

Six Color Printer Characterization Using an Optimized Cellular Yule–Nielsen Spectral Neugebauer Model

Yongda Chen[▲], Roy S. Berns^{▲†} and Lawrence A. Taplin

Munsell Color Science Laboratory, Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, New York, USA

A printer characterization attempts to map, in both directions, corresponding points in colorant and colorimetric spaces. Two limiting approaches are used: analytical models based on a small number of samples, and direct measurement and interpolation requiring many samples. For six-color printers, the former approach often has insufficient accuracy whereas the latter approach requires an excessive number of samples. An intermediate approach was used to characterize a CMYKGO ink jet printer, the Cellular-Yule-Nielsen-Spectral-Neugebauer (CYNSN) model. This model included an optimized Yule-Nielsen n value and one-dimensional look-up tables between digital data and effective area coverage for each colorant. Each colorant was divided into three subspaces, or cells, requiring the selection of two intermediate values and fixed endpoints of 0% and 100% effective area coverage. An optimization was performed that determined these intermediate values by minimizing the maximum spectral error when one-colorant CYNSN models were used to predict 256-step ramps. This technique enabled a considerable reduction of the total number of required samples from several hundreds of thousands to 4,096, the required number of cellular Neugebauer primaries. Of these colors, only 1,024 could be printed; the remainder was non-printable due to inkblots. A third optimization synthesized the spectral properties of the non-printable cellular primaries using weighted spectral regression, the weighting a function of colorant-space location. The CYNSN model based on these three optimizations was able to predict 600 random colors sampling the colorimetric gamut to an average spectral RMS error of less than 0.5% and ΔE_{00} of less than 1.0. The color gamut achievable using the synthesized spectra was 54% larger in colorant space and 15% larger in CIELAB space than that achievable when limiting the CYNSN model to printable cellular primaries.

Journal of Imaging Science and Technology 48: 519–528 (2004)

Introduction

The proliferation of color printers in both home and office environments is a reality. In order to better control and improve color reproduction accuracy, there is a continuing need to characterize a given printer in a colorimetric color space such as CIE XYZ or $L^*a^*b^*$. Developing a printer transfer function, which maps points in colorant space to points in the colorimetric space and the reverse, is known as printer characterization, and is used to populate a color lookup table (CLUT), facilitating efficient color processing via multi-dimensional linear interpolation. There are two limiting approaches to printer characterization. In the first approach, prints are prepared corresponding to a sampling of colorant space. Each sample is measured,

usually using a spectrophotometer, and colorimetric data are assigned. The colorant and colorimetric data are used to populate the CLUT, most commonly using linear or nonlinear interpolation. For example, the seminal research by Hung¹ sampled colorant space at five levels per colorant (ink). For a three-colorant device, 125 samples were printed and measured ($5^3 = 125$). Next nonlinear interpolation was used to populate a CLUT with 35,937 entries ($3^3 = 35,937$). If this approach had been used in the current research, 15,625 samples would need to be printed and measured ($5^6 = 15,625$). Irrespective of how clever one can be in sampling colorant space to reduce the total number of samples, or in recognizing that many colorant amounts result in the identical colorimetric value (particularly colors with significant amounts of black), it is clear that for six-ink printing, a very large number of samples are required. The second approach is to model the printing process analytically, where colorimetric data are predicted from colorant data. Analytical models, such as those summarized by Wyble and Berns,² require only a few samples. A six-color model can be developed using less than 200 samples,³ for example. The model would then be used to populate the CLUT, either using the Hung nonlinear interpolation approach or inverting the model numerically for each CLUT entry.

Intuitively, we expect a tradeoff between the number of samples and CLUT accuracy. A characterization using

Original manuscript received February 12, 2004

▲ IS&T Member

† berns@cis.rit.edu

©2004, IS&T—The Society for Imaging Science and Technology

many thousands of samples usually has higher accuracy than one using hundreds of samples. Simply, there are more colorant space locations with known colorimetric coordinates. Although a perfect model would yield the identical results, color printing models are rarely, if ever, perfect. Predictions are often less accurate than measurements. As stated above, these two approaches are limiting cases. It is certainly legitimate to combine both approaches. That is, colorant space is sampled sparsely creating subspaces. The analytical model is used to predict colorimetric data within the subspace where the model is developed using the known values of the subspace, i.e., the corners of the subspace. This combined approach is quite useful for multi-ink printing and is the subject of this publication.

Analytical Model

The Cellular–Yule–Nielsen–Spectral–Neugebauer (CYNSN) model² was used in this research. The Neugebauer model⁴ is an additive model for multi-ink printing in which a macroscopic colored area is a weighted sum of the individual microscopic colors. The weights are determined from the halftoning algorithm. The microscopic colors are assumed to be identical to the color that results from printing the color uniformly over a macroscopic area. Because of light scattering within the paper, the macroscopic and microscopic equality is false. Yule and Nielsen⁵ found that exponentiating reflectance, in similar fashion to converting reflectance to optical density, greatly improved prediction accuracy. Viggiano⁶ further improved performance by considering the optical mixing over a narrow range of wavelengths. The resulting Yule–Nielsen–Spectral–Neugebauer (YNSN) model is shown below:

$$R_{\lambda} = \left(\sum_i F_i R_{\lambda,i}^{1/n} \right)^n \quad (1)$$

where $R_{\lambda,i}$ is the macroscopic spectral reflectance of the i th color type at 100% area coverage, n is the Yule–Nielsen exponent, and F_i are the fractional area coverages of each microscopic color type. The maximum value of i depends on the number of inks and the halftoning algorithm. For three-color printing that uses rotated screens or frequency modulation, the maximum number is eight, e.g., cyan, magenta, yellow, red, green, blue, black, and paper. That is, three inks printed randomly result in eight unique colors; for six-color printing, it is 64. These colors are known as the Neugebauer primaries. The fractional areas are determined as a product of random variables, shown in Eq. (2). These probabilities when used for printing are attributed to Demichel.⁷

$$F_i = \prod_j \begin{pmatrix} a_j & \text{If ink } j \text{ is in Neugebauer Primary } i \\ (1 - a_j) & \text{If ink } j \text{ is not in Neugebauer Primary } i \end{pmatrix} \quad (2)$$

where a_j is the effective area coverage of ink j . (The term “effective” is used because this area coverage is determined statistically, not optically using reflection microscopy.²) Area coverage is a function of the digital signal, d_j , controlling the amount of ink delivered to the substrate, defined in Eq. (3).

$$a_j = f(d_j) \quad (3)$$

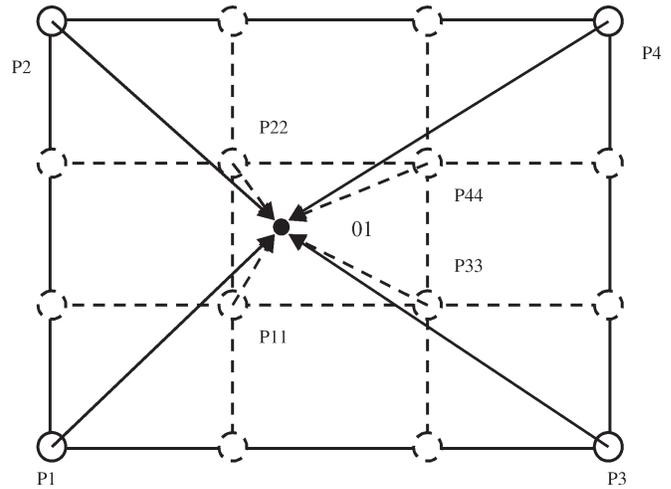


Figure 1. Graphical interpretation of the YNSN and CYNSN models

The n value and the effective area coverage relationships are determined statistically, typically using one-color ramps. See Ref. 2 for an example and below.

