

Accuracy of Various Types of Neugebauer Model

Robert Rolleston and Raja Balasubramanian

Webster Research Center, Xerox Corporation, Webster, New York

Introduction

In characterizing a color hardcopy device, it is necessary to establish the relationship between the input signals that drive the device and the colorimetric response of the device to these signals. Most printing devices are not adequately characterized by a simple linear transformation; hence one is left with the choice of either measuring the printer's colorimetric response at numerous points throughout the input signal space, or deriving a model to predict the printer's response. In this paper, we examine a model based approach to device characterization. In particular, we focus on a model developed by Hans Neugebauer¹ for binary color printers employing a rotated halftone dot screen. In their simplest form, the so called "Neugebauer equations"

are used to predict the broadband reflectance of a halftone pattern printed on paper. In this paper, we investigate the accuracy of the basic Neugebauer equations and several of its modifications. These equations involve some basic constraints² which we will assume to be fulfilled throughout the discussion.

An important application of the Neugebauer model is the calibration of binary color printers. Calibration, which requires a mapping from colorimetric signal space to the printer signal space, is the inverse problem of device characterization. Hence, calibration requires that the Neugebauer model be inverted. Since the Neugebauer equations are nonlinear, the inversion is not trivial, and requires numerical or statistical approaches (see [3] for examples). In this paper, we do not deal with the inverse

calibration problem; rather, we focus only on the forward characterization problem. We evaluate and compare the use of various Neugebauer models to predict the colorimetric response of a Xerox 5775 color printer. This printer uses xerographic technology with four colorants, and has a resolution of 406×1624 dpi.

2. Use of Neugebauer Equations to Model a Printer

A wide range of colors are produced by placing the printer colorants on paper in varying amounts. The Neugebauer equations can be used to predict these colors by interpolating over a small set of known samples. In the case of a four color printer (or a CMYK printer), this known set of samples consists of combinations of the four colorants; cyan (C), magenta (M), yellow (Y), black (K), and the white paper (W). The set of known samples is limited to the colors produced when these colorants are combined in amounts of 0% or 100% area coverage. That is, only the sixteen colors which are produced by placing the minimum and maximum amounts of C, M, Y, and K on the paper are used to predict all other colors.

According to the Neugebauer model, the color produced by a particular set of C, M, Y, K printer values is predicted by a linear combination of the sixteen known (*i.e.* measured) colors. This linear combination is expressed as:

$$X_{cmyk} = \sum_{i=1}^{16} w_i X_i, \quad (2.1)$$

where X_{cmyk} is the broadband reflectance of the unknown color, and X_i are the broadband reflectances of the 16 known colors. We will refer to X_i as the i -th element of the set S_X :

$$X_i \in S_X = \left\{ \begin{array}{l} X_W, X_C, X_M, X_Y, X_K, X_{CM}, X_{CY}, X_{CK}, X_{MY}, X_{MK}, X_{YK}, \\ X_{CMY}, X_{CMK}, X_{CYK}, X_{MYK}, X_{CMYK} \end{array} \right\} \quad (2.2)$$

where X_W is the broadband reflectance of the white paper, X_C is the broadband reflectance of the 100% cyan sample, X_{CM} is the broadband reflectance of the sample with 100% cyan and 100% magenta, *etc.* In most cases a color will be described by a set of three broadband reflectance measurements, X, Y, Z . The two additional color values, Y and Z , are defined by replacing X with Y and Z respectively in Equations (2.1) and (2.2). The broadband reflectances X_i, Y_i, Z_i , can be the reflectances measured through a set of red, green, and blue filters, or a set of cyan, magenta, and yellow filters, or any calculated set of responses, including the XYZ tristimulus values used here. The resulting triplet of values, in this case the XYZ tristimulus values, define a colorimetric specification of the printed sample. Finally, the corresponding weights w_i in (2.1) are given by:

$$w_i \in S_W = \{(1-c)(1-m)(1-y)(1-k), c(1-m)(1-y)(1-k), m(1-c)(1-y)(1-k), y(1-c)(1-m)(1-k), \\ k(1-c)(1-m)(1-y), cm(1-y)(1-k), cy(1-m)(1-k), ck(1-m)(1-y), my(1-c)(1-k), \\ mk(1-c)(1-y), yk(1-c)(1-m), cmy(1-k), cmk(1-y), cyk(1-m), myk(1-c), cmyk\}, \quad (2.3)$$

where c, m, y , and k are the relative dot areas of the four colorants that produce the unknown color. We note from (2.1) and (2.3) that the basic Neugebauer equations are mathematically equivalent to performing a linear interpolation in four dimensional colorant space.

