Open access • Proceedings Article • DOI:10.1117/12.342860

Comparative analysis of color architectures for image sensors — Source link

Peter B. Catrysse, Brian A. Wandell, Abbas El Gamal

Institutions: Stanford University

Published on: 22 Mar 1999 - electronic imaging (International Society for Optics and Photonics)

Topics: RGB color model, Color image, Color depth, Color histogram and Color balance

Comparative analysis of color architectures for image sensors

Peter B. Catrysse*[a], Brian A. Wandell[b], Abbas El Gamal[a]

[a]Dept. of Electrical Engineering, Stanford University, CA 94305, USA

[b]Dept. of Psychology, Stanford University, CA 94305, USA

ABSTRACT

We have developed a software simulator to create physical models of a scene, compute camera responses, render the camera images and to measure the perceptual color errors (CIELAB) between the scene and rendered images. The simulator can be used to measure color reproduction errors and analyze the contributions of different sources to the error. We compare three color architectures for digital cameras: (a) a sensor array containing three interleaved color mosaics, (b) an architecture using dichroic prisms to create three spatially separated copies of the image, (c) a single sensor array coupled with a time-varying color filter measuring three images sequentially in time. Here, we analyze the color accuracy of several exposure control methods applied to these architectures. The first exposure control algorithm (traditional) simply stops image acquisition when one channel reaches saturation. In a second scheme, we determine the optimal exposure time for each color channel separately, resulting in a longer total exposure time. In a third scheme we restrict the total exposure duration to that of the first scheme, but we preserve the optimum ratio between color channels. Simulator analyses measure the color reproduction quality of these different exposure control methods as a function of illumination taking into account photon and sensor noise, quantization and color conversion errors.

Keywords: color, digital camera, image quality, CMOS image sensors

1. INTRODUCTION

The use of CMOS sensors in digital cameras has created new opportunities for developing digital camera architectures. The temporal, spatial and color sampling properties of CMOS sensors make it possible to sample outputs rapidly and to digitize and store data at each pixel[1,2,3]. The large number of design alternatives makes it imperative to develop simulation tools that help us predict the consequences of different designs. Here, we describe an initial implementation of a set of simulation tools. The simulator begins with a physical model of colored targets, ambient illuminants and optics to produce a physical description of the scene that is incident at the camera sensor. Simulations of the camera responses include photon and sensor noise, and several other sensor characteristics.

We have begun using these tools to analyze color reproduction quality of three digital camera color architectures. The first architecture uses a single lens and a conventional sensor array containing three interleaved color mosaics. The second architecture uses dichroic prisms to create three spatially separated copies of the image. The third architecture uses a single sensor array coupled with a time-varying color filter to measure three images sequentially in time. We compare the digital camera architectures by predicting the results at the end of the imaging pipeline, when the camera data are rendered on a simulated display. The color reproduction quality is evaluated by using a perceptual color metric CIELAB (1976) that compares the visible difference between the incident and rendered images.

After describing the basic elements of the simulator, we discuss one aspect of image acquisition: exposure control algorithms. We discuss how these differ between architectures, and we analyze the effect of these differences on both sensor signal-to-noise and on the perceptual fidelity of color reproductions.

* Correspondence: Email: peter@kaos.stanford.edu; Telephone: 650 725 1255; Fax: 650 723 0993

Part of the IS&T/SPIE Conference on Sensors, Cameras, and Applications for Digital Photography • San Jose, California • January 1999
SPIE Vol. 3650 • 0277-786X/99/$10.00

2. COLOR ARCHITECTURES

The most commonly used color architecture is based on a single sensor with a superimposed color filter array (Figure 1a). The color filter array comprises three interleaved color sample mosaics that are applied during the semiconductor processing steps. The individual pixels of the sensor are covered typically with "red", "green", or "blue" filters whose spectral transmissivity along with the other optical elements determine the spectral responsivity of the three color channels. The color filter array is specified both by the spatial layout of the mosaic pattern and by the spectral shape of the three color filters used in the mosaic. In general, the spatial arrangement is important because three sensor images must be interpolated before an image can be rendered. Reproduction errors arising from this interpolation process are called "demosaicing" artifacts. The analysis we report measures only color reproduction error and ignores spatial artifacts, which we expect to address this issue in later reports.

