Measures of Goodness for Color Scanners

Gaurav Sharma and H. Joel Trussell
ECE Dept., North Carolina State University, Raleigh, North Carolina

Abstract

Color errors in scanners arise from two sources: the non-colorimetric nature of the scanner sensitivities and the measurement noise. Several measures of goodness have been used to evaluate scanners based on these errors. In this paper, the trustworthiness of these measures is studied through simulations. A new measure incorporating both the above sources of errors and providing excellent agreement with perceived color error is also presented.

Introduction

The color of an object is specified by its CIE XYZ tristimulus values

$$t_r = A^T L r = A_L^T r, \tag{1}$$

where $\mathbf{t_r}$ is the 3×1 vector of CIE XYZ tristimulus values, $\mathbf{A}$ is the $N \times 3$ matrix of CIE XYZ color matching functions, $\mathbf{L}$ is the $N \times N$ diagonal matrix with samples of the viewing illuminant spectrum along the diagonal, $\mathbf{r}$ is the $N \times 1$ vector of the object reflectance, and $\mathbf{A_L} = \mathbf{LA}$.

Scanner measurements of the object with a K channel scanner can be similarly expressed as

$$u_r = M^T L_s r + \eta = G^T r + \eta, \tag{2}$$

where $\mathbf{u_r}$ is a $K \times 1$ vector of scanner measurements, $\mathbf{M}$ is the $N \times K$ matrix of scanner filter transmittances, $\mathbf{L_s}$ is the $N \times N$ diagonal matrix with samples of the scanning illuminant spectrum along the diagonal, η is the $K \times 1$ measurement noise vector, and $\mathbf{G} = \mathbf{L_s M}$.

To obtain colorimetric data from scanners, it is necessary that the color matching functions be linear combinations of the scanner sensitivities. Since the column spaces of A_L and $\mathbf{G}$ define **the** *human visual illuminant sub-space* (HVISS) and the *scanner visual sub-space* (SVS), respectively, this is equivalent to the requirement that the HVISS be contained in the SVS. The different measures of goodness quantify the fractional "amount" of the HVISS contained in the SVS. In the presence of noise, most of these measures are readily modified to quantify the fractional "amount" of the HVISS recoverable from the scanner measurements.

Measures of Goodness

A "quality factor" for color filters was first proposed by Neugebauer.[7] For each color scanning filter, Neugebauer defined a quality factor as the fraction of its energy lying in the HVISS. Thus if g denotes the sensitivity of a scanner channel, its quality factor is given by

$$q_n(g) = \frac{\left\| P_{A_L} g \right\|^2}{\left\| g \right\|^2}, \tag{3}$$

where $\mathbf{P}_{\mathrm{AL}}$ is the projector onto the HVISS.

The Neugebauer quality factor is limited to the evaluation of one channel at a time. If the sensitivities of the different channels are sufficiently separated in wavelength, the average of the quality factors provides a meaningful measure of goodness for the scanner.[3] However, for more general cases the use of averages could provide misleading results.[11]

Neugebauer's color factor was generalized to multiple filters by Vora and Trussell.[11] The Vora-Measure of goodness corresponds to the normalized sum of Neugebauer Quality factors of an orthogonalized scanner sensitivities. Mathematically, the Vora measure can be expressed as

$$q_v(G) = \frac{tr(P_{A_L} P_G)}{3}, \tag{4}$$

where tr(•) denotes the trace operator[5] and $\mathbf{P_G}$ the projector onto the SVS. The Vora-measure can be used for a scanner with an arbitrary number of channels. It may also be noted that for applications where a general M-stimulus space is considered (for instance multi-illuminant viewing) instead of the typical tristimulus space the Vora-measure is readily modified by replacing the 3 in the denominator by M.

