Special Topic: Leading Edge of Digital Image Systems 2003   Exposition

X3 Sensor Characteristics

Allen RUSH[*] and Paul HUBEL[*]

Abstract The X3 sensor technology has been introduced as the first semiconductor image sensor technology to measure and report 3 distinct colors per pixel location. This technology provides a key solution to the challenge of capturing full color images using semiconductor sensor technology, usually CCD's, with increasing use of CMOS. X3 sensors capture light by using three detectors embedded in silicon. By selecting the depths of the detectors, three separate color bands can be captured and subsequently read and reported as color values from the sensor. This is accomplished by taking advantage of a natural characteristic of silicon, whereby light is absorbed in silicon at depths depending on the wavelength (color) of the light.

Key words: CMOS image sensor, Bayer pattern, photodiode, spatial sampling, color reconstruction, color error

1. X3 Sensor Operation

The X3 sensor is similar in design to other CMOS sensors – a photodiode is used to collect electrons as a result of light energy entering the aperture. The charge from the photodiode is converted and amplified by a source follower, and subsequently read out through a column readout circuit. There are three such detectors, corresponding to blue, green, and red bands of wavelengths.

A conventional Bayer pattern image sensor spatially samples only 25% of the red, 25% of the blue, and 50% of the green color planes. Since a completed image represents all color components (usually in RGB form, although other color spaces are possible), the image acquisition system attempts to calculate the missing color components from the spatially sampled values. There are many approaches to estimating missing colors. Some have been exhaustively studied for accuracy and analysis, such as methods presented by Ramanath et al[1]. Most demosaicing algorithms use a combination of filtering and neighboring color heuristics to determine a best approximation for the missing colors. More computationally intensive approaches incorporate adaptive, gradient, and other image-information-based data-driven algorithms. Errors that occur can be noticed; the extent of the error can be correlated by measuring the ΔE^*_{ab} of the resulting missing color calculation. In a test of five different algorithms and eight different test images, Ramanath reported a ΔE^*_{ab} range of 0 (no noticeable difference) to a worst case of over 65 (a ΔE^*_{ab} of 2–3 is usually considered to create a just noticeable difference, JND). Moreover, all of the approaches failed to adequately estimate images exhibiting high spatial frequencies.

Errors created by approximating missing color have various appearances, depending on the type of error. Moreover, some errors, such as moiré, are minimized by optically low-pass filtering the image, thereby removing the potential for aliasing. In this case, the resulting capture of the image has correspondingly less resolution than the theoretical Nyquist limit. Further attempts to restore the appearance of a sharper image are accomplished by sharpening filters, such as unsharp mask. This processing adds further errors to the result, although in many cases the sharpening has a beneficial perceptual effect. The net result is a series of cause – effect – remedy that adds considerable complexity and cost to the capture system, while the best quality remains unachievable.

By contrast, the X3 sensor reports three color values for each pixel location or spatial sampling site. This sampling approach eliminates the need for demosaicing, and the associated errors that result. Also, when all spatial samples are measured, the task of sharpening can be left to the output-dependent rendering stage (where it should be). The task remains, then, to find a best fit solution for mapping the raw color data to some rendered output color space, with minimum mapping errors.

2. Spectral Response

The spectral response of the X3 sensor is based on the characteristic absorption behavior of wavelengths of light in silicon. This characteristic is shown in Figure 1. Short wavelength light (blue) is more likely to be absorbed near the surface of the
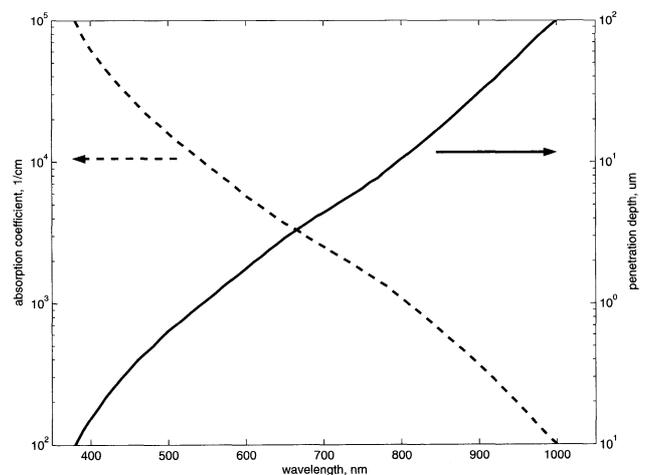