From a geometric viewpoint, the YNSN model performs multi-dimensional linear interpolation across $R_{\lambda}^{1/n}$ space, the interpolation weights calculated using Eq. (2).^{8,9} Hueberger¹⁰ recognized that as an interpolation, performance is always improved by reducing the interpolation area, easily achieved by creating subspaces as described above that he called “cells.” Rolleston and Balasubramanian⁸ evaluated the cellular method when using the YNSN model for printer characterization, that is, the CYNSN model. Improvement was significant. In particular, the cellular approach greatly reduced the need for highly accurate analytical models beyond what was typically achieved using the YNSN model.

The value of the cellular approach is shown graphically in Fig. 1. Here, the outer square and solid circles represent the YNSN model and the point $O1$ is calculated by interpolation from four outer corner points, $P1$, $P2$, $P3$, and $P4$, which are represented by the solid circles, the Neugebauer primaries. The whole figure, including solid and dashed circles, represents the CYNSN model in two dimensions. If each ink is printed at four levels, there are more known values and can be used to create cells (subspaces). The corners of each cell are the cellular Neugebauer primaries, or simply cellular primaries. If the cellular model is used to predict the point $O1$, we can use the nearest four cellular primaries ($P11$, $P22$, $P33$, and $P44$). The accuracy improvement of the cellular model is significant because interpolation is performed in a much smaller subspace. Of course, the cost is that more colors need to be printed and measured. Agar and Allebach¹¹ showed the relationship between prediction error and number of cellular primaries. The accuracy of the cellular model can be improved significantly as more primaries are considered, though as noted by Balasubramanian,⁹ there is a diminishing return.

Iino and Berns^{12,13} used the YNSN model to characterize an ink jet printer and a proofing system for offset printing with good success. Balasubramanian^{9,14}

evaluated the SN (Spectral Neugebauer), YNSN, and CYNNSN models when characterizing a four-color electrophotographic printer. He used a typical black-printer strategy so that the number of samples required to create a four-color model was not excessive. As expected, accuracy improved significantly by adding the Yule–Nielsen n value. The addition of the cellular subspaces resulted in modest incremental improvement. Balasubramanian also investigated using weighted linear regression to optimize the spectral properties of the Neugebauer primaries. Rather than using macroscopic measurements, he optimized the spectral properties of those primaries not containing black ink resulting in the best average colorimetric performance. Samples near the primary in colorant space were weighted more heavily than samples far away. The improvement to the YNSN model was similar to adding the cellular approach. Most recently, Imai, Wyble, and Berns¹⁵ used the YNSN model to characterize a CMYK ink jet printer as part of an end-to-end spectral color reproduction system with reasonable success.

Tzeng¹⁶ and Tzeng and Berns¹⁷ used the YNSN model for six-color proofing using cyan, magenta, yellow, black, orange, and green inks. One goal was to use the proofer to simulate different ink sets. Accordingly a spectral model was desired so that proofs could be produced that were minimally metameric. That is, a printing system defined spectrally either computationally or by direct measurement could be proofed via spectral color reproduction. This required numerically inverting the YNSN model. In order to insure convergence and reduce processing time, Tzeng subdivided the six inks into ten four-ink models. For any color, only four inks would be printed. Model accuracy depended on the particular four inks. Taplin and Berns³ extended Tzeng's research and considered all six inks simultaneously. They used an ink jet printer with a small dot size, an error diffusion halftoning algorithm, and heavyweight art paper. This combination enabled all of the 64 Neugebauer primaries to be printed without ink blotting. The YNSN model was used again with reasonable performance accuracy. More sophisticated optimization algorithms were used for model inversion as well as a continuous-tone model to provide reasonable starting values. The printer was used for a spectral color reproduction system for artwork.¹⁸

In the current research, we were again interested in six-ink ink jet printing, but using common printing conditions, such as a printer with variable dot size and a substrate that could not hold 600% ink. Preliminary experiments revealed that the YNSN model had insufficient accuracy, necessitating the use of the CYNNSN model. Therefore, two limitations had to be overcome. The first was developing a sampling scheme that did not require tens of thousands of printed samples. The second was developing a model that did not reduce color gamut volume. This required estimating the spectral properties of non-printable colors.

Experimental

An Epson Pro 5500 ink jet printer was used in this research. This is a six-ink printer normally equipped with cyan, magenta, yellow, black, light cyan, and light magenta pigmented inks. The light cyan and magenta were replaced with green and orange, that is, a CMYKGO ink set. The environmental temperature was controlled from 22°C to 24°C. Epson's proprietary halftoning was used; images appeared typical of frequency-modulated halftoning resulting in microscopic colors with image

statistics described by Eq. (2). Each color plane had eight bit addressability, that is, 256 levels ($2^8 = 256$). Changing drop size with increasing signal value was also proprietary. We controlled the amount of ink printed on the paper by creating six-plane colored images. All samples were printed on Epson photo quality ink jet glossy paper, model number KA3N20MDK. Printed samples dried for one hour before measurement to ensure colors had reached equilibrium. Each printed page included a custom target to ensure print-to-print repeatability. The spectral measurements were performed using a GretagMacbeth Spectrolino spectrophotometer. The Spectrolino had a 4 mm aperture with a 45/0 annular geometry and measured from 380 to 730 nm in 10 nm intervals. Only data between 400 and 700 nm were used in this research. Colorimetric values were calculated for illuminant D50 and CIE 1931 2° standard observer.

Optimizing the Yule–Nielsen n Value and the Relationship Between Digital Counts and Effective Area Coverage

A nonlinear relationship between the digital count and effective area coverage results from a combination of “optical dot gain,” which is caused by light scattering within the paper, and “mechanical dot gain,” which is caused by the physical spreading of ink on the paper.² In Eq. (1), n is a parameter accounting principally for optical dot gain and Eq. (3) principally describes mechanical dot gain. It should be noted that the relationship between digital signals and effective area coverage varies as n value is varied, reinforcing the empirical nature of n .