A second architecture we consider, uses a prism coated with dichroic mirrors to create three spatially separated copies of the image on three sensor arrays (Figure 1b). The mirrors separate different wavelengths into separate bands that define the "red", "green" and "blue" images. In this design, precise optical alignment is necessary to maintain correspondence between the images from different color channels. An important advantage of this design is that the spatial demosaicing required in the CFA is eliminated. Moreover, with this design every incident photon finds its way to a sensor and can be used in the image reproduction calculations so that the signal-to-noise ratio (SNR) at the sensor level is high. A disadvantage of this design is that one is restricted to using block color sensors, that is sensors with square pass bands.

A third architecture we consider uses a single time-varying color filter in the light path. The filter passes long, middle and short wavelength light sequentially, forming three images on a single sensor array (Figure 1c). This method of creating a color image, called field sequential color (FSC), amounts to making three electronically controlled exposures, one for each color channel. Again, no demosaicing is required.

An interesting feature of FSC and dichroic architectures is that three different exposure times can be used during the acquisition. We will explore the exposure control options and their influence on color reproduction in this paper.

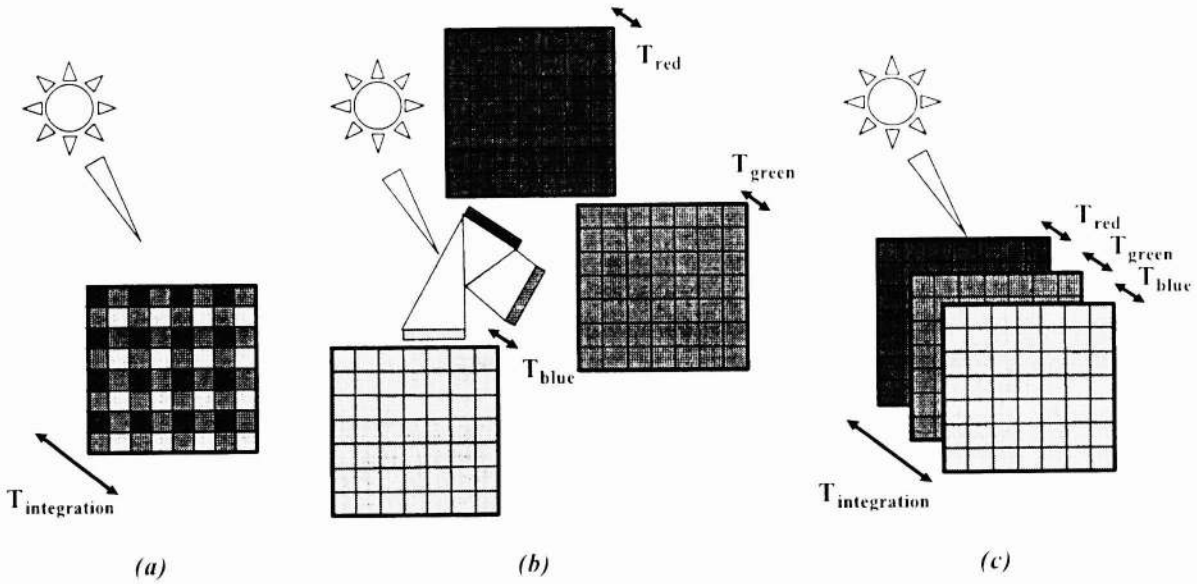

Figure 1: Color architectures: (a) Color Filter Array, (b) Dichroic Prism Filter, (c) Field Sequential Color Filter

For all three architectures, the values in thethree color channels must be converted to display or print values. This step, called color conversion, is a consequence of the need to create a displayed image that stimulates the human cones in approximately the same way as the cones would be stimulated when viewing the original scene. Color conversion is a necessary stage in the imaging pipeline and includes a significant loss of information for any device whose sensor spectral sensitivities are not

within a linear transformation of the human cone spectral sensitivities. For example, in the dichroic design, the color sensors are restricted to be block wavebands, quite unlike the sensors in the eye. This design limitation may influence the accuracy of the color conversion calculations.

All three architectures require the designer to make a series of design tradeoffs concerning spatial, temporal and chromatic image sampling and reconstruction. In the CFA and FSC designs, a significant fraction of the photons are wasted. For example, long-wavelength photons incident on a blue sensor (or during the blue frame of the sequential acquisition) are discarded. The reduction in number of available photons reduces the SNR of the design.