Fig. 1 Absorption coefficient and penetration depth in Silicon, vs. wavelength
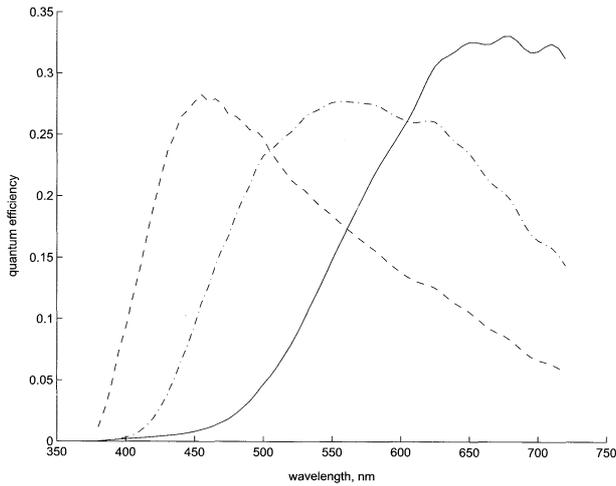
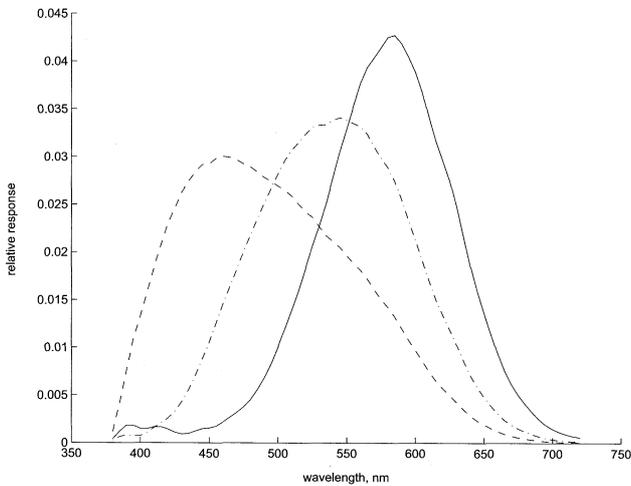Fig. 2   Wavelength vs. quantum efficiency without IR filter

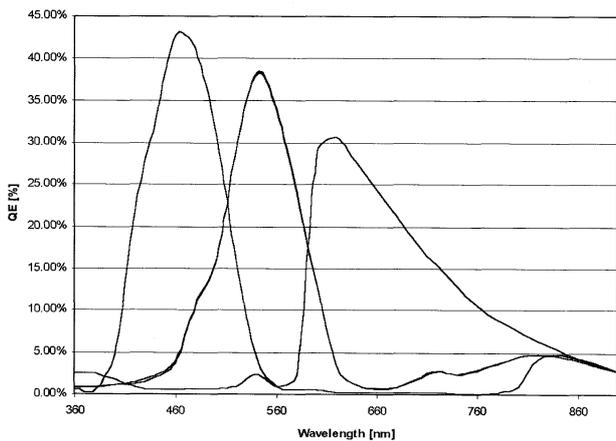Fig. 3   Wavelength vs. spectral sensitivity with a 2mm cm500 IR filter

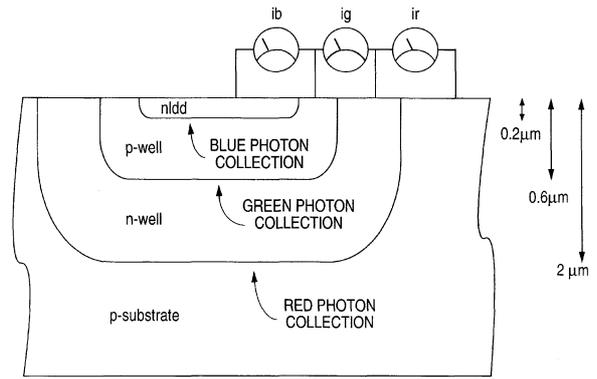Fig. 4   Typical CFA QE (Source: Baer, IEEE CCD and AIS Workshop, 1999)

