Spectral-Based Ink Selection for Multiple-Ink Printing I. Colorant Estimation of Original Objects

Di-Yuan Tzeng and Roy S. Berns
Munsell Color Science Laboratory
Chester F. Carlson Center for Imaging Science
Rochester Institute of Technology, Rochester, New York

Abstract

Research has been initiated to determine a set of six basis colorants which are the best representation of artwork such as paintings. That is, their spectral information can be accurately reconstructed by linear combinations of the six estimated colorants. Since each painting is possibly created by different colorants, the six estimated colorants are image dependent. The clue leading to the six estimated colorants is the six eigenvectors determined from the corresponding spectral measurements. The relationship between the six eigenvectors and estimated colorants is merely the linear transformation (or geometrical rotation). Based on this faith, a constrained-rotation engine using MATLAB was devised to perform the transformation from the eigenvectors to a set of all-positive vectors as the estimated colorants. Once a set of reasonable colorants is uncovered, this set of colorants can be used to synthesize the original artwork with the least metameric effect between the reproductions and originals.

Introduction

A research and development program has been initiated in the Munsell Color Science Laboratory at Rochester Institute of Technology to develop a spectral-based color reproduction system. Research has included multi-spectral capture systems[1-4] and spectral-based printing algorithms.[5-7] The current research is concerned with bridging these analysis and synthesis stages of color reproduction.

Conventional four-color printing systems are limited by sufficient degrees of freedom for tuning the visible region of the spectrum; as a consequence, they are limited, at best, to metameric reproductions. That is, color matches defined for a single observer and illuminant (usually CIE illuminant D50 and the 1931 standard observer) are often unstable when viewed under other illuminants or by other observers. For critical color-matching applications, such as catalog sales and artwork reproductions, the results are usually disappointing due to typical uncontrolled lighting and viewing. Furthermore, the existing multiple-ink printing systems,[8] which all focus on expanding color gamut, do not alleviate metamerism since their separation algorithms are trichromatic in nature. The advantage of more degrees of freedom is not exploited.

The goal of the current research is to minimize metamerism between originals and their corresponding reproductions, thus creating spectral matches. This research extends earlier research performed by Kohler and Berns.[9] Accomplishing this goal requires first estimating the spectral properties of the colorants used to create the original object or set of objects. After the possible colorants are statistically uncovered, they are correlated to an existing ink database for determining an optimal ink set. These two processes comprise the analysis stage of the current research. This contribution is concerned with the first process. (The second process, hopefully, will be presented at CIC'99.)

Approximately Linear Colorant Mixing Space

Kubelka-Munk turbid media theory[10-11] is used as the first-order approximation transforming spectral reflectance factor, R_λ, into an approximately linear space, defined as $(K/S)_\lambda$, where K represents absorption, S represents scattering, and λ is the wavelength within the visible spectrum. The considerable literature and experiential evidence in industries including paints, plastics, and textiles validates the use of $(K/S)_\lambda$ rather than R_λ in these computations. (A rigorous proof is beyond the scope of this article). The opaque form is shown in Eqs. (1) and (2).

$$R_{\lambda, \infty} = 1 + (K/S)_\lambda - \sqrt{(K/S)_\lambda^2 + 2(K/S)_\lambda}, \qquad (1)$$

$$(K/S)_\lambda = (1 - R_{\lambda, \infty})^2 / 2R_{\lambda, \infty}, \qquad (2)$$

where $R_{\infty, \lambda}$ is the spectral reflectance factor of an opaque sample. The transformations for a transparent colorant layer in optical contact with an opaque scattering support is shown in Eqs. (3) and (4).

$$R_\lambda = R_{\lambda, g}\, e^{-2K_\lambda X}, \qquad (3)$$

$$K_\lambda = -0.5 \ln\left(\frac{R_\lambda}{R_{\lambda, g}}\right), \qquad (4)$$

where $R_{\lambda, g}$ is the spectral reflectance factor of an opaque support and X is the thickness of the colorant layer. For simplicity, $(K/S)_\lambda$ as well as K_λ are denoted as Φ_λ for the remainder of this article. Thus, the colorant mixing can be described by linear combinations of individual colorants represented in Φ_λ space, i.e.,

$$\Phi_{\lambda,mix} = \sum_{i=1}^{p} c_i \phi_{\lambda, i}, \qquad (5)$$

where c_i represents the concentration of a basis colorant and $\phi_{\lambda, i}$ is a basis colorant vector normalized to its unit concentration.