In the past, a color quality factor (CQF) has been used in industry to measure the closeness of the HVISS to the SVS. The CQF is defined by measuring how well the color matching functions defined by AL can be fit using the basis vectors defined by $\mathbf{G}$. This measure can be defined as[9]

$$q_c(G) = \min_{i=1,2,3} \frac{\left\| P_G a_i \right\|^2}{\left\| a_i \right\|}, \tag{5}$$

The measures described above relied on notions of distance between subspaces measured in terms of Euclidean distance in a tristimulus space. Since these spaces are perceptually nonuniform the use of a uniform color space could offer a potentially better measure. Such measures are however computationally intensive due to the non-linear nature of uniform color spaces. The linearized CIE L*a*b* space proposed by Wolski et. al.[14] offers a reasonable compromise between computational complexity and perceptual accuracy. A *perceptual measure* based on the linearized CIE L*a*b* space can be defined as[8]

$$q_p(G) = \frac{\tau(G)}{\alpha}, \tag{6}$$

where

$$\tau(G) = \mathrm{vec}(A_L^T)^T S_r (G \otimes I_3). \\ [(G^T \otimes I_3) S_r (G \otimes I_3)]^{-1}. \\ (G^T \otimes I_3) S_r \mathrm{vec}(A_L^T), \tag{7}$$

$$\alpha = \mathrm{vec}\left(A_L^T\right)^T S_r \mathrm{vec}\left(A_L^T\right), \tag{8}$$

vec(•) is an operator that transforms a matrix into a vector by stacking the columns of the matrix one underneath the other in sequence, I_3 denotes the 3×3 identity matrix, $\otimes$ denotes the Kronecker product operator,

$$S_\tau = E\{(rr^T) \otimes (J_F^T(t_r) J_F(t_r))\}, \tag{9}$$

$E\{\bullet\}$ denotes the expectation over the ensemble of objects to be scanned, $\boldsymbol{F}(\bullet)$ denotes the 3×3 (nonlinear) transformation from CIE XYZ to CIE L*a*b* space,[2] and $J_{\boldsymbol{F}}(\mathrm{t_r})$ denotes the Jacobian matrix[6] of the transformation $\boldsymbol{F}(\bullet)$ at $\mathrm{t_r}$.

Of the four measures described above, the first three ignored knowledge of the statistics of the ensemble of scanned reflectance spectra and all four neglected the effects of the measurement noise. By incorporating this information, more comprehensive measures of goodness may be obtained. In order to distinguish these from the measures of the last section, these will be referred to as figures of merit.

Two figures of merit will be considered here. The first is a figure of merit based on an orthogonal color space,

$$q_o(G) = \frac{tr\left(P_{A_L} K_r G \left(G^T K_r G + K_\eta\right)^{-1} G^T K_r\right)}{tr\left(P_{A_L} A_L^T K_r\right)}, \tag{11}$$

where $\mathrm{K}_\eta = E\{\eta\eta^T\}$ is the noise covariance matrix and the other terms are as defined earlier.

The second figure of merit considered is an extension perceptual measure to account for measurement noise.[8] This *perceptual figure of merit* is given by

$$q_{pn}(G) = \frac{\tau_n(G)}{\alpha}, \tag{12}$$

where

$$\tau_n(G) = \mathrm{vec}\left(A_L^T\right)^T S_r \left(G \otimes I_3\right). \\ \left[\left(G^T \otimes I_3\right) S_r \left(G \otimes I_3\right) + S_\eta\right]^{-1}. \\ \left(G^T \otimes I_3\right) S_r \mathrm{vec}\left(A_L^T\right), \tag{13}$$

$$\mathbf{S}_\eta = \mathbf{K}_\eta \otimes E\left\{J_F^T(\mathbf{t_r}) J_F(\mathbf{t_r})\right\} \tag{14}$$

and α is defined in (8).

Experimental Results

In order to examine the trustworthiness of the different measures, their relation to average ΔE^*_{ab} error will be studied through simulations. To test the predictive capabilities of the measures to imperfect filter sets a large number of sets was needed. This was generated by using parameterized mathematical filters. The parameters were varied to obtain a 251 filter sets with three filters per set. For the ensemble of scanner target reflectances a total of 424 reflectances were used. Of these 240 were from the Kodak Q60 photographic scanner target, 64 from the Munsell chart, and 120 from a Dupont Paint catalog.