Fig. 5   A schematic drawing of a sensor stack that captures all of the incident photons, filtering the color components by the wavelength-dependent absorption of silicon

ter array (CFA) type of sensor. In CFA's the filter material is selected to have a sharp cutoff in the transition region, and the bandwidths are selected reduce the overlap between red and blue to a very low level. Figure 2 shows the quantum efficiency of a typical X3 sensor (in this case, a 9um full color pixel). Figure 3 shows the spectral sensitivity response modified by an appropriate IR prefilter. Figure 4 shows the spectral response of a typical CFA type array. While the spectral response of each of the bands is more narrow than the corresponding X3 band, only one of the bands is recorded for each pixel location in a CFA array. Also, in Figure 4 the overall quantum efficiency per spatial sample of the sensor must be reduced due to the CFA pattern (2/3 of the incident light is absorbed by the filters and not converted into image signal).

The mechanism for achieving the spectral response for the three channels is a stack of detectors that corresponds in depth to the wavelength depth dependency shown in Figure 1. A schematic of this detector stack is shown in Figure 5. The detectors are very similar to photodiodes used in standard CMOS image sensors (CIS). The major difference is the vertical stacking approach that allows them to collect electrons at the specified depths.

Since the aperture is unobstructed, all of the light that enters the photodetector area is subject to conversion into measured color image data. The total quantum efficiency of the photodetector is the sum of each of the spectral response channels, and is broadband in nature – covering approximately 350nm to 800nm. By contrast, the quantum efficiency of a CFA photodetector is limited to the band that corresponds to the color filter covering the aperture.

3. Color Reconstruction

Image data is captured through a lens system and bandlimiting filters, usually IR cutoff. In addition, UV cutoff filters will help to establish an effective prefilter together with the IR filter. Raw data representing the output values of blue, green, and red photodiodes are combined to form tristimulus color values, in a selected color space. The color space can be one of any number of color spaces: sRGB, CIELAB, etc. For the X3 sensor, several functions need to be completed prior to color conversion. A linearization step and a dark frame subtraction create a data set that is linear in response to light input. The linear data can sub-

silicon, while long wavelength light is more likely to be absorbed ~3um deep in the silicon. This leads to the relatively broad spectral bandwidth of the detectors, and the larger amount of spectral overlap between the color channels, relative to a typical color fil-
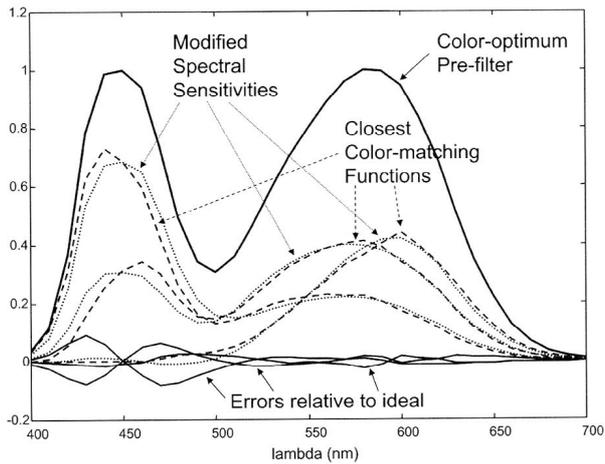
Fig. 6 Comparison of the closest linear combination of cone sensitivities with a combination of the spectral sensitivities of the Foveon X3 sensor and optimized pre-filtering

Fig. 7 Luminance linearity and noise for 6 gray patches of Macbeth chart

Table 1

Camera	Sensor	Metamerism index
HP715	Sony ICX-252	85
MegaVision S2	Dalsa FTF3020	88
Sigma SD9	Foveon X3–F7	91
HP618	Sony ICX-284	92

sequently be converted to another color space using standard 3×3 matrix algebra.

The Luther-Ives condition requires that the matrix solution for complete color accuracy is a nonsingular transformation of the CIE color matching functions [3]. A set of color matching functions can be generated by finding a matrix that maps the spectral response onto either a standard observer response (for example the CIE 1931 XYZ tristimulus functions) or an output rendering color space such as sRGB. The matrix can be found using techniques such as least square matrix inversion. Depending on the error constraint goals, the matrix may yield either a balanced error or biased error, favoring some metamers over others. A set of color matching functions for the sensor is shown in Figure 6, along with the associated errors for each of the three converted color coefficients. In this particular case, an optimal pre-filter was calculated and simulated into the optical path.