Both n value and the digital counts to effective area coverage relationship were determined by optimization, using single-colorant ramps. The candidate objective functions were spectral root mean square (RMS), weighted spectral RMS, or color difference.¹⁹ Because of the spectral nature of this research, spectral RMS of the difference between measured and predicted spectral reflectance was used as the objective function, shown in Eq. (4):

$$\sqrt{\frac{\sum_{\lambda=1}^i (R_{\lambda,meas} - R_{\lambda,pred})^2}{i}} \quad (4)$$

where $R_{\lambda,meas}$ and $R_{\lambda,pred}$ are measured and predicted spectral reflectance factor, respectively. For the wavelength range of 400 – 700 nm in 10 nm increments, $i = 31$.

A target was printed consisting of single colorant ramps. For each ink, all 256 levels were printed and measured. The spectral reflectance factor measurements of each ink at 100% effective area coverage and the substrate are plotted in Fig. 2. The optimization employed a direct search where n value varied from 1 to 10 in 0.5 unit intervals and from 10 to 20 in 1 unit intervals. For a given n , the effective area coverage for each sample was determined by minimizing spectral RMS error using Eqs. (1) and (2). (For single-colorant samples, Eqs. (1) and (2) reduce to the well-known Murray–Davies equation for $n = 1$.^{2,20}) The average performance for the entire sample set as well as when grouped by colorant are plotted in Fig. 3. The chromatic inks followed the expected trend: increasing n value improved spectral-estimation accuracy, the improve-

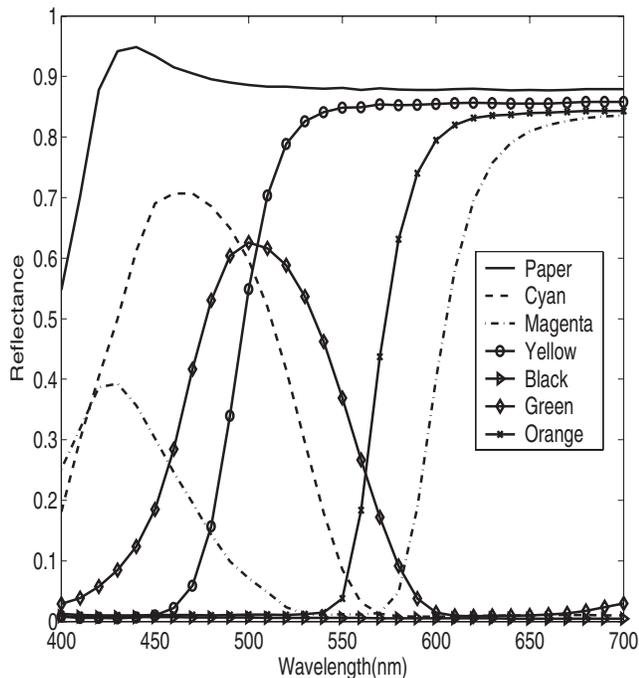


Figure 2. Spectral reflectance factor measurements of each colorant at maximum signal value and the paper substrate

ment significant initially. For black ink, changing n had a negligible effect. As a pigmented ink, its spectral reflectance was nearly flat; consequently model accuracy was largely independent of n value. The majority of the curves leveled off at $n = 10$. Given the potential for numerical error with small values for exponents, i.e., $1/n$, an n value of 10 was selected and used for the remainder of this research.

The effective area coverages for $n = 10$ were used to define the nonlinear relationships, i.e., Eq. (3). Because the ramps were printed corresponding to each digital value, the results were used to define six lookup tables (LUTs), plotted in Fig. 4. Typically, much fewer samples are evaluated and the LUT is populated using either an analytical function modeling the nonlinear relationship or piece-wise linear or nonlinear interpolation. The LUTs were different for each colorant because of differences in ink spreading (“mechanical dot gain”) and optical properties. Figure 4 indicates that the chromatic colorants reached the 100% effective area coverage at the maximum digital value of 255. For the black colorant, this occurred at a value of 200. That is, between digital values of 200 and 255, the spectral data of black colorant are indistinguishable.

Optimization of Cellular Primary Positions

One of the main factors that decrease prediction accuracy is an uneven error distribution when using the CYNSEN model with uniform sampling in colorant space. That is, large errors occur in some cells of a CYNSEN model. The error distribution is determined by the accuracy of the YNSN model within a subspace. Lack of accuracy is caused by a lack of linearity via Eq. (1). A highly nonlinear relationship will result in large interpolation errors. Agar and Allebach¹¹ proposed an iterative CYNSEN method to keep the average predicted color difference error in each cell less than a given value

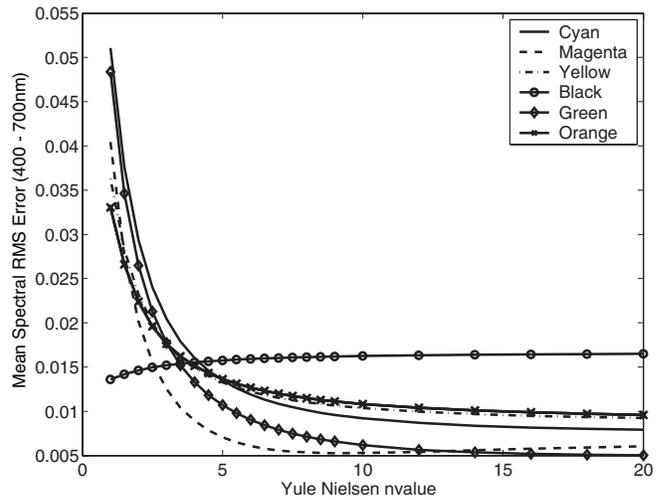


Figure 3. Average spectral RMS error versus n value for each colorant ramp and the entire test target

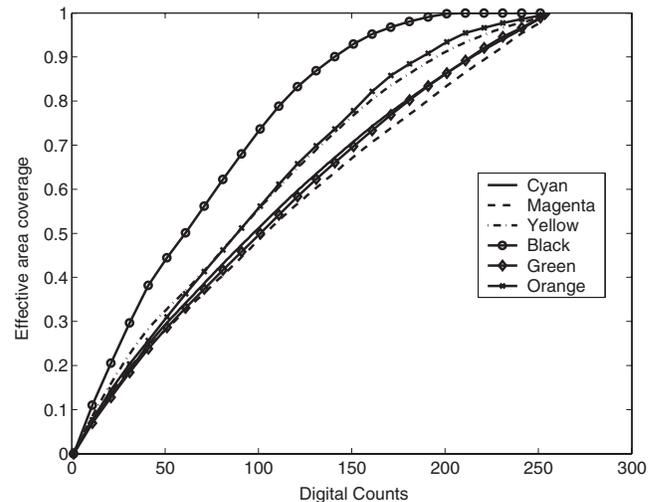


Figure 4. Digital counts versus effective area coverage for each colorant

as a cost function. This method decreased the unacceptable errors through continuously dividing large error areas into smaller cells by adding cellular primaries so that a finer interpolation was performed until the prediction error was less than the cost function. This method inevitably increases the number of cellular primaries and computation complexity because of the resulting non-uniform grid structure.

