Evaluating the quality of EDOF in camera phones

Kalin Atanassov, Sergio Goma,

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

ABSTRACT

Extended Depth of Focus technologies are well known in the literature, and in recent years this technology has made its way into camera phones. While the fundamental approach might have significant advantages over conventional technologies, often in practice, it turns out the results can be accompanied by undesired artifacts that are hard to quantify. In order to conduct an objective comparison with the conventional focus technology, new methods need to be devised that are able to quantify not only the quality of focus but also the artifacts introduced by the use of EDOF methods. In this paper we propose a test image and a methodology to quantify focus quality and its dependence on the distance. Our test image is created from a test image element that contains different shapes to measure frequency response.

Keywords: EDOF, Extended Depth of Focus , cell phone camera.

1. INTRODUCTION

In the past years the consumer market for cellular phones, digital cameras and camcorders, has enjoyed considerable growth. Current technology trends point towards the migration of the current PC multimedia experience towards similar support on personal portable handheld devices such as cellular phones. It is true that cellular phones with digital cameras and video recording capabilities have become ubiquitous today, and we expect the trend to continue with the introduction of high definition video and increase in sensor resolution.

End product expectations on behalf of users also grow as the technology progresses, placing increased weight on improving the quality and reliability of image processing algorithms. In addition, the development of new sensors in the pursuit of increased resolution, process or overall quality improvement, requires careful benchmarking and tuning of existing image processing algorithms for the new parameters and changes introduced.

In a desire to reduce the cost of camera modules used in camera phones, one of the solutions explored in the market is extended depth of focus (EDOF). Although EDOF methods are well known in the literature for quite some time[1][2][3][4][5], only in recent years this technology has made its way into camera phones (one of the first commercially available camera phone implementing EDOF to date is PalmPre[4]). While the fundamental approach might have significant advantages over conventional technologies, often in practice when a new technology is mass-produced, it turns out the results can be accompanied by undesired artifacts that are hard to quantify.

As EDOF technology is not commonly used in everyday photography, currently there aren't any common methods available to quantify the quality of EDOF or a methodology to estimate the potential artifacts introduced. In EDOF case, if current SFR techniques[7] are used as for a conventional optics system, the results could be misleading as the assumptions used in a conventional optics system are no longer valid. In order to do a objective comparison with the conventional focus technology, new methods need to be devised that are able to quantify not only the quality of focus or the depth of focus but also the artifacts introduced by the use of EDOF methods and the variation of the quality of focus over the whole range.

In this paper we propose a test chart and a methodology to quantify focus quality and its dependence on the distance and we use this methodology to estimate and quantify the quality of a commercially available EDOF solution (PalmPre[4]).

Image Quality and System Performance VII, edited by Susan P. Farnand, Frans Gaykema, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 7529, 75290K · © 2010 SPIE-IS&T · CCC code: 0277-786X/10/$18 · doi: 10.1117/12.839101

The paper is organized as follows: we start by describing the complex problem of evaluation EDOF and how we developed our target chart, then, we describe the data measurements we perform on the captured image. Next, we present different representations and interpretations of the results, ending with a discussion of the observed artifacts and conclusions.

2. METHODOLOGY

2.1 Problem description

Generally, a commercial implementation of an EDOF technology can be described as a combination of specially designed lens coupled with additional post-processing and offers at post-processing time options to control the focus; implicitly with no mechanical movement with the purpose of achieving a wider range of focus than regular lenses – or with the same range of focus but offering an equivalent wider aperture. Intuitively, a wider aperture would offer significant advantages as the current camera phones are using small pixels that are light-starving. However, as the small pixels inherently have high noise numbers associated, common sense (in this case error propagation) suggests that any post-processing will likely degrade further those numbers, raising legitimate questions about the practical use of any post-processing required to take any advantages of a wider aperture. In this context, as with any post-processing technique used with high noise data, care needs to be taken in order to balance potential advantages. Currently, in camera phones EDOF technology is used to achieve a wider depth of focus, rather than a wider aperture, but the concerns remain the same: what about the price of using EDOF in those conditions?

Even so, questions arise as how good is the focus and how big is the range. The dependence of the system performance on the content is also another factor that needs to be investigated. Most EDOF algorithms utilize some form of adaptive signal processing which raises additional questions about potential artifacts introduced. Based on our experience and the results we present later in this paper, we note that:

- analyzed EDOF implementations are scene dependant and decoupling the technology in order to analyze its effects might not be desirable – likely an 'overall' evaluation as suggested here can be used;
- it is not enough to estimate the depth of focus achieved, but also the focus quality variation over the range covered – as it is shown, the focus quality is not the same for the whole range covered;
- it is not enough to estimate the quality of focus only on b/w edges, but on color as well, since the assumption made on the optical design that all the visible range will behave similar is not valid in the evaluated EDOF technologies;
- possible artifacts associated with EDOF post-processing need to be evaluated;
- noise measurements need to accompany the increased depth of focus gained to be able to estimate EDOF impact;

As a consequence, a good evaluation methodology of EDOF needs to be able to:

1) observe the above issues preferable in a single captured image;
2) compare EDOF with fixed and adjustable focus cameras in order to quantify a potential EDOF gain;
3) offer some feedback that might be used to tune the EDOF modules.

