Resolution and Light Sensitivity Tradeoff with Pixel Size

Joyce Farrell, Feng Xiao*, Sam Kavusi
Stanford University, Stanford, CA
*Agilent Technologies, Santa Clara, CA

ABSTRACT

When the size of a CMOS imaging sensor array is fixed, the only way to increase sampling density and spatial resolution is to reduce pixel size. But reducing pixel size reduces the light sensitivity. Hence, under these constraints, there is a tradeoff between spatial resolution and light sensitivity. Because this tradeoff involves the interaction of many different system components, we used a full system simulation to characterize performance. This paper describes system simulations that predict the output of imaging sensors with the same dye size but different pixel sizes and presents metrics that quantify the spatial resolution and light sensitivity for these different imaging sensors.

1. INTRODUCTION

Digital cameras are now a standard feature in cellular phones, driving the market for CMOS imagers that can fit within a small form factor. Given that the size of a CMOS imaging sensor array is fixed, the only way to increase sampling density and spatial resolution is to reduce pixel size. But reducing pixel size reduces the light sensitivity. Hence, under these constraints, there is a tradeoff between spatial resolution and light sensitivity.

For any fixed process technology and pixel architecture, decreasing pixel size will decrease pixel performance. Without compensating technologies, smaller pixels have lower dynamic range, lower fill factor, worse low light sensitivity, higher dark signal, and higher non-uniformity. Mobile imaging applications have driven innovations in image sensor technologies that significantly compensate for the expected degradation in performance with decreasing pixel sizes. Process modifications including improved micro-lenses, pinned photodiode, dual –gate oxide, floating diffusion, circuit techniques such as device sharing, and active reset, compensate for the many factors that would otherwise reduce performance.

System simulation is a valuable tool for evaluating the complex set of factors involved in optimizing the image quality of CMOS imaging sensors. In this paper, we use the ***Image Systems Evaluation Toolkit (ISET)*** to quantify the relationship between spatial resolution and light sensitivity for imaging sensors with the same dye size but different pixel size. *ISET* is an integrated suite of Matlab software routines designed to simulate the entire image processing pipeline of a digital camera [1]. The analysis described in this paper extends the work previously published by Chen et al [2] and Catrysse and Wandell [3] who also simulated the optical blurring, sensor transduction and image processing pipelines for CMOS imaging sensors with the same dye size but varying pixel sizes. We extend their analysis to color imaging sensors and consider the role other sensor parameters, such as quantum efficiency, voltage swing, conversion gain and different noise sources, as well as pixel size and fill factor, have upon the spatial resolution and light sensitivity of an imaging sensor.

2. SIMULATIONS

The Image Systems Evaluation Toolkit (*ISET*) is organized around four key software modules: **Scene**, **Optics, Sensor**, and **Processor**. Each module includes a variety of specialized tools and functions that help the user set parameters, experiment with alternative designs and component properties, and calculate relevant metrics.

The **Scene** module represents the input to the simulated digital camera as a multidimensional array describing the spectral radiance (photons/sec/nm/sr/m^2) at each pixel in the sampled scene. The **Optics** module converts the scene

radiance data into an irradiance image at the sensor. The conversion of scene radiance image to irradiance is determined by the properties of the simulated optics. The **Sensor** module manages the transformation from irradiance to sensor signal. This transformation includes an extensive model of both optical and electrical properties of the sensor and pixel. The **Processor** module transforms the electron count into a digital image that is rendered for a simulated color display. This module includes algorithms for demosaicing, color conversion to a calibrated color space, and color balancing.

In our simulations, we use synthetic scene data designed for specific image quality metrics. For example, we use an image of a slanted edge to calculate the spatial modulation transfer function (using the ISO 12233 method) and we use a Lambertian surface uniformly illuminated with a D65 light to calculate sensor dynamic range, SNR and photometric exposure. The optical image is calculated using a diffraction-limited model of a lens with an f-number of 2.8. We simulate the effects of an optical diffuser that filters out signals above the Nyquist frequency limit of the imaging sensor. This is accomplished using a Gaussian filter with full-width half maximum equal to the pixel width. We simulate a color imaging sensor with a dye size of 512 x 512 microns and with system spectral quantum efficiency equal to those measured for a Nikon D100 digital camera. The peak quantum efficiency is 0.65 and the pixel fill factor is 50%. All other sensor parameters, such as pixel size, read noise, dark voltage, photoresponse non-uniformity levels (prnu), voltage swing and conversion gain were derived from Rhodes [4] and are listed in Table 1 (below). The image processing pipeline is based on bilinear demosaicing and "Gray World" color balancing that scales the sensor values so that their mean is the same as that produced by an 18% neutral gray target. The optical, sensor and image processing parameters are important only in the sense that they define viable design solutions. The purpose of this paper is to describe a methodology for evaluating potential sensor designs.