In the simulations we describe below, we ask: How important are these physical considerations when evaluating digital camera architectures? Conventional sensor measures, such as pixel SNR and quantum efficiency, place fundamental limits on subsequent image reconstruction. However, color architectures should not be compared at the sensor level because, in the end, the quality of the reproduction depends on the appearance of the displayed image, not the acquired image. Thus, these camera architectures must be compared by analyzing how well the system samples the image information needed to render an accurate reproduction of the scene. We use the color metric CIELAB (1976) to measure the perceptual similarity of the incident and rendered image.

3. SIMULATION METHODS

The software simulation divides the imaging pipeline into six parts: scene modeling, imaging optics, color architecture, sensor modeling, color processing and color reproduction evaluation. The software is implemented as a series of MATLAB routines. Each of the parts is a self-consistent module and inter-module communication takes place by making extensive use of data structures to accept and pass parameters. Each module can be modified to allow for improvements in the complexity of the physical models and algorithms without affecting the remaining parts in the pipeline. We summarize the computations and models used in the current implementation.

3.1. Scene model

The scene consists of planar objects with diffuse reflectance functions. The physical parameters used to model the scene[4] are the power spectral distribution of the ambient light $L(\lambda)$ and the surface reflectance of the object surfaces at each location $S(x_o,\lambda)$. We assume that the ambient light is generated by an extended source of radiation and we characterize this source by the radiance L in watts per square meter per steradian. This describes the flux per unit of projected area per unit solid angle (Figure 2). The spectral power distribution $L(\lambda)$ is then defined as the distribution of the radiance L with respect to the wavelength λ

$$L(\lambda) = \frac{\partial L}{\partial \lambda}.$$

The reflectance functions describe the fraction of ambient light scattered at each wavelength. The color signal $C(x_o,\lambda)$ is computed as the spectral product of ambient light and the reflectance functions. (We do not incorporate specular reflections or other geometric factors such as secondary reflections.) The color signal $C(\lambda)$ at point x_o and time t is denoted as

$$C(x_o,t,\lambda) = L(x_o,t,\lambda)S(x_o,t,\lambda).$$

Many radiance meters measure the light signal in terms of the energy or power. For our purpose it is more useful to have a photon or quantum description for the color signal. The conversion from energy can be obtained by dividing the energy expressions by hc/λ. The fundamental quantity, $L_q(\lambda)$, represents the number of photons per unit time, per unit spectral interval, per unit projected area, per unit solid angle. The color signal in photons is computed from

$$C_q(x_o,t,\lambda) = L_q(x_o,t,\lambda)S(x_o,t,\lambda).$$

3.2. Imaging optics model

We model the imaging optics using the effective f-number $f_{eff}^{\#}$, which represents the ability of the system to provide (large) image irradiance. The smaller the f-number the more irradiance the optics can deliver to the image. The effective f-number is defined[1] as the ratio of the effective focal length efl of the system to the diameter D of the clear aperture

$$f_{eff}^{\#} = \frac{efl}{D} = \frac{1}{2\tan\theta_i}.$$

The optics transforms the photonic radiance $C_q(\lambda)$ from the scene into a photonic irradiance $E_q(\lambda)$ of the image plane,

$$E_q(x_i,t,\lambda) = \pi \sin^2\theta_i T(\lambda) C_q(x_o,t,\lambda) \cong \frac{\pi}{4}\frac{1}{\left(f_{eff}^{\#}\right)^2} T(\lambda) C_q(x_o,t,\lambda),$$

where the last equality is valid for $f_{eff}^{\#} >> 1/4$. For the calculations described here, we assume the optics free of transmission losses. This means $T(\lambda) = 1$ for all visible wavelengths. We have not yet adjusted for off-axis behavior and ignore the spatial resolution characteristics of the optics since we are in this work focussing on color reproduction only.

