Image quality of oversampling cameras

Juha Alakarhu, Samu Koskinen, Eero Tuulos, Nokia Corporation, Camera Technologies,
juha.alakarhu@nokia.com, samu.koskinen@nokia.com, eero.tuulos@nokia.com,

Abstract

This paper considers the oversampling cameras and their image quality. The formulas for analyzing the low light performance are developed and they are compared to subjective testing results. In addition to the developed formulas, the key findings show that there is a significant amount of spatial information available above Nyqvist frequency of the camera system that can be captured with an oversampling camera. We also show that the low light performance is essentially defined by image sensor rather than the pixel size. Nokia 808 PureView product implements an oversampling camera.

1. Introduction

In imaging, oversampling means utilizing higher resolution image sensor than the camera output image resolution. Many smartphones use larger sensor than display resolution, but in oversampling cameras, scaling to smaller resolution is done inside camera and it can be done in earlier stages, potentially with higher quality.

Utilizing extremely high resolution camera and implementing zoom without upscaling is first time implemented in Nokia 808 PureView smartphone. The original driver for implementing oversampling was the desire to implement a high quality zoom in form factor that fits into a smartphone. This approach provides many advantages over optical zoom: full aperture size and macro distance are available throughout the zoom range, zooming is silent and can be smoother than mechanical movement, structure is simpler and more reliable, and it is possible to utilize a large image sensor providing excellent image quality. In addition to the zoom, oversampling can provide significant image quality advantages.

Small pixels or high resolution have been often claimed to compromise the image quality. These comparisons usually mistakenly consider one pixel performance rather than the whole camera performance. This paper considers the image quality of oversampling camera system. Many of the results are applicable to any contemporary camera because they usually records image files that are larger than the display resolution [2].

Sensor technology evolution provides smaller and smaller pixels, and oversampling provides a natural way to utilize the developing pixel technology. It also creates an interesting opportunities for utilizing the pixels in new ways, such as single frame HDR, phase detection for faster autofocus, or spectral sensing.

2. Scaling partially correlated pixels

Let us start with a simple black and white camera where the pixel values are uncorrelated. When N pixels are combined into one, the signal-to-noise improvement *I* can be calculated using well known formula.

$$I(N) = \frac{SNR_C}{SNR_P} = \frac{\dfrac{\sum_i^N s_i}{\sqrt{\sum_i^N n^2}}}{\dfrac{s}{n}} = \sqrt{N} \qquad (1)$$

The formula represents the best possible case. In a practical system, the pixel values are partially correlated due to crosstalk or signal processing that is applied before downscaling. We add an “uncorrelation factor” *U* to the above formula and get

$$I(N,U) = U\ (\sqrt{N} - 1) + 1 \qquad (2)$$

If the pixel values are uncorrelated, *U* is given value 1, and the formula becomes the same as Formula 1. If the values are fully correlated, downscaling has no effect in signal-to-noise ratio, and the function gets value 1. This is a simple way to estimate the downscaled SNR, as long as we know the value for *U*. The uncorrelation factor can be obtained by practical testing and generalizing the results to wider range of camera systems. More details of this will be considered in Section 5.

The formula can be further extended by making the factor *U* a function of the scaling factor. We can expect that pixels are more correlated when they are close to each other (small scaling factors) and uncorrelation factor reaches 1 with large enough scaling. A simple linear estimation is as follows:

$$U(N) = \min(1, (N\frac{(1-A)}{K-1} + A - \frac{1-A}{K-1})) \qquad (3)$$

Where value of the uncorrelation factor with minimal downscaling is *A*, and U reaches 1 with kernel size K. Now the whole formula becomes:

$$I(N,A,K) = \min(1, (N\frac{(1-A)}{K-1} + A - \frac{1-A}{K-1}))(\sqrt{N} - 1) + 1 \qquad (4)$$

Most of digital cameras have larger resolution than the display, and the above scaling formulas can be applied. However, image processing and compression usually makes the pixels more correlated, which makes the scaling less effective, so resulting lower uncorrelation factor. In addition, in this case, the scaler algorithms are defined by the viewer rather than the camera. In an oversampling camera, scaling is applied at least in two phases, inside the camera and the viewer. The low light performance can be calculated by cascading the formula.

