Color Printer Characterization Adjustment for Different Substrates

Mark Shaw,*[1] Gaurav Sharma,[2,†] Raja Bala,[2] Edul N. Dalal[2]

[1] Xerox Engineering Systems, 4000 Burton Drive, San Jose, CA 95054

[2] Xerox Innovation Group, Xerox Corporation, 800 Phillips Rd., Webster, NY 14580

Received 2 May 2002; revised 22 October 2002; accepted 6 November 2002

Abstract: The use of multiple substrates in color printers requires color characterization for each of the individual substrates. A full re-characterization for each substrate is measurement and labor intensive. In this article, a variety of methods are proposed and evaluated for determining the color characterization for a new substrate based on a complete characterization on a reference substrate and a small number of additional measurements for the new substrate. This saves significant time and effort in comparison to the traditional method of repeating the color characterization for each new substrate. The methods developed and tested include model-based approaches based on Beer's law, Kubelka-Munk theory, and Neugebauer equations; and an empirical technique based on principal component analysis. Results indicate that the model based techniques offer only a small improvement over direct use of the reference characterization, whereas, the empirical technique offers a more significant improvement with as few as 26 measurements on the new substrate. © 2003 Wiley Periodicals, Inc. Col Res Appl, 28, 454–467, 2003; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/col.10198

Key words: color reproduction; printer color characterization; multiple substrates; Kubelka-Munk; Neugebauer; principal component anaylsis; Beer's Law

INTRODUCTION

Color calibration of a printer is typically a two-step process. In the first step, the printer response is characterized by printing a number of color patches with known device control values, measuring the colors obtained, and generating a characterization function that maps device control values, such as *CMYK*, to corresponding colors specified in a device independent color space, such as CIELAB. In the second step, the characterization function is inverted to determine the device control values required to produce a color specified in device independent color space. The final color correction that inverts the characterization function is often implemented as a 3D look-up table that maps from a device independent color space (e.g. CIELAB) to the device control values (e.g. *CMYK*).

Several factors influence the printer response and consequently, the printer characterization function. Some factors are often unpredictable, in that they produce random variations in the response rather than systematic changes. Examples of such factors include printer drift and location-sensitivity. The printer may change daily or even hourly depending on the environmental conditions and throughput. In addition, the response may depend on spatial location (non-uniformity). Typically, the impact of these factors is minimized through careful design of the printing system. However, even if one were able to build a printer that did not drift, that had uniform location sensitivity, with a stable, predictable set of colorants, there is still one variable that cannot be controlled—the print substrate. The substrate can influence the color characteristics of the print in several ways. Certain substrate properties such as surface roughness can affect mechanical interactions between colorants and substrate, in turn affecting the resulting color. For example, inkjet colorants usually achieve more saturated colors when printed on coated paper than on uncoated paper, because in the former case, a fraction of the ink is absorbed into the paper fiber. Secondly, optical properties play an important

* Correspondence to: Mark Shaw, HP, 8153 W. Ringbill Lane, Boise, ID 83714; (e-mail: mark.q.shaw@hp.com).

† Currrent address: ECE Dept., Hopeman 417, University of Rochester; Rochester, NY 14627

role. The most obvious property is the spectral reflectance of the substrate. Other effects include optical scattering or diffusion of light within the substrate. These effects have been modeled by other researchers[1–4]

Since the printer response can vary considerably with different printing substrates, a printer characterization is specific to the substrate. If the user decides to use a substrate different from that used in the characterization, the color output error will depend upon the magnitude of the differences between the properties of the substrate used in the characterization and that used for printing.

One approach is to ignore the change in substrate and use the original characterization for the new substrate. This approach is far from satisfactory and will in some cases yield large color errors. The other straightforward alternative is to repeat the entire characterization procedure for each new substrate. This approach is both measurement and labor intensive and can be prohibitively costly in a system supporting many different print substrates.

The intermediate approach explored in this article is to assume an accurate characterization on a reference substrate and to correct this for each new substrate with a small number of additional measurements. As mentioned earlier, many characteristics of the substrate affect the characterization including color, weight, morphology, dimensional stability, coating technology, to name but a few. Of the substrate characteristics listed, that which constitutes probably the most important in color characterization and which will be the focus of this work, is the substrate's spectral reflectance characteristic.

MINIMAL-EFFORT COLOR CHARACTERIZATION OF A NEW SUBSTRATE

Various approaches were developed and implemented to re-characterize the printer with a small effort when the substrate is changed. Each of these methods is based on techniques that predict the reflectance of a patch with specified device control values on the test substrate using the known reflectance of the patch on the reference substrate with the same device control values, in conjunction with characteristics of the test substrate. These techniques include model-based approaches such as Beer's law, Kubelka-Munk theory and Neugebauer equations, and an empirical regression approach based on Principal Component Analysis.