In the current research, we opted for a simpler approach. We assumed that the prediction error distribution in one-dimensional colorant space affects the error distribution in six-dimensional colorant space. This assumption was verified by exploring the correlation between one- and six-dimensional error distributions. Accordingly, it was a simple and efficient method to diminish unacceptable errors by optimizing the one-dimensional positions of the cellular primaries for each colorant.

TABLE I. Selected Digital Counts and Area Coverages of Every Ink for Nodes Corresponding to the Minimum of the Maximum Spectral RMS Errors

	Digital signal				Effective area coverage			
Cyan	0	50	168	255	0	0.30	0.77	1.00
Magenta	0	56	173	255	0	0.31	0.75	1.00
Yellow	0	49	162	255	0	0.32	0.81	1.00
Black	0	56	126	255	0	0.48	0.85	1.00
Green	0	43	182	255	0	0.25	0.81	1.00
Orange	0	64	186	255	0	0.38	0.90	1.00

TABLE II. Selected Digital Counts and Area Coverages of Every Ink for Nodes Corresponding to the Minimum of the Maximum Spectral RMS Errors

	Digital signal				Effective area coverage			
Cyan	0	32	102	255	0,00	0.20	0.51	1.00
Magenta	0	30	93	255	0,00	0.19	0.46	1.00
Yellow	0	29	97	224	0,00	0.22	0.53	1.00
Black	0	21	65	190	0,00	0.21	0.53	1.00
Green	0	41	122	255	0,00	0.24	0.57	1.00
Orange	0	25	79	255	0,00	0.18	0.45	1.00

TABLE III. Selected Digital Signals and Area Coverages of Every Ink Evenly Spaced by Effective Area Coverages

	Digital signal				Effective area coverage			
Cyan	0	53	124	255	0	0.33	0.67	1.00
Magenta	0	61	144	255	0	0.33	0.67	1.00
Yellow	0	53	122	255	0	0.33	0.67	1.00
Black	0	35	91	255	0	0.33	0.67	1.00
Green	0	60	138	255	0	0.33	0.67	1.00
Orange	0	51	114	255	0	0.33	0.67	1.00

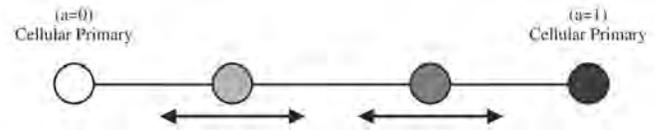


Figure 5. Graphical interpretation of determining the cellular primary inner positions for a single colorant

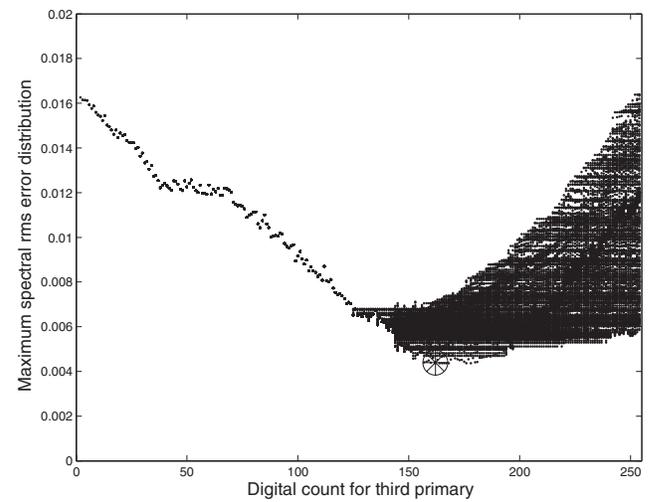
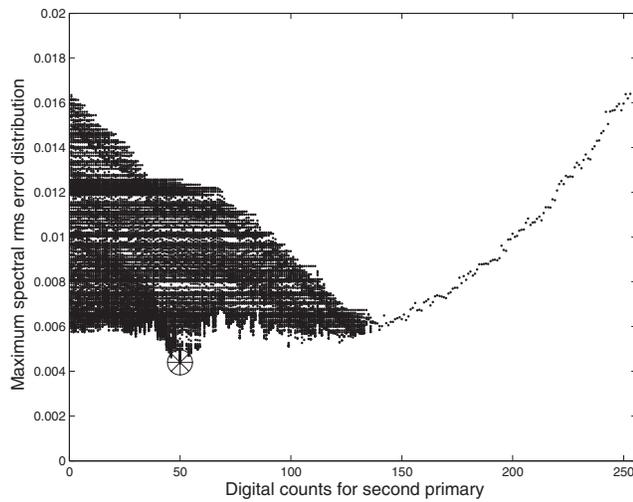


Figure 6. The maximum spectral RMS compared with the two middle cellular primary positions for the cyan ink. The circle star located the minimum value. Thus the second cellular primary corresponds to 50 and the third cellular primary corresponds to 168

Our sampling scheme was to create three subspaces for each colorant, increasing the number of primaries from two to four. The endpoints were fixed, corresponding to 0% and 100% effective area coverage, the usual Neugebauer primaries. The task was to determine the positions of the two inner cellular-primary positions, as shown in Fig. 5. There are

$$\frac{254!}{(254 - 2)!} = 32, 131$$

positions that the two inner points might have. For each combination, a CYNSEN model was created, and the spectral reflectance spectra of all 256 levels were predicted. The maximum RMS spectral difference between measured and predicted spectra was used as the error metric. The two inner positions resulting in the best performance were selected. This was repeated for each colorant. As an example, Fig. 6 shows the maximum spectral RMS errors for all combinations of

the middle two positions for cyan. The left and right figures show the maximum spectral RMS error distributions for the second and third positions, respectively. Their digital signals were chosen corresponding to the minimum of maximum RMS error, shown by a circled star in Fig. 6. The digital signals and effective area coverages of the optimized positions are listed in Table I.

Another two CYNSEN models with different cellular positions were created to evaluate the effectiveness of the new method. One model used positions corresponding to equal lightness, L^* , spacing while the other had equal effective-area-coverage spacing, which is equivalent to equal density spacing. The digital signals and area coverages for these additional models are shown in Tables II and III. Figure 7 shows the prediction error distribution for cyan for the three different spacings. We can find that the distribution of spectral RMS errors for the model with optimized cellular positions was more uniform than the other two. The performance for the evenly spaced effective area

TABLE IV. Comparison of Performance of Cellular Model on One-Color Ramps with Optimized Nodes and the Cellular Nodes Equally Spaced by Lightness and Effective Area Coverage

Positions of inner cellular-primary locations	ΔE_{00}		Spectral RMS	
	Mean	Maximum	Mean	Maximum
Even spaced by lightness, L^*	0.8	7.0	0.5%	1.8%
Even spaced by effective area coverage, a	0.6	5.5	0.3%	1.1%
Optimized	0.3	3.2	0.2%	0.6%

TABLE V. Comparison of Performance of the CYNSEN Model with Inner Position Cellular Primary Spacing as Listed