For each filter set noisy scanner measurements of the target ensemble were simulated using (2). These measurements and the actual XYZ values from (1) were converted to CIE L*a*b* space and the average ΔE^*_{ab} error was computed. Simulations were performed for signal-to-noise ratio (SNR) values of 40, 50 and 60 dB, where the SNR was defined as

$$SNR(dB) = 10 \log_{10}\left(\frac{tr\left(G^T K_r G\right)}{\sigma_\eta^2}\right). \tag{15}$$

Using the results for the different filter sets, scatter plots of the different measures vs. the average ΔE^*_{ab} error were made. These are presented in Figures 1-6.

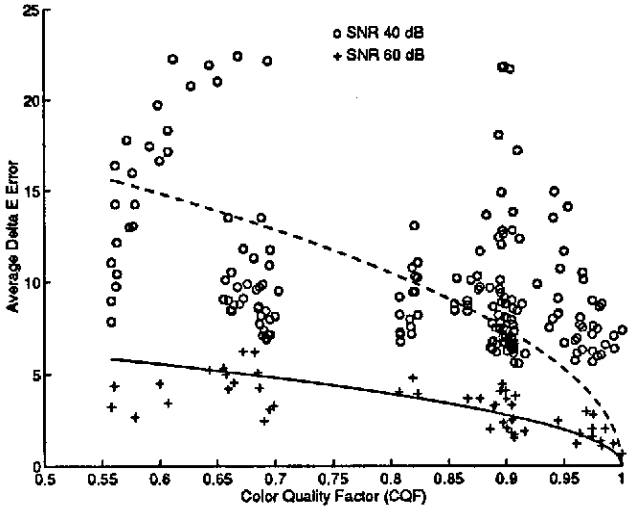

Figure 1. CQF vs. Average ΔE^*_{ab}.

For an ideal measure of scanner goodness, the points in the scatter diagram should lie along a smooth monotonic curve. Consider the scatter plots in Figures 1-4 for the measures that ignored the noise statistics. Even at a relatively high SNR of 60 dB the CQF (Figure 1) performs extremely poorly, with points being widely scattered. At the same

SNR, the average Neugebauer quality factor (Figure 2) is somewhat better, and the Vora-Measure (Figure 3) is significantly better particularly in the region corresponding to high measures, where the scatter points are close to a monotonic curve. However, for filter sets with lower measures there is considerable spread in points on the scatter diagram. The perceptual measure (Figure 4) performs ideally at a 60 dB SNR with the scatter points lying extremely close to a monotonic curve. However, as the noise level increases, all of these measures perform poorly. At 40 dB SNR, the scatter plots for all measures are widely spread out and no clear functional relation is apparent between the measures and the average ΔE^*_{ab} error, even for the perceptual measure that was based on a linearized CIE L*a*b* space.

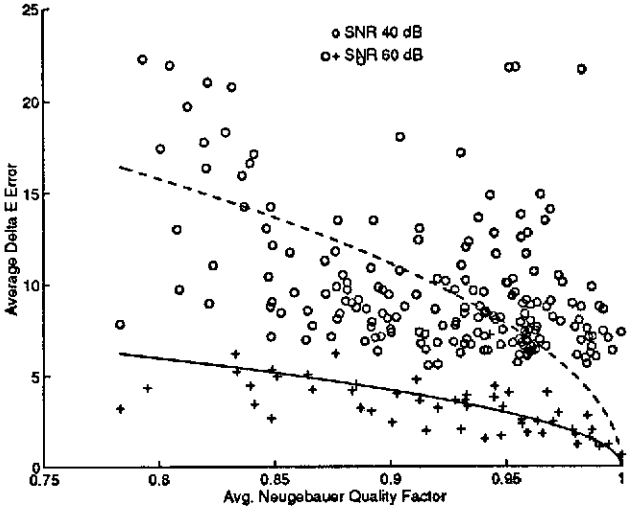

Figure 2. Avg. Neugebauer quality factor vs. Avg. ΔE^*_{ab}.

