A Model for Calculating the Potential ISO Speeds of Digital Still Cameras based upon CCD Characteristics

Richard L. Baer and Jack Holm
Hewlett-Packard Company, Palo Alto, California

Abstract

This paper presents an algorithm that can be used to predict the ISO speed of a digital camera, based upon the characteristics of the CCD. The model is based upon the ISO TC42/WG18 standard for the specification of speed in electronic and digital still cameras. The model uses an analytical description of a Planckian radiator and the measured quantum efficiencies of the CCD to determine the number of CCD electrons that are produced in each color channel per lux-second of focal plane exposure. The CCD signal and noise are passed through white balance and color correction matrices, which are calculated from the measured CCD quantum efficiencies, in order to determine the overall signal to noise ratio of the system. The ISO speed is determined by solving for the focal plane exposure required to reach a specified signal to noise ratio. The model has been used to investigate the effects of various CCD parameters on the potential ISO noise-based speed of a digital camera.

Introduction

The ISO speed model is based upon International Standard ISO 12232, Photography – Electronic still-picture cameras – Determination of ISO speed.[1] The standard describes a procedure for determining the noise-based speed range of a digital camera. This speed has the same interpretation as the ISO speed for photographic film, in that it specifies the correct nominal exposure conditions (f-stop, exposure period) for a given scene brightness.

The ISO standard describes speeds that are based on the focal-plane illumination required to achieve a specific mid-tone signal-to-noise ratio. Other methods of speed determination have been developed,[2] but the advantage of this method is that it specifies the image quality and makes speed comparisons among different cameras meaningful.

Our model uses the measured CCD characteristics to estimate the potential ISO noise-based camera speed. The calculation has three steps: evaluation of the CCD responsivity, determination of the signal to noise (S/N) ratio as a function of focal plane exposure, and determination of the ISO speed from the S/N ratio. These steps are described in the succeeding sections of this paper, followed by an example of their application, and our conclusions.

CCD Electrons/lux•sec Calculation

The pixels of a CCD capture electrons in proportion to the focal plane exposure, in lux-seconds. The response is linear up to the point of saturation, and may be characterized by a single set of responsivity values, one for each color channel, for a given illuminant. The responsivities can be calculated directly from the CCD absolute quantum efficiency versus wavelength curves. These curves may be obtained from the CCD manufacturer, or measured with a monochrometer.

Lux is a photometric measure of illumination, based upon the sensitivity of the human eye. If the illumination source generates photons with wavelengths that fall outside the range of human vision but within the range of the CCD, the ISO speed values that are calculated will be meaningless. For example, if the illumination consisted of IR radiation the lux value to which the responsivity would be normalized would be zero. Silicon CCDs do not respond strongly wavelengths of light that are below the range of human vision because the photon absorption depth is too short. However, they do respond to longer wavelengths in the near IR, from about 800 to 1100 nm. Photographic systems usually filter out IR in order to reproduce scenes as human observers perceive them. This is especially important in color photography, where the IR response causes inaccurate color rendition. For these reasons, we modeled a 1 mm thick CM-500 absorptive IR filter in our ISO speed calculations.

The analysis begins by modeling the illumination source as a Planckian radiator. This model was chosen because it has a simple analytical representation and because the color temperature can be varied. Any other illuminant could be incorporated in the model by including either its analytical representation or a set of sample values.

The spectral irradiance density of the black-body illuminant in watts/($m^2 \cdot$Hz), as a function of wavelength (λ), appears in Equation 1. α is an arbitrary constant which is normalized out in the analysis. It includes the conversion from radiant exitance to spectral irradiance.

$$\Phi_e(\lambda) = \frac{\alpha}{\lambda^5} \cdot \left(e^{hc/kT\lambda} - 1\right)^{-1} \tag{1}$$

The number of CCD electrons generated by the illuminant is calculated in Equation 2. The integral is performed for each color channel, resulting in a vector with one element for each

CFA color. The units of N_e are electrons/(m^2·sec). In this expression QE(λ) is the vector of quantum efficiencies for each color channel and IR(λ) is the infrared filter transmission.

$$\overline{N_e} = \int \overline{QE}(\lambda) \cdot IR(\lambda) \cdot \Phi_e(\lambda) / (hv) \cdot d\lambda \qquad (2)$$

The focal plane illuminance associated with the illuminant is calculated in equation 3, based on the photopic response. The units of *l* are lux. In this expression ν(λ) is the standard luminosity function.