Positions of inner cellular primaries	Spectral RMS		ΔE_{00}		MI under D50		MI under A	
	Mean	Max	Mean	Max	Mean	Max	Mean	Max
Even spaced by L^*	0.7%	4.4%	1.2	6.2	0.2	1.2	0.2	1.3
Even spaced by AC	0.5%	3.2%	0.9	4.8	0.1	1.0	0.1	1.1
Optimized	0.4%	2.7%	0.8	2.7	0.1	0.7	0.1	0.8

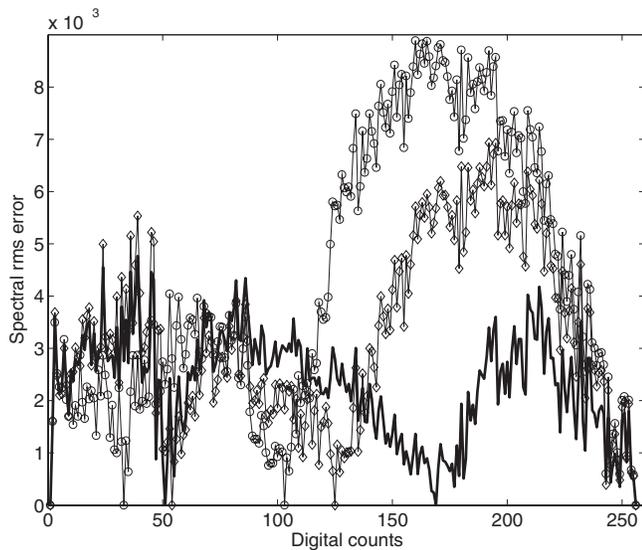


Figure 7. Distribution of spectral RMS errors with different primaries for cyan: the solid line with circles is evenly spaced L^* , the solid line with diamonds is evenly spaced effective area coverage, and the thick solid line is optimized positions

coverage was better than the evenly spacing lightness, because the effective area coverage compensates part of the nonlinearity of the model. Table IV shows the summary of performances of these three CYNSEN models predicting all the single-colorant colors.

In order to evaluate the performance of the integrated models, three six color CYNSEN models were created with cellular primary positions as described above. The printed primaries were constrained to eliminate inkblots caused by excessive ink applied to the substrate. Thus, these models were incomplete models, not fully utilizing the system's color gamut. A target was created of 200 samples, all of which could be predicted using the three CYNSEN models. The spectral reflectance of these samples were measured and compared with predicted reflectance. Table V lists the performance of the each model. ΔE_{00} was calculated for illuminant $D50$ and the 2° standard observer. The predicted spectra were parametrically corrected²¹ such that a perfect match was

obtained for $D50$. A $CIEDE2000$ was calculated for illuminant A and used as a metameric index (MI under A). (See Ref. 22 for a numerical example.) The correction was also performed for illuminant A and metamerism evaluated for $D50$ (MI under $D50$). The former MI penalizes long wavelength lack of fit to a greater extent than shorter wavelengths, and vice-versa for the later MI. The optimized primary spacing resulted in the best performance in all cases. The reduction in maximum error is notable. Given the level of accuracy achieved with this technique, more complex approaches were not considered.

Estimating Physically Non-Realizable Cellular Primaries

Excessive amounts of ink result in inkblotting, in which ink puddles on the surface of the substrate. For the printing system used in this research, it was not possible to print all six colorants at 100% ink amount. We found that temperature and humidity had a large effect on the maximum ink amount, hence their control during all the experiments. A simple target was adopted to estimate the approximate edge of the printable area. For each of the 63 Neugebauer primaries, a ramp was printed decrementing ink amounts in steps of one digital count. For example, for the cyan and magenta (blue) primary, $d_c = d_m = 255, 254, 253, \dots, 1$. There were 6 one-colorant ramps, 15 two-colorant overprint ramps, 20 three-colorant overprint ramps, 15 four-colorant overprint ramps, 6 five-colorant overprint ramps and 1 six-colorant overprint ramp. A portion of the target is shown in Fig. 8. Bars were printed corresponding to every ten digital counts. The critical values for inkblots of each Neugebauer primary ramp were determined visually. We looked for puddles and bronzing. Table VI shows the summary of the inkblot critical values as a function of the number of colorants. As expected, as more chromatic inks were overprinted together, the blots appeared at lower digital counts with maximum ink amount varying from 222% to 325%. Some ink combinations that included black ink had high inkblot critical values because the black ink darkened the prints and obscured the inkblot phenomenon. The final critical value for inkblots was chosen as 530 for digital signals, 220% effective area coverage. This maximum effective area coverage was

TABLE VI. Summary of Digital Signal and Effective Area Coverage of All the Overprint Ramps for Determining Inkblot Critical Values

	Digital signal			Area coverage		
	Mean	Maximum	Minimum	Mean	Maximum	Minimum
2 ink overprints	510	510	510	200%	200%	200%
3 ink overprints	563	570	530	252.6%	269.7%	221.9%
4 ink overprints	574	580	560	288.4%	305.3%	258.8%
5 ink overprints	550	550	550	301.1%	308.8%	282.7%
6 ink overprints	570	570	570	324.5%	324.5%	324.5%

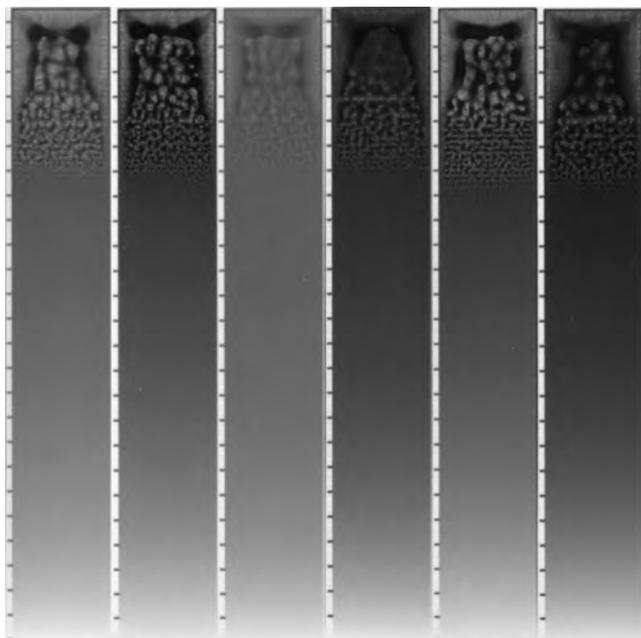


Figure 8. Portion of the target used for detecting inkblots

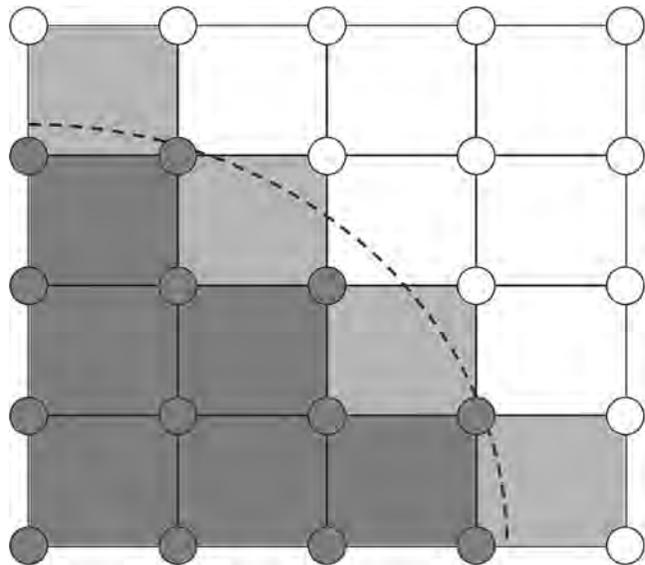


Figure 9. Graphical interpretation of the inkblot limitation in the cellular Yule–Nielsen modified Neugebauer model in two dimensions

considerably larger than typical ink jet printing, often determined by colorimetric and cost criteria.