As an example, a still life painting of a floral arrangement was produced with six independent acrylic paints. Each paint was applied on a paper stock at a thickness achieving opacity and measured spectrally using a Gretag SPM 60 spectrophotometer. Each reflectance vector had thirty-one components: 400 nm - 700 nm at 10 nm bandwidths and intervals. These reflectance vectors were transformed to Φ_λ. Ideally, one needs thirty-one "spectral colorants" with 10 nm bandwidth absorption and scattering properties at the sampled wavelengths in order to reconstruct the measured sample spectra. Realistically, colorants do not have such narrow band properties. Furthermore, reproducing a color by mixing thirty-one colorants is highly impractical for any real coloration process. Fortunately, chromatic stimuli are not originally created by such spectral colorants; hence, their Φ_λ do not span the entire thirty-one dimensional Φ_λ space. Instead, they are distributed in a lower dimensional Φ_λ subspace. If an original painting was only painted by, for example, six independent colorants, then, ideally, the measured set of Φ_λ should be distributed only in a six-dimensional subspace of Φ_λ space.

Principal Component Analysis

Principal component analysis (PCA) can provide a measure to statistically determine the dimensionality of the sample population.[12] Various research in color science applications had shown its prominent success and progress.[13-15] The underlying assumption is that the set of sampled Φ_λ vectors are multivariate-normally distributed in a thirty-one dimensional Φ_λ space. The in-depth discussions for the importance of the multivariate normality of the sample space which impacts the accuracy of PCA are currently under study by the present authors and, hopefully, will be published in the near future.[16]

The linear combinations of the first p eigenvectors should describe the entire set of Φ_λ if the original was created by p colorants, i.e.,

$$\Phi_{\lambda, sample} = \sum_{i=1}^{p} b_i e_{\lambda, i} ,$$

$$(6)$$

where $e_{\lambda, i}$ is the ith eigenvector and b_i is the corresponding coefficient to reconstruct a sample. Rewriting equation (6) in matrix form:

$$\Phi = \mathbf{EB}. \qquad (7)$$

E is the matrix of the first six eigenvectors and **B** is the coefficient matrix to reconstruct the sample population, Φ. Figure 1 shows the first six eigenvectors which were obtained from the still life painting explaining the most sample variations (99.98%) in Φ_λ space.
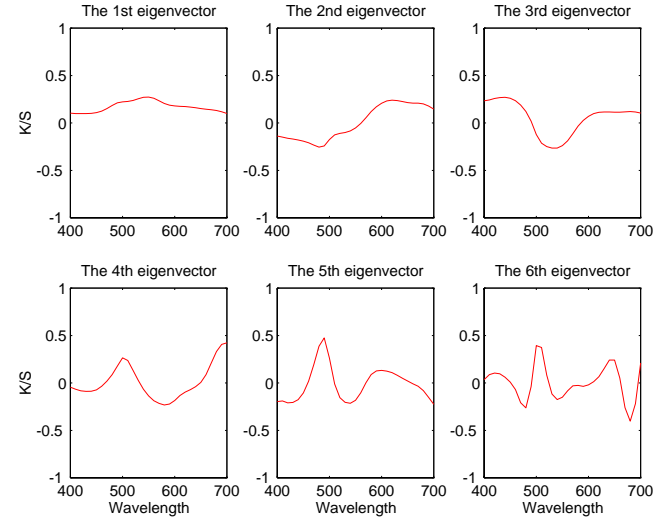

Figure 1. The six eigenvectors obtained from the still life painting.