Due to nonlinearities in the printer, the dot areas c, m, y, k are not necessarily linearly related to the digital values C, M, Y, K input to the printer. Various empirical techniques have been developed to estimate dot areas.^{4,5} With our approach, the dot areas were obtained by measuring the colorimetric response of the printer to samples of pure colorants, and choosing the dot areas that minimized the perceived colorimetric difference between the measured and predicted colors. Sixteen samples were printed and measured of each of the cyan, magenta, yellow, and black colorants, with area coverages ranging from 0% (*i.e.* plain paper) to 100% (*i.e.* solid area coverage). The Neugebauer equations (2.1) were then used to predict the color of each of the measured patches. When using the Neugebauer equations along a line of pure colorant, the prediction reduces to a simple one dimensional linear interpolation as a function of the colorant dot area. The cyan dot area $c(i)$ for the i -th sample was chosen to minimize a color error metric $\Delta E_{c(i)}$ between the measured color and predicted color of that sample. In order for the error metric to be perceptually meaningful, we chose $\Delta E_{c(i)}$ to be the CIE $L^*a^*b^*$ ΔE error. This error minimization process was repeated for all samples, $1 \leq i \leq 16$, and for all colorants c, m, y, k .

To test the accuracy of the models described in this paper, 1000 test patches randomly distributed in CMYK space were printed and measured. Each of the models were then used to predict the colorimetric values of each of the 1000 samples. The performance of a model was quantified by the mean, maximum value, and standard deviation of the CIE $L^*a^*b^*$ ΔE errors between predicted and measured colors. For ease of comparison, these results are tabulated in Sec. 6. The average error incurred in modeling the printer response with the basic Neugebauer equations (2.1) and (2.3), was $\Delta \bar{E} = 7.4$, which is an unacceptable result.

3. Yule-Nielsen Modified Neugebauer Equations

One of the first modifications made to the Neugebauer equations was an attempt to account for the scattering of light within the paper substrate.⁶ Due to such scattering effects, light which enters the paper through an area with no colorant, may leave the paper through an area which is covered with colorant, and light which enters the paper through an area which is covered by colorant may leave the paper through an area which is not covered by colorant. This scattering has the result of producing a nonlinear relationship between the relative area covered with colorant, and the relative reflectance of the same area. For example; a halftone pattern with a 50% dot area coverage will produce a relative reflectance less than the average of the 0% and 100% coverage reflectances.

A simple way to model the nonlinear relationship between area coverage and reflectance is with the use of a power law,^{6,7} though other more sophisticated charac-

terizations of light scattering have also been proposed.⁸ With the power law model, the Neugebauer equations discussed above in Section 2, can be modified to operate in a nonlinear reflectance space. The predicted broadband reflectance is now given by:

$$X_{cm\bar{y}k}^{1/n} = \sum_{i=1}^{16} w_i X_i^{1/n} \quad (3.1)$$

where n is often referred to as the Yule-Nielsen factor. The measured broadband reflectances X_i are the same as given above, but note that they are raised to the power of $1/n$ before performing the weighted sum.

The incorporation of the nonlinear relationship adds another parameter n into the Neugebauer equations. The Yule-Nielsen factor can be chosen to correlate with some known physical properties of the paper⁹ or it can be used as a free parameter to minimize some cost function.¹⁰ For all the examples in this paper, the value of n was chosen to minimize an error function in the following manner. For a fixed n , the optimal dot areas were obtained using the procedure described in Sec. 2. This procedure yielded CIE L*a*b* ΔE values for each of the pure tone samples. The average of these ΔE 's gave a composite error metric $\Delta \bar{E}(n)$ for the given n . This process was repeated for a sweep of values of n , $1.0 \leq n \leq 8.0$, and the n was picked that minimized $\Delta \bar{E}(n)$.

With the addition of the Yule-Nielsen modification, the evaluation of the Neugebauer equations becomes equivalent to doing a linear interpolation in a nonlinear space. While this modification to the Neugebauer Equations noticeably reduces the errors, this model is still a poor predictor of printer response. In our experimentation, the average error from this model was $\Delta \bar{E} = 4.5$.