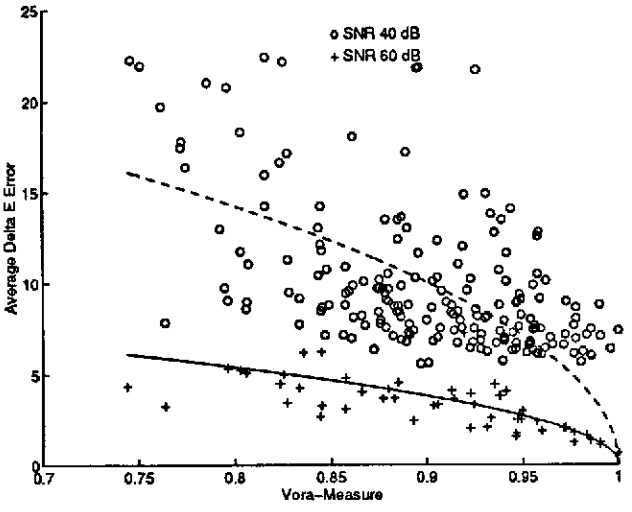

Figure 3. Vora-measure vs. Average ΔE^*_{ab}.

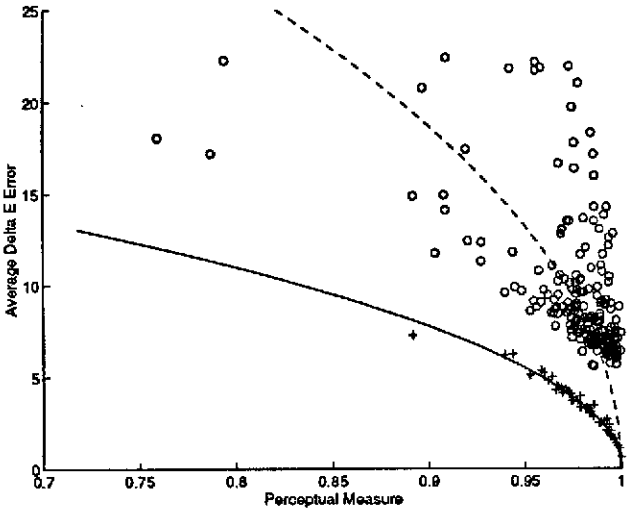

Figure 4. Perceptual Measure vs. Average ΔE^*_{ab}.

Figures 5 and 6 contain the scatter plots for the orthogonal color space figure of merit and the perceptual figure of merit, respectively. These figures of merit account for measurement noise in their formulation and therefore capture the trade-off between the colorimetric quality of the scanner and the noise performance[10] in a continuous fashion. At lower SNRs, the values of these figures of merit are also lower and the corresponding points are shifted to the left on the scatter plots. Both the figures of merit perform better than the measures discussed in the last paragraph. The perceptual figure of merit, however, performs exceedingly well in comparison to all the other measures and the points on the scatter diagram in Figure 6 all lie very close to a smooth monotonically decreasing curve.

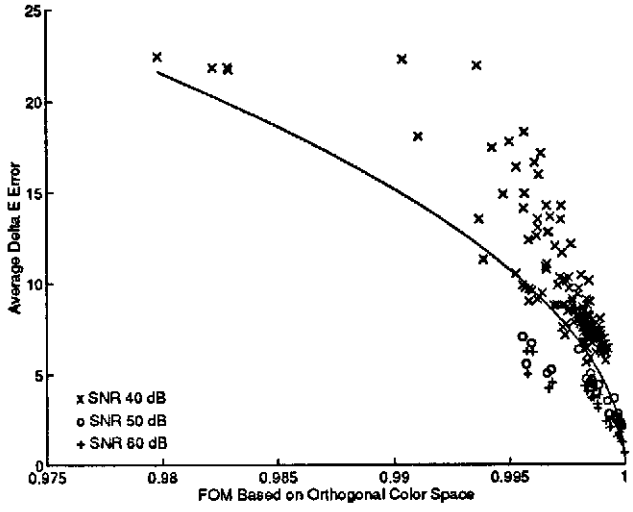

Figure 5. Orthogonal Color Space Based figure of merit vs. Average ΔE^*_{ab}.

