Spatially Varying Color Correction (SVCC) Matrices for Reduced Noise

SukHwan Lim and Amnon Silverstein
Hewlett-Packard Laboratories
Palo Alto, California, USA

Abstract

Color space transformation (or color correction) needs to be performed in typical imaging devices because the spectral sensitivity functions of the sensors deviate from the ideal. Several researchers have shown that when the color channels are correlated, color correction can result in sensor noise amplification.[1-4] In this paper, we describe a color correction method that significantly alleviates the problem of noise amplification. The key idea is to use spatially varying color correction (SVCC) that adapts to local image statistics. We show experimental results that illustrate the reduction of noise when color correction is performed.

I. Introduction

The spectral sensitivity functions (or spectral responsivity) of the 3 or more color channels in digital imaging devices do not match those of the desired output color space (e. g. CIE-XYZ, sRGB, NTSC). Thus, it is necessary to transform the raw color images into the desired color space, which is usually performed using a linear transformation matrix. For sensors with R, G and B color channels, color correction is typically performed by multiplying a 3x3 matrix with the vector formed by the R, G and B values at each pixel. i. e.,

$$\begin{bmatrix} R_{out} \\ G_{out} \\ B_{out} \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} R_{in} \\ G_{in} \\ B_{in} \end{bmatrix}$$

The main differences among the linear transformation methods are the constraints they use to derive the color correction matrix. One method is to obtain the color correction matrix by solving the least-squares problem that minimizes the sum-of-squared-difference between the ideal and color-corrected spectral sensitivity function. Although this method minimizes the color error in the color-corrected R, G and B values, the 3x3 multiplication may amplify the image sensor noise. This becomes a major concern when the spectral sensitivity functions of the image sensor have high correlation between them. For example, Barhoeffer et. al.[1] have shown that some sensors with cyan, magenta, yellow, green (CMYG) filter set suffer from this noise amplification.

Several authors have investigated the color estimation error trade-offs.[1-4] Barnhoeffer et. al.[1] explored the trade-off between mean color deviation and the amplification of noise. The trade-off was described mathematically and a new methodology for choosing an appropriate transformation was proposed. Vora et. al.[2] showed that the noise amplification is related to the degree of orthogonality of the filters and noise reduction comes at the cost of color saturation. In these approaches, the trade-off is performed by choosing the optimum color correction matrix for the entire image. We argue that by loosening the constraint of having a fixed color correction matrix for the entire image, a better trade-off can be obtained. In this paper, we describe a spatially varying color correction (SVCC) method that achieves a better trade-off between color fidelity and image sensor noise amplification. The method first estimates the 2nd order statistics of local image regions and computes the optimum color correction matrix for each local image region. Note that this color correction method is optimum in a mean-squared-error sense.

The organization of this paper is as follows. Section II describes how the optimum color correction is obtained from the 2nd order local image statistics and shows how it may be implemented in an imaging system. Section III shows some experimental results that illustrate the improvement from using the proposed method.

II. Spatially Varying Color Correction Method

In this section, we describe the new color correction (transformation) method that alleviates noise amplification. In Subsection II-A, we first give a derivation on how to obtain a color correction matrix for each local image region assuming that 2nd order local image and noise statistics are known. In Subsections II-B and II-C, we then describe how to practically implement this and we provide possible extensions to the baseline approach.

A. Derivation of the Method

In this subsection, we describe how each color correction matrix is computed assuming the local image statistics have already been estimated. Assume that we have the color correction matrix $C_{NOMINAL}$ that minimizes color error (but does not take sensor noise into consideration). This color correction may have large off-diagonal elements and suffer from severe noise amplification. We describe how to vary this matrix from image region to image region in order to solve the problem of noise amplification with minimum sacrifice of color fidelity.

