone plate
- ***Saturation MS(Siemens star- B&W) -*** Average over the squared saturation for Siemens star
- ***Hue Deviation (zone plate – CL) -*** Computed as described above using horizontal high pass filter due to flat horizontal lines by design (in the polar plot)
- ***Hue Deviation (Siemens star – CL)-*** Computed as described above using vertical high pass filter due to flat vertical lines by design (in the polar plot)
- ***Luma variance (zone plate – CL) -*** Zone plate has variance of 160 which also is very close to the means that all demosaic methods produce. The metric is just luma variance
- ***Luma High Freq.(Zone plate –B&W) -*** As described above , using B&W zone plate polar figure, project(average) in alpha dimension. The result is a chirp signal (a signal with increasing frequency) . Maximum chirp frequency is at the end and is roughly a mid frequency. The response at that frequency computed as described in the paper provides an estimation of the resolution capabilities

We computed those metrics for all 4 cases tested with the chart in Figure 2 (original element, original element without white balance, original element with white balance and noise and original element without white balance and with noise) and the results are summarized in the following tables:

	CAM_SG	MATLAB	ORIG	CAM_CX	AHD	BILN
Saturation MS (zone plate - B&W)	0.0322	0.0454	0.0001	0.0752	0.0397	0.0641
Saturation MS (Siemens star- B&W)	0.0603	0.0830	0.0001	0.1114	0.0806	0.3443
Hue Deviation (zone plate – CL)	2.6946	2.5962	0.3716	9.3715	6.1193	1.6061
Hue DEviation (Siemens star – CL)	2.9160	2.8934	0.6677	6.6572	2.5235	4.3550
Luma variance (zone plate – CL)	108.4446	82.6162	0.1189	217.8510	117.0780	4.2168
Luma High Freq. (Zone plate –B&W)	1.2158	1.2506	1.0000	1.4295	1.1828	0.1983

	CAM_SG	MATLAB	ORIG	CAM_CX	AHD	BILN
Saturation MS (zone plate - B&W)	0.0008	0.0021	0.0000	0.0001	0.0012	0.0113
Saturation MS (Siemens star- B&W)	0.0049	0.0093	0.0000	0.0014	0.0082	0.0583
Hue Deviation (zone plate – CL)	1.9606	2.3968	0.3673	11.8527	2.9190	1.3099
Hue DEviation (Siemens star – CL)	6.1822	7.0857	0.6548	7.8055	6.8575	7.4853
Luma variance (zone plate – CL)	78.3075	89.6316	0.0319	114.3675	58.8574	58.1978
Luma High Freq. (Zone plate –B&W)	1.0633	1.1174	1.0000	1.0971	0.9375	0.3137

Figure 8 Left: Original Test Pattern Results – Right: Test Pattern without WB

	CAM_SG	MATLAB	ORIG	CAM_CX	AHD	BILN
Saturation MS (zone plate - B&W)	0.0035	0.0049	0.0040	0.0012	0.0040	0.0136
Saturation MS (Siemens star- B&W)	0.0126	0.0185	0.0107	0.0045	0.0160	0.0633
Hue Deviation (zone plate – CL)	7.8484	3.5917	1.3182	50.1400	4.2412	2.5834
Hue DEviation (Siemens star – CL)	4.9070	5.7567	1.0528	8.3842	5.6851	5.7725
Luma variance (zone plate – CL)	129.9537	127.1790	19.6487	158.4354	97.2811	75.7086
Luma High Freq. (Zone plate –B&W)	1.0643	1.1149	1.0000	1.0884	1.0106	0.3125

	CAM_SG	MATLAB	ORIG	AHD	CAM_CX	BILN
Saturation MS (zone plate - B&W)	0.0618	0.0681	0.0409	0.0855	0.0649	0.0855
Saturation MS (Siemens star- B&W)	0.1148	0.1255	0.0725	0.1276	0.1262	0.3484
Hue Deviation (zone plate – CL)	30.4175	26.5568	3.3941	66.1227	32.3603	18.0463
Hue Deviation (Siemens star – CL)	7.8570	7.0399	2.4707	13.8935	8.0778	6.6472
Luma variance (zone plate – CL)	250.1378	201.6743	70.2498	346.4383	223.3952	62.9969
Luma High Freq. (Zone plate –B&W)	1.2402	1.2632	1.0000	1.4569	1.1939	0.2093

Figure 9 Left: Test Pattern with noise – Right: Test Pattern without WB and with nosie

An evaluation of a demosaic algorithm using as a reference picture a picture that has no noise and is white balanced, is expected to produce results consistent with the values presented in Figure 8 - left. This is the usual case – further, to test if the same algorithm would perform the similarly in a real situation, the evaluated algorithm needs to maintain its performance in all remaining quadrants of the chart proposed in Figure 2 – in other words, the same metrics in presence of noise or absence of white balance need to be correlated in all the quadrants of the chart, this is an indication that the demosaic algorithm is robust and will deliver the expected performance in a real situation.

4. CONCLUSIONS

For a demosaic algorithm, while the problem is well defined, the evaluation of performance is usually done through visual inspection (or MSE) of processed reference pictures that are white balanced and noise free. This approach does not test the robustness of a demosaic algorithm and can provide conclusions that might not correlate well with the real-case situation, especially in the challenging case of cameraphones. To address this situation, we propose a test chart, along with accompanying metrics that provide input about a demosaic algorithm in a real-case condition and we show how it can be used to estimate the performance and robustness of a demosaic algorithm. The features we have are a regular black and white pattern and a luma constant pattern in which any variation in the luma component is an indication of a demosaic artifact. We use both a chirp pattern (zone plate) and a Siemens star and we measure artifacts in hue and saturation on the black and white pattern and in luma in the constant luma pattern. We also found that consistent results in presence of noise and on a chart that is not white balanced are a good indication of demosaic robustness.

REFERENCES

[1] Dubois E.,"Frequency Domain Methods for Demosaicking of Bayer-Sampled Color Images", SPLetters, No. 12, pp. 847-850, (2005)

[2] Ramanath R., Snyder W.E., Bilbro G.L., Sander III, W.A.,"Demosaicking methods for Bayer color arrays" J. Electron. Imaging 11, pp306, (2002)

[3] Trussell H.J., Hartwig R.E."Mathematics for demosaicking", IEEE Transactions on Image Processing vol. 11 issue 4, pp. 485-492, (2002)

[4] Longere P., Xuemei Z., Delahunt P.B., Brainard D.H.,"Perceptual assessment of demosaicing algorithm performance", Proceedings of the IEEE vol. 90 nr. 1, pp123-132, (2002)

[5] Hirakawa,K., Parks T.W.,"Adaptive Homogeneity-Directed Demosaicing Algorithm", ICIP03, 669-67, (2003)

[6] Lu W., Tan Y-P.,"Color filter array demosaicking New method and performance measures", IEEE Transaction on Image Processing Vol.12 Issue 10, pp1194-1210, (2003)

[7] Li X., Gunturk B., Zhang L.,"Image Demosaicing: A Systematic Survey", PROC SPIE 6822, pp6822IJ, (2008)

[8] Trussell H.J., "A MMSE Estimate for Demosaicking", ICIP01, 358-361, (2001)