Considering the above, especially in a single captured image, the first thought point toward a slanted chart along a wide focus range, which is the solution we picked – with additional details. As in a 'classic optical' camera design, where SFR curves[7] are used to quantify MTF of an optical system, for EDOF we expanded this logic to a family of SFR curves distributed along the focus range and the primary colors, resulting for each color in a 3D graph showing the variation of SFR along the focus range which we used as the basis of our focus quality metric for EDOF.

Further, as our chart is slanted, we used a feature that is scaling and rotation invariant – in our case a Siemens star.

2.2 Test Chart

As discussed above, a special test chart needs to be designed to be able to compare focal capabilities of normal and EDOF systems and provide important information for a subsequent data processing step. A practical consideration is that the test chart should not take a lot of space and should be easily deployable. The proposed EDOF test chart contains 4x12 grid of perspective compensated Siemens stars. Every row has 12 stars that have the same properties. There are 4 rows with the following properties:

1) black and white stars;

2) flat stars with no red (the red plane is flat at a fixed offset);

3) flat stars with no green (the green plane is flat at a fixed offset);

4) flat stars with no blue (the blue plane is flat at a fixed offset);

For camera phones, we used a perspective adjustment that assumes a camera view angle of 60° and camera position at 30cm from the right end of the chart. The test chart is approximately 3x9 feet with the closest circles center is at 30cm while the furthest one is 220cm from camera. Under the above assumptions the perspective is adjusted such as after capture all Siemens stars will appear as circles, as shown in Figure 1.

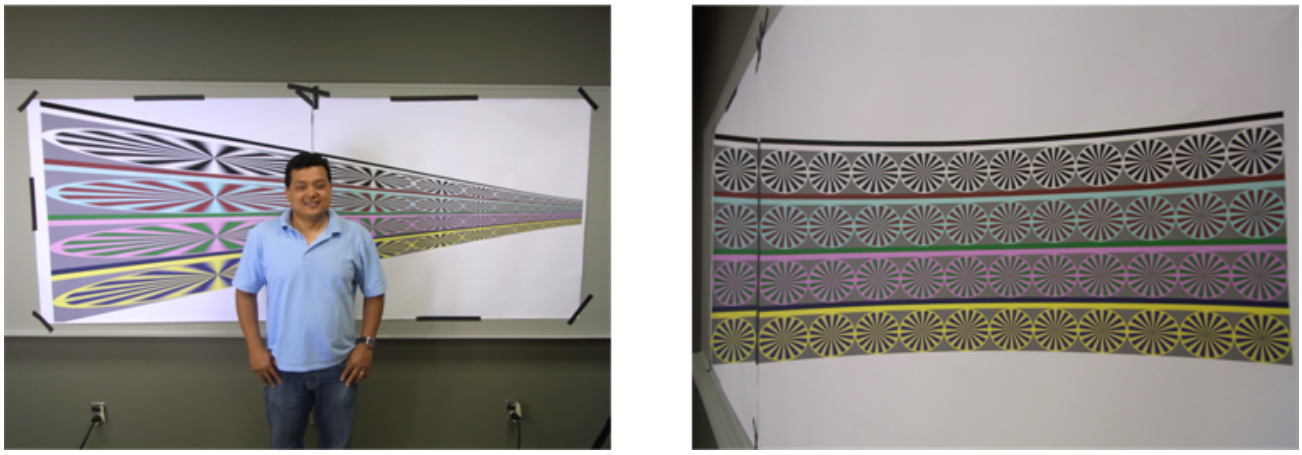


Figure 1: Left: Test Chart with perspective adjustment; Right: Captured image of the test chart

2.3 Proposed Metrics

Similar with the traditional SFR methodology recommended in[7], the data extracted from our test chart will estimate various frequency responses but in addition with the SFR methodology, it will do it for each set of stars/distances and at each distance for each color component. To further quantify the EDOF capabilities and trade-offs, two more metrics are used to estimate the noise component and potential artifacts.