Table 1: Sensor parameters derived from Rhodes et al [4]

Pixel size (um)	Peak voltage (V)	Conv. gain (V/e-)	Well-cap (e-)	Dark voltage e-/sec	Read noise (e-)	PRNU (st. dev. % gain change)	DSNU
2	2	125	16,000	7	6	0.03	0
2.6	1.8	64	28,150	11	7	0.02	0
2.9	1.6	64	25,000	10	7	0.01	0
3.2	1.4	42	33,333	12	8	0.01	0
5.2	1.2	32	37,500	22	10	0.01	0

3. SENSOR PERFORMANCE

The metrics for characterizing spatial resolution and light sensitivity are performance summaries derived from the complete system modulation transfer function (MTF) and the system signal-to-noise ratio (SNR), respectively. These measures depend, in turn, on a variety of sensor properties, including read noise, dark noise, conversion gain, and voltage swing. All of these parameters vary with pixel size and influence the MTF and SNR.

3.1. Light sensitivity

Well capacity declines with pixel size. This has consequences both for high and low levels of scene luminance and for short and long exposure durations. Small pixels have fewer photons incident at their aperture than large pixels and saturate at lower photometric exposure values. These properties have implications for sensor dynamic range and signal-to-noise ratios.

The sensor dynamic range (DR) is calculated using the formula,

$$DR = 20\log(\frac{E_{\max}}{E_{noise}}), \quad E_{noise} = \sqrt{\sigma_D^2 + \sigma_R^2 + \sigma_{DSNU}^2}$$

In this case, $E_{\max}$ is the well-capacity in electrons and E_{noise} is the noise contributed by the variances of the dark noise, read noise, and dark-signal non-uniformity. The well-capacity declines with pixel size and thus reduces the dynamic range. However, as Rhodes et al [4] pointed out, advances in technology have make it possible to reduce the noise levels associated with small pixels which, in turn, increases the dynamic range. For example, higher conversion gain helps reduce E_{noise}. The combined effect is to maintain the dynamic range across pixel size at a level of roughly 45-48 dB for a 16 millisecond exposure duration for the simulated color imaging sensors.

The sensor signal-to-noise ratio (SNR) is calculated using the formula

$$SNR(E_s) = 20\log(\frac{E_s}{\sqrt{\sigma_s^2 + \sigma_R^2 + \sigma_{DSNU}^2 + \sigma_{PRNU}^2}})$$

All of the quantities are measured in units of electrons. The term E_s is the signal level, and $\sigma_s^2 (= E_s)$ is the shot noise variance; $\sigma_R^2, \sigma_{DSNU}^2, \sigma_{PRNU}^2$ are the variances of the read noise, dark signal non-uniformity and photo-response non-uniformity, respectively.

Figure 1 plots the sensor SNR for color imaging sensors with pixels described in Table 1. The peak SNR declines with pixel size mainly because well-capacity (maximum number of electrons prior to saturation) decreases with pixel size from about 37,500 (5.2um) to 16,000 (2um) electrons. The reduction in well-capacity alone, without any other noise contributions, would produce a decline in the peak SNR of more than 3dB. The SNR drop is larger, however, because the effect of noise is more significant when added into lower signal levels. Thus, even though the technology advances described in Rhodes [4] result in lower levels of read noise and dark voltage, their impact remains high because these noises are superimposed on relatively low signal levels. The peak SNR of the smallest pixel is approximately 8dB lower than that of the largest pixel.