Colorant Estimation

As shown in Figure 1, the thirty-one components of the eigenvectors are often bipolar; consequently, they are not a set of physical colorants. Furthermore, their corresponding coefficient vectors are also bipolar not representing physical concentrations. Real colorants should have all-positive Φ_λ as their vector components, and the corresponding concentrations should be all-positive. Since an original was created by mixing a set of existing physical colorants at different concentrations and the color mixing operation is mathematically described by Eq. (5), the sampled Φ are distributed in an all-positive space. Rewriting Eq. (5) in matrix form:

$$\Phi = \phi \mathbf{C}, \qquad (8)$$

where ϕ is the matrix of the basis colorants and **C** is the concentration matrix to reconstruct the sample population, Φ. Notice that Eq. (8) can be equated with Eq. (7) in order to obtain the relationship between the eigenvectors and the ϕ_λ of the basis colorants used for creating the original painting. Based on this vision, the relationship between the eigenvectors and the physical basis colorants is merely a linear transformation, or a geometric rotation. Since

$$\Phi = \mathbf{EB} = \phi \mathbf{C}, \qquad (9)$$

this implies that

$$\phi = \mathbf{EBC}^- = \mathbf{EM}, \qquad (10)$$

where **C⁻** stands for the pseudo-inverse of the concentration matrix and **M** is the representation of the matrix product of **B** and **C⁻**. The linear transformation from eigenvectors to physical basis colorants should result in two important properties. First, the rotated eigenvectors should be a set of all-positive vectors. Second, the concentration matrix should have all non-negative entries. These two constraints should result in colorant spectra that are very similar or linearly related to the actual colorants.

This constrained rotation was previously performed by Ohta.[17] His research goal was to estimate the spectral density curves of an unknown dye set for photographic materials using only the spectra of color mixtures such as ANSI IT8 targets. A Monte Carlo method was also used to help identify the most likely dye set. It was a three dimensional vector transformation. This research extends the challenge to six dimensions.

In the current analysis, a constrained-rotation engine using MATLAB as the calculation platform was devised to solve the problem. Since the ultimate goal of this research is to identify a set of printing inks that minimize metamerism between a set of objects and their printed reproduction, the dimensionality is limited to six, corresponding to six printing stations. If the dimensionality of the original Φ is greater than six, or if there is appreciable spectral measurement error, residual errors will result. Hence, goodness metrics are required. The spectral accuracy was quantified by an index of metamerism that consists of both a parameric correction[18] for D50 and the use of CIE94[19] under illuminant A. The colorimetric accuracy is calculated using CIE94 under D50 for the 1931 observer.

Justification of Eigenvector Reconstruction Without Sample Mean

In practice, the measured samples often reveal more than p dimensions due to measurement noise and limitations in the validity of the Kubelka-Munk transformations. Given the p limited dimensions for spectral reconstruction, one should employ Eq. (6) together with the sample mean for better accuracy.[12] That is,

$$\Phi_{\lambda, \text{ sample}} = \sum_{i=1}^{p} b_i e_{\lambda, i} + \Phi_{\lambda, \text{ sample mean}} \quad . \quad (11)$$

The Existence of a sample mean for spectral reconstruction poses several difficulties for this research. First, the sample mean is only a statistical result which specifies the average Φ_λ behavior for the set of samples. The sample mean does not represent any physical colorant. Second, in Eq. (11), the sample mean is acting as an offset vector which impedes the equality relationship in Eq. (9). Since the eigenvectors are the only clue leading to a set of possible colorants, the sample mean must be excluded for maintaining the rotation relationship between eigenvectors and the set of possible colorants which is specified by Eq. (11). Finally, the confidence for excluding the sample mean is that if the dimensionality of sample population is approximately the constrained number of dimensions, then the sample mean approximately resides in the reconstructed sample population. That is, the sample mean can be

approximately expressed as a linear combination by the limited number of eigenvectros. Henceforth, the sample mean in Eq. (11) can be excluded without significant error, i.e., Eq. (6) will be used.