The frequency response metrics analyzed are:

- **Frequency response at DC** – related to the dynamic range of the sensor with some spatial component due to lens shading. The horizontal stripes (dark and bright) above every star are used to get the true (no edge enhancement) DC response as a maximum level difference for every star and color component (R, G, and B). The spread between the two levels is then used to normalize the subsequent higher frequency measurements;
- **Frequency response at frequencies up to Fs/2** – computed for every star and every color channel and normalized with respect to the DC level. The power at a particular frequency is evaluated at a circle with certain

radius around the center of the star[10] (as known, there is a direct correspondence between the radius in pixels and digital frequency). The ideal Siemens star would contain a single (and fixed) digital frequency at circles with arbitrary radius as long the number of sampling points is constant over the different radii. Due to geometrical lens distortion that may or may not be compensated, the power is measured within a narrow band around desired frequency. For that particular example every circle has exactly 16 periods which is determined by the star design. This was selected to provide a reasonable granularity and range for frequency measurements that might be affected by the geometry of the test conditions. It has to be noted that the frequency response measured in this manner is the average of the frequency responses in the azimuthal dimension.

- **Frequency response drop off** - signifies roughly how much frequency was there to start with (let by the lens). Many camera pipelines contain an edge enhancement block that does not create high frequencies but rather create the illusion of them by enhancing present mid frequencies (ie. sharpness). The inflection point of the frequency response can be used as a gauge for how much frequency was passed by the lens and what part is due to post processing

The metrics used to quantify the noise and introduced artifacts are:

- **Adjacent circle MSE** - the MSE between two concentric circles ideally should be the noise if there were no artifacts present. If the center of circle is placed correctly and the radius of circle is big enough, the geometrical distortion will matter less and a measured bigger MSE can be interpreted as an indication that artifacts are present.
- **Remaining noise in a circle** - all concentric circles contains exactly 16 full periods (based on the star design). The radius of the circle can be related to the digital frequency while the energy in that band represents the response to that frequency. The energy out of the band that is significantly above the noise level represents how much energy was artificially introduced by an EDOF processing and is related to the amount of artifacts.

2.4 Data Processing

There are 4 color planes that are considered separately: luma, red, green, and blue. Additionally there are 4 types of stars, black and white, constant red, constant green and constant blue. Between those, there are 16 scenarios in which EDOF performance is evaluated.

The frequency response is computed from a star in the following manner. Firstly, the centers of the stars are obtained semi-automatically by prompting the user to select centers of stars around boundaries and then, further refining of the center positions is done by minimizing the variance along the radial lines from the center. Once centers are obtained, a contrast difference is computed automatically using two patches from the dark and bright stripes above each star (one difference is computed per star). The above computations are performed for the R, G, B, and luma channels independently. The actual frequency response is computed as follows:

First, convert to polar coordinate as:

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

where *C(x,y)* is the Cartesian image and *P(r,α)* is the polar image. It needs to be noted that Cartesian coordinates will not be even multiples of pixels. To mitigate the effects of that, bilinear or higher order interpolation can be used to produce coordinates that are not on pixel boundaries. The resolutions for $\boldsymbol{r}$ and $\boldsymbol{\alpha}$ can be picked arbitrary while the maximum of r is determined by the image size with $\boldsymbol{\alpha}$ between *0* and 2π. For that particular example $\boldsymbol{\alpha}$ range is sampled with 512 points while $\boldsymbol{r}$ has a resolution of one pixel. By design, every star contains 16 periods. When converted to polar coordinates those will appear flat in the radial direction and a sinusoid in $\boldsymbol{\alpha}$ direction - for this particular example, the sinusoid will appear at Fs/32 since $\boldsymbol{\alpha}$ is sampled with 512 points and all sinusoids are comprised of 512/16=32 points.

Next, the Fourier transform is computed in the $\boldsymbol{\alpha}$ dimension:

$$P_F(r,\mathrm{A}) = \underset{\alpha}{\Im}(P(r,\alpha))$$

Ideally, only the power in the desired (and very narrow) frequency band should be measured. In reality the measurement should also include the wider frequency band to accommodate any geometrical or perspective distortion that might be present and would cause warping of the stars.

The power at every radius point is computed as:

$$Q(r) = \sum_{A=d-\Delta}^{d+\Delta} \left| P_F(r, \mathrm{A}) \right|^2$$

where ***d*** is to the frequency bin corresponding to ***Fs/32*** and ***Δ*** is the range (in bins) for the power computation.

Lastly, the power at a given radius is converted to a power at digital frequency as:

$$S(f) = Q\left(\frac{16}{2\pi f} \right)$$

Since radius ***r*** has ***2πr*** points which span over 16 sinusoid periods, there are ***2πr/16*** points per sinusoid which corresponds to 16/***2πr*** digital frequency ***f*** .