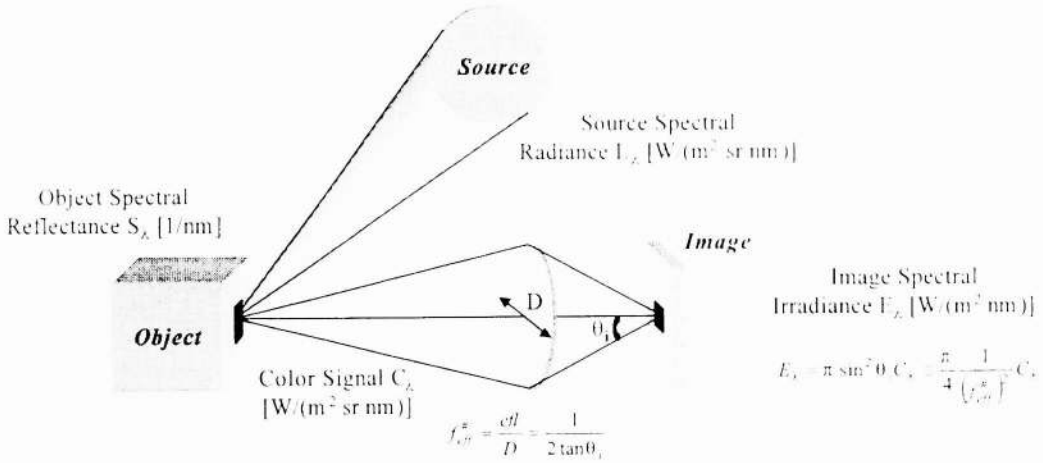

Figure 2: Imaging geometry and physical quantities used to describe the photon distribution incident at the sensor array.

3.3. Color responsivity

The various wavelengths of the irradiance at the image sensors, E_q are differentially transmitted or absorbed by the color filters, sensors, and other elements in the imaging path. In all of the color architectures, the wavelength dependent effects are combined into a spectral responsivity function, $R_k(x,t,\lambda)$ for each color channel $k = 1,2,3$. This function combines the dimensionless spectral transmittances of the color filters with the spectral quantum efficiency of the sensor, which has units of electrons per photon. The number of photon-generated electrons at image location x_i is then calculated by integrating over the wavelength

$$\rho_k(x_i,t) = \int_{\lambda} E_q(x_i,t,\lambda) R_k(x_i,t,\lambda)\, d\lambda.$$

[1] The f-number of an optical system is usually defined by the relation f# = f/D where f is the (second) focal length of the system and D is the diameter of the clear aperture. For an object at infinity, the effective f-number is equal to the f-number, although for objects at finite distances the effective f-number is larger than the f-number.

Finally, the total channel response is then given by integrating over the color channel exposure time T_k and the pixel area A

$$\rho_k(x_i, T_k) = \int_0^{T_k} \rho_k(x_i, t) A dt .$$

3.4. Sensor model

We model the CMOS image sensor using a simplified photocurrent-to-output-voltage sensor model followed by an n-bit uniform quantizer. In our model, the photodiodes in the CMOS sensor are operated in current integration mode. For a single color channel, this mode of operation can be modeled by a photocurrent $i_{ph} = \rho(x,t)A$, due to the total channel response, and a dark current i_d. The sum of both currents is integrated on a capacitor C_d for an exposure duration T and produces an accumulated charge

$$Q = \int_0^T (i_{ph} + i_d) dt$$

This charge is subsequently read out, converted into a voltage V and quantized. The sensor has a finite charge capacity of q_{max} electrons and that the linear charge-to-voltage amplification is given by g (Figure 3).

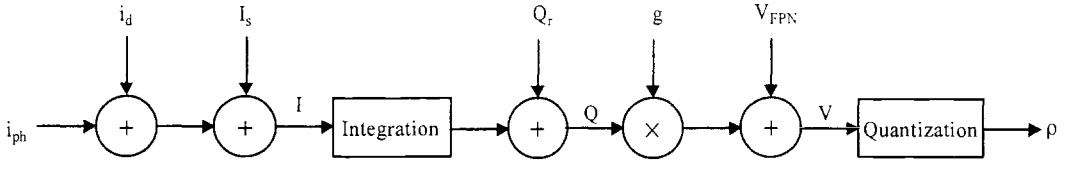

Figure 3: CMOS sensor noise model for one of the three color channel

The intrinsic noise is shot noise in the form of current passing through the diode[‡]. This shot noise has two components, a photocurrent and dark current component and is represented by the current source $I_s(t)$. If the number of detected photons is large, this noise can be modeled by a signal dependent Gaussian distribution with a variance equal to the mean of the generated signal $i_{ph}+i_d$. Besides shot noise, there is readout circuitry noise Q_r, which includes input referred amplifier noise and reset noise for CMOS APS. We assume that this noise can be modeled as additive Gaussian signal-independent noise with zero mean and standard deviation σ_r. The final charge accumulated on the capacitor therefore becomes,

