

Methods to Reduce the Amplification of Random Noise in the Color Processing of Imager Data

Brent McCleary
Pictos Technologies
Newport Beach, California

Abstract

In the image processing of a digital camera system, a colorimetric matching transformation is implemented to minimize the error between the sensor output transformed color responses and the human vision responses. The color matching transform does not usually attempt to minimize the resulting noise of the color corrected data. The CMOS sensors used in embedded applications typically exhibit high levels of random noise and cross-talk (the loss of photons or electrons from a pixel to neighboring pixels). This results in a trade-off between color error (e.g., de-saturated colors) and the amount of noise in the image. This is especially true as the size of pixels are progressively reduced, which leads to highly de-saturated (muted) images that require severe color corrections and have lower signal-to-noise ratios. In this paper, methods of color correction are presented which exploit the fact that the human visual system is most sensitive to color errors in lower spatial frequencies than in higher spatial frequencies. Thus more accurate (and usually more complex) color correction algorithms should be applied to the lower spatial frequency data, and less accurate (and usually simpler) algorithms can be applied to the higher spatial frequency data. Three methods are developed and analyzed which use segmentation of sensor data in the spatial frequency domain: the discrete cosine transformation (DCT), the discrete wavelet transformation (DWT), and a simple low pass/high pass filter. Metrics for noise performance, color accuracy, and image sharpness are provided. The novel approach developed in this paper is the use of variable complexity color corrections applied within the frequency transform domains. This allows the process of compressing image data to be used to reduce the amplification of uncorrelated pixel noise, while still achieving accurate color enhancement and maintaining image sharpness.

Background

Color image sensors use a layer of color filters bonded to the silicon in order to sample ranges of wavelengths of light in the visible spectrum (380nm to 740nm). Typically, camera systems use three-color primaries to color-match colors of arbitrary visible wavelengths. In order to achieve low cost and compact designs, many camera systems use a

single sensor with a Bayer¹ color filter array (CFA). The Bayer pattern (Fig. 1) is the most popular CFA, and uses a checkerboard pattern with alternating rows of filters that subsamples and mosaics the colors of the captured image. Since the human vision system is most sensitive to green luminance in defining image quality, the Bayer pattern has twice as many green pixels as red and blue pixels. In order to form a color image, an interpolation or de-mosaicing algorithm must be used to produce an RGB triplet at each pixel.

Blue	Green Even	Blue	Green Even
Green Odd	Red	Green Odd	Red
Blue	Green Even	Blue	Green Even
Green Odd	Red	Green Odd	Red

Figure 1. Typical Bayer Filter Pattern

A color space matrix transform (such as a 3x3 matrix color correction) multiplies pixel signal data from separate pixels together and performs a linear combination of all the components. The red, green, and blue data come from separate pixels (and are interpolated to form RGB triplets), which have uncorrelated random noise. Thus, the noise data from the pixels are also multiplied and summed. The resulting random noise is increased due to the uncorrelated noise summing up according to a square root law.² The colorimetric matching transformation typically attempts to minimize the error between the sensor output transformed color responses and the human vision responses. No attempt is usually made to minimize the resulting noise of the color corrected data.

The CMOS sensors used in embedded applications suffer from many types of noise while operating under room temperature conditions. It is usually possible for the image back end processor to remove most of the fixed-pattern noise (time-invariant noise). The temporal noise

that CMOS sensors encounter include³: capacitive reset (kT/C) noise, dark current time-varying noise, Johnson (thermal or white) noise, and 1/f noise (frequency-dependent). For the cost-effective CMOS active pixel architecture, the kT/C noise is usually the largest source of temporal noise. This noise is generally uncorrelated from pixel to pixel, thus if pixels values are summed, then the variances will also sum (pixels values are independent normal random variables). The reset field-effect transistor switch generates the kT/C noise, where k is Boltzmann's constant, T is the temperature, and C is the capacitance of the attached load. As the size of the pixel photodiode is reduced due to the demands of embedded applications, the capacitance also decreases. This increases the kT/C noise, as well as decreasing the full-well signal level. Additionally, the signal itself has photon shot noise, in which its variance equals its mean. Smaller pixels have a smaller maximum signal to shot noise ratio. These effects reduce the sensor's dynamic range and signal-to-noise ratio (SNR). Thus, the level of noise in the imager data is a parameter that must be addressed during the image processing and compression.