Even when the optimal pre-filter is not used, the color reproduction can be extremely accurate. Using IR and UV interference filters designed at the right cut-off wavelengths, we were able to achieve a metamerism index of 91 (using the draft standard technique-ISO 17321). Table 1 shows a comparison of several sensor technologies using the new metamerism technique. Unlike the previous technique (ISO draft standard ISO17321 WD4) as used in a previous publication [6], this index increases with more accurate color. Since perfect color reproduction is scaled to a value of 100, any sensor that achieves a metamerism index above 90 will be capable of providing outstanding color reproduction (although this will be also dependent on the rendering techniques employed).
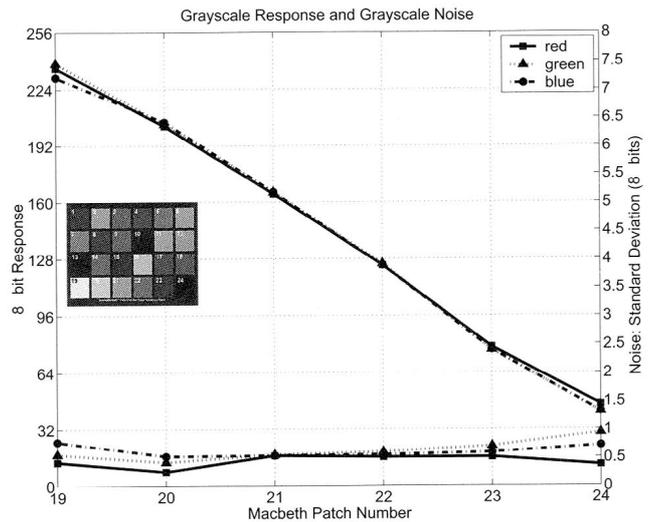
4. Experimental Results

Test images using the Sigma SD9 camera were taken and data was collected for both noise and spectral accuracy. The SD9 is the first commercial digital camera to incorporate the X3 technology. It uses a sensor with 9um pixel size, in a configuration of $2268 \times 1512 \times 3$, measuring 25mm on the diagonal.

The tests were performed using a standard studio setup, target, and lighting conditions. The illuminant was a studio flash, 5000K. The ISO setting was 100, the exposure was 1/125@f/8. The lens was a Sigma 28–70 zoom, at 48mm. White balance was set to sunlight.

As a simple measure for color fidelity, a Macbeth chart was used as a target, and images were measured in both RGB color space and $L^*a^*b^*$ color space. For each square of the Macbeth chart, a 32×32 pixel region was measured.

The matrix solution for converting raw image color data to rendered sRGB color contains some high gain coefficients off axis. In the test measurements, the matrix was

$$\begin{bmatrix} 0.36 & 0.11 & -0.003 \\ -0.55 & 1.50 & -1.08 \\ 0.39 & -1.43 & 2.43 \end{bmatrix}$$

The result of applying the matrix to the raw (linearized, dark subtracted) input data increases the chroma and luminance noise. The rendering software subsequently reduces this noise, and the completed image contains a combination of low visual noise and a high degree of color accuracy. The noise results for luminance (gray) are shown in Figure 7, which correspond to the 6 patches of gray on the Macbeth chart. The standard deviation shown averages less than 1 count (all values are based on 8 bit per color RGB color space). The maximum error occurs at about $\sigma = 1.4$. The curve also indicates a high degree of neutral chroma response (gray).

Similar measurements were taken for the color patches of the chart. In this case, the noise measurements also include a calculation for ΔE^*_{ab}. The values for both standard deviation and ΔE^*_{ab} compare favorably with other systems with similar pixel size and resolution.

