COLOR FILTER SELECTION FOR COLOR CORRECTION IN THE PRESENCE OF NOISE

Michael J. Vrhel and H. Joel Trussell
Electrical & Computer Engineering Dept.
North Carolina State University
Raleigh, NC 27695-7914

ABSTRACT In this paper the effect of the noise on the number of effective channels (color filters), used to record a color image, is investigated. Transmittances of color filters are calculated which minimize the mean square error that occurs when estimating, from the recorded data, the colors in the image under a collection of viewing illuminants. Since the results indicate that a significant improvement in color correction accuracy is achieved by using four channels, there is good reason to consider using four-tuples for representation of colorimetric information.

1 INTRODUCTION

In the reproduction of a color image, the original image is initially recorded using a device which measures the reflected or transmitted energy of the image in a number of different wavelength bands. Sampling the visible spectrum (400nm - 700nm) at N wavelengths, the recording process can be represented using vector notation. In this case, if P color filters are used to obtain P different wavelength bands, then the recording process can be represented using

$$\mathbf{c}_j = \mathbf{N}^T \mathbf{L} \mathbf{f}_j + \mathbf{u}_j \tag{1}$$

where $\mathbf{c}_j$ is the recorded P-stimulus value at pixel j, $\mathbf{u}_j$ is additive signal independent noise, $\mathbf{f}_j$ is an N-dimensional vector representing the spectral reflectance of an object at pixel j, $\mathbf{L}$ is an NxN diagonal matrix representing the spectral power distribution of the illumination, and $\mathbf{N} = [\mathbf{n}_1, \mathbf{n}_2, ...\mathbf{n}_P]_{N \times P}$ where $\mathbf{n}_i$ represents the spectral transmittance of the ith filter.

The reproduction is standardly viewed and compared with the original image under an illuminant that is different from the imaging illuminant. Typically there will be a small number of viewing illuminants under which the two images may be compared. From the recorded data it is desired to estimate the colors of the original image under each of the viewing illuminants. This work focuses on the effects of the noise which limit the accuracy of estimation of this information.

2 PROBLEM FORMULATION

If the matrix $\mathbf{A}$ contains the CIE color matching functions, then the CIE tristimulus value

$$\mathbf{t}_{ij} = \mathbf{A}^T \mathbf{L}_i \mathbf{f}_j \tag{2}$$

quantifies the color of the object with spectral reflectance $\mathbf{f}_j$ under illuminant $\mathbf{L}_i$[3]. Clearly the color of an object with spectral reflectance $\mathbf{f}_j$ under illuminant $\mathbf{L}_i$ is uniquely determined by the projection of $\mathbf{f}_j$ onto the range space of $\mathbf{L}_i\mathbf{A}$.

The data obtained in the recording process is a function of the system noise, the color filters, the imaging illuminant, the sensor spectral sensitivity and nonlinearities, and the spectral reflectances in the original image. The spectral sensitivity response of the sensors can be combined with the transmittance of the color filters to give an overall system spectral sensitivity. For this work, it is assumed that the brightness of the image and illuminant are such that the device is operating in its linear region.

When designing a color imaging device there is usually limited control over the sensor characteristics. The overall system spectral sensitivity is obtained by careful selection of the color filters and imaging illuminant. The problem of how to select the color filters and imaging illuminant can be formulated as an optimization problem. Assume there are K viewing illuminants. The original image may be compared to the reproduction under any one viewing illuminant. The cost func-

0-7803-0946-4/93 $3.00

tion is the mean square error between the three dimensional vectors which describe the colors of the original image under each of the K viewing illuminants and the estimated three dimensional vectors.

Mathematically the cost function, which is minimized with respect to the color filters/imaging illuminant matrix $\mathbf{LN}$, is

$$\epsilon[\mathbf{LN}] = \sum_{i=1}^{K} \eta_i^2 E\{||\mathcal{P}_i(\mathbf{c}) - \mathbf{O}_i^T\mathbf{f}||^2\} \qquad (3)$$

where $\mathcal{P}_i$ is the illumination color correction transformation for the ith illuminant, $\mathbf{c}$ is the data obtained from the output of the color filters, the η_i^2 are used to provide a weighting for the cost of color errors, and the expectation operator is taken over the reflectance spectra $\mathbf{f}$ and the additive noise. The Nx3 dimensional matrices $\mathbf{O}_i \quad i = 1...K$ are constructed to have orthonormal columns with the range space of $\mathbf{O}_i$ the same as the range space of $\mathbf{L}_i\mathbf{A}$. The orthonormalization is performed to remove weightings caused by magnitude differences in the illuminants. Note that $\mathbf{O}_i^T\mathbf{f}$ quantifies the color of spectral reflectance $\mathbf{f}$ under illuminant $\mathbf{L}_i$. For simplicity the color filters/imaging illuminant matrix will be denoted by $\mathbf{G} = \mathbf{LN}$.