Testing the Constrained-Rotation Engine by a Virtual Sample population

The constrained-rotation engine was first tested for a virtual sample population with three thousand random mixtures created by linear combinations of the six acrylic-paint spectra in Φ_λ space. These six spectra, plotted in Figure 2, were carefully chosen and verified to be independent colorant vectors, i.e., no one colorant vector can be expressed as the linear combination of any other five colorant vectors.
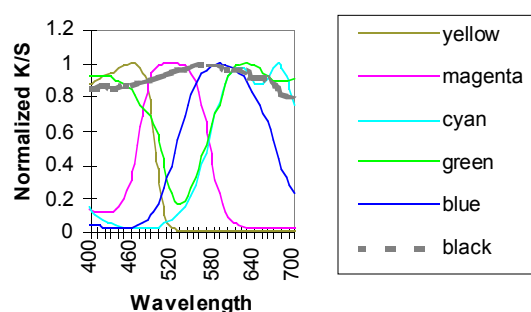

Figure 2. The six acrylic-paints used for generating for the virtual sample population.

Hence, the virtual sample population is ensured to be six dimensional. The corresponding concentration vectors were randomly generated from uniform distributions. Thus, the resulted population of linear combinations of six uniform distributions is approximately multivariate normally distributed (convolution of six uniform distributions is approximately normal).

Given that real sample populations can be confounded by processes and measurements, the idea of using the virtual sample population to test the constrained-rotation engine is to provide a noise free sample population. This ensures that the rotated eigenvectors with all-positive vector components as the estimated colorant spectra should be identical or linearly related to the six acrylic-paint spectra if the proposed vector transformation theory expressed as Eq. (10) is valid.

Six-eigenvector reconstruction without a sample mean vector, based on Eq. (6), yielded approximately zero spectral errors, hence, zero colorimetric errors since full dimensionalty was employed. Then, an arbitrary set of six colorant vectors were used as the initial values for the constrained-rotation engine. The resultant all-positive eigenvectors as the set of estimated colorant vectors are identical to the original six acrylic-paint spectra. In addition, another set of six block spectra evenly spaced within 400 nm to 700 nm representing an initial colorant vectors was utilized and the resulted estimated colorant vectors were also identical to the

six acrylic-paint spectra. This is surprising since the vector transformation can not be unique; multiple solutions should exist. Whereas, those solution are linearly related with each other since they all are the linear transformations of the six eigenvectors. Thus, the first test shows that the constrained-rotation engine is able to converge to an all-positive representation of the eigenvectors.

Colorant Estimation for the Kodak Q60C Target

The second verification was performed on the Kodak Q60C, a photographic reflection target that is a precursor to ANSI IT8 target. Three eigenvector reconstruction should yield low spectral and colorimetric errors corresponding to the fact that it is manufactured by three dyes. Kulbeka-Munk transformation for transparent material was used to transform reflectance factor to absorption. The spectral and colorimetric accuracy, based on the three-eigenvector reconstruction, is shown in Table I. Ideally, this is a three dimensional problem. Whereas, the spectral and colorimetric accuracy is confounded by the manufacturing, processing and measuring noise, and the model accuracy limitations of Kulbeka-Munk theory.

Table I. The Spectral and Colorimetric Accuracy of the Three-Eigenvector Reconstruction for the Kodak Q60C.

	CIE94	Metameric Index
Mean	0.48	0.19
Std Deviation	0.2	0.17
Maximum	1.12	1
Minimum	0	0

Uncovering the set of all-positive eigenvectors as the estimated dye spectra of the Q60C was preceded by using the first eigenvectors of cyan, magenta, and yellow ramps of the Q60C as the initial colorant vectors. The first eigenvector of each ramp, denoted as local eigenvector, is the first statistical estimation of the real dye spectrum.[20] By this approach, the advantage is to get a close solution and help expedite the rotation process. The estimated dye spectra (thick lines) and the local eigenvectors (dotted lines) are plotted in Figure 3. Since the all-positive eigenvectors representing the estimated dyes are an exact linear transformation of the first three eigenvectors, denoted as global eigenvectors determined from the Kodak Q60C target, the spectral and colorimetric performance of estimated dyes is the same as that of global eigenvectors.