3. Line noise

Scaling also affects the line noise. The ratio for pixel to line noise can be derived, assuming both noises are Gaussian:

$$R = R_O \frac{\sqrt{N}}{\sqrt{\sqrt{N}}} = R_O \frac{\sqrt{N}}{\sqrt[4]{N}} = R_O \sqrt[4]{N} \qquad (5)$$

This shows that the random to line noise ratio drops in scaling if the line noise is Gaussian. Line noise can be also individual outlier lines or wider bands. In the former case, scaling effectively reduces the line noise, but in the latter case the line noise can stay almost the same.

The visibility of the line noise depends on both pixel to line and signal to line noise ratios. Downscaling improves the signal to line noise ratio even if the pixel to line noise ratio can degrade. Figure 1 illustrates a case where the line to random noise ratio is 10, and SNR 1 when no scaling is applied.

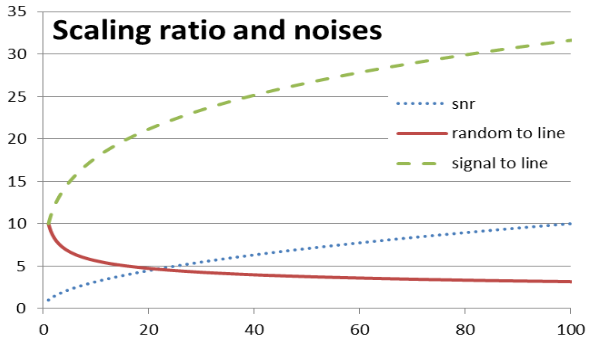

Figure 1. Signal to noise, random to line and signal to line noise ratios, as a function of scaled pixels.

4. Large vs. small pixel

Let us consider the low light performance of two cameras with the same sensor size, one using small pixels and oversampling and the other using larger pixels without oversampling. We assume ideal downscaling, and two main noise sources (photon shot and noise floor). The following formulas show the SNRs with and without the noise floor, for large pixel:

$$\frac{Ms}{\sqrt{n_f^2 + n_{pM}^2}} = \frac{Ms}{\sqrt{n_f^2 + Ms}} \approx \sqrt{sM}$$

and small pixel when scaled to the same resolution:

$$\frac{s\sqrt{M}}{\sqrt{n_f^2 + n_p^2}} = \frac{s\sqrt{M}}{\sqrt{n_f^2 + s}} \approx \sqrt{sM}$$

Here, M is the ratio of the pixel areas, s is the number of signal electros, and n_f the noise floor.

In photon shot noise limited area, we can neglect n_f. With 2e noise floor, photon shot noise starts dominating already at 4e. Practical difference between small and large pixels could be even smaller as small pixels tend to have lower noise floor.

Naturally, also the pixel quantum efficiency curves affect the low light performance. Improving the spectral response, i.e., higher quantum efficiency and lower crosstalk have been also keys in achieving good SNR10 [6].

To summarize previous sections, the low light performance difference between large and small pixel systems depends on

- The scaler uncorrelation factor.
- Line noise amplitude and type compared to signal and random pixel noise.
- Pixel noise floor.
- Pixel quantum efficiency curves.

By designing these properly, the noise and low light performance is defined almost entirely by the sensor size rather than the pixel size.

5. Measurements

To illustrate the use of the scaling formulas, a subjective evaluation to determine the uncorrelation factor U for a simple simulated use case was conducted. The test was done by applying quality ruler methods [7]. Two sets of flat grey images were generated for the subjective evaluation. Both sets have the same image processing, but the reference image set has images with different levels of noise, and the evaluation set has images with the same amount of noise but with different levels of scaling.