$$l = 680 \int v(\lambda) \cdot \Phi_e(\lambda) \cdot d\lambda \qquad (3)$$

The result of equation 2 is normalized by results of equation 3 to get the CCD responsivity vector (η_p), in units of electrons/(lux·sec·m^2). η_p is the photopic quantum efficiency to a particular illuminant. The product of η_p, the pixel area, and the focal plane exposure yields the number of CCD electrons.

$$\overline{\eta_p} = \overline{N_e} / l \qquad (4)$$

Signal to Noise Ratio Calculation

The CCD responsivity values that were calculated in the previous section can be used to compute the signal to noise ratio in the image as a function of focal plane exposure. The signal level is calculated in Equation 5, where M_w is the white balance matrix, M_c is the color correction matrix, M_y is the matrix that transforms RGB to luminance, *A* is the pixel area and H is the focal plane exposure in lux·seconds. In the monochrome case, the η_p vector collapses to a scalar and the matrix transformations are unnecessary.

$$S = \overline{\overline{M_y}} \cdot \overline{\overline{M_c}} \cdot \overline{\overline{M_w}} \cdot \left(A \cdot \overline{\eta}_p \cdot H \right) \qquad (5)$$

M_w is a diagonal matrix that balances the levels of all the color channels. It is derived by selecting a set of multipliers that equalize the levels of all the elements in the CCD responsivity vector to the level of channel 1. The color of the illuminant defines the white point for any color temperature, so this method works for any source distribution. If one color channel is less responsive than the others, the white balance matrix will boost its level causing the channel with the lowest responsivity to strongly influence the overall S/N ratio. For this reason the color filters should be designed to produce balanced responsivities under normal lighting conditions.

The color correction matrix transforms the color channels of the sensor to standard RGB. The matrix is calculated from the CCD quantum efficiency curves. The derivation of the color correction matrix is outside the scope of this work, but can be found in References 3, 4 and 5. In general, the color correction matrix will be highly diagonal if the color filters on the CCD are spectrally narrow. If the color filters are spectrally broad then matrix will contain large off-diagonal elements that will amplify the noise more than the signal.[6] The color correction matrix has the most potential of any of the image processing operations to impact the ISO speed.

The S/N ratio is calculated for a super-pixel that contains a contribution from each of the separate color channels in the color filter array. If the CCD has three color channels, then three pixels contribute to each super-pixel. This is a good model of simple demosaicing schemes like nearest-neighbor interpolation. However, it doesn't represent more complicated demosaicing schemes very well, which achieve higher resolution at the expense of signal to noise ratio.

The incoming photon flux obeys Poisson statistics, so the temporal variance of the number of CCD electrons per pixel is equal to the mean. Consequently the responsivity vector can also be used as the input to the noise calculation by interpreting it as noise power rather than mean signal. The square of the read noise (in electrons) can be added to the shot noise terms to get the total noise power. The inclusion of read noise only makes a significant difference in the noise-based speeds calculated for low mid-tone signal to noise ratios. Otherwise the shot noise of the photon flux is dominant.

The effects of dark current can be lumped in with the amplifier read noise, since the dark current level isn't dependent on the level of illumination. There are three important sources of noise associated with the dark current: 1) the shot noise of the mean dark current accumulated during the readout period, 2) the dark current accumulated during the exposure period (fixed pattern noise) and 3) the shot noise of the exposure period dark current. The readout dark current itself isn't important because it is highly uniform and can be eliminated by black level clamping. The shot noise of the readout dark current increases linearly across the frame so it can be represented in the model by the average value in the middle of the frame. In modern CCDs, the average dark current accumulated during the exposure period is very low (~20 nA/cm^2 at room temperature). However the distribution is very broad and it is not well described by a statistic like the standard deviation. We have only included the exposure period dark current in calculations of the maximum useful exposure period. The shot noise of the exposure period dark current is only a relevant consideration when dark-frame subtraction is used.

The total luminance noise is calculated in Equation 6, where N_r is the total read noise (in electrons). The net transformation matrix that is obtained by cascading the white balance, color correction and RGB to luminance matrices is squared in order to sum the powers of the noise terms. The noise level of the image is determined by taking the square root of the noise power.

$$N_Y = \sqrt{\left(\overline{\overline{M_y}} \cdot \overline{\overline{M_c}} \cdot \overline{\overline{M_w}} \right)^2 \cdot \left(A \cdot \overline{\eta_p} \cdot H + N_r^2 \right)} \qquad (6)$$

The total noise, including the effects of chrominance noise, can be determined from the luminance channel (Y) and two chrominance signals, (R-Y) and (B-Y) using Equation 7 (taken from Reference 7). In this equation σ_c is the total noise, σ_Y is the luminance noise and the chrominance channel noises are $\sigma_{(R-Y)}$ and $\sigma_{(B-Y)}$. The constants c_1 and c_2 have been determined by a set of noise perception experiments, and are approximately $c_1 \approx 0.64$ and $c_2 \approx 0.16$.