The color correction of CMOS sensor images is of particular importance since these sensors typically exhibit high degrees of cross-talk. Cross-talk is the loss of electrons or photons from a pixel to neighboring pixels. There are both optical and electrical sources of cross-talk. Photons can pass through the CFA of one pixel at an oblique angle such that it is collected by an adjacent pixel, which has a different color filter. After photons have created carriers in a photodiode, they may diffuse to the depletion region of an adjacent pixel. Photons of longer wavelength (e.g., red light) penetrate deeper into the photodiode. The carriers created by these photons are more likely to diffuse to neighboring pixels. Therefore, there will be a larger degree of cross-talk from red pixels to green pixels than other color combinations. The amount of cross-talk will increase as the size of pixels is progressively reduced (e.g., less than 4 μm x 4 μm). This can lead to highly de-saturated (muted) images that require severe color corrections. Cross-talk will also reduce the sharpness of an image. In Ref. [4] a color correction method was developed which takes into account the color matching error and the noise error in processed images. The color error and the noise error levels can be balanced to produce an optimal image. This trade-off can be extremely important in digital camera systems where the sensor/system SNR levels can be very low (e.g., less than 20). However, the color quality will suffer when the SNR levels become low.

The color correction of pixel data is commonly done by using a 3x3 color correction matrix ([CC]) multiplication with the pixel color data. Higher order polynomial equations can produce lower errors in color-matching than those achievable in 1st order equations (e.g. 3x3 matrix), but add more computational complexity. A 'color matching optimized' 3x3 color correction matrix [CC] can be determined using a target with known

calibrated colors under known illumination and applying polynomial regression. Polynomial regression is based on the assumption that the correlation between color spaces can be approximated by a set of simultaneous equations (which is usually a good approximation). Sample points in the source color space are selected and their color specifications in the destination space are measured. A 3x3 color correction matrix which links the source and destination color specifications can then be defined.⁵ The 3x3 color correction matrix [CC] can be broken down into a 3x3 saturation matrix [Sat] and a 3x3 white balance matrix [WB] that only has terms on its main diagonal:

$$\begin{vmatrix} \text{Sat11} & \text{Sat12} & \text{Sat13} \\ \text{Sat21} & \text{Sat22} & \text{Sat23} \\ \text{Sat31} & \text{Sat32} & \text{Sat33} \end{vmatrix} \begin{vmatrix} \text{WB1} & 0 & 0 \\ 0 & \text{WB2} & 0 \\ 0 & 0 & \text{WB3} \end{vmatrix} = \begin{vmatrix} \text{CC11} & \text{CC12} & \text{CC13} \\ \text{CC21} & \text{CC22} & \text{CC23} \\ \text{CC31} & \text{CC32} & \text{CC33} \end{vmatrix} \quad (1)$$

In order to maintain the white balance of an image, the saturation matrix [Sat] must have the coefficients of each of its rows sum to one. Thus, as the magnitude of the off-diagonal saturation terms increase, the on-diagonal terms of [Sat] must also increase. This means that for sensors that have more cross-talk, the [Sat] coefficients will have larger values, and the SNR of the processed images will decrease.

Noise Reduction Methods

The first pixel noise reduction method described attempts to take advantage of the DCT operation that is performed in camera systems which store images in JPEG format. If much of the image processing is done in firmware, then we have the opportunity to modify the processing to incorporate minor changes that may result in minimal increases in computational time. In fact decreases in processing time and computational complexity can occur due to a reduction in the number of mathematical operations performed. Assume that the image process chain can be represented by the simplified set of operations shown below:

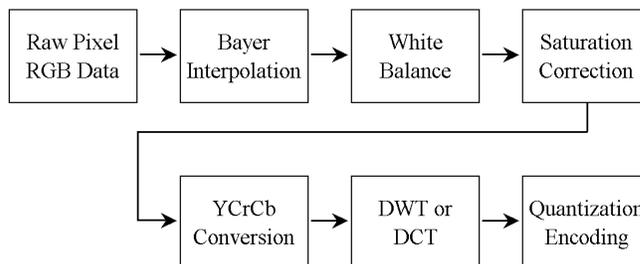


Figure 2. Simplified Image Processing Operation Path

The DCT and color correction transforms are both linear, thus permutations between these operations can be

performed.⁶ Then we can modify the image process path to become:

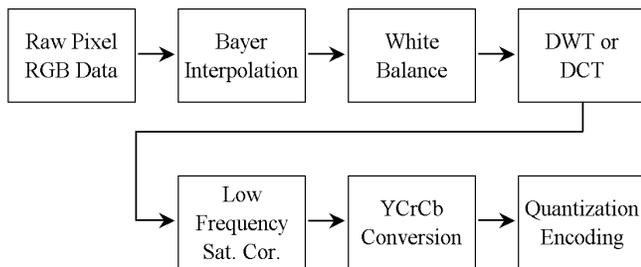


Figure 3. Modified Image Processing Operation Path

When the DCT operation is performed, 8 by 8 pixel blocks are formed. The 64 terms in the blocks represent DC to high frequency terms. We can then apply the full 3x3 color correction in the DCT domain to the low frequency terms only ([Low Frequency Saturation Correction]). All of the pixels prior to the DCT have had a white balance applied to them. The color space conversion [YCrCb Conversion] is then performed after the saturation operation, with the standard processing following. This methodology was implemented by applying the full 3x3 color correction (white balance followed by saturation correction) to three sets of coefficients within each 8x8 block: the lowest frequency coefficient (DC term) only, the four lowest coefficients, and the 16 lowest coefficients. This corresponds to performing a full color correction to 1/64, 1/16, and 1/4 of the image data, respectively. Furthermore, the selection of the coefficients that receive the full 3x3 color correction can be implemented real-time based upon the SNR of the captured image. One drawback of this method is that 8x8 hue blocking artifacts can occur. This is due to the low frequency color values of an 8x8 pixel block being used for all of the pixels within that block. A filter that slides across the image does not have this problem. However, methods have been developed which can remove this blocking artifact.⁷

The next method considered is the DWT color processing method. It is similar to the DCT method in that most of the color correction process is moved from the 'image processing space' to the DWT space (refer to Fig. 4). Wavelet transform methods exploit redundancies in scale to reduce information stored in the wavelet transform domain.⁸ The DWT is used in JPEG2000, so future camera systems may have the DWT in the image process path. When the DWT step is already a part of the image processing, little or no additional cost is incurred. As with the DCT method, decreases in processing time and computational complexity can occur since fewer mathematical operations are performed. The overall complexity of the algorithm (processing cost) can be controlled by selecting the number of subbands (resolutions) that have the higher complexity color corrections (e.g., 3x3 color correction matrix multiplication) applied to them. Refer to Fig. 4 to see how

the image is partitioned in coefficient frequency subbands. Other partitioning schemes may be used. A quarter of this 2-D transform has its coefficients from a high-pass filter (H) operating on the image's rows followed by a high-pass filter operating on the image's columns. This block of coefficients is indicated by HH1 in the lower right-hand corner of Fig. 4. Another quarter of the 2-D transform has its coefficients from a low-pass filter (L) operating on the image's rows followed by a high-pass filter operating on the image's columns, and is indicated by LH1 in the upper right-hand corner of Fig. 4. The upper left-hand corner of the 2-D transform is subdivided into smaller blocks. The wavelet transform is performed on successively smaller number of coefficients until the final coefficient in the upper left-hand corner has only low-pass operations performed on it. The higher frequency resolutions will only be white balance processed, which will result in reduced noise compared to the standard 3x3 color correction method. The wavelet and color correction transforms are both linear, thus we can use permutations between these operations.