$$Q(x,T) = Q_r(x) + \int_0^T (i_{ph}(x,t) + i_d(x,t) + I_s(x,t)) dt$$

This charge is converted into a voltage and is subject to fixed pattern noise, $V(x) = g \cdot Q(x,T) + V_{FPN}(x)$, before being quantized to yield the digital color response of one particular color channel $\rho(x)$. We will model the fixed pattern noise V_{FPN} as additive signal-independent noise as well[§].

3.5. Color processing

To evaluate the effect of the various scene elements and the camera model on color reproduction, as seen by the user, the camera response must be rendered. For the simulations here, we have used a simple CRT display model as an output device. For the color conversion step, we applied a linear transformation, found using conventional methods[5], to convert the (linear) camera responses into the (linear) RGB display levels.

[‡] Any measurement of an optical signal will exhibit uncertainty due to the quantum nature of light. The photocurrent generated by the photodiode is the result of a Poisson counting process of the incident photons during integration and consequently the uncertainty shows up as photocurrent shot noise.

[§] This ignores gain FPN, which is signal dependent, but we will compensate for this through choice of the FPN parameter value.

Unless the camera spectral responsivities are within a linear transformation of the CIE (1931) XYZ functions, this color conversion step introduces error into the imaging pipeline. We describe the relative contribution of this conversion error below and compare it with other sources of reproduction error.

3.6. Color reproduction evaluation

Because the simulated targets are large uniform patches, it is possible to use the CIELAB (1976) metric to measure perceptual errors[6]. When applying the simulation to images with more complex spatial structure we plan to use a spatial extension, S-CIELAB[7].

The units of perceptual error defined for CIELAB and S-CIELAB are called ΔE_{ab}. The average value of this measure that is accepted in printing applications has been found to be approximately a ΔE_{ab} of six[8] units. Another study[9] reports on the perceptibility tolerance for pictorial images and found the average ΔE_{ab} to be 2.15 units. As a rule of thumb, colors are difficult to distinguish if the ΔE_{ab} value between two colors is less than three.

4. SIMULATIONS

4.1. Illuminations and surfaces.

We simulated ambient illuminants as blackbody radiator sources whose color temperatures ranged between 3000K and 7000K. These sources span ambient lights similar to natural daylight (D65) and artificial illuminants such as tungsten filament lamps that are in common use today [10].

We simulated diffuse surfaces using the reflectance functions in the Macbeth Color Chart. This is a collection of 24 colored squares, some of whose reflectance functions are similar to natural objects of special interest, such as human skin, foliage and blue sky. Because the squares match in reflectance they can be used to calculate the color signal under any illumination. Completing the chart are the additive primary colors blue, green and red; the subtractive primaries yellow, magenta and cyan; and a neutral series ranging from white to black

4.2. Imaging optics parameters.

The effective f-number $f_{eff}^{\#}$ of the system is chosen to be 1.4. The spectral transmittance of the optics is $T(\lambda) = 1$ for all visible wavelengths.

4.3. Color architecture parameters.

The color filter array (Figure 1a) architecture is specified both by the spatial layout of the mosaic pattern and by the spectral shape of the three color filters used in the mosaic. We apply a Bayer pattern for the color mosaic and choose block color filters with cutoffs at 490nm and 610nm. For the field sequential color architecture (Figure 1c) we use a single time-varying color filter in the light path with filter spectral shapes identical to the ones for the CFA architecture. Finally, we model the dichroic prism architecture, which creates three spatially separated copies of the image on three sensor arrays (Figure 1b).

4.4. Image sensor parameters

We model the image sensor using a simplified photocurrent-to-output-voltage sensor model followed by an n-bit uniform quantizer as described in the previous section. The main parameters values used in this model are $q_{max} = 131072$ *electrons*, $\sigma_r = 20$ *electrons*, $i_d = 10$ *electrons/(s* μm^2*)*, $n = 8$, and *FPN <1% full output swing.*

4.5. Exposure control models

The main differences we have explored between the color architectures are exposure control algorithms. For the CFA architecture, one must choose a single exposure duration that applies simultaneously to the three interleaved color filter arrays. Dichroic prisms and FSC architectures can use different exposure times in each color channel.