The frequency response computed in this fashion should be normalized with respect to DC corresponding to the contrast range between bright and dark patches from above every star. The final transform is normalized by the ideal power at the sinusoid assuming there has been no attenuation (purely dependent on the amplitude):

$$p_{ideal} = \frac{\left(\frac{L_{bright} - L_{dark}}{2} \right)^2}{2} \qquad S_N(f) = \frac{S(f)}{p_{ideal}}$$

3. RESULTS

We applied the proposed methodology with 4 different cameras (one DSLR, one DSC and two camera phones). The cameras under evaluation were: Nikon DSLR, Ricoh DSC, Palm-pre (with EDOF), and Phone camera 1 (this is a camera phone using the same silicon sensor used in Palm-Pre camera phone, except it is packaged in a module with a conventional focus system. The 4 cameras had the following properties:

- Palm-pre EDOF is turned on by default. At picture taking time the algorithm decides which channel is the sharpest and uses it as a reference to bring "sharpness" to others – in the figures named 'PALM';
- Cam1-small aperture, automatic focus with the focus fixed in the midrange of the chart – in the figures named 'Cam1';
- Ricoh –middle range aperture automatic focus with the focus fixed in the midrange of the chart – in the figures named 'RICOH';
- Nikon- big aperture, automatic focus, F-number adjusted for max range with the focus fixed in the midrange of the chart – in the figures named 'NIKON';

Each camera was used to take one picture of the chart which is further analyzed according to the methodology described in the previous chapter.

The following format is used to display the frequency plot of the EDOF/Normal images: X-axis - frequency in terms of cycle/pixel, Y-axis - distance from camera to the middle of the star. In order to meaningfully compare performance for cameras with different resolutions, the frequency unit is cycles per spatial angle. The maximum number that a camera can resolve is specified by a combination of distance from lens, sensor size and resolution. In that regard, the value of ***cut_off*** signifies the maximum frequency that can be resolved with that sensor (the response after that in reality does not exist and is set to 0)

3.1 Frequency response results

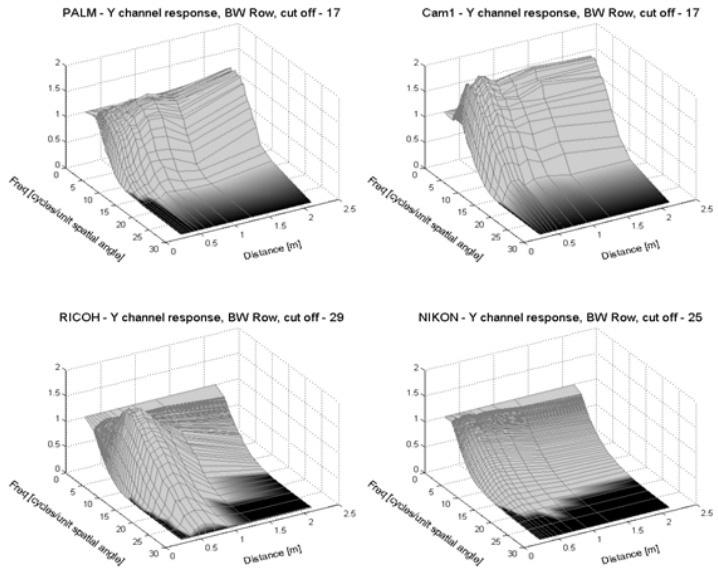


Figure 2: Y channel frequency response on B/W star - all distances

The NIKON plot offers the expected behavior of a camera. RICOH shows added sharpening in post-processing and the fact that the camera had focused at around 0.6m distance. On camera phones, the over sharpening is obvious, regardless if they have EDOF or not. However, it is interesting to note the sharp drop of the EDOF based camera phone (PALM) in contrast to the one not having EDOF (FFA).