4. Spectral Yule-Nielsen Modified Neugebauer Equations

Because colored inks do not have constant reflectances throughout the visible spectrum, it has been argued that broadband reflectance measurements are inappropriate for a Neugebauer equation model of a color printer.^{11, 12} Dealing with the full set of spectral data for the printer model adds considerable computational complexity to the problem. Both the models discussed above are computationally quite fast. The first requires only a four space linear interpolation for X , Y , and Z . The second requires a power law evaluation to map each of the three color values back into a linear colorimetric space. The power law evaluation of the measured color values is done once during the calibration process.

In this investigation, the spectral data was measured over the range of 400 nm–700 nm at 10 nm increments. With this type of spectral resolution, operating with spectral data requires on the order of 30 data values (instead of the three broadband reflectance color values discussed above) for each of the measured sample points. Evaluation of a predicted spectral reflectance curve will therefore involve on the order of 30 four space linear interpolations. In addition, there will be some type of summation over the spectral data to obtain a colorimetric triplet of values.

The use of a spectral data does not change the weights of the Neugebauer Equations, but the three broadband color values (referred to as X , Y , and Z above) are now replaced with a wavelength dependent sampling of the reflectance curves. The linear combination used to predict the spectral reflectance is given by:

$$R(\lambda)_{cm\bar{y}k}^{1/n} = \sum_{i=1}^{16} w_i R_i(\lambda)^{1/n} \quad (4.1)$$

where the weights w_i are the same as in (2.3). The spectral samples $R_i(\lambda)$ are the i -th member of the set \mathbf{S}_R given by

$$R_i(l) \in \mathbf{S}_R = \{R_W(l), R_C(l), R_M(l), R_Y(l), R_K(l), R_{CM}(l), R_{CY}(l), R_{CK}(l), R_{MY}(l), R_{MK}(l), R_{YK}(l), R_{CMY}(l), R_{CMK}(l), R_{CYK}(l), R_{MYK}(l), R_{CMYK}(l)\} \quad (4.2)$$

where $R_W(\lambda)$ is the spectral reflectance of the white paper, $R_C(\lambda)$ is the broadband reflectance of the 100% cyan sample, $R_{CM}(\lambda)$ is the broadband reflectance of the sample with 100% cyan and 100% magenta, *etc.* After this predicted spectral reflectance curve is calculated, three different weighted sums of the reflectance are calculated to yield a colorimetric triplet specification of the color. In this case the three weighted sums use the color-matching functions of the 1931 CIE Standard Colorimetric Observer to obtain tristimulus values X , Y , and Z .

Using spectral information in the Yule-Nielsen modified Neugebauer Equations led to a significant improvement in accuracy. In this case the average error was $\Delta \bar{E} = 2.7$, an error that may be considered to be within the noise level of the system.

5. Cellular Neugebauer Equations

Up to this point the Neugebauer equations have been limited to using only the set of 16 samples obtained from the various combinations of 0% and 100% colorant. With the addition of partial dot area coverages, the Neugebauer equations can be modified to make use of more than the small set of 16 sample prints. The addition of these partial overprint samples is equivalent to partitioning the CMYK space into rectangular cells and employing the Neugebauer equations within each cell. Hence, this model is referred to as the cellular Neugebauer model.¹³

Using only 0% and 100% dot area coverages with four colorants resulted in $2^4 (=16)$ known sample points in the model. In this section we will discuss the effects of using the 0%, 50%, and 100% colorant signals and their combinations, which will require the use of $3^4 (=81)$ known sample points, and the effects of using 0%, 25%, 50%, 75%, and 100% colorant signals and their combinations, which will require the use of $5^4 (=625)$ known sample points. As before, the cellular Neugebauer equations will have the geometric interpretation of linear interpolation in 4-space, but now the interpolation is within smaller subcells, rather than within the entire CMYK space. It should be noted that the intermediate colorant levels are not limited to those chosen here. It is also not required that the same cellular division occur along each of the four colorant

axis; rather it is possible to choose the cellular division to lie on a non-cubic (but rectangular) lattice.

While Heuberger *et al.*¹³ extended the basic broadband Neugebauer model to the cellular case, to our knowledge, there has been no attempt to incorporate the Yule-Nielsen correction or spectral measurements into a cellular framework. In the following subsections, we will describe the cellular approach as introduced in [13], and also investigate the effect of employing Yule-Nielsen correction, and spectral measurements.