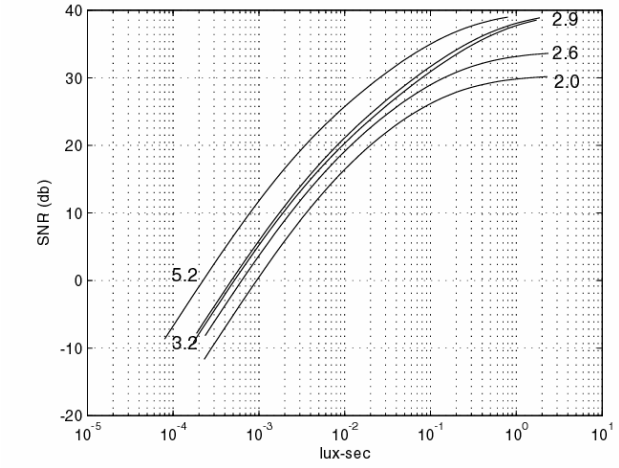

Figure 1: SNR plotted for imaging sensors with different pixel size.

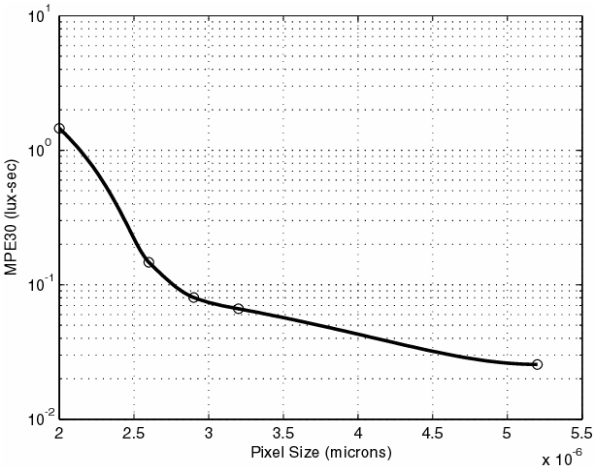

Figure 2: The minimum photometric exposure to produce an SNR of 30 (MPE30) plotted as a function of pixel size.

Even ideal sensors with no electronic sources of noise will generate visible noise at low light levels. This is because photon noise will be a significant source of image contrast. The results of visual psychophysical experiments reported by Xiao et al [5] show that sensor SNR must be 30dB or greater to render photon noise invisible. We use these results to define a psychophysical threshold for the detection of luminance noise and apply this threshold to quantify the light sensitivity of imaging sensors with different pixel size. Our luminance noise visibility metric is based the minimum photometric exposure that will capture an image of a uniform field with less than 3% noise variation or, conversely, an SNR > 30db. We refer to this metric as MPE30.

Figure 2 plots the MPE30 for the simulated color imaging sensors as a function of the pixel size. The decline in the MPE30 value means that at any scene luminance level the exposure duration required to reach a sensor SNR of 30db decreases with increasing pixel size. Or, equivalently, for any exposure duration the scene luminance required to reach a sensor SNR of 30 db decreases with increasing pixel size

3.2. Spatial resolution

The Modulation Transfer Function (MTF) describes the ability of an imaging system to capture image contrast over a range of spatial frequencies. We measured the system MTF using the ISO standard ISO 12233 [6]. This method measures the system response to a slanted edge. The lines passing through the edge measure the step response in a variety of different phase relationships to the pixel sampling grid. These responses are combined to estimate the system MTF.

Figure 3 plots the MTF of the green sensor channel for the simulated color imaging sensors. Sensors with smaller pixels can capture and preserve higher spatial frequency details, as expected. These curves are summarized by a single value, the spatial frequency at which the amplitude falls to 50% of the highest amplitude (MTF50), in Figure 4.

System properties other than pixel size influence the sensor MTF and MTF50 as well. For example, the choice of optics and demosaicing algorithm will determine the sensor MTF. Another important factor to consider is camera motion. Camera motion introduces blur and reduces the system MTF. Camera motion is greatest when the button that initiates a photo is depressed. This motion is also greater for smaller cameras and increases with exposure duration [7]. Camera motion is, then, a significant problem for small high-resolution cameras and camera phones and it limits their effective imaging resolution [7].