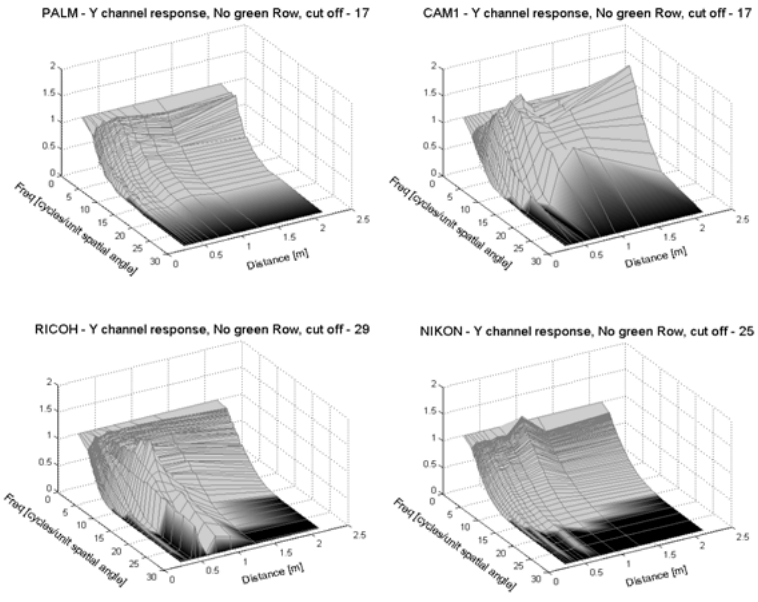


Figure 3: Y channel frequency response on no green star - all distances

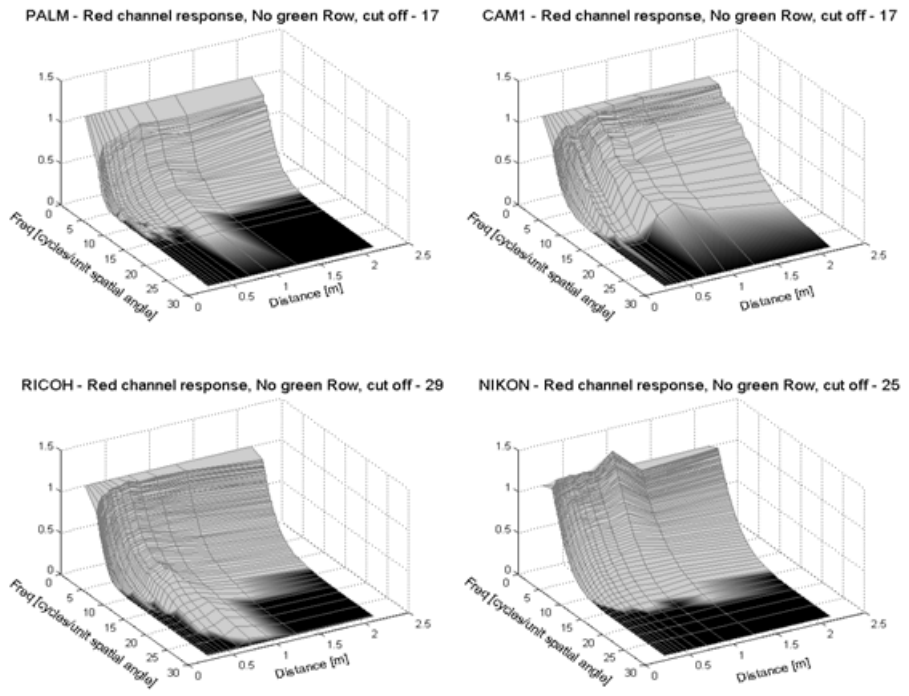


Figure 4: RED channel frequency response on no green star - all distances

Figure 3 and Figure 4 are showing that the sharpening for camera phones and RICOH appears to be done in Y domain, while for NIKON is present in the RED channel. Again, the drop in frequency for EDOF camera phone (PALM) is notable compared with the 'regular' camera phone. The shadow at the bottom of each curve is meant to illustrate the 10% limit.

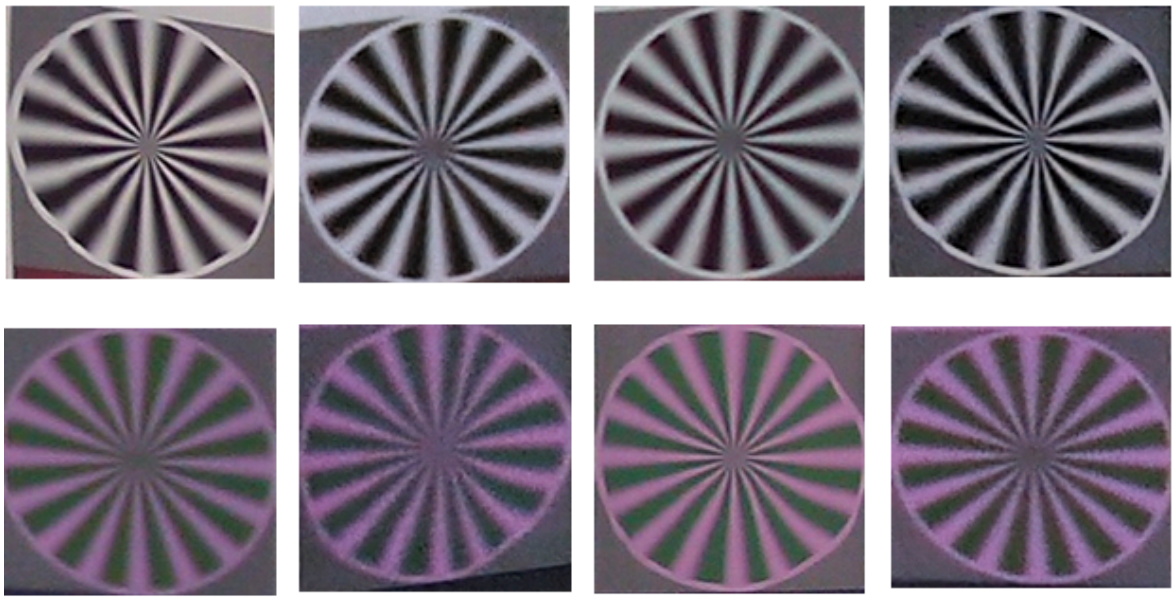


Figure 5: Close and Far on B/W stars (above) and no green stars (below) - first is FFA next PALM (left to right), left pair is close, right pair is far (both rows)

In Figure 5 it can be seen a visual confirmation of the observations made on the frequency graphs: the frequency response for the EDOF module does not appear to have higher frequency content than the one without EDOF, but the noise presence in the EDOF case is noticeable.