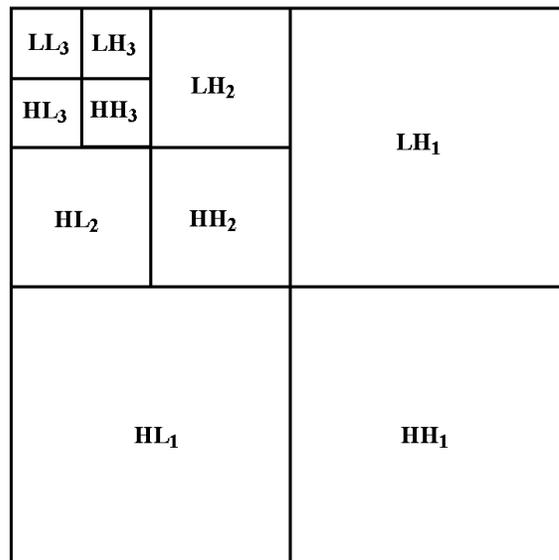


Figure 4. 2-D Wavelet Transform

A study in Ref. [9] used human psycho-visual appreciation to weigh the importance of the DWT coefficient blocks. In Ref. [9], the importance of each wavelet block in terms of the overall reconstructed image quality was ranked. Based upon that study, the selection of the resolution planes that have their color correction matrices simplified (e.g., reduced to a white balance only) as the image SNR decreases is ordered as follows:

$$HH_1, LH_1, HL_1, HH_2, LH_2, HL_2, HH_3, LH_3, HL_3, LL_3$$

In the DWT processing, I recursively applied a 9 tap low pass filter and a 7 tap high pass filter (9-7 Daubechies/Antonini filters).¹⁰ Three sets of subbands that

received the full 3x3 saturation correction after white balancing were implemented: the lowest frequency subband (LL₃) only, the lowest four frequency subbands (LL₃, LH₃, HL₃, HH₃), and the lowest seven frequency subbands (LL₃, LH₃, HL₃, HH₃, HH₂, LH₂, and HL₂). As with the DCT corrections used, this corresponds to performing a full color correction to 1/64, 1/16, and 1/4 of the image data, respectively. Omitting the full 3x3 color correction of the high frequency components can result in too large a loss of edge detail. In such cases, the white balanced high frequency components can be amplified (without amplifying and summing uncorrelated noise) to give better edge detail.

The last method explored uses a simple low-pass/high-pass sliding filter (LPF/HPF). It is in fact a simplification of the DWT method. The low pass data is obtained by summing the RGB values of the local area of color-interpolated pixels and obtaining the average value. By selecting an area which has a number of pixels that is a factor of 2, it can be implemented without the use of any multiplications; simply sum and shift or shift during accumulation. Line buffers would be needed to store the image data for filtering. The high pass data is then simply the difference between the pixel data and the low pass value. Thus, in this process the pixel data is first color interpolated and then white balanced. The pixels then have the saturation correction applied to the low-pass filtered data and then the low and high frequency components are summed. The high frequency components can be amplified (without amplifying uncorrelated noise) if better edge detail is desired. The simple equations for these operations are shown below:

$$RGB' = Factor*[WB][RGB-RGBave] + [Sat][WB][RGBave], \text{ or}$$

$$RGB' = Factor*[WB][RGB] + ([Sat] - Factor*[I])[WB][RGBave], \quad (2)$$

where:

$$RGBave_{ij} = Average(RGB_{k-l,j}) \quad k,l = -(n-1), \dots, 0, \dots, n \quad (3)$$

and 'Factor' is a scalar term used to amplify the high frequency information, 'RGB' is the red, green or blue pixel data, and [I] is the identity matrix. In order to match the scale of the DWT and DCT methods implemented, the LPF/HPF method was performed using local pixel area window sizes of 8x8, 4x4, and 2x2 pixels. This corresponds to applying a full color correction to 1/64, 1/16, and 1/4 of the image data, respectively.