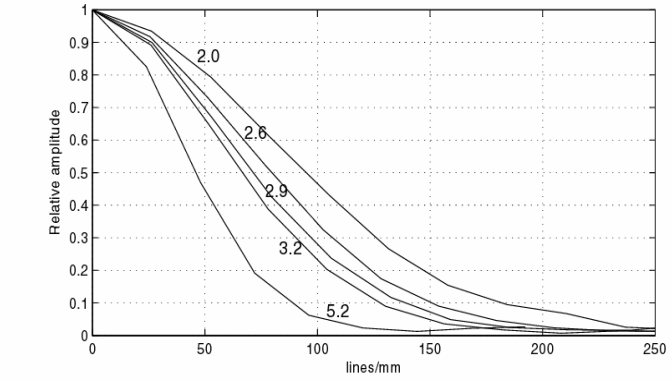

Figure 3: Luminance MTFs for imaging sensors with different pixel size.

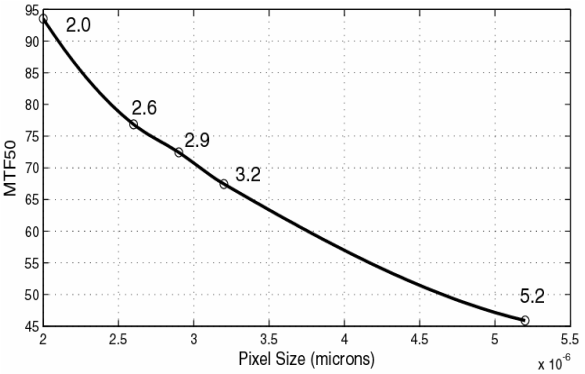

Figure 4: Spatial frequency for which the amplitude of the luminance MTF (see Figure 3) falls to 50%, plotted as a function of pixel size.

4. IMAGE QUALITY TRADEOFFS

The tradeoff between spatial resolution and light sensitivity is illustrated in Figure 5 which plots the MTF50 against the reciprocal of MPE30 for each of the simulated color imaging sensors. Higher image quality is associated with higher values of MTF50 and higher values of 1/MPE30. Smaller pixels are associated with high values of MTF50 but low values of 1/MPE30. Larger pixels are associated with high values of 1/MPE30 but low values of MTF50. There is no obvious optimal point along the tradeoff function depicted in Figure 5.

Chen et al [2] used the SCIELAB metric [8] to define the "optimal" pixel size for a given sensor dye size. In their analysis, an imaging sensor is "optimal" if it produces an image that is similar to the image that would be captured by a hypothetical ideal imaging sensor with high spatial resolution (small pixel size), high dynamic range and no sensor noise. Chen et al [2] use the SCIELAB metric to quantify the similarity or, conversely, the visible difference between an ideal monochrome imaging sensor and each of the lower resolution monochrome imaging sensors. We extend their analysis to color imaging sensors by calculating the SCIELAB difference between images captured by a simulated "ideal" color imaging sensor and all other simulated color imaging sensors. The simulated "ideal" color imaging sensor has 1 micron pixels with read noise, dark voltage, prnu, dsnu values set to 0, peak QE and pixel fill factor equal to 1, and the maximum settings for conversion gain (125 volts/electron) and voltage swing (2).

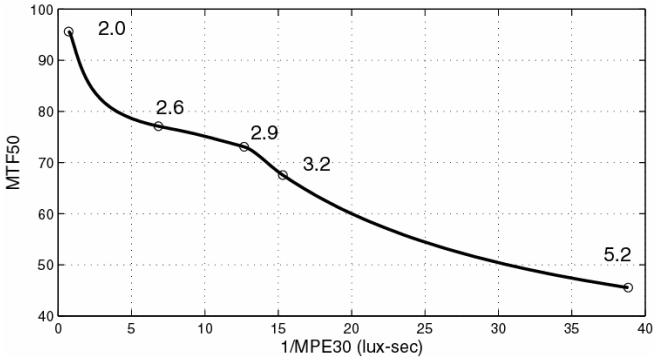

Figure 5: MTF50 plotted against 1/MPE30 for each of the simulated color imaging sensors.

Once again, we create an image of a noise-free slanted edge as our original scene data. We chose this image to highlight uncorrelated image noise visible in the uniform areas and blur visible along the edge. We set the scene luminance to 60 cd/m^2 and the exposure duration to 1/60 second. All other sensor properties remain the same (see the section on Simulations above).