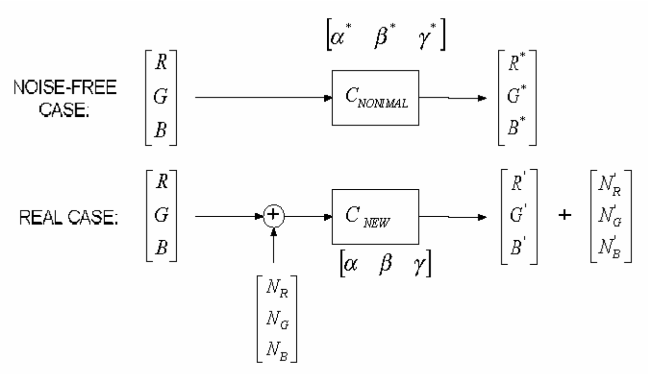

Figure 1. Model used for the derivation

The model we use for our derivation is shown in Figure 1. An ideal, noise-free case would be when there is no noise and we use $C_{NOMINAL}$ to perform color correction. Since there is no noise, $C_{NOMINAL}$ would still be optimum. However, when noise is present $C_{NOMINAL}$ may amplify noise and be sub-optimum in mean-squared-error sense. In the real case, we compute spatially varying C_{NEW} to alleviate noise amplification. The definitions of the symbols are given as follows.

- R, G and B: Noise-free red, green and blue values before color correction is applied.
- R^*, G^* and B^*: Noise-free red, green and blue values after color correction ($C_{NOMINAL}$) is applied.
- R', G' and B': Noise-free red, green and blue values after color correction (C_{NEW}) is applied.
- N_R, N_G and N_B: Noise in red, green and blue channels before color correction is applied.
- N'_R, N'_G and N'_B: Noise in red, green and blue channels after color correction (C_{NEW}) is applied.
- α^*, β^* and γ^*: Weights for the green channel in the nominal matrix. (i.e., the second row of $C_{NOMINAL}$).
- α, β and γ: Weights for the green channel in the new matrix. (i.e., the second row of C_{NEW}).

Using the symbols defined, Figure 1 can be summarized as follows.

$$\begin{bmatrix}R'\\G'\\B'\end{bmatrix} = C_{NEW}\begin{bmatrix}R\\G\\B\end{bmatrix}, \begin{bmatrix}R^*\\G^*\\B^*\end{bmatrix} = C_{NOMINAL}\begin{bmatrix}R\\G\\B\end{bmatrix} \text{ and } \begin{bmatrix}N'_R\\N'_G\\N'_B\end{bmatrix} = C_{NEW}\begin{bmatrix}N_R\\N_G\\N_B\end{bmatrix}.$$

The objective is to obtain C_{NEW} (or α, β and γ) that minimizes expected sum of color error and amplified noise. In other words, the objective is to estimate C_{NEW} that minimizes the expected difference between the outputs of the "Noise-free case" and the "Real case" as illustrated in Figure 1. Consider color correction coefficients $[\alpha\ \beta\ \gamma]^T$ for the G channel, which correspond to the second row of color correction matrix. Other channels can be derived similarly. We wish to minimize f, the expected value of the sum of color error and output noise.

$$f = E[(G' - G^* + N'_G)^2], \tag{1}$$

where $E[]$ is the expected value. Optionally, one can weigh NG differently than shown in Equation (1) where the weight was equal to 1. Higher weight on N_G would put more emphasis on the noise amplification while sacrificing color fidelity. In our derivation, the weight is set to 1 for simplicity. Since $G^* = \alpha^* R + \beta^* G + \gamma^* B$, $G' = \alpha R + \beta G + \gamma B$ and $N'_G = \alpha N_R + \beta N_G + \gamma N_B$, Equation (1) can be re-written as

$$f = E[((\alpha-\alpha^*)R + (\beta-\beta^*)G + (\gamma-\gamma^*)B + \alpha N_R + \beta N_G + \gamma N_B)^2] \tag{2}$$

Equation (2) can be simplified by assuming that N_R, N_G, N_B, R, G and B are uncorrelated. Further assuming that N_R, N_G and N_B have zero means and standard deviations of σ_R, σ_G and σ_B, we obtain

$$\begin{aligned} f &= (\alpha-\alpha^*)^2 E[R^2] + (\beta-\beta^*)^2 E[G^2] + (\gamma-\gamma^*)^2 E[B^2] \\ &+ 2(\alpha-\alpha^*)(\beta-\beta^*)E[RG] + 2(\beta-\beta^*)(\gamma-\gamma^*)E[GB] \\ &+ 2(\gamma-\gamma^*)(\alpha-\alpha^*)E[RB] + \alpha^2\sigma_R^2 + \beta^2\sigma_G^2 + \gamma^2\sigma_B^2 \end{aligned} \tag{3}$$