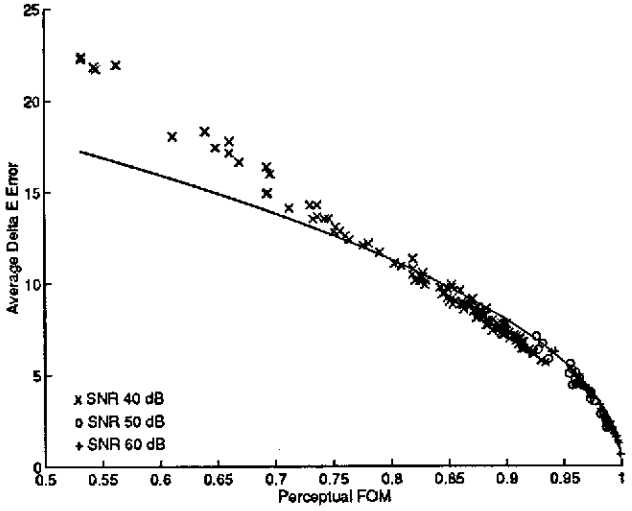

Figure 6. Perceptual figure of merit vs. Average ΔE^*_{ab}.

Conclusions

In this paper, the capabilities of different measures of goodness for predicting the perceived color error in scanners (quantified as the average ΔE^*_{ab} error) was examined through simulations. Several existing measures and a couple of new measures were considered in the comparisons. It was demonstrated that measures that ignored noise and provided quantitative estimates of the non-colorimetric nature of the scanner sensitivities performed poorly in the presence of noise. The figures of merit that incorporated knowledge of noise statistics performed significantly better, with the new perceptual figure of merit providing close agreement with average ΔE^*_{ab} error for a wide range of SNRs and across filters with considerable variation in colorimetric quality.

References

1. P. Chen and H. J. Trussell, "Color Filter Design for Multiple Illuminants and Detectors," *Proc. Third IS&T/SID Col. Imag.Conf.*, Nov. 1995, pp. 67-70.
2. CIE, *Colorimetry*, 2nd Ed., CIE Publication 15.2, Central Bureau of the CIE, Paris, 1986.
3. P.G. Engeldrum, "Color Scanner Calorimetric Design Requirements," *Proc. SPIE*, vol. **1909**, 1993, pp. 75-83.
4. J. E. Farrell and B. A. Wandell, "Scanner Linearity," *J. Electronic Imaging*, vol. **2**, no. 3, July 1993, pp. 225-230.
5. G. H. Golub and C. F. VanLoan, *Matris Computations*, 2nd Ed, The Johns Hopkins University Press, 1989.
6. J. R. Magnus and H. Neudecker, "Matris Differential Calculus with Applications in Statistics and Econometrics," Wiley, 1988. 315
7. H. E. J. Neugebauer, "Quality Factor for Filters Whose Spectral a Transmittances are Different from Color Mixture Curves, and its Application to Color Photography," *JOSA*, Vol. **46**, No. 10, Oct 1956, pp 821-824.
8. G. Sharma and H. J. Trussell, "Figures of Merit for Color Scanners and Cameras," submitted to *IEEE Trans. Image PTOC.*
9. H. J. Trussell, G. Sharma, P. Chen, and S. A. Rajala, "Comparison of Measures of Goodness of Sets of Color Scanning filters," *Proc. IEEE MDSP WI:JP. 1996, pp.* 98-99.
10. G. Sharma and H. J. Trussell, "Color Scanner Performance Tradeoffs", *Proc. SPIE*, v. **2658**, pp. 270-278.
11. P. L. Vora and H. J. Trussell, "A Measure of Goodness of a Set of Color Scanning Filters," *JOSA-A* V. **10**, no. 7, 1993, pp. 14991508.
12. M. J. Vrhel and H. J. Trussell, "Filter Considerations in Color Correction", *IEEE Trans. Image Proc.*, vol. **3**, No. 2, pp. 147-161, March 1994.
13. M. J. Vrhel and H. J. Trussell, "Optimal Color Filters in the Presence of Noise," *IEEE Trans. Image Proc*., vol. **4**, No. 6, pp. 814-823, June 1995.
14. M. Wolski, C. A. Bouman, and J. P. Allebach, "Optimization of Sensor Response Functions for Colorimetry of Reflective and Emissive Objects," *Proc. IEEE ICIP-95*, pp. II 323-326.
15. G. Wyszecki and W. S. Styles, "Color Science: Concepts and Methods", *Quantitative Data and Formxlac*, Second edition, John Wiley & Sons, Inc., 1982.