Depending on the nature of the sensor, it is possible to use several different types of algorithms for establishing the exposure duration. For the simulations reported here we used a very simple method, the white-patch-exposure method. For the CFA architecture we set the exposure time so that the maximum of the three color channels is near saturation. For the FSC and dichroic architectures we set the exposure times so that the response to white is at maximum in all three channels.

Figure 4 shows the relative response of the three sensor classes to a white color patch under several ambient illuminants, ranging from a tungsten yellow bulb (3000K) to a blue-sky (7000K). Observe that for the CFA architecture, in which there is only one exposure duration, only the green channel responds optimally across the entire range of illuminants. The red and blue channels never achieve their maximum response and hence their SNR is less than optimum. Across this choice of illuminants the red sensor response decreases and the blue increases, but both stay below the green.

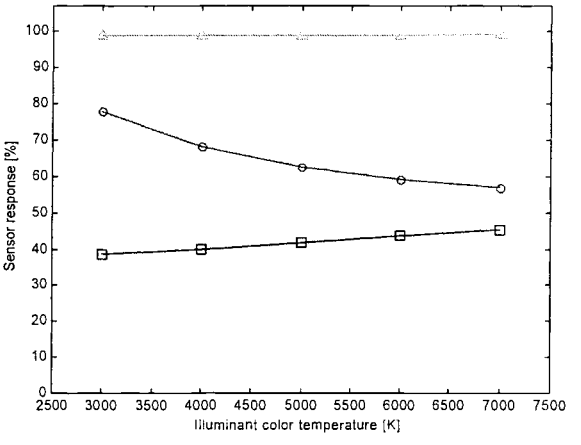

Figure 4: Relative response of the three sensor classes to a white surface as a function of illumination and exposure control (circles: "red" channel; triangles: "green" channel ; squares: "blue" channel)

When using the FSC architecture, each color channel has its own exposure duration and thus the sensor SNR for a white (or gray) surface is independent of the illuminant and optimal in terms of SNR. This balance is desirable feature because it serves to optimize SNR in all three color channels. The visual significance of this difference in exposure control will depend on several properties of the camera as well as the surface and illumination conditions.

5. RESULTS

5.1. SNR results

Figure 5a shows the combined SNR for the three color channels of the imaged object as a function of the illuminant color temperature for the CFA and FSC architectures using three exposure control algorithms. These values were calculated using a CMOS sensor model in combination with the block spectra defined earlier. The open triangles show the FSC architecture with exposure durations set to maximize SNR to a white surface. The circles show the CFA with conventional exposure control. The squares show FSC with total exposure equal to the CFA exposure duration.

The FSC architecture (triangles) improves SNR over CFA (circles) by about 1-3dB, particularly for the 3000K and 7000K illuminants. The improved SNR for the FSC, however, comes at the expense of a prolonged exposure time (4x). If the total exposure time is set equal to the CFA exposure time (squares), the SNR remains balanced in the three channels but causes an overall decrease in SNR by 10dB across illuminants. In some applications, such as studio photography or color microscopy, the increased exposure duration might be acceptable. In other applications involving moving images, the additional exposure time will adversely affect image quality. When using block sensors, dichroic prism architecture is similar to FSC except for some differences in the exposure time. The total exposure time is now determined by the slowest channel, not the sum of the three channels.