To minimize f in Equation (3), we take partial derivatives of f with respect to α, β and γ, and set them to be zero. We then obtain three equations that can be summarized in matrix form as

$$Cor\begin{bmatrix}\alpha\\\beta\\\gamma\end{bmatrix} = (Cor - CorNN)\begin{bmatrix}\alpha^*\\\beta^*\\\gamma^*\end{bmatrix} \tag{4}$$

where Cor is the correlation matrix of $[R + N_R \quad G + N_G \quad B + N_B]^T$ and $CorNN$ is the correlation matrix of $[N_R\ N_G\ N_B]^T$. Note that $[R + N_R \quad G + N_G \quad B + N_B]^T$ are the pixel values that we can measure while $[R\ G\ B]^T$ are the noise-free pixel values that we do not have access to. From Equation (4), the α, β and γ that minimizes the sum of color error and output sensor noise can be simplified as

$$\begin{bmatrix}\alpha\\ \beta\\ \gamma\end{bmatrix} = (Cor)^{-1}(Cor - CorNN)\begin{bmatrix}\alpha^*\\ \beta^*\\ \gamma^*\end{bmatrix}$$

Similar derivation can be applied to R and B channels and by combining them, we obtain

$$C_{NEW} = C_{NOMINAL}(Cor - CorNN)^T(Cor^{-1})^T \qquad (5)$$

Equation (5) shows how to vary the color correction (or transformation) matrix based on the correlation matrix of the pixel intensity values (with noise) and the variances of the noise. The correlation matrix of the pixel intensity values can be estimated by computing average values of $[R + N_R \quad G + N_G \quad B + N_B]^T[R + N_R \quad G + N_G \quad B + N_B]$. Although we assumed that the noise values and the pixel intensity values are uncorrelated in our derivation, the variances of the noise do depend on the intensity. This is because the image sensor noise is the sum of the shot noise and readout noise and the variance of shot noise for each channel depends on the pixel intensity values.

The simplest way to use Equation (5) would be to apply it to the whole image (i. e. estimate *Cor* of the entire image and apply C_{NEW}). This would result in a color correction matrix similar to LMMSE solution described in References 3 and 4. However, the real merit of using Equation (5) can be seen when different values for C_{NEW} are applied to smaller set of pixels. Since the 2^{nd} order image statistics (i.e. correlations) are not stationary throughout the image and vary from one local image region to another, it is advantageous to apply different color correction matrices to different local image regions. To maximally benefit from having different color correction matrices, the size of the local image regions should be small enough such that the pixel values within the local region have similar 2^{nd} order image statistics but large enough for accurate estimation of the correlations. Equation (5) provides a way to adapt the color correction matrices to alleviate noise amplification problem given a set of pixels in a local image region.

B. Baseline Implementation

The block diagram of the method is shown in Figure 2. The first block is optional since the variance of the image sensor noise can be obtained from the sensor data specifications. Even in this case, however, the noise variance needs to be computed from the average value of the local image region because of shot noise component in the image sensor noise. If color correction is performed after image processing or when sensor specifications cannot be obtained, the image noise variance can be estimated from methods described in References 5 and 6. The next step is to divide the image into local regions and estimate 2^{nd} order statistics of the image regions. After obtaining the correlation matrix of R, G and B pixel values (with noise), the color correction matrix for the local image region can be obtained using Equation (5).

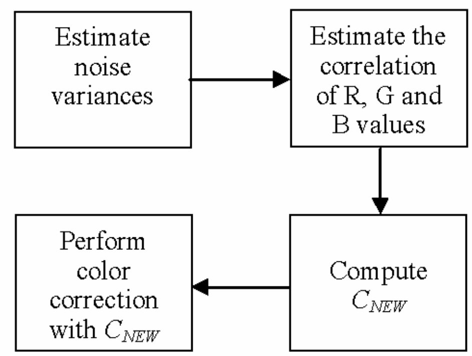

Figure 2. Block diagram of the new color correction method

There are many ways to implement region-based image processing. The simplest way that is commonly used in compression standards such as JPEG or MPEG is to divide the image into non-overlapping blocks. This is very attractive in terms of implementation because the algorithm does not require additional frame memory for implementation. Unlike these block-based compression standards our color correction method does not suffer from blocking artifacts as can be seen in Section III. This is because C_{NEW} tries to minimize color error as well as noise, which makes it more robust to blockiness artifact. To choose optimum block sizes, we applied the new method while we varied the block sizes and monitored the mean-squared-error after color correction. Although optimum block size depends on the image content and statistics, a block size of 8 by 8 seemed to achieve the best results for typical images. Images with more high frequency content generally require smaller block sizes.