The above analysis located the maximum color gamut and the maximum-gamut printable cellular primaries. If the CYNSN model were constrained to these positions, a considerable portion of the printing system’s color gamut would not be used. This is diagrammed in Fig. 9 where the small circles represent the cellular primaries. The filled circles represent the printable primaries while the open circles represent the unprintable primaries. The dashed line represents the true printable boundary. Positions within the dark area can be predicted by the CYNSN model. Positions within the light-gray area cannot be predicted by the CYNSN model because the interpolation endpoints cannot be printed.

In the cited research, this limitation was not addressed: All the cellular primaries as well as the Neugebauer primaries could be printed. One solution would be to add cellular primaries at the edge of the color gamut. For six-colorant printing, this would require a considerable number of additional samples and was considered impractical. Instead, we synthesized spectra for these unprintable colors in similar fashion to Balasubramanian.^{9,14} He described a technique to optimize the Neugebauer primaries using weighted spectral regression. This method used a set of measured printed samples as training samples to predict the Neugebauer primaries. The YNSN model shown in Eq.

(1) can be described using vector-matrix notation for m samples:

$$\mathbf{R}^* = \mathbf{F}\mathbf{P}^* \quad (5)$$

where

$$\mathbf{R}^* = \begin{bmatrix} R_{1,\lambda=400}^{1/n} & R_{1,\lambda=410}^{1/n} & \cdots & R_{1,\lambda=700}^{1/n} \\ R_{2,\lambda=400}^{1/n} & R_{2,\lambda=410}^{1/n} & \cdots & R_{2,\lambda=700}^{1/n} \\ \vdots & \vdots & \ddots & \vdots \\ R_{m,\lambda=400}^{1/n} & R_{m,\lambda=410}^{1/n} & \cdots & R_{m,\lambda=700}^{1/n} \end{bmatrix},$$

$$\mathbf{P}^* = \begin{bmatrix} R_{P1,\lambda=400}^{1/n} & R_{P1,\lambda=410}^{1/n} & \cdots & R_{P1,\lambda=700}^{1/n} \\ R_{P2,\lambda=400}^{1/n} & R_{P2,\lambda=410}^{1/n} & \cdots & R_{P2,\lambda=700}^{1/n} \\ \vdots & \vdots & \ddots & \vdots \\ R_{P64,\lambda=400}^{1/n} & R_{P64,\lambda=410}^{1/n} & \cdots & R_{P64,\lambda=700}^{1/n} \end{bmatrix}$$

and

$$\mathbf{F} = \begin{bmatrix} F_{1,1} & F_{1,2} & \cdots & F_{1,64} \\ F_{2,1} & F_{2,2} & \cdots & F_{2,64} \\ \vdots & \vdots & \ddots & \vdots \\ F_{m,1} & F_{m,2} & \cdots & F_{m,64} \end{bmatrix}$$

\mathbf{R}^* is an $(m \times 31)$ matrix containing the predicted spectral reflectances of m printed samples. (The asterisk is used to denote reflectance factor exponentiated by $1/n$.) \mathbf{P}^* is a (64×31) matrix of the spectral reflectances of the 64 Neugebauer primaries. \mathbf{F} is an $(m \times 64)$ matrix of fractional area coverages, calculated using Eqs. (2) and (3).

Generally, the Neugebauer primaries and fractional area coverages are used to predict spectral reflectance. Alternatively the reflectances of a set of measured samples, \mathbf{R}^* , combined with their fractional area coverages, \mathbf{F} , can predict the 64 Neugebauer primaries, \mathbf{P}^* , using least squares regression, shown in Eq. (6):

$$\mathbf{P}^* = \mathbf{F}^- \mathbf{R}^* \quad (6)$$

where \mathbf{F}^- is the pseudoinverse of \mathbf{F} . Now the optimized primaries, \mathbf{P}^* , rather than their direct measurements, comprises the complete YNSN model. This model, however, is dependent on the training samples. That is, different training sample sets result in YNSN models with slightly different spectral primaries. Therefore an optimized model can provide a slightly better prediction for the samples near to the training samples than the samples far from them.

Caution must be exercised when optimizing the Neugebauer primaries in order to predict the cellular Neugebauer primaries for multi-ink printers, because there is often a large amount of spectral redundancy.^{23,24} That is, different ink combinations achieve nearly identical spectral reflectance. This could lead to an ill conditioned pseudoinverse operation. Therefore localized fitting was performed to solve this problem using weighted linear regression, the weights dependent on colorant position. Within colorant space, training samples located far from a sample to be predicted will be given negligible consideration in the regression. The weights were the reciprocal of the distance from every training sample to the unrealizable cellular primary to be predicted, calculated by Eq. (7).

$$w_d = \frac{1}{\sum_{\lambda=1}^6 (a_{cell,i} - a_{t,i})^2} \quad (7)$$

where $a_{cell,i}$ is the effective area coverage of a non-printable cellular primary and $a_{t,i}$ is the effective area coverage of a training sample. The reciprocal of the distances from the predicted primary to all training samples comprised a diagonal $(m \times \mu)$ weighting matrix:

$$\mathbf{w} = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_n \end{bmatrix}$$

Incorporating the weighting, Eq. (6) becomes Eq. (8):

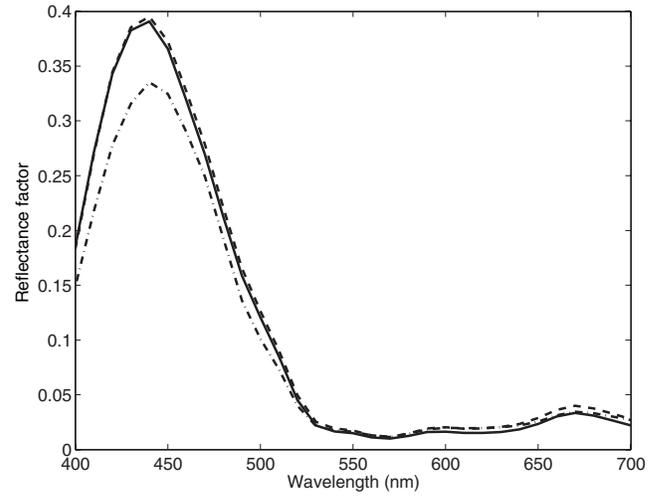


Figure 10. Measured (solid line), predicted without weighting (dash-dot line), and predicted with weighting (dashed line) spectra for the cyan and magenta overprint cellular Neugebauer primary.