The simulated sensors generate images with sizes ranging from 512x512 (for the "ideal" sensor) to 98x98 (for the sensor with 5.2 micron pixels). The SCIELAB metric compares images with the same size. To create same size display images, we down-sample the simulated sensor images to produce 98x98 images. We justify the down-sampling by noting that most digital camera images are too large to view on a typical 100 dpi computer display. Consequently, image rendering software must down-sample the high resolution digital camera images in order to view the entire image on a display. Moreover, a 98x98 image will subtend a visual angle of 4.8 degrees when it is displayed on a 100 dpi display and viewed at a distance of 12 inches. From the point of view of observers, therefore, a 98x98 image is larger than the 2 degree fovea over which they have their highest visual acuity.

Figure 6 illustrates the 98x98 images produced by imaging sensors with different pixel sizes. When displayed on a 100 dpi display the images acquired by small pixel sensors (e.g. 2 micron) have visible noise in the uniform bright areas; the images acquired by large pixel sensors (e.g. 5.2 micron) have blurred edges.

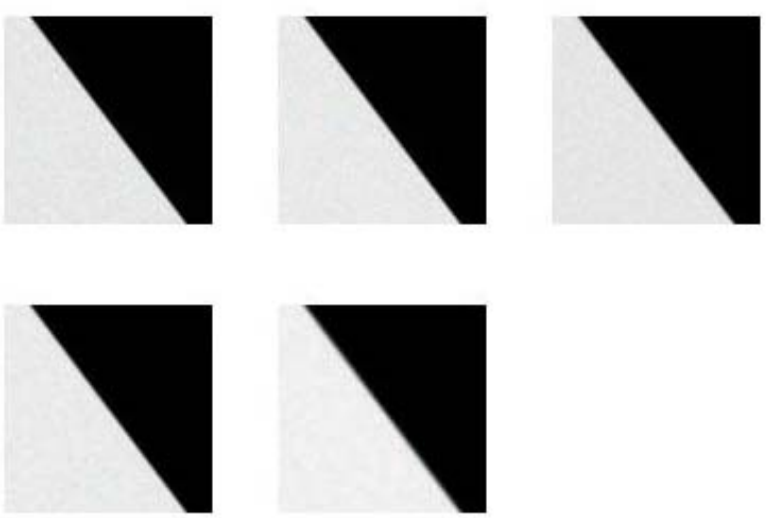

Figure 6: Images captured by sensors with same optics and dye size but with different pixel sizes.

To calculate the SCIELAB difference between the images generated by the "ideal" color imaging sensor (with 1 micron pixels and no noise) and each of the other sensor images, we must first calculate the CIE Standard Observer XYZ values that would be measured when the images are displayed on a computer monitor. We assume that the images are displayed on a standard 100 dpi sRGB display [9] and that observers are sitting 12 inches from the display. The XYZ images, along with the viewing parameters (i.e. viewing distance and display resolution), are sent to the ISET function that calculates the S-CIELAB color image difference.

Figure 7 shows SCIELAB color differences between displayed images generated by the "ideal" color imaging sensor and each of the other sensor images**.** The difference images highlight the type of distortion that will be visible in the images of the slanted edge captured by each of the simulated color imaging sensors. For example, the difference images show that uncorrelated noise will be visible in images captured by color imaging sensors with small pixels (e.g. 2.0 microns) and that the blurred edge will be visible in images captured by sensors with large pixels (e.g. 5.2 microns).