Performance Criteria

The performance of each method to reduce noise in the color processing was calculated and compared with that of standard processing. Specifically, the color corrected accuracy, signal-to-noise ratio, and sharpness of processed images were compared. The test image used for the performance measurements included a MacBeth Color Checker chart and an ISO Resolution Chart for Electronic

Still Cameras. The image was taken using an f/2.8 lens with 100 lux lighting at an exposure time of 48ms, which corresponds to 5.7 EV (photometric exposure value). The sensor used was a CMOS imager with 4µm x 4µm pixels. This test set-up provides a good subject to observe the color correction-image sharpness-noise trade offs.

The image of a MacBeth color checker chart was used for the color error and SNR calculations. All 24 color squares were used for color matching error calculations. The total error in color matching was measured by using the equation:

$$Color\ Error = \sum_{i=1}^{24} [\sqrt{Red_error(i)^2 + Green_error(i)^2 + Blue_error(i)^2}] \quad (4)$$

where:

$$Red_error(i) = Weighting_Factor(i) * (Red\ MacBeth\ Square\ Ideal\ Value - Imager\ Red\ Corrected\ Value) \quad (5)$$

Green_error(i) and Blue_error(i) were calculated in the same manner and all the weighting factors were set to 1 for this study. The color error measured was then used to calculate the normalized color accuracy. The colors errors from applying a simple white balance matrix (WB Color Error) and the optimized 3x3 color correction matrix (CC Color Error) are used in the following equation to produce the normalized color accuracy:

$$Normalized\ Color\ Accuracy = (WB\ Color\ Error - Measured\ Color\ Error) / (WB\ Color\ Error - CC\ Color\ Error) \quad (6)$$

A normalized color accuracy of 100% corresponds to the best 3x3 color correction possible, and a value of 0% corresponds to a simple white balance alone.

The degree of noise in the processed images was measured by looking at the .44 optical density gray square of the MacBeth chart (square number 22). The values of the red, green, and blue pixels were used to calculate the luminance values per pixel. The ratio of the mean of the pixel luminance values to the standard deviation of the pixel luminance values was used for the SNR parameter.

The effects of the noise reduction color processing methods on the sharpness of an image were measured by calculating the modulation transfer function (MTF) when each method was used. The MTF was found at 1/2 Nyquist using the equation from Ref. [11]:

$$MTF = (S_{MAX} - S_{MIN}) / (S_{MAX} + S_{MIN}) \quad (7)$$

where S_{MAX} and S_{MIN} are the maximum and minimum modulation levels after processing, respectively. Clipping of the data at the high and low ends was avoided.

The processing of the images was minimized in order to isolate the effects of the color processing algorithms. Only DC black level offset, simplified color filter pattern interpolation (Bayer), and the specialized color corrections were performed.

Discussion of Results

The normalized color accuracy (Eq. (6)), SNR, and MTF values at $\frac{1}{2}$ Nyquist frequency (Eq. (7)) are shown in the plots of Figures 5, 6, and 7 for the DWT, DCT, and LPF/HPF methods. The standard color processing method of applying the color correction in the spatial domain is also shown for comparison. Three data points are used for the standard processing method: full optimal 3x3 color correction [CC] matrix, 50% saturation correction (the [Sat] matrix's off-diagonal terms are reduced by 50%), and white balance [WB] correction only. The DWT, DCT, and LPF/HPF methods each have data points for 3 color correction resolutions. These resolutions correspond to applying a full color correction to 1/64, 1/16, and 1/4 of the image data for each of the proposed methods. The plots of SNR and MTF versus normalized color accuracy (Figs. 5 and 6) show that the proposed color correction methods produce color accuracy that are as good as an optimal 3x3 standard correction and are fairly independent of SNR and MTF. In contrast, the standard method of color correction results in SNR decreasing with color accuracy and MTF increasing with color accuracy. Color accuracy of the MacBeth color checker measurements depend only on the low frequency sensor data, thus the three non-standard color correction methods are designed to produce good results for this test. This also matches the way humans perceive color. As more higher frequency data is used in the calculations, the SNR will decrease and the MTF will increase. The SNR decreases due to amplification of pixel uncorrelated noise. The MTF increases due to correction of signal lost to adjacent pixels through cross-talk. The plot of SNR versus MTF in Fig. 7 shows that all of the color correction methods follow approximately the same trade-off between SNR and MTF.