$$\mathbf{P}^* = (\mathbf{wF})^- (\mathbf{wR}^*) \quad (8)$$

The spectral reflectances of the 64 Neugebauer primaries, \mathbf{P}^* , were optimized for one specific cellular Neugebauer primary, \mathbf{P}_{cell} . So the spectral reflectance of \mathbf{P}_{cell} can be predicted by Eq. (9):

$$\mathbf{P}_{cell} = \mathbf{F}_{cell} \mathbf{P}^* \quad (9)$$

where \mathbf{F}_{cell} is the fractional area coverages of the non-printable cellular Neugebauer primary to be predicted, expressed by a (1×64) matrix and \mathbf{P}_{cell} is expressed by a (1×31) matrix:

$$\mathbf{F}_{cell} = [F_1 \ F_2 \ \cdots \ F_{64}]$$

$$\mathbf{P}_{cell} = \begin{bmatrix} R_{cellP,\lambda=400} & R_{cellP,\lambda=410} & \cdots & R_{cellP,\lambda=700} \end{bmatrix}$$

A training set of samples were printed and measured, consisting of 1624 colors uniformly distributed in colorant space within the printable gamut. The optimization method described above was used to synthesize the non-printable cellular primaries. The spectral reflectance of a cyan–magenta overprint sample was predicted in order to evaluate the importance of incorporating the colorant–location weighting, \mathbf{w} . Figure 10 compares the measured, predicted without weighting (Eq. (6)), and predicted with weighting (Eq. (8)) spectra. Clearly, the weighting was a key element in the optimization.

The sampling scheme for the CYNSN model was to divide each colorant into three subspaces. This corresponded to 4,096 cellular primaries ($4^6 = 4,096$). Among them, 3,072 could not be printed because of inkblots. Thus the weighted least squares regression described above was used to synthesize these non-printable cellular primaries one by one. We also synthesized the printable 1,024 cellular primaries in

TABLE VII. Performance Comparison of the Two CYNSEN Models. See Text for Explanation of the Two Models

Number of synthesized primaries	Spectral RMS		ΔE_{00}		MI under D50		MI under A	
	Mean	Max	Mean	Max	Mean	Max	Mean	Max
3,072	0.4%	2.6%	1.0	3.3	0.2	1.0	0.2	1.2
4,096	0.5%	3.1%	1.0	5.9	0.2	1.6	0.2	1.6

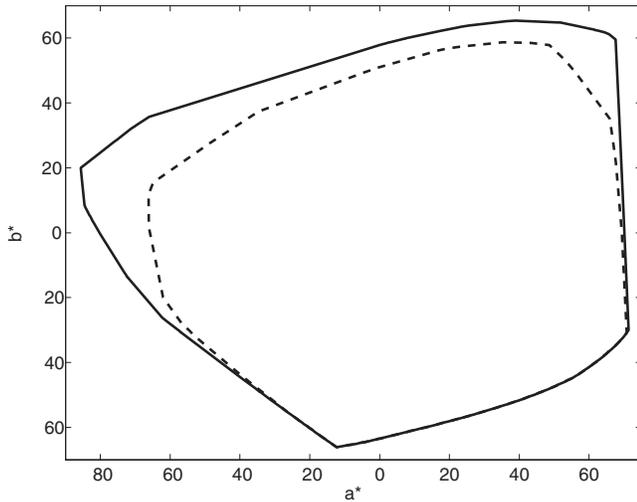


Figure 11. Color gamut comparison using only printable cellular primaries (dashed line) and using both printable and synthesized cellular primaries (solid line) at $L^* = 40$

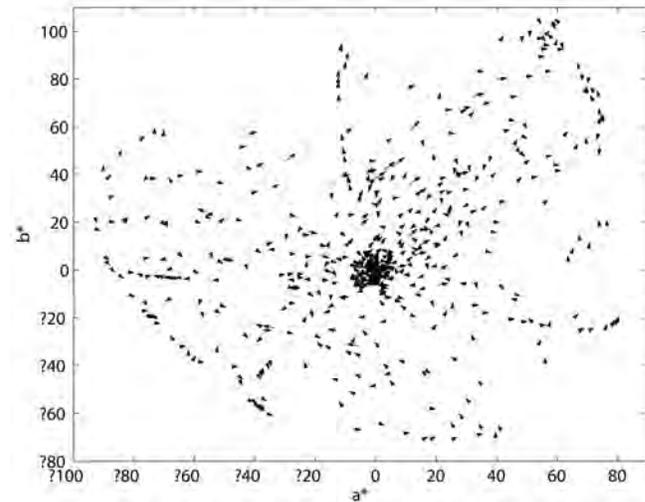


Figure 12. The prediction error in a^*b^* subspace of CIELAB for model based on synthetic and measured spectra

order to evaluate the plausibility of the synthetic spectra.

The colorant-space volumes using the CYNSEN model composed of only printable cellular primaries and of both printable and synthesized primaries were calculated. Adding the synthetic cellular primaries increased the volume by 54% in colorant space and 15% in CIELAB. These different proportions were a result of metamerism; different colorant combinations mapped to identical CIELAB coordinates. Figure 11 is a plot of the two color gamuts at $L^* = 40$.

Two cellular models were created. The first used synthesized spectra for all 4,096 cellular primaries. The second used the synthesized spectra for the 3,072 non-printable cellular primaries and the measured spectra of the 1,024 printable cellular primaries. A test target of 600 samples randomly distributed within the entire colorimetric gamut was printed and measured. The two CYNSEN models were used to predict spectral reflectance from digital data, the results shown in Table VII. The performance of the two models was nearly identical, on average. The maximum errors were smaller for the model using measured and synthesized spectra. This was the expected result since this model contained a greater proportion of “ground truth,” actual measurements. A T-test was performed to determine if the slight difference was of statistical significance. For an $\alpha = 0.05$ the model performance was not statistically different. This indicated that the synthesized spectra were reasonable.

The performance of the CYNSEN model composed of measured and synthesized cellular primaries is plotted in Fig. 12 in which the tail of each arrow locates the measured print and the arrowhead locates the predicted print. For most of the colors, there are not any tails, indicating good performance.

Conclusions

When developing a spectral based model for multi-ink printers, there is a tradeoff between the number of samples required to create the model and model accuracy. This is a particular issue for printers with more than four colorants such as CMYKGO or CMYKRGB. In this research, the YNSN model was used with the cellular enhancement, necessary to achieve sufficient prediction accuracy. Because of the tradeoffs between the number of printed samples, the amount of consumables, measurement time, processing time, spectral accuracy, and colorimetric accuracy, three optimizations were performed to achieve acceptable spectral accuracy with a reasonable number of printed samples.