The reference images have known SNR on sensor level, after which the Bayer data was processed by applying an advanced demosaicing [4], typical color correction matrix in halogen condition and finally sRGB gamma [5]. In this test, the scaled images are generated by applying bicubic scaling after image processing and jpeg compression. The scaled images are cropped to the same size as the reference images (500x500) for easier subjective comparison.

The analysis was done by 10 experts. The results from the subjective testing show that the uncorrelation factor is roughly constant with typical scaling factors ranging from 1 to 10 as shown in Figure 2.

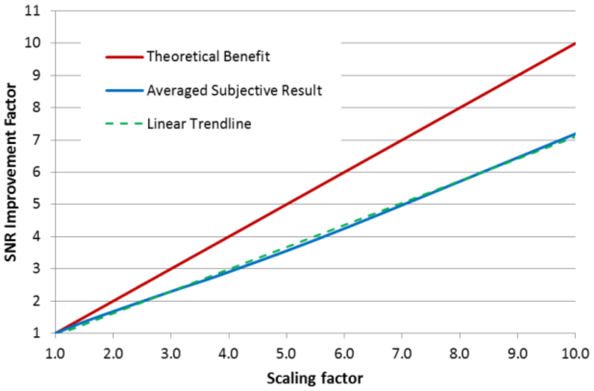

Figure 2. Normalized results of the subjective evaluation of the uncorrelation factor

The slope of the averaged subjective result determines the uncorrelation factor which is illustrated with a linear trendline in the graph. The slope of the trendline is 0.68 which is therefore also the uncorrelation factor. It means that the benefit of the downscaling in terms of noise is roughly 2/3 of the theoretical one when applied to normal images where noise has been correlated, for example, by demosaicing and a color correction matrix.

The result naturally apply to this specific simulated image use case only, the value needs to be calculated for different scalers.

Against our expectations, the results are linear and constant uncorrelation factor *U* can be applied through wide range of scaling ratios. It seems the advanced formula (4) is not needed.

6. Sharpness

Typically the goal in optimizing the camera resolution is to match the pixel size with smallest resolution element that the optical system is capable of producing. In terms of sampling theorem, the pixel pitch defines the spatial sampling frequency, $f_{S,}$ and thereby Nyquist frequency, $f_N = ½*f_S$ of the imaging system. Nyquist frequency defines the frequency above which aliasing can happen, but it doesn't yet tell at what frequency image details can be resolved. Optics is typically matched with sensor resolution so that it is capable of reproducing image details at spatial frequency that corresponds to sensor f_N / 2. The sampling done by the image sensor itself is not ideal, as pixels are not point elements but have a certain area over which the obtained signal is averaged. In effect this non-ideality acts as low-pass filter.

When capturing images of objects with repetitive patterns higher than f_N frequency defined by pixel pitch of the image sensor, aliasing happens and this can be seen as moiré effect. This is evident in luminance channel at full sensor resolution, but with a Bayer pattern sensor the sampling of R,G,B color channels is lower than full sensor resolution and color moiré can result at even lower frequency.

This leads to difficult tradeoff with the sensor-optics matching: Reasonable goal is to have as good or sharp optics as possible, but this is challenging if at the same time there shouldn't be frequencies above f_N to avoid aliasing. In many SLR camera systems optical low-pass filters (OLPF) are used to achieve this effect, but for when prioritizing high resolution it has been left out from some models like Nikon D800E [3].

The problem with OLPF is that it adds cost, thickess and also reduces sharpness, making it not preferred for mobile phones. Aliasing can be reduced by making the optics less sharp, but in this case the nominal resolution of the sensor is not achieved. As a summary, it is very difficult or impossible to capture images that would have true 5 megapixel resolution with a mobile camera that has only 5 megapixels. In this case "true" refers to a situation where for an example line pairs projected to an image plane at exactly f_N spatial frequency could be reproduced accurately in the final image and without moiré.