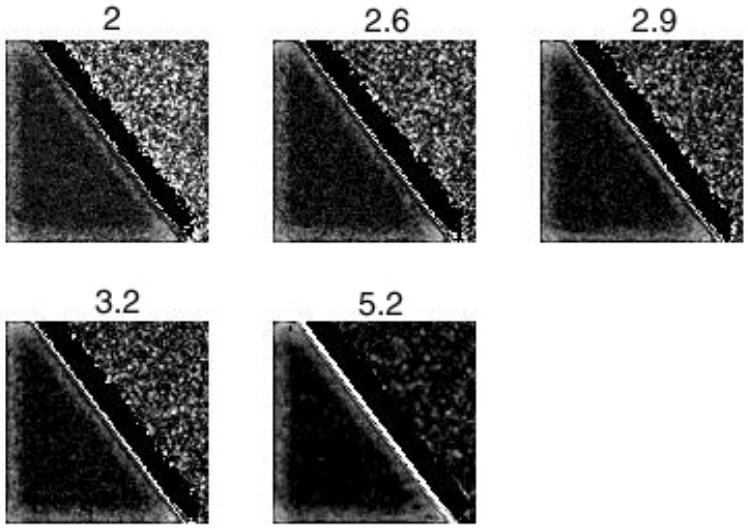

Figure 7: SCIELAB color differences between images captured by a simulated "ideal" high-resolution noise-free imaging sensor and a lower-resolution sensor for a 1/60 second exposure duration. Each of the difference images is labeled according to the size of the sensor pixel.

Figure 8 shows a histogram of the SCIELAB differences shown in Figure 7. Figure 9 plots the 95th percentile and median values for each of these images as a function pixel size. These figures show that the noise generated by imaging sensors with small pixels (e.g. 2 microns) is slightly more visible than the blur generated by sensors with larger pixels (e.g. 5.2 microns).

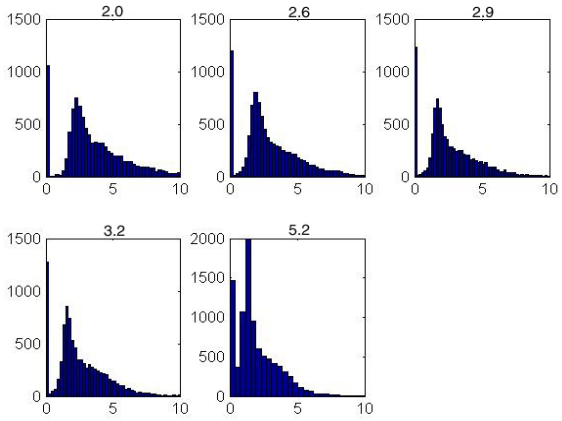

Figure 8: Histogram of SCIELAB differences between images captured by the hypothetical "ideal" sensor and each of the simulated color imaging sensors for exposure durations of 1/60 second.

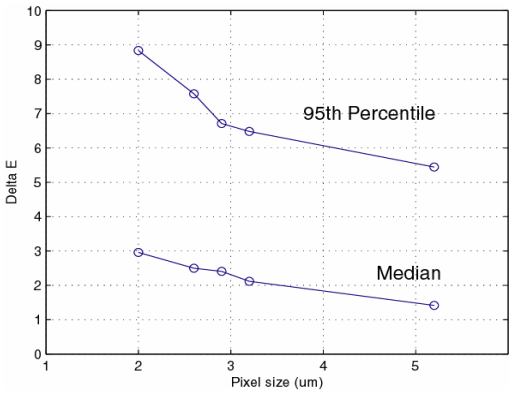

Figure 9 : The 95th percentile and the median of the SCIELAB differences shown in Figure 8.

The image noise and blur are less visible at longer exposure durations. Figure 10 shows a histogram of the SCIELAB differences between the simulated "ideal" high-resolution noise-free imaging sensor and each of the lower-resolution color imaging sensors for exposure durations of 1/15 seconds. The fact that the SCIELAB values are low (< 5 ΔE) suggest that for these images the image noise and blur are barely visible. The histogram plots are similar indicating that the image noise and blur are equally visible.

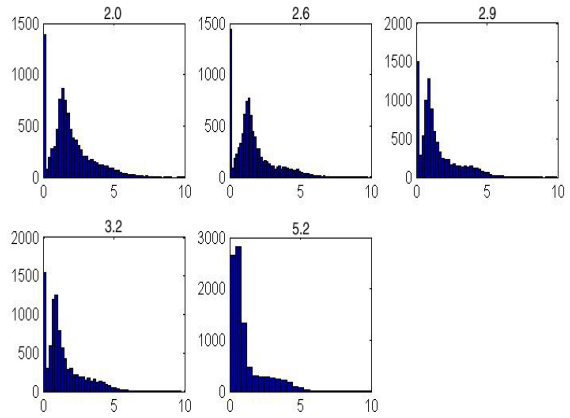

Figure 10 : Histogram of SCIELAB differences between images captured by the hypothetical "ideal" sensor and each of the simulated color imaging sensors for exposure durations of 1/15 msec.