It was found that the local eigenvectors were not the exact transformation of the global eigenvectors. The spectral reconstructibility by local eigenvectors was worse than that of the estimated dyes as expected. Furthermore, the broader

absorption bandwidths of the local eigenvectors symbolize the possible impurity contamination during the manufacturing, processing and measuring. On the contrary, the all-positive eigenvectors as the estimated dye spectra showing narrower absorption bandwidths may be close to the real dye spectra based on the support of low spectral and colorimetric errors.
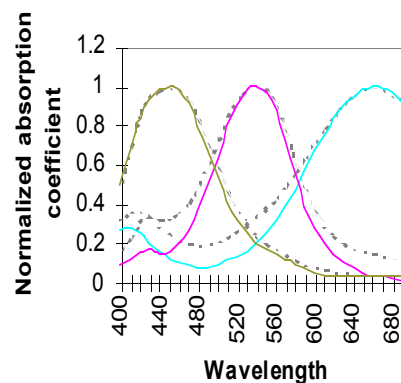

Figure 3. The all-positive eigenvectors as the estimated dye spectra (thick line) and the local eigenvectors (dotted lines).

The testing for the proposed colorant-estimation engine favors the sense of reverse engineering, i.e., uncovering the spectral structures of real colorants. However, for the current research applications, it needs only one reasonable set of colorant spectra which can be used to search through the existing ink database or for a colorant chemist to synthesize the exact inks. Once one exact or similar set of inks is selected, spectral-based printing process can utilize the selected ink set to fulfill the least metameric reproduction. Hence, it is not critical for the proposed colorant estimation engine to converge to the very real colorant spectra which were used to manufacture the colored objects.

Table II. The Spectral and Colorimetric Accuracy of the Six-Eigenvector Reconstruction for the Still Life Painting.

	CIE94	Metameric Index
Mean	0.21	0.18
Std Deviation	0.14	0.16
Maximum	0.75	0.95
Minimum	0.02	0.01

Colorant Estimation for the Still Life Painting

The final verification for the constrained-rotation engine was performed by spectral measurements of the still life

painting mentioned previously. The painting was painted by six independent acrylic-paints whose Φ_λ spectra are plotted in Figure 2. One hundred and twenty-six samples were obtained to represent the entire Φ_λ space of the painting whose six eigenvectors plotted in Figure 1 explaining 99.98% of total variation indicated that this sample population is approximately six dimensional. The spectral and colorimetric accuracy of the six-eigenvector reconstruction is specified in Table II.

Initially, the colorant estimation was intended to directly rotate the six eigenvectors to one set of all-positive representations. The resultant colorant spectra are plotted in Figure 4 and show that there is a colorant (thin dotted line) with various absorption bands across the visible spectral region and the reasonable appearance of the rest of the five colorants. Several sets of colorant vectors were used as the initial estimation for the constrained-rotation engine. The resulting sets of estimated colorants all possessed the similar spectral properties. These initial attempts did not reveal the existence of a neutral colorant judged by the lack of a flat spectrum.
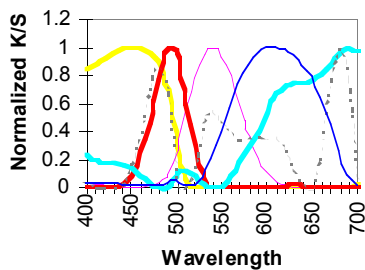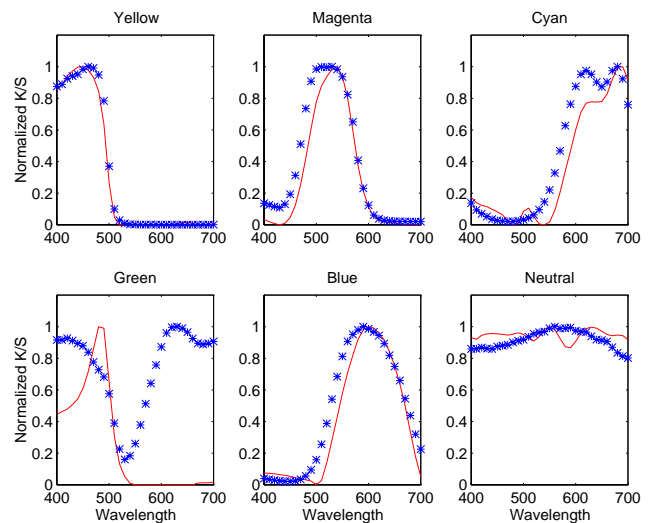