Typically, in (mobile phone) cameras the image that is produced by the optics is not bandwidth limited so that no frequencies above the f_N frequency wouldn't be present. An example MTF curve of high-quality optics designed for Nokia is shown in Figure 3. In this figure three points are marked, the f_N / 2 frequency for 1.4 µm pixel and for 3.8 µm pixel and in furthermore the f_N frequency for 3.8 µm pixel. These points have been marked based on the example calculation where we downscale 38 Mpix image to 5 Mpix, and the corresponding pixel sizes would be 1.4 µm for 38 Mpix and 3.8 µm for 5 Mpix. It can be seen that by oversampling we can record details with good accuracy (f_N /2 of 1.4 µm pixel, ~180 lp/mm) at frequency that would be clearly above the highest possible frequency without aliasing (f_N of 3.8 µm pixel, ~130 lp/mm) for the "final" pixel size in 5 Mpix image.

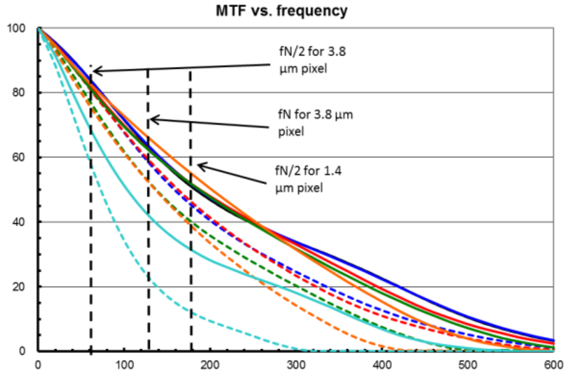

Figure 3. MTF plot of high-quality mobile phone optics

As a conclusion, by using oversampling of the optics it is possible to capture pixel sharp images without aliasing.

7. Product

Nokia 808 PureView product uses oversampling – based camera system. It has 41.5Mpix 1.4µm image sensor, 5-element F2.4 lens, and a mechanical shutter [1].

Oversampling the optics makes it possible to use also higher frequencies reproduced by the optics in the image creation. For an example with Nokia 808 Pureview the pixel size of camera is 1.4µm and sensor resolution for 4:3 images is 38 Mpix, but in default automatic mode the camera outputs 5 Mpix images. With this approach, excellent quality 5 Mpix images without moiré can be obtained by using oversampling. This is illustrated in Figure 4 where Nokia 808 Pureview 5 Mpix image is compared with another 5 Mpix high-end mobile phone camera.

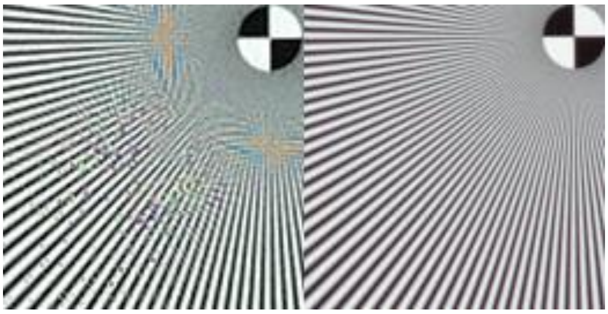

Figure 4. Comparison of 5M high-end cameraphone image without overasampling (left) and Nokia Pureview 808 5Mpix with oversampling (right).

Full resolution comparison with Canon EOS 1Ds Mark III with an L series lens set to 28 mm F2.8 (21.5 Mpix) shows that Nokia 808 Pureview can capture more details. This shows how good sharpness it is possible to capture by a well-designed smartphone camera, further reinforces that there is much information that can be captured by oversampling.