The three proposed color correction methods offer the ability to select the optimal SNR-MTF trade-off while always providing excellent low frequency color accuracy. The standard color correction method cannot achieve this since it operates in the spatial domain. As the signal decreases or as the cross-talk increases, the selection of the optimal SNR-MTF trade-off becomes more critical in determining the overall image quality.

Samples from a low light (low SNR) still life image are provided in Figs. 8 and 9 to show the performance of the color corrections. All three methods can result in some 'softening' of the image due to a reduction in the amount of cross-talk correction applied. However, the hue random noise is attenuated. The strong high frequency edges are also maintained. We can also amplify the high pass portion of the data in order to improve our detail information. It is seen that the DCT method introduces some undesirable 8x8 hue blockiness, which could be eliminated with further processing. Another benefit of this color correction methodology is the reduction of color filter pattern interpolation aliasing. In areas that exhibit a sharp transition from light to dark, a zippering effect of alternating hue can occur along the edge. Simplistic CFA

interpolation algorithms applied to the high frequency data can produce this effect. When the standard 3x3 color enhancement is then applied to the color interpolated data, this effect is amplified. Applying the majority of the color enhancement to the low frequency image data terms only and applying a white balance to the high frequency image data terms attenuate this effect.

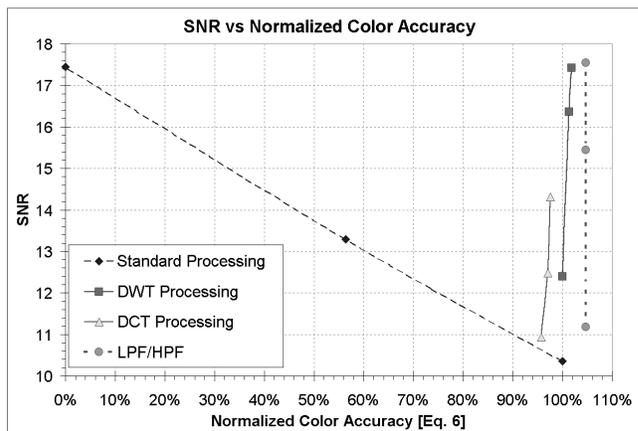


Figure 5. SNR versus Normalized Color Accuracy

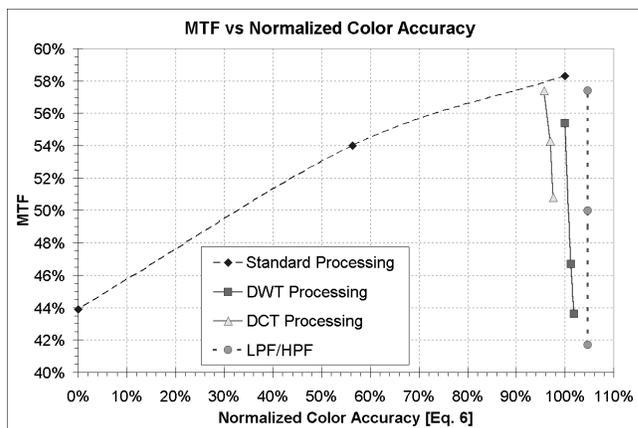


Figure 6. MTF versus Normalized Color Accuracy

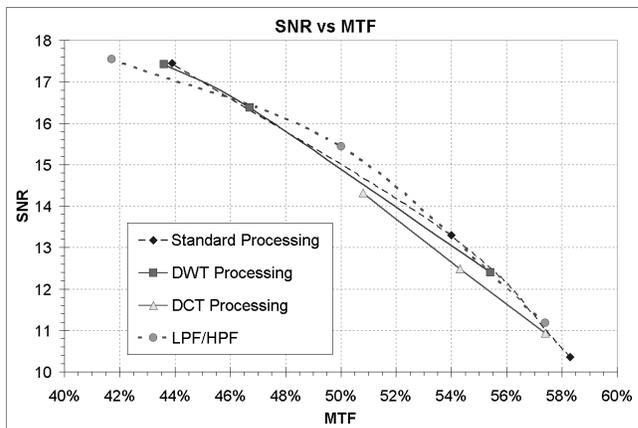


Figure 7. SNR versus MTF



Figure 8. Flower Image Detail. Top: Standard 3x3 [Sat] correction, Bottom: DWT with 3x3 [Sat] correction applied to the lowest four frequency subbands only (LL3, LH3, HL3, HH3).