5.1 Cellular Broadband Neugebauer Equations

Suppose a set of C, M, Y, K digital values result in a set of dot areas c, m, y, k , which may be represented as a point in 4 space. This point will fall in a rectangular cell which is bounded by the lower and upper extrema, denoted $C_b, C_u, M_b, M_u, Y_b, Y_u, K_b, K_u$, along each of the 4 axes. Mathematically, we may specify C_l and C_u as being the two points along the cyan axis that satisfy the constraints:

$$0 \leq C_l \leq c < C_u \leq 1; \quad C_b, C_u \in I_c, \quad (5.1)$$

where I_c is the set of points along the cyan axis that specify the cellular division. Analogous definitions hold for the magenta, yellow and black coordinates. In order to perform the interpolation within a given cell, we need to normalize the dot area values c, m, y, k to occupy the interval [0, 1] within that cell. The normalized cyan value is given by:

$$c' = \frac{c - C_l}{C_u - C_l}, \quad (5.2)$$

with analogous expressions for $m', y',$ and k' . The Yule-Nielsen modified Neugebauer equation (3.1) for the X component then becomes:

$$X_{cmyk}^{1/n} = \sum_{i=1}^{16} w_i' (X_i')^{1/n}, \quad (5.3)$$

where

$$X_i \in \mathbf{S}_{X'} = \left\{ \begin{array}{l} X_{C_l M_l Y_l K_l}, X_{C_u M_l Y_l K_l}, X_{C_l M_u Y_l K_l}, X_{C_l M_l Y_u K_l}, X_{C_l M_l Y_l K_u}, X_{C_u M_u Y_l K_l}, X_{C_u M_l Y_u K_l}, X_{C_u M_l Y_l K_u}, \\ X_{C_l M_u Y_u K_l}, X_{C_l M_u Y_l K_u}, X_{C_l M_l Y_u K_u}, X_{C_u M_u Y_u K_l}, X_{C_u M_u Y_l K_u}, X_{C_u M_l Y_u K_u}, X_{C_l M_u Y_u K_u}, X_{C_u M_u Y_u K_u} \end{array} \right\}, \quad (5.4)$$

are the set of broadband X reflectances of the 16 cornerpoints of the cell in 4 space that contains the point $cmyk$. The weights w_i' are given by (2.3) with c, m, y, k being replaced respectively by c', m', y', k' , as given in (5.2). Definitions analogous to (5.3) and (5.4) hold for the remaining two broadband reflectances Y and Z.

There is a noticeable improvement in performance when using a cellular model. The improvement is proportional to the number of cellular divisions (or equivalently, inversely proportional to the size of the cells). This is not surprising, as the cellular Neugebauer model is geometrically equivalent to linear interpolation between the cornerpoints of the cells, and smaller cells yield a smaller interpolation error. With Yule-Nielsen correction, the

average error in the model's prediction dropped from $\Delta\bar{E} = 4.5$ in the non cellular case to $\Delta\bar{E} = 3.1$ in the cellular case with 3^4 nodes, and $\Delta\bar{E} = 2.6$ in the cellular case with 5^4 nodes.

5.2 Cellular Spectral Neugebauer Equations

If we use spectral, rather than broadband, measurements to predict a color, then we have the cellular spectral Neugebauer equation:

$$R(\lambda)_{cmyk}^{1/n} = \sum_{i=1}^{16} w_i' R_i'(\lambda)^{1/n}, \quad (5.5)$$

where

$$R_i'(\lambda) \in S_{R'} = \left\{ \begin{array}{l} R_{C_l M_l Y_l K_l}(\lambda), R_{C_u M_l Y_l K_l}(\lambda), R_{C_l M_u Y_l K_l}(\lambda), R_{C_l M_l Y_u K_l}(\lambda), R_{C_l M_l Y_l K_u}(\lambda), R_{C_u M_u Y_l K_l}(\lambda), \\ R_{C_u M_l Y_u K_l}(\lambda), R_{C_u M_l Y_l K_u}(\lambda), R_{C_l M_u Y_u K_l}(\lambda), R_{C_l M_u Y_l K_u}(\lambda), R_{C_l M_l Y_u K_u}(\lambda), \\ R_{C_u M_u Y_u K_l}(\lambda), R_{C_u M_u Y_l K_u}(\lambda), R_{C_l M_u Y_u K_u}(\lambda), R_{C_l M_l Y_u K_u}(\lambda), R_{C_u M_u Y_u K_u}(\lambda) \end{array} \right\}, \quad (5.6)$$

are the set of spectral reflectance functions of the 16 cornerpoints of the cell containing the point $cmyk$, and w_i' is as defined in Sec. 5.1.