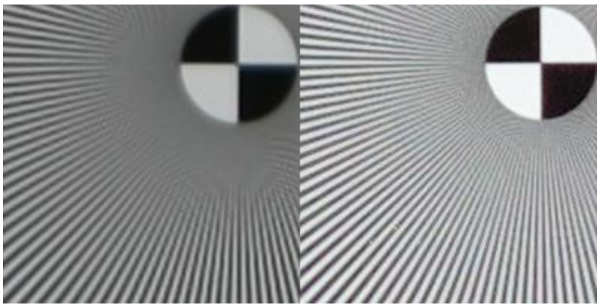

Figure 5. Comparison of 21.5 Mpix high-end SLR image (left) and Nokia Pureview 808 38Mpix image (right).

8. Discussion

The results show that, when implemented properly, oversampling-based camera provides superior image quality over traditional camera with better sharpness, lack of processing and aliasing artefacts, lossless zooming, the same aperture and macro distance throughout the zoom range, and silent zoom. The low light performance is essentially defined by the sensor size, technology and optics aperture, i.e., the oversampling does not have major impact on it.

The key requirements for the best results are small pixels with good spectral response and optics that provides enough sharpness for enabling the zoom. The results provided in this paper show that it can be done in smartphone size.

Oversampling provides exciting opportunities for the future. The large amount of pixels can be used for new features, such as single frame HDR, phase detection for faster autofocus, or spectral sensing, just to mention few.

Oversampling provides also good challenge for the future pixel development. The resolution of future cameras can be extremely high. 808 PureView uses 1.4µm pixel size, 1.1µm is already in mass production, and 0.9µm in the horizon [10]. It is possible to build small pixels without sacrificing image quality. Back-side illuminated sensors can provide high quantum efficiency even with small pixel size. The pixels also need to have extremely well handled crosstalk. For example DTI technology [9] can provide perfect electrical isolation between the pixels. Looking further in to the future, we can go to completely different type of approaches, where the solid state imager can start to behave more like an analog film [8].

9. Conclusions

The image quality of oversampling cameras, including low light performance and sharpness, has been considered in this paper. The results show that in well-designed dystem low light performance is essentially defined by the optics, sensor size and spectral response, rather than the pixel size. The results also show that smartphone optics has significant amount of higher than Nyqvist information that can be utilized by oversampling. Oversampling camera has been implemented in Nokia 808 PureView. Comparison of this product to high-end 5MP smartphone reveals how much more information oversampled 5MP photo has, and how much more pleasing it is visually. We also compare the 808 PureView in full resolution mode to a high-end SLR. This demonstrates how much information very advanced smartphone camera with high resolution sensor can capture.

References

[1] J. Alakarhu et al., *"PureView Imaging Technology"*, Nokia 2012

[2] E. Funatsu, *"Small pixel CIS Technology and Image Quality Evaluation Method"*, Image sensors 2012, London.

[3] Nikon D800 / D800E OLPF information, Nikon web pages. Accessed on 27.3.2013 http://imaging.nikon.com/lineup/dslr/d800/features01.htm#a12

[4] E. Chang et al., "*Color Filter Array Recovery Using a Threshold-based Variable Number of Gradients*" Proceedings of SPIE, January 1999

[5] M. Stokes et al., "*A Standard Default Color Space for the Internet – sRGB, Version 1.10*", November 5, 1996

[6] J. Alakarhu, *"Image sensors and Image Quality in Mobile Phones,"* in Proc. 2007 International Image Sensor Workshop

[7] J. Elaine et al., "*Softcopy quality ruler method: Implementation and validation*", Proceedings of SPIE, January 2009

[8] E. Fossum, "*Gigapixel Digital Film Sensor (DFS) Proposal*" in Proc. 2005 International Image Sensor Workshop

[9] A. Tournier et al., *"Pixel - to - Pixel isolation by Deep Trench technology: Application to CMOS Image Sensor"* in Proc. International Image Sensor Workshop (IISW), Hokkaido, Japan, 2011

[10] S.G.Wuu et al., *"A Leading-Edge 0.9µm Pixel CMOS Image Sensor Technology with Backside Illumination: Future Challenges for Pixel Scaling (Invited)",in Proc.* IEDM, San Francisco 2010