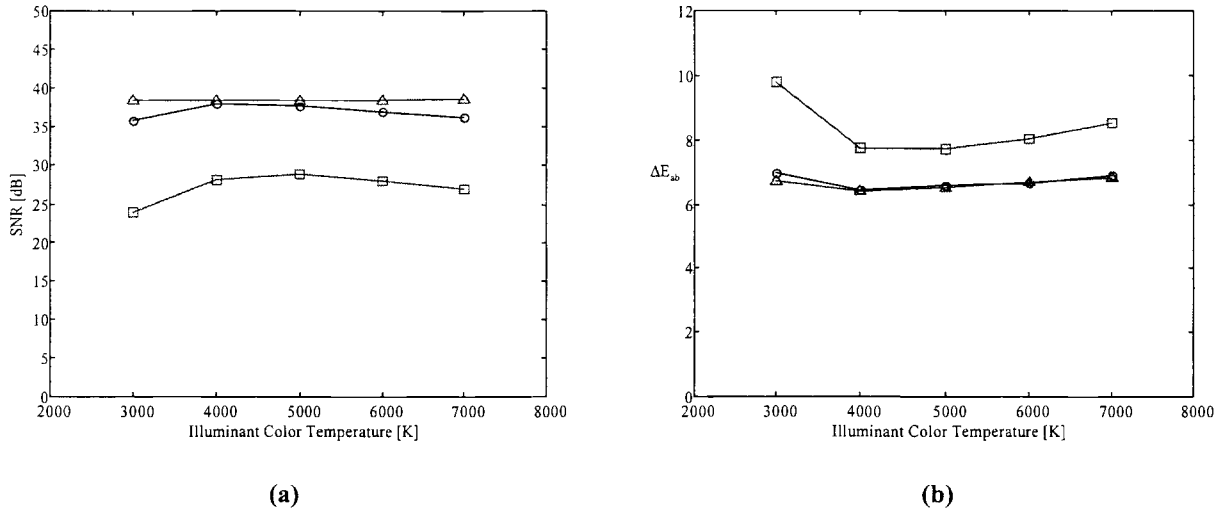

Figure 5: SNR (panel a) and perceptual error ΔE_{ab} (panel b) for exposure control algorithms for the CFA and FSC architectures. The SNR and perceptual error values represent the average for all surfaces in the Macbeth Color Chart. (circles: CFA timing duration = 28 ms; triangles: FSC timing duration = 111 ms; squares: FSC balanced timing duration = 28 ms).

5.2. Perceptual reproduction accuracy

Figure 5b shows the results of the simulation again, but this time the error is measured in perceptual units, ΔE_{ab}. With a single exposure setting, the average color reproduction error of the CFA is 7 units for all ambient illumination color temperatures. The improvement when using separate color exposures for FSC is less than 1 unit. Simulations of the dichroic prism architecture match the behavior of the FSC architecture. If we restrict the total exposure time for FSC by a factor of 4, making it the same as the CFA, perceptual error increases only 2-3 units. The small difference in perceptual quality compared to the large (10dB) difference in SNR is not immutable. Rather, it depends on the overall quality of the imaging system. In this simulation regime, the image data are of relatively high quality compared to the visual sensitivity and reducing the signal to noise by shortening the exposure duration only produces a modest loss of color fidelity. Under other conditions, a 10dB improvement in sensor SNR can result in a substantial change to the perceived image quality.

5.3. Error contributions from different sources

Simulation tools are well suited to assist the designer in isolating the major noise contributors because sources of error can be independently removed from the simulation. To understand the contributions of noise sources we started with a perfect system and added the various noise sources sequentially. Figure 6 shows, in the form of a cumulative perceptual error in ΔE_{ab} chart, how error grows as we add photon shot noise, CMOS noise, quantization noise (8 bit) and color conversion artifacts. With every step the error increases, but the main sources of error are the CMOS noise and the color conversion step. The color conversion step is highly sensitive to the choice of color filters and this suggests that further experiments with filter design could yield important benefits.