Figure 5. The estimated colorants (solid lines) and the original colorants (astroidal lines) used for the still life painting.

The neutral colorant with an approximately flat spectrum can be approximated by the linear combination of the rest of the five estimated colorants. The lack of a neutral colorant indicated that the rest of the five estimated colorants did not explain sufficient spectral variation. Since the current research aims to uncover one neutral and five chromatic colorants for printing processes, the approach has to constrain the assumption of the existence of the neutral colorant. Hence, the colorant estimation for the still life painting was proceeded by: first, estimate the neutral colorant using linear regression to fit the perfect flat spectrum by the six eigenvectors. Second, rotate the most significant five eigenvectors to their all-positive representations. The resulted estimated colorants should explain a higher degree of spectral variation once the neutral dimension is constrained. The spectral and colorimetric accuracy of the resulted six estimated colorants is shown in Table III and their spectral curves (solid lines) are simultaneously plotted with the six original colorants (astroidal lines) used for the still life painting in Figure 5.

Evaluation and Analysis

The constrained-rotation of the six eigenvectors obtained from the still life painting yielded a reasonable set of estimated colorants. Judging from them, most colorants have similar spectral properties to the original which were utilized to create the still life painting. Whereas, the spectral property similar to green is absent in this set of estimated colorants. Instead, the constrained-rotation process gave out a spectrum equivalent to a yellow colorant. This can be attributed to the sampling error due the usage of larger aperture size of spectrophotometer which violates the additive assumption in Φ_λ space.

Consider the field of view of a spectrophotometer shown in Figure 6. Once the spectrophotometer samples at

Figure 4. The six all-positive eigenvectors as the estimated colorants for the still life painting.

Table III. The Spectral and Colorimetric Accuracy of the Six Estimated Colorants for the Still Life Painting.

	CIE94	Metameric Index
Mean	0.22	0.21
Std Deviation	0.16	0.18
Maximum	0.92	1.01
Minimum	0.02	0.01

a spot where two or more contiguous colors are within its field of view, the reading is equivalent to the additive result of the spectral reflectance factors confined by the spot whose spectral reflectance is contributed by that of the two or more colors.
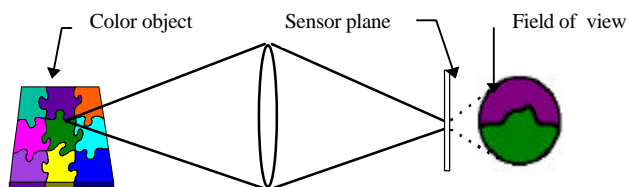

Figure 6: The possible field of view of a spectrophotometer.

Furthermore, the additive operation in reflectance space undergoing a nonlinear transformation such as Eqs. (2) and (4) results in the additive operation undefined in Φ_λ space. As a consequence, the behavior of samples in Φ_λ space are not predictable by the linear model of Eq. (5). This type of sampling error can be reduced when the spectral reflectance factor of a color object is estimated by a high resolution CCD camera with very narrow field of views for each pixel.

The under-sampling of green and over-sampling of yellow-orange color are the other source of errors which cause the estimated colorants do not agree with the original colorant. Once the sample gamut is approximately uniform, i.e., each color has approximately equal probability of occurrence in the sample population, this type of error is minimized.