$$\sigma_c = \sqrt{\sigma_Y^2 + c_1\sigma_{(R-Y)}^2 + c_2\sigma_{(B-Y)}^2} \qquad (7)$$

The color noise is calculated in Equation 8. In this equation $M_{(R-Y)}$ is the transformation matrix from RGB to (R-Y) and $M_{(B-Y)}$ is the transformation matrix from RGB to (B-Y).

$$\begin{aligned} N_c = \Big\{ & \left[A \cdot \overline{\eta}_p \cdot H + N_r^2\right] \\ & \Big[\left(\overline{\overline{M}}_y \cdot \overline{\overline{M}}_c \cdot \overline{\overline{M}}_w\right)^2 + c_1\left(\overline{\overline{M}}_{(R-Y)} \cdot \overline{\overline{M}}_c \cdot \overline{\overline{M}}_w\right)^2 \\ & + c_2\left(\overline{\overline{M}}_{(B-Y)} \cdot \overline{\overline{M}}_c \cdot \overline{\overline{M}}_w\right)^2\Big]\Big\}^{1/2} \end{aligned} \qquad (8)$$

The signal to noise ratio is calculated by dividing the signal level from Equation 5 by the noise level from either Equation 6 or Equation 8, depending on whether luminance noise is being considered.

Noise-based ISO Speed

The ISO noise-based speed is defined as:

$$S_{noisex} = 10 / H_{S/Nx} \qquad (9)$$

where S_{noisex} is the ISO noise speed at a mid-tone signal to noise ratio of x, and $H_{s/Nx}$ is the corresponding focal plane exposure in lux-seconds. In order to determine the speed, we solve for the level of illumination required to obtain a specific signal to noise ratio and derive the focal plane exposure from it.

An analytical expression for the speed can be derived in the monochrome case, where only the luminance channel is present. The monochrome speed appears in Equation 10. In this equation S/N_x is the mid-tone signal to noise ratio upon which the speed is based.

$$S_{noisex} = \frac{20A\eta_p}{S/N_x{}^2}\left(1 + \sqrt{1 + \frac{4N_r^2}{S/N_x{}^2}}\right)^{-1} \qquad (10)$$

In the color case we have used numerical means to calculate the speed by implementing the model in *Mathematica* and used its "Solve[ ]" function.

Example of ISO Speed Calculation

We have used the techniques described in this paper to analyze the ISO speed of the SONY ICX084AK, a VGA-resolution progressive-scan interline-transfer CCD that has been used in many digital cameras. This device has square 7.4 µm pixels. The quantum efficiency curves that we measured with our monochrometer system are shown in Figure 1. The response of the CCD to near IR radiation demonstrates the importance of including an infrared filter in the model (and the camera).

The CCD responsivity was calculated at a color temperature of 5500°K. The η_p values that we obtained were {420, 1050, 1100} electrons/lux·sec·µm^2, in the red, green and blue channels respectively. For comparison, the total photon flux incident on the sensor after the infrared filter is 8380 photons/lux·sec·µm^2. The low responsivity of the red channel degrades the ISO speed potential of the sensor at this color temperature. This CCD would have a higher speed potential if the red channel response were higher to compensate for the effects of the infrared filter.

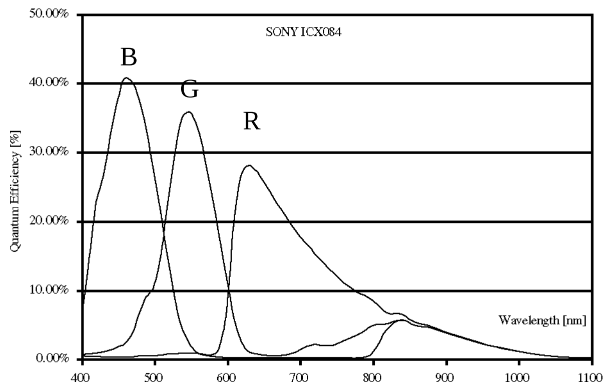

Figure 1. Quantum Efficiency vs. Wavelength

In the ISO standard, the specific mid-tone signal-to-noise ratios of 10 and 40 are used to represent "acceptable" and "excellent" image quality. These values are based on subjective tests performed during the development of the standard. These are the S/N values that we have used to determine a potential ISO speed range.

Based only upon the noise in the luminance channel, we calculate an ISO speed range of 394 to 3126 (S/N = 40 and S/N = 10 respectively). The speed range calculated for the green channel, using the analytical expression for the monochrome case, is 320 to 2760. If color noise is considered, the speed range falls to 102 to 1023.