It can be shown that the linear minimum mean square error estimator of the three dimensional vectors $\mathbf{d}_{ij} = \mathbf{O}_i^T\mathbf{f}_j$ is

$$\hat{\mathbf{d}}_{ij} = \mathcal{P}_i(\mathbf{c}_j) = \mathbf{O}_i^T\mathbf{K}_f\mathbf{G}[\mathbf{G}^T\mathbf{K}_f\mathbf{G} + \mathbf{K}_u]^{-1}[\mathbf{c} - \bar{\mathbf{u}} - \mathbf{G}^T\bar{\mathbf{f}}] + \mathbf{O}_i^T\bar{\mathbf{f}} \qquad (4)$$

where $\mathbf{K}_f$ is the covariance matrix of the reflectance spectra, $\mathbf{K}_u$ is the covariance matrix of the noise, $\bar{\mathbf{f}}$ is the mean of the reflectance spectra, and $\bar{\mathbf{u}}$ is the mean of the noise[2]. The sensor measurements for the P channels are performed independently on similar channels. Therefore, it is assumed that $\mathbf{K}_u = \mathbf{I}\sigma^2$.

Substituting (4) into (3), it can be shown that minimizing ϵ with respect to $\mathbf{G}$ is equivalent to maximizing

$$\zeta = \text{Trace}[\mathbf{S}\mathbf{S}^T\mathbf{K}_f\mathbf{G}[\mathbf{G}^T\mathbf{K}_f\mathbf{G} + \mathbf{I}\sigma^2]^{-1}\mathbf{G}^T\mathbf{K}_f] \qquad (5)$$

where $\mathbf{S} = [\mathbf{O}_1\eta_1, \mathbf{O}_2\eta_2, ..., \mathbf{O}_K\eta_K]$. Using additional algebraic manipulations ζ can be written as

$$\zeta = \text{Trace}[\mathbf{K}_f^{\frac{1}{2}}\mathbf{S}\mathbf{S}^T\mathbf{K}_f^{\frac{1}{2}}\mathbf{U}\mathbf{\Lambda}[\mathbf{\Lambda}^2 + \mathbf{I}\sigma^2]^{-1}\mathbf{\Lambda}\mathbf{U}^T] \qquad (6)$$

where $\mathbf{G} = \mathbf{K}_f^{-\frac{1}{2}}\mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$, $\mathbf{V}^T = \mathbf{V}^{-1}$, $\mathbf{U}^T\mathbf{U} = \mathbf{I}_{p\times p}$, and $\mathbf{\Lambda} = \text{Diag}[\lambda_1, ..., \lambda_P]$. Note that ζ is independent of the unitary matrix $\mathbf{V}$.

The total power collected by the recording device is the integral of the image intensity over the duration of the measurement. Since the intensity of the image and the measurement time are limited by practical considerations, it is necessary to include a constraint on the signal power. The constraint is

$$E\{||\mathbf{G}^T\mathbf{f}||^2\} = \rho \qquad (7)$$

where ρ is a constant. The constraint on the signal power results in a constraint on the $\lambda_i \quad i = 1, ..., P$ which is

$$\text{Trace}[\mathbf{\Lambda}^2] = \rho - ||\mathbf{G}^T\bar{\mathbf{f}}||^2 = \kappa \qquad (8)$$

where κ represents the power in the unknown portion of the signal. Note that the total signal power is fixed, while the distribution of the signal power over the channels is not fixed. As the number of channels, P, is increased, the signal power is divided among a larger number of channels, and additional noise is introduced. This results in a decreased signal-to-noise ratio, $\frac{\kappa}{P\sigma^2}$, as the number of channels is increased.