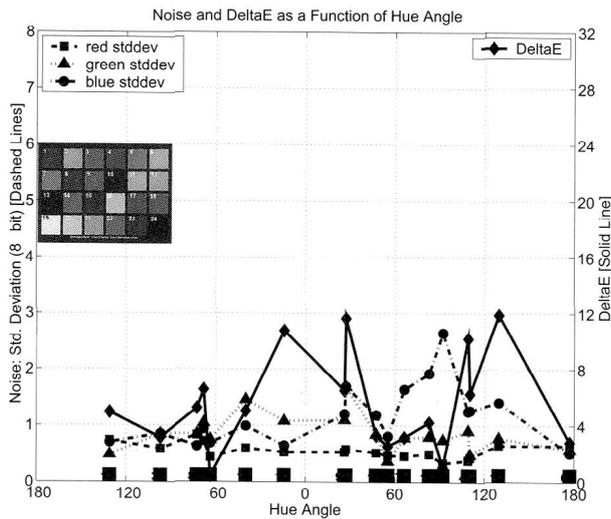Fig. 8   Chroma noise as a function of hue angle, with associated ΔE^{*}_{ab}

Fig. 9   Chroma noise plotted in $a^{*}b^{*}$ coordinates with target values indicated in colored squares

The color accuracy of the target is determined by comparing the measured result with target values for the illuminant (5000K). These are shown in Figure 8. The axes are a^{*} and b^{*}, and the colored squares indicate the target values in $a^{*}b^{*}$ color space for each of the measured squares. The distance between the circles (rendered result) and the colored square is an indication of the magnitude of the error. Again, this compares favorably with other sensors in both maximum error (about 16) and average ΔE^{*}_{ab}.

5. Conclusion

The X3 sensor technology has been developed to overcome several key limitations of CFA (Color Filter Arrays) technology. By measuring three distinct color values at each pixel location, an X3 sensor can provide much more accurate information about a scene without the need to interpolate color for full color restoration. The mechanism for achieving this is a stack of three photodetectors that capture wavelengths of light that correspond to the depth of absorption in the silicon. The spectral characteristics of the X3 sensor response are similar in breadth to the human cone sensitivities, and have been shown to be an extremely accurate linear transform of these fundamentals. The primary difference is seen in the relative breadth and spectral overlap of the X3 responses compared to those of the CFA. The predicted color accuracy is represented by a metamerism index. With the selected
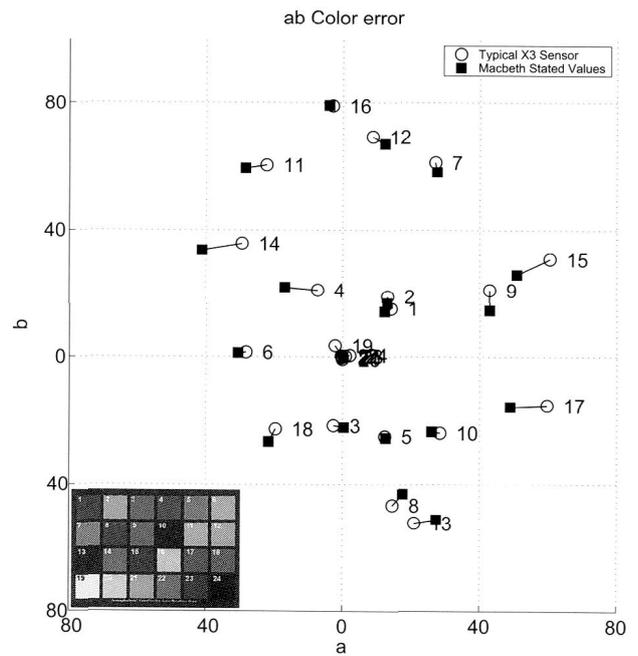
color transformation matrix, the metamerism index is among the best of the collection of sensors sampled. The experimental results (of rendered, sRGB test images), the noise and color accuracy compare very well with similarly configured sensor systems.

References

1) Ramanath *et al.*, "Demosaicing Methods for Color Filter Arrays", Journal for Electronic Imaging, July 2001.

2) Kemeny *et al.*, "Multiresolution Image Sensor", in IEEE Transactions on Circuits and Systems for Video Technology, Aug 1997, pp. 575-581.

3) Sharma *et al.*, "Color Imaging for Multimedia", Proceedings of the IEEE, June 1998.

4) P. L. Vora, H. J. Trussel, "Mathematical Methods for the Analysis of Color Scanning Filters", IEEE Transactions on Image Processing, Feb. 1997.

5) R. L. Baer, "CCD Requirements for Digital Photography", IS&T 2000 PICS Conference.

6) P. Hubel, R. Lyon, "Eyeing the Camera: into the Next Century".