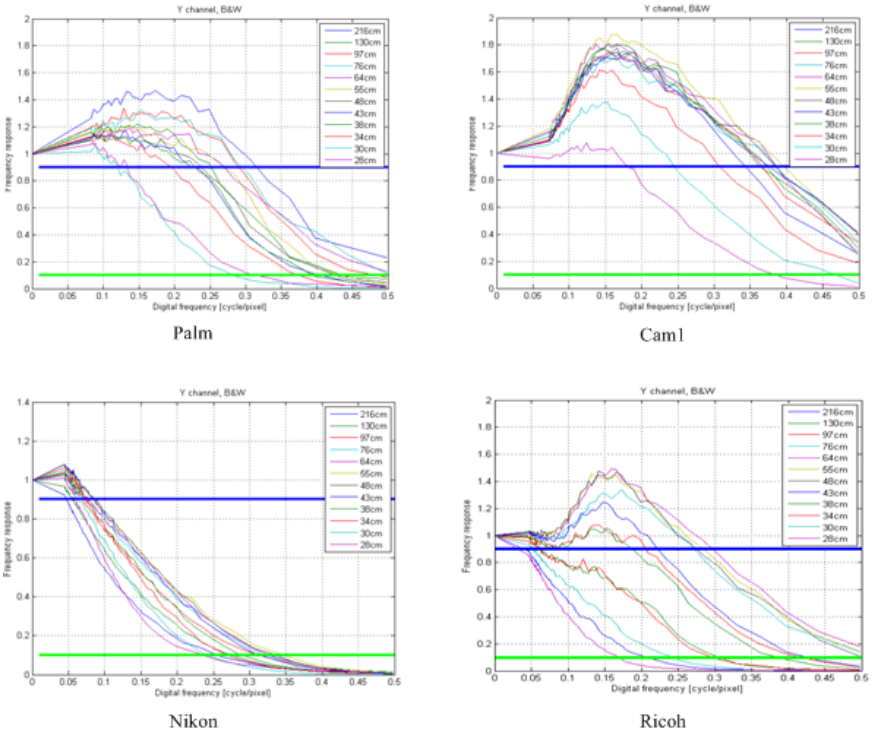


Figure 6: Frequency response graphs for all distances

The frequency response graphs in Figure 6 present the same data as Figure 2, but here the spread of focus curves is more evident, though their distribution in depth is not. However, it can be noted the significant over sharpening used by both camera phones and RICOH, which could account in the increased perceived sharpness of those cameras. Those graphs show that in EDOF case, the frequency response curves aren't grouped together as expected, further they appear to falloff faster than in the case of a regular camera phone module with conventional focus technology that is used in a fixed focus position.

3.2 Noise measurements and results

Frequency response is one of the parameters that can be used to describe the EDOF system performance. Adjacent circle MSE and out-of-band energy are two other parameters that can be used to assess the presence of artifacts[9]. The latter are illustrated in the following plots (due to the fact they are correlated, cameras with similar aperture/noise level are used for comparison). As indicated above, using circles with different radius and same number of points, allows the measurements to happen at the same frequency band. Any energy out of the band is considered indicative of the amount of artifacts present. The noise power is computed in a similar fashion to signal power with the only difference being that summation is performed over the non-signal bands:

$$N(r) = \sum_{A \notin [d-\Delta : d-\Delta]} \left| P_F(r, \mathrm{A}) \right|^2 \qquad S_N(f) = N\left(\frac{16}{2\pi f}\right)$$