Figure 9. Flower Image Detail. Top: DCT with 3x3 [Sat] correction applied the 16 lowest frequency coefficients, Bottom: LPF/HPF with 3x3 [Sat] correction applied to 4x4 averaged data.

Conclusions

Three separate methods to reduce the amplification of random noise in the color processing of imager data have been described. These three methods of color correction exploit the fact that the human visual system is most sensitive to color errors in lower spatial frequencies than in higher spatial frequencies. The three color correction methods utilized transforms to process colors in the frequency domain, namely: the discrete cosine transformation (DCT), the discrete wavelet transformation (DWT), and a simple low pass/high pass filter. A summary of the results are provided in Figs. 5, 6, and 7, with a set of sample processed images shown in Figs. 8 and 9. These results showed that the three methods of color correction produce low frequency color accuracy that are as good as applying an optimal 3x3 color correction matrix. The proposed color correction methods allow the control of the SNR-MTF trade-off.

References

1. B.E. Bayer, "Color image array", US Patent 3971065, 1976
2. E. Hill, "Color Capture, Color Management and the Problem of Metamerism: Does Multispectral Imaging Offset the Solution?", IS&T/SPIE, vol. 3963, Jan. 2000, pgs. 2-14
3. T. Zarnowski, T. Vogelsong, and J. Zarnowski, "Inexpensive Image Sensors Challenge CCD Supremacy", Photonics Spectra, May 2000, pgs. 188-190
4. Y. P. Tan and T. Acharya, "A Method for Color Correction with Noise Consideration", IS&T/SPIE, vol. 3963, Jan. 2000, pgs. 329-337
5. H. R. Kang, "Color Technology for Electronic Imaging Devices", SPIE Optical Engineering Press, vol. PM28, Jan. 1997, pgs. 55-63
6. R. Balasubramanian, R. V. Klassen, and R. de Queiroz "Color Correction of JPEG Compressed Images", Proc. IS&T/SID's 5th Color Imaging Conf., Scottsdale, AZ, Nov. 1997, pgs. 83-87
7. J. Chou, M. Crouse, and M. K. Ramchandran, "A Simple Algorithm for Removing Blocking Artifacts in Block-Transform Coded Images", IEEE Signal Processing Letters, vol: 5 Issue: 2, Feb. 1998, pgs. 33-35
8. S. Welstead, "Fractal and Wavelet Image Compression Techniques", SPIE Press, vol. TT40, Nov. 1999, pgs. 131-153
9. F. Truchetet, B. Joanne, F. Perot, and O. Laligant, "High-Quality Still Color Image Compression", Optical Engineering/SPIE, Vol. 39(02), Feb. 2000, pgs.409-414
10. M. Antonini, M. Barlaud, P. Mathieu, L. Daubechies, "Image Coding Using Wavelet Transform", IEEE Transactions on Image Processing, vol: 1 Issue: 2, Apr 1992, pgs. 205 -220
11. J. Janesick, "Scientific Charge-Coupled Devices", SPIE Press, vol. PM83, Jan. 2001, pgs. 338-370

Biography

Brent McCleary is a Ph.D. candidate in the Department of Electrical Engineering at the University of Southern California. He received his Masters of Science degree in Computer Engineering from USC in 1995, and Masters of Science degree in Aerospace Engineering from USC in 1991. He is a CMOS sensor systems engineer with the Conexant Systems spin-off Pictos Technologies. His research interests include image processing methods for CMOS sensors with small pixels. He is a member of SPIE and IEEE.