For spectral models, there is no great improvement when using the cellular approach, provided the optimum Yule-Nielsen factor is chosen in each case. The choice of Yule-Nielsen value is more critical for the non-cellular model. The average ΔE error for 16, 81, and 625 sample points was similar, and was less than 3.1. This is probably

the most interesting result of the paper, and implies that the accuracy of the spectral Yule-Nielsen Neugebauer models (whether noncellular or cellular) is within the noise level of the system.

6. Experimental Results

We divide our results into 12 cases as follows:

Case 1: Basic Neugebauer Equations;

- Case 2: Yule-Nielsen modified Neugebauer Equations ($n = 2.5$);
- Case 3: Spectral Neugebauer Equations;
- Case 4: Spectral, Yule-Nielsen modified Neugebauer Equations ($n = 5.5$);
- Case 5: Cellular Neugebauer Equations, 3^4 cellular division;
- Case 6: Cellular Yule-Nielsen modified Neugebauer Equations, 3^4 cellular division ($n = 7.0$);
- Case 7: Cellular Neugebauer Equations, 5^4 cellular division;
- Case 8: Cellular Yule-Nielsen modified Neugebauer Equations, 5^4 cellular division ($n = 6.0$);
- Case 9: Cellular spectral Neugebauer Equations, 3^4 cellular division;
- Case 10: Cellular spectral Yule-Nielsen modified Neugebauer Equations, 3^4 cellular division ($n = 7.0$);
- Case 11: Cellular spectral Neugebauer Equations, 5^4 cellular division;
- Case 12: Cellular spectral Yule-Nielsen modified Neugebauer Equations, 5^4 cellular division ($n = 6.0$).

For each of the above 12 models, the average, maximum, and standard deviation of the CIE $L^*a^*b^*$ ΔE values for the 1000 random samples were calculated, and are shown in Table 6.1. We remark that *Case 1* and *Case 3* are in fact equivalent, because the former involves integrating a spectral reflectance function to 3 tristimulus values and then performing a linear interpolation; while the latter involves a linear interpolation in reflectance space followed by integration to tristimulus values. Since integration and linear interpolation are both linear operators, they are commutable, and the net results are identical.

Figure 6.1 shows plots of the average ΔE for the same 1000 samples as a function of Yule-Nielsen value n for the broadband Neugebauer models. Figure 6.2 shows similar plots for the spectral Neugebauer models. From Fig. 6.1, we see that for the non cellular case, there is a strong dependence of ΔE on n . As the number of cells increases, ΔE decreases as a whole, and the dependence on n becomes weak. From Fig. 6.2, we see that for the spectral models, again the dependence on n decreases as the number of cells increases. However, unlike the broadband case, the non cellular spectral model, with an appropriate choice of Yule-Nielsen value, yields a performance equivalent to the cellular spectral models. It must be emphasized that in the spectral case, the choice of Yule-Nielsen value is crucial to the success of the non-cellular model.

7. Conclusions

We have investigated the use of the Neugebauer model in predicting the colorimetric response of a printer. Several variations of the basic model that have been proposed in the literature were evaluated and compared. In addition, the cellular model was extended to include Yule-Nielsen correction and spectral measurements. As is seen in the experimental results, introduction of a Yule-Nielsen correction always improves model performance. If broadband measurements are used to perform the Neugebauer interpolation, considerable improvement may be obtained by using a cellular model. If spectral measurements are used, however, there is little to be gained by going to a cellular

framework. This shows that a fairly accurate printer model can be generated with only a few spectral measurements.