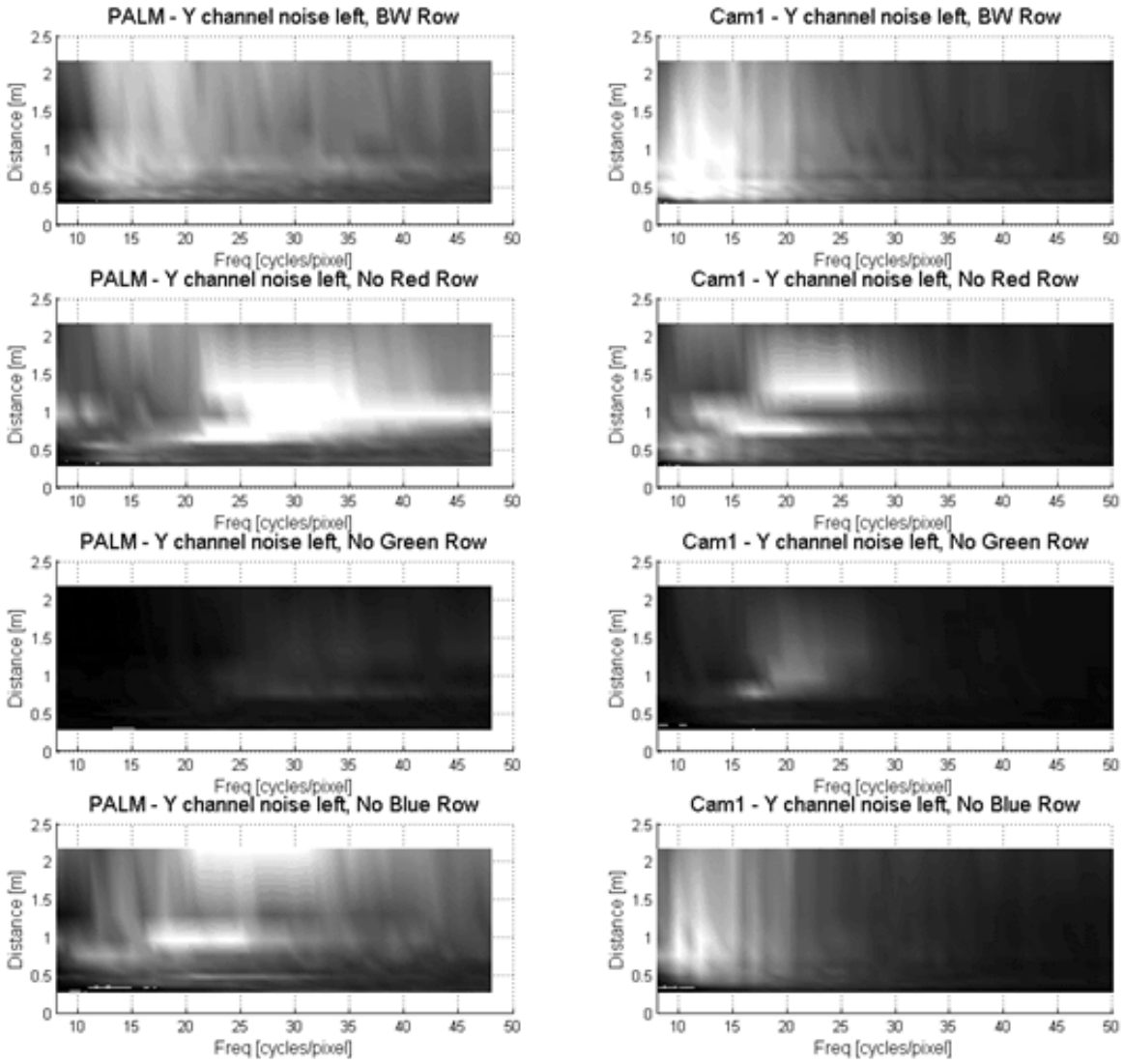


Figure 7: Noise measurements: out-of-band energy, for each channel over the distance – brighter = more noise

Another metric that can be used to indicate discontinuities in the radial dimension, is adjacent circle variance. In order to make measurements more robust, circle with sufficient big radii should be used. The adjacent circle variance is computed as:

$$AV(r) = \sum_{\alpha} \left(P(r+1,\alpha) - P(r,\alpha) \right)^2$$

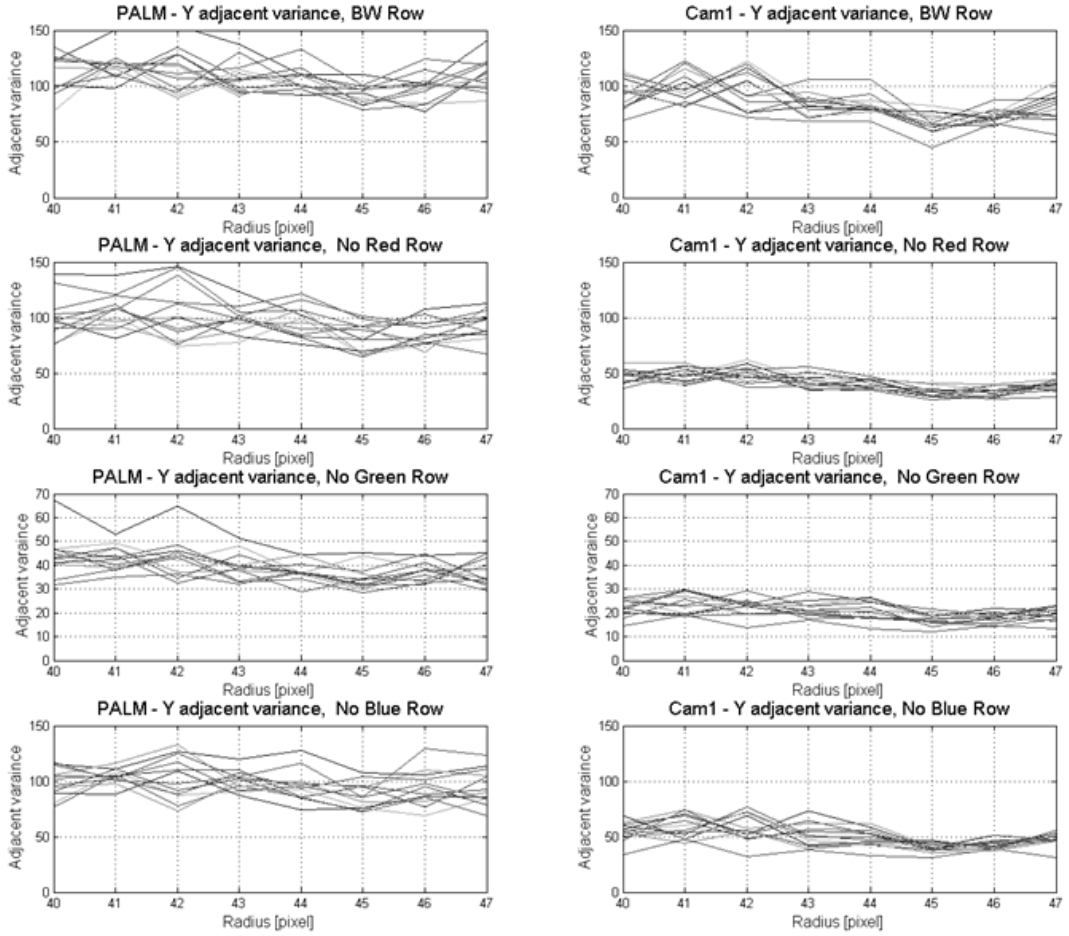


Figure 8: Noise measurements: Adjacent circle variance – higher number = more noise

Figure 7 and 8 confirm the observations made for Figure 5 regarding the increased noise for the EDOF camera phone. Also, interesting noise variations are observed for the EDOF camera phone over the whole distance, whereas for the regular focus module the noise variation over the distance is lower and somehow consistent over the whole distance (close-far).

4. CONCLUSIONS

EDOF is definitely a technology that can be used to extend the capabilities of the classical lens design. In this paper we proposed a methodology that can be used to analyze the potential benefits and drawbacks of EDOF for camera phones. We then applied the proposed metrics in 4 different cases: DSLR, DSC and two camera phones: one with EDOF and one using the same sensor packaged in a conventional focus system. We show how EDOF quality can be estimated and evaluated using our proposed metrics and our results show that the EDOF evaluated here does not show significant depth of focus increase for the range tested compared to the conventional focus module focused around the middle of the scene and the artifacts associated with EDOF in this case are noticeable.

REFERENCES

[1] Ren Ng, "Fourier slice photography", ACM SIGGRAPH July, 2005

[2] Ashok Veeraraghavan, Ramesh Raskar, Amit Agrawal, Ankit Mohan, Jack Tumblin, "Dappled Photography: Mask Enhanced Cameras for Heterodyned Light Fields and Coded Aperture Refocusing", ACM SIGGRAPH 2007

[3] Bradley, A.P., Wildermoth, M., Mills, P. , "Virtual Microscopy with Extended Depth of Field ", Digital Image Computing: Techniques and Applications, 2005.

[4] Frédéric Guichard, Hoang-Phi Nguyen, Régis Tessières, Marine Pyanet, Imène Tarchouna, Frédéric Cao, "Extended depth-of-field using sharpness transport across color channels", SPIE, January 2009

[5] E. R. Dowski and G. E. Johnson, "Wavefront coding: a modern method of achieving high-performance and/or low-cost imaging systems", SPIE, 1999.

[6] D. Williams and P. D. Burns, "Measuring and Managing Digital Image Sharpening", Proc. IS&T Archiving Conference, 2008

[7] ISO 12233, Photography – Electronic still picture cameras – Resolution measurements, ISO, 1998

[8] Guichard F. , "Advances in Camera Phone Picture Quality", Photonics Spectra, November 2007

[9] Uwe Artmann and Dietmar Wueller, "Noise reduction versus spatial resolution", SPIE January 2008

[10] C. Loebich, D. Wueller, B. Klingen, A. Jaeger, "Digital camera resolution measurement using sinusoidal Siemens stars", SPIE Electronic Imaging Conference 2007