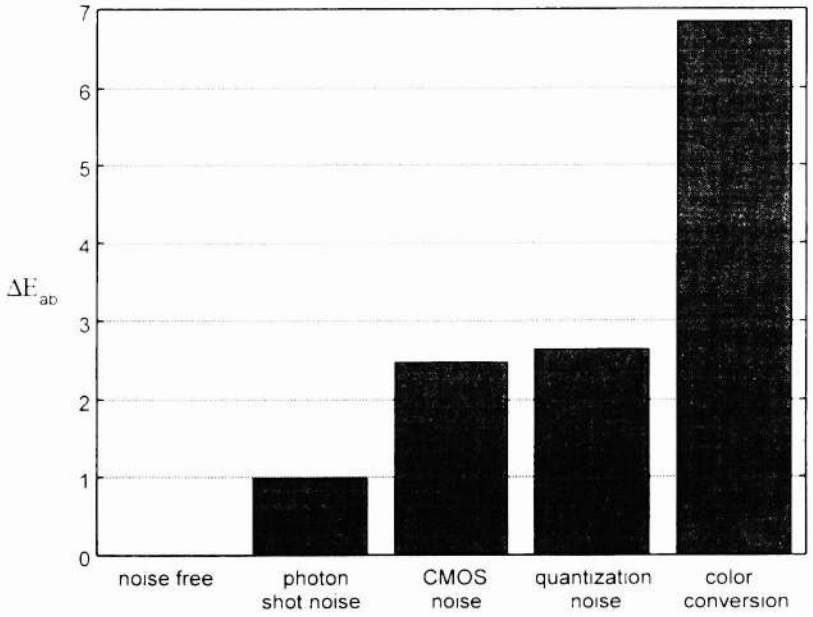

Figure 6: Cumulative perceptual error in the imaging pipeline. The average perceptual error of all surfaces in the Macbeth Chart measured under a D65 illuminant is shown. Each bar represents the rendered error as that step and all the steps to the left are included in the simulation.

6. DISCUSSION AND CONCLUSIONS

We have described a preliminary implementation of simulation tools designed to assist in the development and evaluation of digital camera architectures. These tools include methods of tracking the flow of photons to the sensor array, how the sensor converts these photons to electrons, and finally how these signals are converted back to the photons in a displayed color image. Here, we have reported the consequences of using different exposure controls. We have emphasized the color reproduction differences between CFA and FSC specifying the temporal disadvantages of FSC and the potential SNR advantages. We measured system performance using both sensor SNR and the CIELAB metric. The simulations shown here include an example of a 10dB improvement in SNR that corresponds to only a 1-2 ΔE_{ab} improvement. Hence, a rather large SNR can correspond to a very modest perceptual improvement. In other simulations we have performed comparable SNR improvements produce much larger perceptual differences suggesting that it is important to evaluate system performance using perceptual metrics.

From our analysis of the error propagation in the color imaging pipeline, we find that the main sources of perceptual error in the modeled system are due to the CMOS noise and the color conversion. In the future, we will incorporate spatial factors to understand how the elimination of demosaicing provides a spatial advantage for FSC. We anticipate using these simulation tools broadly in conjunction with experimental measurements of CMOS sensors we have designed, implemented and are testing in the lab.

ACKNOWLEDGMENTS

This work was supported by the Programmable Digital Camera project at Stanford University (http://smartcamera.stanford.edu/pdc.html). Peter Catrysse acknowledges the support from Hewlett-Packard through the Fellow/Mentor/Advisor (FMA) Program and is an "Aspirant" of the Fund for Scientific Research – Flanders (Belgium) in connection with the Department of Applied Physics and Photonics of the Vrije Universiteit Brussel (VUB).

REFERENCES

[1] E.R. Fossum, "CMOS image sensors: Electronic camera-on-a-chip," *IEEE Transactions on Electron Devices,* **44**, No. 10, pp. 1689-1698, 1997.

[2] H.S. Wong, "CMOS Image Sensors – Recent advances and device scaling considerations," *IEEE IEDM Technical Digest*, pp. 8.5.1-8.5.4, 1997.

[3] A. El Gamal, B.A. Fowler, D.X.D. Yang, "Pixel level processing: why, what, how?" *Proceedings of SPIE*, **3650**, 1999 (this volume)

[4] B.A. Wandell, "The Synthesis and Analysis of Color Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* **PAMI-9**, No. 1, pp. 2-13, 1987.

[5] B.A. Wandell, *Foundations of Vision,* Sinauer Associates, Sunderland, 1995.

[6] G. Wyszecki and W.S. Stiles, *Color Science: Concepts and Methods. Quantitative Data and Formulae*, Wiley, New York, 1982, p.166

[7] X. Zhang, B.A. Wandell, "A spatial extension of CIELAB for digital color image reproduction," *Journal of the Society for Information Display,* **5**, No. 1, pp. 61-63, 1997.

[8] S. Stamm, "An investigation of color tolerance," *TAGA Proceedings*, pp. 156-173, 1981.

[9] M. Stokes, M.D. Fairchild, and R.S. Berns, "Precision requirements for digital color reproduction," *ACM Transactions on Graphics,* **11**, No. 4, pp. 406-422, 1992.

[10] R.W.G. Hunt, *The Reproduction of Colour*, Fountain Press, Kingston-upon-Thames, 1995, p. 196