It can be shown that the optimal value for the matrix $\mathbf{U}$ is the matrix containing the P eigenvectors associated with the P largest eigenvalues of the matrix $\mathbf{K}_f^{\frac{1}{2}}\mathbf{S}\mathbf{S}^T\mathbf{K}_f^{\frac{1}{2}}$, in which case the matrix $\mathbf{G}$, which minimizes ϵ in (3), can be computed from the solution to the following optimization problem:
Maximize

$$\sum_{i=1}^{P} \delta_i \frac{\gamma_i}{\gamma_i + \sigma^2} \qquad (9)$$

with respect to the γ_i, subject to

$$\gamma_i \geq 0 \qquad \sum_{i=1}^{P} \gamma_i = \kappa$$

where the δ_i are the P largest eigenvalues of $\mathbf{K}_f^{\frac{1}{2}}\mathbf{S}\mathbf{S}^T\mathbf{K}_f^{\frac{1}{2}}$, $\delta_1 \geq \delta_2 \geq ..... \geq \delta_P$, and $\lambda_i^2 = \gamma_i$.

From the first-order necessary conditions and algebraic manipulations, it can be shown that the optimal γ_j can be calculated using

$$\tilde{\gamma}_{j,M} = (\kappa + M\sigma^2)\frac{\sqrt{\delta_j}}{\sum_{i=1}^{M}\sqrt{\delta_i}} - \sigma^2 \quad j = 1, ...P \qquad (10)$$

where the optimal γ_j are

$$\gamma_j^* = \tilde{\gamma}_{j,M} \quad j = 1, ...M \quad \gamma_j^* = 0 \quad j = M+1, ...P \quad .$$

and M is the number of nonzero γ_j^*.

A problem with (10) is that it depends on knowledge of M, the number of non-zero γ_j^*. This information is not known *a priori*. It can be shown that if a value R, which is not the number of positive γ_i^*, is used for M in (10), then the conditions $\tilde{\gamma}_{j,R} > 0 \quad j = 1, ...R$ and $\tilde{\gamma}_{j,R} = 0 \quad j = R+1, ...P$ are not satisfied. In other words these conditions are satisfied only if $R = M$, the number of nonzero γ_j^*.

Therefore, the γ_j^* values can be calculated by the following process;
1) From (10) calculate $\tilde{\gamma}_{r,r}$ for $r = 1, ...P$.
2) The largest value of r for which $\tilde{\gamma}_{r,r} > 0$ is M, the number of nonzero γ_i^*, and

$$\gamma_j^* = \tilde{\gamma}_{j,M} \quad j = 1, ...M \quad \gamma_j^* = 0 \quad j = M+1, ...P \quad .$$

From the above results, an optimal set of filters which minimize ϵ will be contained in the set

$$C_G = \{ \mathbf{G} \mid \mathbf{G} = \mathbf{K}_f^{-\frac{1}{2}} \mathbf{U}^* \mathbf{\Lambda}^* \mathbf{V}^T \quad \mathbf{V}^T = \mathbf{V}^{-1} \} \quad (11)$$

where the columns of $\mathbf{U}^*$ are the P eigenvectors associated with the P largest eigenvalues of the matrix $\mathbf{K}_f^{\frac{1}{2}} \mathbf{S}\mathbf{S}^T \mathbf{K}_f^{\frac{1}{2}}$, $\Lambda^* = \text{Diag}[\lambda_1^*, ...\lambda_P^*]$, and $\lambda_i^* = \sqrt{\gamma_i^*}$.

3 PHYSICAL CONSTRAINTS

An optimal set of realizable filters will be any set in the intersection of C_G and $C_n = \{ \mathbf{G} \mid \mathbf{G} \geq \mathbf{0} \}$. A unitary transformation $\mathbf{V}$ may not exist which results in a nonnegative set of filters, in which case the intersection of C_G and C_n is empty. If the intersection is empty, then it is of interest to find a matrix $\mathbf{G}$ in C_n which is a minimum distance from C_G with respect to some measure. This can be formulated as the following optimization problem:
Minimize

$$\frac{\sigma_{\max}(\mathbf{V})}{\sigma_{\min}(\mathbf{V})} \quad (12)$$

with respect to $\mathbf{V}$, subject to

$$\mathbf{K}_f^{-\frac{1}{2}} \mathbf{U}^* \mathbf{\Lambda}^* \mathbf{V}^T \geq \mathbf{0} \qquad \text{Trace}[\mathbf{V} \mathbf{\Lambda}^* \mathbf{\Lambda}^* \mathbf{V}^T] = \kappa$$

where $\sigma_{\max}(\mathbf{V})$, $\sigma_{\min}(\mathbf{V})$, are the maximum and minimum singular values of the matrix $\mathbf{V}$ respectively. Standard constrained optimization methods can be used to obtain a solution to the above problem[1].