The first optimization determined the best Yule-Nielsen n value that minimized spectral prediction error for 256-level ramps for the six colorants, CMYKGO. Once the global n value was found, one-dimensional look-up tables were created that related input digital signals with effective area coverages.

The second optimization determined the best one-dimensional positions of the four cellular primaries for each colorant. Fixing the minimum and maximum ink amounts at 0% and 100%, respectively, the two inner cellular-primary locations were determined that minimized the maximum spectral error for all 256 possible ink amounts. Each colorant was optimized independently. We assumed that minimizing errors in one dimension would be correlated with six dimensions. This was verified using a target of 200 samples. The performance was sufficiently high that additional cellular locations were unnecessary.

Dividing each colorant into three subspaces resulted in 4,096 cellular primaries. Of these, only 1,024 could be printed. The remaining primaries would have

inkblots, caused by excessive amounts of ink. The third optimization synthesized the spectral properties of the 3,072 non-printable colors. The YNSN model and weighted spectral regression was used; the weighting was a function of colorant-space location. The color gamut achievable using the synthesized spectra was 54% larger in colorant space and 15% larger in CIELAB space than that achievable when limiting the CYNNSN model to printable cellular primaries.

Following both optimizations, the complete model predicted the spectral properties of 600 random colors sampling the colorimetric gamut to an average accuracy of less than 0.5% RMS and $1\Delta E_{00}$. This level of performance was considered sufficient for printer characterization.

There are a number of areas for future research. The first is to explore minimizing the number of samples required for printer characterization. In the current research there were six 256 step one-colorant ramps and 1,024 cellular-primary samples, around 2,500 samples taking into account sample redundancy. Reducing the number of steps from 256 to under 20 would be an obvious improvement. We imagine that the total number could easily be less than 1,000.

We used the YNSN and weighted regression in order to predict non-printable primaries. It would be interesting to use Kubelka–Munk turbid media theory to predict the non-printable 64 Neugebauer primaries²⁵ followed by the YNSN model to predict the non-printable cellular primaries and compare performance.

This research used CMYKGO colorants. The accuracy of the CYNNSN model depends on the local linearity, ultimately determined by the optical properties of the printing system. When the Yule–Nielsen n value was evaluated for each individual colorant, different optimal values resulted, indicating that model performance depends on the spectral properties of the individual colorants. This research should be repeated using different colorant sets, different substrates, and different printing technologies. ▲

Acknowledgment. This research was supported by the Seiko–Epson Corporation. Their financial support and equipment and consumable donations are sincerely appreciated.

References

1. P. C. Hung, Colorimetric calibration in electronic imaging devices using a look-up-table model and interpolations, *J. Electronic Imag.* **2**, 53 (1993).
2. D. R. Wyble and R. S. Berns, A critical review of spectral models applied to binary color printing, *Color Res. Appl.* **25**, 4 (2000).
3. L. A. Taplin and R. S. Berns, Spectral color reproduction based on a six-color ink jet output system, *Proc. IS&T/SID Ninth Color Imaging Conference*, IS&T, Springfield, VA, 2001, pp. 209–213.
4. H. E. J. Neugebauer, Die theoretischen Grundlagen des Mehrfarbendrucks, *Z. Wiss. Photograph.* **36**, 73 (1937).
5. J. A. C. Yule and W. J. Nielsen, The penetration of light into paper and its effect on halftone reproductions, *Proc. TAGA* **3**, 65 (1951).
6. J. A. S. Viggiano, The color of halftone tints, *Proc. TAGA* **37**, 647 (1985).
7. M. E. Demichel, *Procédé*, **26**, 17–21, 26–27 (1924).
8. R. Rolleston and R. Balasubramanian, Accuracy of various types of Neugebauer model, *Proc. IS&T/SID Color Imaging Conference*, IS&T, Springfield, VA, 1993, pp. 32–37.
9. R. Balasubramanian, Optimization of the spectral Neugebauer model for printer characterization, *J. Electron. Imaging* **8**, 156 (1999).
10. K. J. Heuberger, Z. M. Jing and S. Persiev, Color transformations and lookup tables, *Proc. TAGA/ISCC*, 863–881, (1992).
11. A. U. Agar and J. P. Allebach, An iterative cellular YNSN Method for color printer characterization, *Proc. Sixth IS&T/SID Color Imaging Conference*, IS&T, Springfield, VA, 1998, pp. 197–200.
12. K. Iino, R. S. Berns, Building color management modules using linear optimization I. Desktop color system, *J. Imaging Sci. Technol.* **42**, 79 (1998).
13. K. Iino and R. S. Berns, Building color management modules using linear optimization II. Prepress system for offset printing, *J. Imaging Sci. Technol.* **42**, 99 (1998).
14. R. Balasubramanian, The use of spectral regression in modeling halftone color printers, *Proc. IS&T/OA Optics and Imaging in the Information Age*, IS&T, Springfield, VA, 1996, pp. 372–375, (1996)
15. F. H. Imai, D. R. Wyble and R. S. Berns, A feasibility study of spectral color reproduction, *J. Imaging Sci. Technol.* **47**, 543, (2003).
16. D. Y. Tzeng, *Spectral based color separation algorithm development for multiple-ink color reproduction*, Ph.D. Dissertation, Rochester Institute of Technology, Rochester, New York, 1999.
17. D. Y. Tzeng and R. S. Berns, Spectral based six-color separation minimizing metamerism, *Proc. IS&T/SID Eighth Color Imaging Conference*, IS&T, Springfield, VA, 2000, pp. 342–347.
18. See www.art-si.org for more details and a listing of publications.
19. F. H. Imai, M. R. Rosen and R. S. Berns, Comparative study of metrics for spectral match, *Proc. IS&T First European Conference on Color in Graphics, Imaging, and Vision*, IS&T, Springfield, VA, 2002, pp. 492–496.
20. A. Murray, Monochrome reproduction in photoengraving. *J. Franklin Inst.* **221**, 721 (1936).
21. H. S. Fairman, Metameric correction using parametric decomposition, *Color Res. Appl.* **12**, 261 (1987).
22. R. S. Berns, *Billmeyer and Saltzman's Principles of Color Technology*, John Wiley and Sons, New York, 2000, p211.
23. M. Rosen, E. Hattenberger and N. Ohta, Spectral redundancy in a six-ink ink jet printer, *Proc. IS&T PICS Conference*, IS&T, Springfield, VA, 2003, pp. 236–243.
24. M. Rosen, E. Hattenberger and N. Ohta, Spectral redundancy in a six-ink ink jet printer, *J. Imaging Sci. Technol.* **48**, in press (2004).
25. D. Y. Tzeng and R. S. Berns, Spectral reflectance prediction of ink overprints by Kubelka–Munk turbid media theory, *Proc. TAGA*, Rochester, NY, 682–697, (1999).