The summary of the baseline procedure is given below.

1) Divide the image into non-overlapping 8 by 8 blocks.
2) For each block, compute the correlation matrix (*Cor*) of the R, G and B channels and estimate the correlation matrix (*CorNN*) of the image sensor noise.
3) Compute the color correction matrix using the correlation matrices *Cor* and *CorNN*.
4) Apply the newly calculated color correction matrix C_{NEW} to all the pixels in the block.
5) Proceed onto the next block and repeat the steps 2), 3) and 4)

C. Extension of the Baseline Approach

There are several ways to extend the baseline approach. One extension would be to use other metric than mean-squared-error in RGB space. For example, optimizing for smallest error in CIELAB space may yield better perceptual results. Another extension would be to apply the method to a 4 (or higher) color imaging system. In the previous sections, we described the method assuming 3 color channels for the image sensor. Although this is true for most image sensors today, this method can also be used to convert more than 3 color channels to the standard R, G and B color space. Since each color channel will have different noise statistics, the proposed color conversion matrix will naturally choose the color channels that have lower noise than the others. Thus, this method can be used to adaptively weight the color channels depending on the noise characteristics. For example, if the sensor has cyan, magenta, yellow and green color channels, the proposed color-correction matrix can be used as a vehicle to choose (and weigh) the color channels that minimizes sensor noise and color error.

In the "Step 1)" of the baseline approach (II-B), the image is divided into multiple regions. Instead of dividing the images into 8 by 8 blocks, it is more logical to group the pixels that have similar statistics. One way would be to group the pixels that have similar colors using clustering algorithms or vector quantization algorithms and then calculate the color correction matrix for that region. This could potentially give better results than using non-overlapping rectangular blocks but would be more complex to implement. One extreme case of this would be to divide the image according to "photometric similarity" rather than "geometric proximity". In other words, we could have a look-up-table of several color correction matrices based on the pixel values.

In the case when calculating the inverse of *Cor* matrix is too complex to implement, the computational complexity of the proposed method can be reduced by using numerical algorithms such as conjugate gradient or steepest descent method. Initial starting point for C_{NEW} matrix can just be $C_{NOMINAL}$ or the C_{NEW} matrix of the adjacent block.

III. Experimental Results

A. Experimental Setup

To test the effectiveness of the SVCC method, we used hyperspectral images obtained from Reference 7. Each hyperspectral image consists of 31 monochrome image planes, corresponding to wavelengths between 400nm to 700nm in 10nm steps. The hyperspectral images were subsampled by a factor of 4 (both horizontally and vertically) to reduce the immense size of data. Hyperspectral images allow us to simulate arbitrary color filters instead of being pinned to a specific color filter. Figure 3 shows the spectral response of a set of color filter arrays (including quantum efficiency of the image sensor). The spectral response that has high overlap between color channels was chosen to illustrate the effectiveness of our method. Recall that when color channels are highly correlated, the resulting color correction matrix will have a high condition number. Also, the hyperspectral images have high bit depth and extremely low noise, which facilitate quantitative analysis and extensive testing. We simulated the image capture process of an ordinary consumer digital camera with the image sensor noise model described in Reference 8. The image sensor noise is the sum of shot noise, readout noise and fixed pattern noise regardless of the type of the image sensors. We chose typical image sensor parameters, which are listed below.