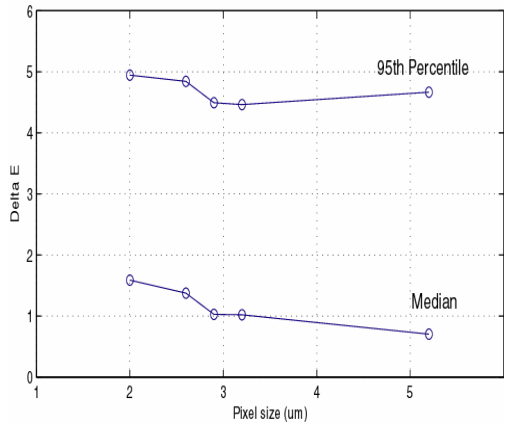

Figure 11 : The 95th percentile and the median of the SCIELAB differences shown in Figure 10.

5. SUMMARY

This paper describes simulations that predict the output of color imaging sensors with the same dye size but different pixel sizes and electrical properties. Image quality is quantified with respect to light sensitivity and spatial resolution as a function of pixel size. Uncorrelated image noise is visible in images generated by sensors with small pixels (e.g. 2 micron); image blur is visible in images generated by sensors with larger pixels (e.g. 5.2 microns). We introduce a new

metric (MPE30) to quantify the visibility of uncorrelated image noise, and we use the MTF50 metric to quantify the amount of image blur. The SCIELAB metric quantifies how the visibility of these two types of image distortions (uncorrelated noise and image blur) tradeoff. At short exposure durations (e.g. 1/60 second), image noise is more visible than image blur. At longer exposure durations (e.g. 1/15 second), these two different types of image distortions are equally visible. Although these two types of image distortions can be equated for visibility that does not imply that the two types of distortions are equally preferred. In the future, we will perform human psychophysical measurements based on images generated by the simulated color imaging sensors, and we will have observers indicate their preference. These experiments should allow us to superimpose iso-preference curves on the image quality metrics and thus choose a preferred system design.

The ISET software scripts used in our analyses are available at www.imageval.com

ACKNOWLEDGMENTS

The authors thank the team at ImagEval, particularly B. A. Wandell, for useful discussions.

REFERENCES

1. J. E. Farrell, F. Xiao, P. Catrysse, B. Wandell, A simulation tool for evaluating digital camera image quality. *Proceedings of the SPIE*, Volume. 5294, p. 124-131, 2004.

2. T. Chen, P. Catrysse, A. El Gamal, and B. Wandell, How Small Should Pixel Size Be? *Proceedings of the SPIE*, Volume 3965, pp. 451-461, 2000.

3. P. Catrysse and B. Wandell, Roadmap for CMOS image sensors: Moore meets Planck and Sommerfeld, *Proceedings of the SPIE*, Volume. 5678, p. 1-13, 2005.

4. H. Rhodes, G. Agranov, C. Hong, U. Boettiger, R. Mauritzson, J. Ladd, I. Karasev, J. McKee, E. Jenkins, W. Quinlin, I. Patrick, J. Li, X. Fan, R. Panicacci, S. Smith, C. Mouli, and J. Bruce , CMOS imager technology shrinks and image performance.;In *2004 IEEE Workshop on Microelectronics and Electron Devices*, pp. 7 – 18, 2004.

5. F. Xiao, J. E. Farrell, B. Wandell, Psychophysical thresholds and digital camera sensitivity: The thousand photon limit. *Proceedings of the SPIE* , Volume 5678, p. 75-84, 2005.

6. Williams, D. and Burns, P. D, Low-Frequency MTF Estimation for Digital Imaging Devices using Slanted Edge Analysis, *Proceedings of the SPIE*, Volume 5294, 2004. *(*http://www.i3a.org/downloads_iso_tools.html)

7. F. Xiao, A. Silverstein & J. Farrell, Camera-Motion and Effective Spatial Resolution, International Congress of Imaging Science, Rochester, NY, May 2006.

8. X. Zhang and B. Wandell, A spatial extension of CIELAB for digital color reproduction. *Society for Information Display Symposium Technical Digest*, Volume 27, p. 731-734, 1996.

9. "Basic sRGB Math." <http://www.srgb.com/basicsofsrgb.htm>