Since the intersection of C_G and C_n could be empty, it may be more appropriate to impose the nonnegativity constraint on the initial optimization problem. The optimization problem is:
Maximize

$$\text{Trace}[\mathbf{S}\mathbf{S}^T \mathbf{K}_f \mathbf{G} [\mathbf{G}^T \mathbf{K}_f \mathbf{G} + \mathbf{I}\sigma^2]^{-1} \mathbf{G}^T \mathbf{K}_f] \quad (13)$$

with respect to $\mathbf{G}$, subject to

$$\mathbf{G}^T \geq \mathbf{0} \qquad \text{Trace}[\mathbf{G}^T \mathbf{K}_f \mathbf{G}] = \kappa$$

The above optimization problem is difficult to solve and large scale nonlinear programming techniques may be necessary. The set of filters obtained from the optimization problem in (12) could be used as an initial starting point in a numerical solution to the above problem.

4 SIMULATIONS

To test the performance of the filters, experiments were performed in which the tristimulus values of reflectances under several illuminants were estimated from simulated recorded data obtained from a single set of filters at various signal power-to-sensor noise ratios.

An ensemble of 343 spectral reflectances from a color copier were recorded using a spectraradiometer. A sampling width of 10nm was used which resulted in N=31 samples between 400nm and 700nm. Sets of 3,4,5,6, and 7 color filters were calculated for the copier data set and the illuminants; CIE incandescent illuminant A, CIE daylight illuminant D65, and CIE fluorescent illuminant F2, at signal power-to-sensor noise ratios of 30dB, 35dB, 40dB, 45dB, 50dB, and ∞dB. Since three illuminants were used, $K = 3$ in (3). Uniform weighting of the errors under each viewing illuminant was used so that $\eta_i^2 = 1 \quad i = 1, ...K$ in (3). Nonnegative sets of filters were calculated using the optimization problem in (12). The minimization was performed using a commercial scientific software library optimization subroutine. The subroutine uses a sequential quadratic programming algorithm in which the search direction is the solution of a quadratic programming problem.

The recorded data using the filter set $\mathbf{G}$ is

$$\mathbf{c}_j = \mathbf{G}^T \mathbf{f}_j \quad j = 1, 2....343 \; . \quad (14)$$

The CIE tristimulus values for each reflectance ensemble were calculated using each of the three viewing illuminants. This can be written as

$$\mathbf{t}_{ij} = \mathbf{A}^T \mathbf{L}_i \mathbf{f}_j \quad j = 1, 2....343 \quad i = 1, 2, ...3 \; . \quad (15)$$

The CIE tristimulus values were estimated from the recorded data using a LMMSE estimator. The CIE

color difference measure ΔE_{Lab} is often used as a measure of perceptual color difference[4]. The ΔE_{Lab} values between $\mathbf{t}_{ij}$ and the estimate $\hat{\mathbf{t}}_{ij}$ were calculated. As a rule of thumb, a ΔE error larger than three is perceptually noticeable.

Table 1 contains the results for signal power-to-sensor noise ratios of 30dB, 40dB, 50dB, and ∞dB. In the table, ΔE_{avg} denotes the average ΔE value of the set and MSE denotes the error defined by (3). The nonnegative sets of filters are denoted as N3 to N7 and the optimal filters, computed using (9), are denoted as O3 to O7, where the number indicates the number of filters. The filters labeled CMF's refer to a set of filters with transmittances equal to the CIE color matching functions. In the design of commercial desk-top scanners, attempts are often made to fabricate these filters. The filters labeled $\mathbf{A}^T\mathbf{L}_i$ denotes the use of filters $\mathbf{A}^T\mathbf{L}_i$, to obtain the tristimulus values $\mathbf{t}_{ij}$. The filters $\mathbf{A}^T\mathbf{L}_i$ are ideal in the noiseless case.