*Well capacity: 40000 electrons
*Sensor readout noise: 60 electrons
*Conversion gain: 25 μ V/e

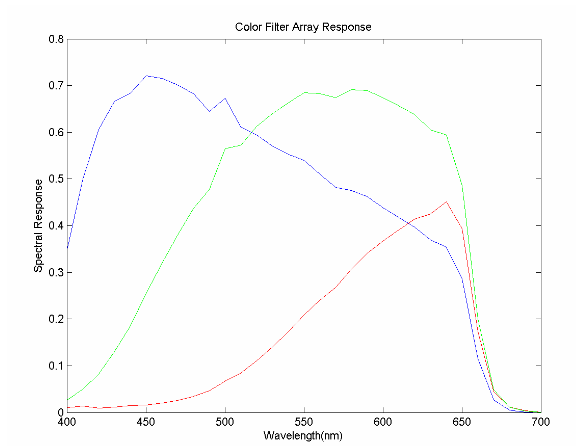

Figure 3. Spectral response of the color filter used for our simulations

When the entire capture process is simulated without adding any noise, the ideal noise-free color corrected image (i. e., the "Noise-free case" in Figure 1) can be obtained and used as the ground truth image. The "Real-case" images resulting from different color correction methods can be quantitatively compared by computing the mean-squared-error difference with the ground truth image.

B. Results

From the spectral response shown in Figure 2, we computed the color correction matrix that transforms the raw R, G and B values to sRGB space. The color correction matrix with the least color error ($C_{NOMINAL}$) is

$$C_{NOMINAL} = \begin{bmatrix} 3.7423 & -0.9882 & 0.1377 \\ -2.1828 & 1.8432 & -0.5416 \\ 0.7365 & -1.4519 & 1.4612 \end{bmatrix}$$

Note that the high off-diagonal element results in high noise amplification. This is mainly because of the high correlation between the color channels. Figures 4, 6 and 8