The saturation speed[1] of the sensor can be determined from the η_p vector and the pixel area, as shown in Equation 11. In this equation N_{sat} is the full well capacity of the CCD and MAX( ) is a function that returns the largest element of its vector argument.

$$S_{sat} = 78 / H_{sat} = 78 \cdot MAX\left(\overline{\eta_p}\right) \cdot A / N_{sat} \qquad (11)$$

Based upon our experimental determination that the full-well capacity is 12,000 electrons, the saturation speed is 400. Since this is higher than the SNR=40 color-noise based speed, the CCD is incapable of reaching excellent image quality and the actual speed range is 400 to 1023.

Conclusions

It is instructive to inspect the equation for monochrome noise speed in order to discover some of the fundamental scaling laws. The noise-based speed increases linearly with pixel area (A) and quantum efficiency ($\propto\eta_p$). While the pixel area is bounded only by economic constraints, the quantum efficiency is limited to 100%. Modern state-of-the-art CCDs have peak quantum efficiencies of 30%-50% (without color filters) and they are unlikely to exceed the ~70% peak efficiencies of large area silicon detectors. Unfortunately, manufacturers are continuing to shrink the pixel dimensions of their CCDs at a greater rate than they are improving their quantum efficiencies. Considering both of these factors, the ISO speeds of digital cameras are likely to fall in the future.

The model also identifies the fundamental speed versus resolution tradeoff. The (speed×resolution) product is linearly proportional to the area of the sensor. The only way to improve the photographic quality of a digital camera in terms of both resolution and speed is to increase the sensor area. A good metric for the utility of an image sensor in digital photography is the product of resolution and ISO speed, which is proportional to the product of the sensor area and the peak quantum efficiency.

The impact of read noise is also apparent in the monochrome speed equation. Read noise significantly decreases the speed when $N_r \approx S/N$ ratio. Modern CCDs typically have read noise levels of 10 to 20 electrons (depending on various factors like the electronic bandwidth). Upon this basis, improvements in CCD read noise (and dark current) could significantly enhance the upper limit (SNR=10) noise speeds of future digital cameras, although they are unlikely to significantly effect the lower limit (SNR=40) speeds.

The human visual system is insensitive to high frequency information in chrominance. Television systems take advantage of this by dedicating less bandwidth to the chrominance signal than the luminance signal. The same method can be used in digital still photography to increase the ISO speed. Spatial averaging can be used to decrease the resolution of the chrominance channel. The ISO speed asymptotically approaches the luminance-only speed as the radius of the blur circle is increased. In our example, the potential ISO speed increase that could be achieved by spatial averaging is approximately a factor of three.

The spectral distributions of the color filters have a significant effect on the ISO speed. We have used this model to examine the tradeoffs in terms of speed and color saturation for primary and complementary color filters. This work can be found in Reference 5.

This model does not include some image-processing elements like tone correction and advanced demosaicing methods. Consequently the predictions of the model are not expected to correspond exactly with measured speed values from complete digital cameras. These elements could be included to achieve greater accuracy.

References

1. ISO 12232: Photography – Electric still-picture cameras – Determination of ISO speed, (1998)
2. M. Kriss, "A Model for Equivalent ISO CCD Camera Speeds", *Proceedings, International Congress on Imaging Science,* Antwerp, Belgium, 1998.
3. P. M. Hubel, J. Holm, G.D. Finlayson, "Illuminant Estimation and Colour Correction*", Proceedings, Colour Imaging in Multimedia – CIM 98,* Derby, England, p. 97-105, 1998.
4. P. M. Hubel, J. Holm, G.D. Finlayson, M.S. Drew, "Matrix Calculations for Digital Photography", *Proceedings, IS&T/SID Fifth Color Imaging Conference: Color Science, Systems, and Applications,* p. 105-111, 1997.
5. J. Holm, "Issues Relating to the Transformation of Sensor Data into Standard Color Spaces", *Proceedings, IS&T/SID Fifth Color Imaging Conference: Color Science, Systems, and Applications,* p. 290-295, 1997.
6. R. L. Baer, W.D. Holland, J. Holm, P. Vora, "A comparison of Primary and Complementary Color Filters for CCD-based Digital Photography", *Proceedings IS&T/ SPIE Symposium on Electronic Imaging,* (in press), San Jose, CA, 1999.
7. J. Holm, "The Photographic Sensitivity of Electronic Still Cameras", *J. of the Society of Photographic Science and Technology of Japan*, Vol. 5**9**, No. 1, p. 117-131, 1996.