Since the sample gamut of the still life was carefully controlled to be as uniform as possible, the lack of green colorant is mainly caused by the violation of the additive assumption in Φ_λ space due to the large field of view of the spectrophotometer.

Conclusions

An algorithm was developed for the colorant estimation of original objects through vector analysis and principal component analysis. The relationship between basis colorants and eigenvectors is elucidated by performing a constrained linear transformation. Since the basis colorants used for creating original objects can be statistically uncovered with sufficient accuracy, the color reproduction at the synthesis stage gains the maximum capability to spectrally reconstruct a sample from the original. Therefore, the metamerism between the reproduction and original is minimized.

Acknowledgments

References

1. P.D. Burns, Analysis of Image Noise in Multispectral Color Acquisition, Ph.D. Dissertation, Rochester Institute of Technology, 1997.
2. P.D. Burns and R.S. Berns, Error propagation in color signal transformations, *Color Res. Appl.* **22**, 280-289 (1997).
3. P.D. Burns and R.S. Berns, Modeling colorimetric error in electronic image acquisition, *Proceedings OSA annual meeting*, in press (1997).
4. P.D. Burns and R.S. Berns, Analysis of multispectral image capture, *Proc. 4th IS&T/SID Color Imaging Conference,* 19-22 (1996).
5. K. Iino and R.S. Berns, A spectral based model of color printing that compensates for optical interactions of multiple inks, AIC Color 97, *Proc. 8th Congress International Colour Association*, 610-613 (1997).
6. K. Iino and R.S. Berns, Building color management modules using linear optimization I. desktop color system, *J. Imag. Sci. Tech.* **42**, 79-94 (1998).
7. K. Iino and R.S. Berns, Building color management modules using linear optimization II. prepress system for offset printing, *J. Imag. Sci. Tech.,* in press (1998).
8. A. Di Bernardo and C. Matarazzo, Hi-Fi Color, *American Ink Maker*, 37-54, April 1995.
9. T. Kohler and R. S. Berns, Reducing Metamerism and Increasing Gamut Using Five or More Colored Inks, *IS&T's 3rd Technical Symposium on Press, Proofing & Printing*, 24-28 (1993).
10. P. Kubelka and F. Munk, Ein Beitrag zur Optik der Farbanstriche, *Z. Tech. Phys. (German)* **12**, 593-601 (1931).
11. P. Kubleka, New Contribution to the Optics of Intensely Light-Scattering Materials. Part I*, J. Opt. Soc. of Am.* **38**, 448-457 (1948).
12. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*, 3rd Ed. Prentice Hall, New York, 459-486, 1992.
13. J. L. Simonds, Application of Characteristic Vector Analysis to Photographic and Optical Response Data, *J. Opt. Soc. of Am.* **53**, No. 8, 968-974, (1963).
14. D. B. Judd, D. L. MacAdam, and G. Wyszecki, Spectral Distribution of Typical Daylight as a Function of Correlated Color Temperature, *J. Opt. Soc. of Am.* **54**, No. 8, 1031-1040, (1964).
15. T. Jaaskelainen, J. Parkkinen, and S. Toyooka, Vector-Subspace Model for Color Reproduction, *J. Opt. Soc. of Am.* **7**, No. 4, 725-730, (1990).
16. D. Y. Tzeng and R. S. Berns, Proper Space for Linear Modeling Techniques, in progress (1998).
17. N. Ohta, Estimating Absorption Bands of Component Dyes by Means of Principal Component Analysis, *Anal. Chem.* **45**, 553-557 (1973).
18. H. Fairman, Metameric Correction Using Parameric Decomposition, *Color Res. Appl.* **12**, 261-265 (1987).
19. CIE Technical Report 116, Industrial Color-Difference Evaluation, CIE, Vienna, 1995.
20. R. S. Berns and M. J. Shyu, Colorimetric Characterization of A Desktop Drum Scanner Using A Spectral Model, *J. Elec. Imag.* **4**, 360-372 (1995).