When increasing from three to four filters there is a large decrease in MSE and ΔE_{avg} at SNRs greater than 40dB. This occurs for both the optimal filters and the nonnegative filters. At 30dB there is little improvement when a fourth filter is included. For the nonnegative filters, using more than three filters at 30dB results in an increased MSE and ΔE_{avg} values. Using the CIE color matching functions as filters resulted in the largest errors. To obtain the tristimulus values $\mathbf{t}_{ij}$, the filters $\mathbf{A}^T\mathbf{L}_i$ are often considered optimal. In the absence of noise the filters, $\mathbf{A}^T\mathbf{L}_i$, produce no color errors. In the presence of noise however the filters, $\mathbf{A}^T\mathbf{L}_i$, are not optimal as evident when comparing N3 with $\mathbf{A}^T\mathbf{L}_i$ at 30dB.

References

[1] D. G. Luenberger, Introduction to Linear and Nonlinear Programming, Addison–Wesley, Reading, MA, 1973.

[2] W. K. Pratt, Digital Image Processing, 2nd Ed., John Wiley and Sons, New York, NY, 1991.

[3] H. J. Trussell "Applications of Set Theoretic Methods to Color Systems", Color Res. Appl., Vol. 16, No. 1, pp 31-41, Feb. 1991.

[4] G. Wyszecki and W. S. Stiles, Color Science: Concepts and Methods, Quantitative Data and Formulae, 2nd Ed., John Wiley and Sons, New York, NY, 1982.

SNR		30dB			40dB			50dB			∞ dB		
illum		D65	A	F2	D65	A	F2	D65	A	F2	D65	A	F2
O3	ΔE_{avg}	4.80	5.09	4.75	2.64	2.61	2.28	2.29	2.15	1.87	2.25	2.06	1.84
	MSE	4.90E-3			1.41E-3			1.06E-3			1.02E-3		
O4	ΔE_{avg}	4.32	4.65	4.48	1.40	1.57	1.73	0.55	0.69	1.06	0.35	0.49	0.95
	MSE	4.14E-3			5.36E-4			1.75E-4			1.35E-4		
O5	ΔE_{avg}	4.32	4.62	4.45	1.41	1.47	1.56	0.54	0.50	0.77	0.34	0.17	0.59
	MSE	4.07E-3			4.39E-4			7.40E-5			3.35E-5		
O6	ΔE_{avg}	4.32	4.63	4.41	1.37	1.47	1.39	0.44	0.47	0.44	3.1E-2	2.8E-2	2.1E-2
	MSE	4.05E-3			4.08E-4			4.11E-5			3.61E-7		
O7	ΔE_{avg}	4.32	4.63	4.41	1.37	1.47	1.40	0.44	0.46	0.44	8.6E-3	7.5E-3	4.5E-3
	MSE	4.05E-3			4.08E-4			4.08E-5			1.40E-8		
N3	ΔE_{avg}	7.29	7.39	7.37	3.12	3.03	2.87	2.29	2.15	1.87	2.25	2.06	1.84
	MSE	9.37E-3			1.86E-3			1.10E-3			1.02E-3		
N4	ΔE_{avg}	7.55	7.97	7.56	2.45	2.57	2.56	0.86	0.96	1.22	0.35	0.49	0.95
	MSE	9.73E-3			1.13E-3			2.35E-4			1.35E-4		
N5	ΔE_{avg}	8.78	8.71	8.55	2.99	2.84	2.88	0.93	0.83	1.09	0.34	0.17	0.59
	MSE	1.15E-2			1.29E-3			1.42E-4			3.35E-5		
N6	ΔE_{avg}	7.80	7.26	7.75	3.12	2.82	3.05	0.90	0.83	0.87	3.1E-2	2.8E-2	2.1E-2
	MSE	9.88E-3			1.34E-3			1.28E-4			3.61E-7		
N7	ΔE_{avg}	7.80	7.26	7.75	3.12	2.82	3.05	1.21	1.17	1.11	8.6E-3	7.5E-3	4.5E-3
	MSE	9.88E-3			1.63E-3			2.15E-4			1.40E-8		
CMF	ΔE_{avg}	9.09	9.60	10.42	2.96	4.29	4.34	1.17	3.37	3.28	0.75	3.24	3.08
	MSE	1.31E-2			2.60E-3			1.52E-3			1.40E-3		
AL	ΔE_{avg}	9.05	10.80	9.35	2.87	3.43	2.98	0.91	1.08	0.94	0.0	0.0	0.0

Table 1: Simulation Results