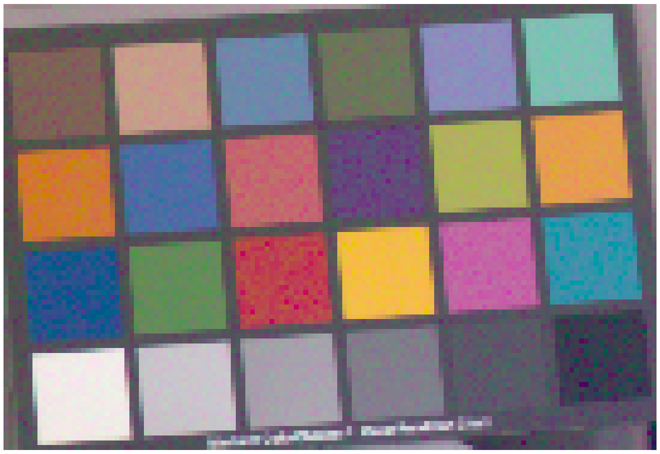

Figure 4: After color correction with $C_{NOMINAL}$

Figure 5: After color correction with C_{NEW}

Figure 6: After color correction with $C_{NOMINAL}$

Figure 7: After color correction with C_{NEW}

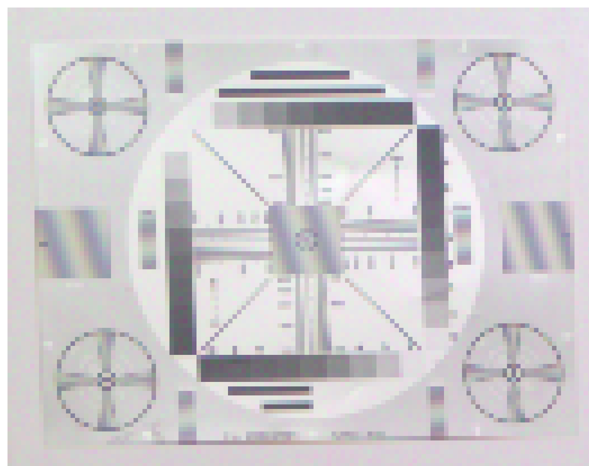

Figure 8: After color correction with $C_{NOMINAL}$

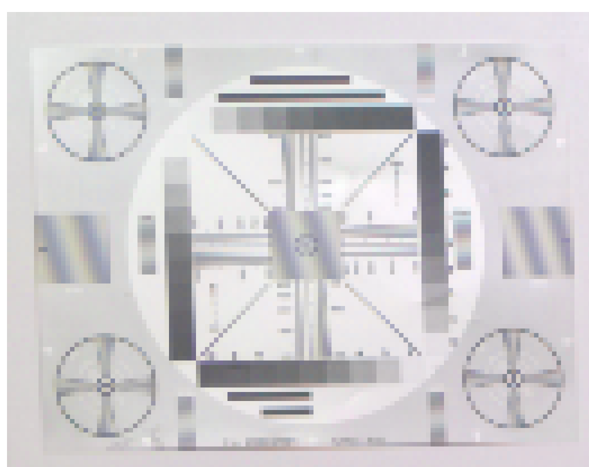

Figure 9: After color correction with C_{NEW}

illustrate the zoomed-in parts of the color corrected image using the conventional method ($C_{NOMINAL}$). Figures 5, 7 and 9 illustrate the zoomed-in parts of the color corrected image using the new SVCC method (C_{NEW}).

To measure the noise reduction quantitatively, we computed the mean-squared-error between the "Noise-free case" image (ground truth) and the "Real-case" images using both the conventional and the SVCC method. The mean-squared-error (MSE) which includes both the color error and amplified noise was 6.69 Digital Number (DN) using the conventional method ($C_{NOMINAL}$) while the MSE for the SVCC method was 3.8 DN.

IV. Summary

The paper described the SVCC method that performs color correction (or transformation) without excessive noise amplification even when color channels are highly correlated. The key idea was to locally optimize color correction matrices instead of having a single matrix for the entire image. It is worthwhile to note that the SVCC method can be applied to any color transformation such as RGB to YCbCr transformation. Also, the SVCC method can be used to choose the color channels adaptively when certain channels have lower fidelity than the others. As was briefly discussed in Section II-C, we believe the SVCC method can be improved by a more sophisticated choice of "local regions" than a fixed block.

Acknowledgements

The authors would like to thank David Brainard and Stephen Pollard for helpful discussions and providing data for the experiments.

References

1. U. Barnhofer, J. DiCarlo, B. Olding and B. A. Wandell, "Color Estimation Error Trade-offs", in Proceedings of the SPIE Electronic Imaging Conference, Vol. 5017, San Jose, CA, 2003.
2. P. Vora and C. Herley, "Trade-offs Between Color Saturation and Noise Sensitivity in Image Sensors", in Proceedings of the IEEE International Conference on Image Processing, Vol. 1, pp. 196-200, 1998.
3. G. Sharma and H. J. Trussell, "Figures of Merit for Color Scanners", IEEE Transactions on Image Processing, Vol. 6, No. 7, pp. 990-1001, Jul 1997
4. M. J. Vrhel, H. J. Trussell, "Filter Considerations in Color Correction", IEEE Transactions on Image Processing, Vol 3, No. 2, pp. 147-161, Mar 1994
5. D. L. Donoho and I. M. Johnstone, "Ideal Spatial Adaptation via Wavelet Shrinkage," Biometrika, Vol. 81, pp. 425-455, 1994.
6. D. L. Donoho and I. M. Johnstone, "Adapting to Unknown Smoothness via Wavelet Shrinkage", Journal of the American Statistical Association, Vol. 90, No. 432, pp. 1200-1224, December 1995
7. 7. David Brainard, http:// color. psych. ucsb. edu/ hyperspectral/ (Hyper-spectral Image Data)
8. 8. A. El Gamal, EE392B Classnotes: Introduction to Image Sensors and Digital Cameras, http:// www. stanford. edu / class/ ee392b, Stanford University, 2001

Biography

SukHwan Lim received the B. S. degree from Seoul National University, Seoul, Korea, in 1996, the M. S. degree and Ph. D. degree in Electrical Engineering from Stanford University in 1998 and 2003, respectively. He is currently a member of technical staff at Hewlett-Packard Laboratories. His research interests include image sensors and digital signal processing architectures for smart imaging devices. He is also interested in image/ video processing algorithms for imaging devices such as demosaicing, image restoration, optical flow estimation and video stabilization.

Amnon Silverstein received his B. A. degree in Computer Science from Oberlin College in 1987, a M. S. degree in Physiological Optics from the University of California at Berkeley in 1994 and a Ph. D. in Vision Science from the University of California at Berkeley in 1999. Since 1996 he has worked at Hewlett-Packard Laboratories in Palo Alto, California. His work has focused on image quality assessment, digital image processing and applications of motion sensing to imaging devices.