Table 6.1 Average, maximum, and standard deviation of CIELAB $L^*a^*b^*$ ΔE errors for the 12 cases.

case	ΔE_{avg}	ΔE_{max}	σ
1	7.414	15.52	2.644
2	4.528	11.47	1.901
3	7.414	15.52	2.644
4	2.700	9.709	1.341
5	4.194	11.16	1.771
6	3.059	9.493	1.513
7	2.845	8.764	1.372
8	2.600	8.395	1.309
9	4.194	11.16	1.771
10	3.065	10.05	1.544
11	2.845	8.764	1.372
12	2.615	8.708	1.339

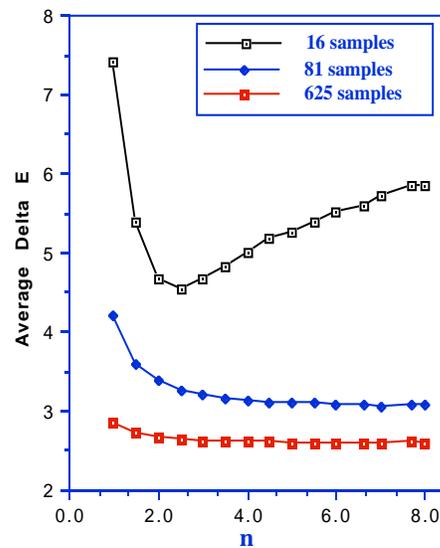


Figure 6.1 Average ΔE as a function of Yule-Nielsen correction n for broadband Neugebauer models with $2^4 = 16$, $3^4 = 81$, and $5^4 = 625$ samples.

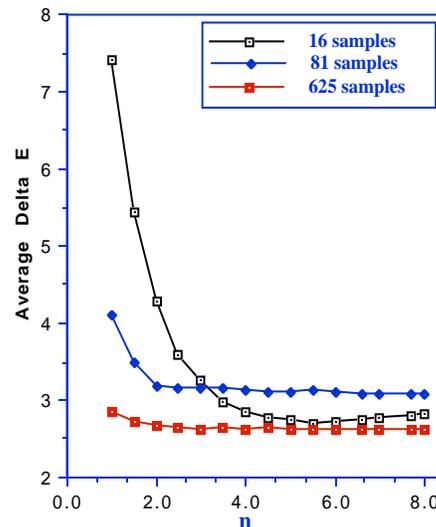


Figure 6.2 Average ΔE as a function of Yule-Nielsen correction n for spectral Neugebauer models with $2^4 = 16$, $3^4 = 81$, and $5^4 = 625$ samples.

8. References

1. H. E. J. Neugebauer, "Die Theoretischen Grundlagen des Mehrfarben-edruckes," *Zeitschrift Wissenschaften Photography*, 1937, pp. 73-89.
 2. *Principles of Color Reproduction*, John A. C. Yule, John Wiley & Sons, New York, 1967.
 3. *Neugebauer Memorial Seminar on Color Reproduction*, Kazuo Sayanagi, Editor, SPIE Vol. **1184**, 1989.
 4. I. Pobboravsky and Milton Pearson, "Computation of Dot Area Required to Match a Colorimetrically Specified Color Using the Modified Neugebauer Equations," *1972 TAGA Proceedings*, pp. 65-77.
 5. A. Froslev-Nielsen, "Some comments on the formulation and determination of dot area in halftone prints on paper," *Advances in Printing Sci. Tech.*, W. H. Banks Editor, 1969, pp. 141-155.
 6. J. A. C. Yule and W. J. Nielsen, "The Penetration of light into Paper and its Effect on Halftone Reproduction," *1951 TAGA Proceedings*, pp. 65-76.
 7. F. R. Clapper and J. A. C. Yule, "Reproduction of Color with Halftone Images," *1955 TAGA Proceedings*, pp. 1-12.
 8. Martin Maltz, "Light Scattering in Xerographic Images," *J. Applied Photographic Engineering*, Vol. **9**, No. 3, June 1983, pp. 83-89.
 9. F. R. Ruckdeschel and O. G. Hauser, "Yule-Nielsen effect in printing: a physical analysis," *Applied Optics*, Vol **17**, No. 21, November 1978, pp. 3376-3383.
 10. Milton Pearson, "n Value for General Conditions," *1980 TAGA Proceedings*, pp. 415-425.
 11. J. A. Stephen Viggiano, "The Color of Halftone Tints," *1985 TAGA Proceedings*, pp. 647-661.
 12. J. A. Stephen Viggiano, "Modeling the Color of Multi-Colored Halftones," *1990 TAGA Proceedings*, pp. 44-62.
 13. K. J. Heuberger, Z. M. Jing, and S. Persiev "Color Transformations and Lookup Tables," *1992 TAGA/ISCC Proceedings*, Vol **2**, pp. 863-881.
-