Beer's Law

Beer's law[5] states that the absorption of light is proportional to the number of absorbing molecules in its path and describes simple subtractive colorant mixing on a wavelength by wavelength basis. For a single colorant, the transmittance ($t_{(\lambda)}$) of a colorant layer is predicted as

$$t_{(\lambda)} = 10^{-\varepsilon_{(\lambda)}cx} \tag{1}$$

where λ denotes wavelength and $\varepsilon_{(\lambda)}$, c and x represent, respectively, the extinction coefficient at wavelength λ, the concentration, and thickness for the colorant. The term $\varepsilon_{(\lambda)}$ cx is referred to as the spectral density of the colorant layer. The reflectances observed when the colorant layer is printed on two different substrates are given by

$$R_{1(\lambda)} = t^2_{(\lambda)} \cdot Rp_{1(\lambda)} \quad R_{2(\lambda)} = t^2_{(\lambda)} \cdot Rp_{2(\lambda)}$$

where $Rp_{1(\lambda)}$ and $Rp_{2(\lambda)}$ are the reflectances of the substrates and $R_{1(\lambda)}$ and $R_{2(\lambda)}$ are the reflectances observed with the colorant layer printed on the substrates. Using the above relation, the reflectance for the new substrate ($R_{2(\lambda)}$), can be obtained from the reflectance spectrum ($R_{1(\lambda)}$) of that colorant mixture printed on a known substrate ($Rp_{1(\lambda)}$), and the reflectance of the new substrate ($Rp_{2(\lambda)}$) as

$$R_{2(\lambda)} = \left(\frac{R_{1(\lambda)}}{Rp_{1(\lambda)}}\right) \cdot Rp_{2(\lambda)} \tag{2}$$

While the relation above was derived for a single colorant layer, it generalizes to multiple colorants under the assumption that the colorant layers combine additively in spectral density. It further generalizes to halftone tints under the assumption of a spectral Neugebauer model (this is applicable with or without an exponential Yule-Nielson correction, which is discussed in a later section). While this is not a direct implementation of Beer's law, assumptions are made analogous to Beer's law in that the colorants are transparent, the colorants have zero scattering, and that the toner thickness and colorant concentration do not vary with a change in substrate. As long as the assumptions hold, it is possible to predict the reflectance of any colorant mixture on a new substrate ($R_{2(\lambda)}$), given only the reflectance spectrum ($R_{1(112)}$) of that colorant mixture printed on a known substrate ($Rp_{1(\lambda)}$), and the reflectance of the new substrate ($Rp_{2(\lambda)}$). In this method, it is necessary to measure only the reflectance spectrum of each new paper.

Kubelka-Munk (KM) Theory

When a colorant is not completely transparent, the Beer's law assumption of zero scattering does not hold. Therefore, one must build a model that attempts to take into account a colorant's absorption and scattering properties. This is an important consideration when modeling xerographic processes. The printer toners are not completely transparent and typically each toner has its own scattering coefficient.

Kubelka and Munk derived a color mixing model which describes the reflectance and transmittance of a color sam-

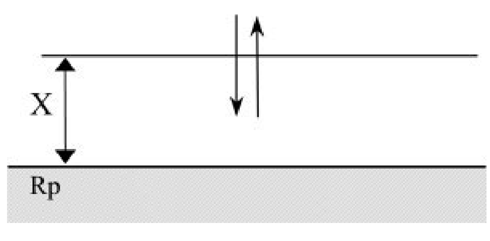

FIG. 1. The Kubelka-Munk model.

ple in terms of the absorption and scattering ($K_{(\lambda)}$ and $S_{(\lambda)}$) coefficients of the colorant material. The model considers a colorant layer of thickness X that both absorbs and scatters light that is passing through it,[6] which is placed in optical contact with a background of reflectance $Rp_{(\lambda)}$ (Fig. 1).

The Kubelka-Munk (KM) model assumes that the light is scattered in only two directions, up and down, with the change in light flux in the two directions described by two simultaneous differential equations involving the colorants' absorption and scattering coefficients. The effective reflectance of the sample can be determined by solving the differential equations and is given by:

$$R_{(\lambda)} = \frac{\left(\frac{R_{p(\lambda)} - R_{\infty(\lambda)}}{R_{\infty(\lambda)}}\right) - R_{\infty(\lambda)}\left(R_{p(\lambda)} - \frac{1}{R_{\infty(\lambda)}}\right) \times \exp\left[S_{(\lambda)}X\left(\frac{1}{R_{\infty(\lambda)}} - R_{\infty(\lambda)}\right)\right]}{R_{p(\lambda)} - R_{\infty(\lambda)} - \left(R_{p(\lambda)} - \frac{1}{R_{\infty(\lambda)}}\right)\exp\left[S_{(\lambda)}X\left(\frac{1}{R_{\infty(\lambda)}} - R_{\infty(\lambda)}\right)\right]} \quad (3)$$

where $R_{(\lambda)}$ is the reflectance of the colorant layer on the substrate, $Rp_{(\lambda)}$ is the reflectance of the substrate, $(R_{\infty(\lambda)})$ is the reflectance of a colorant sample of an infinitely thick sample of the colorant, $S_{(\lambda)}$ is the scattering coefficient of the colorant, and X is the thickness of the colorant layer. Equation 3 is commonly referred to as the exponential form of the KM equation.[7]

The parameter R_∞ represents the reflectance of an infinitely thick colorant, that is, a thickness such that any further increase in thickness has no effect on the reflectance of the sample. R_∞ is directly related to the absorption and scattering coefficients of the colorant, shown in Eq. 4.

$$R_{\infty(\lambda)} = 1 + \frac{K_{(\lambda)}}{S_{(\lambda)}} - \sqrt{\left(\frac{K_{(\lambda)}}{S_{(\lambda)}}\right)^2 + 2\left(\frac{K_{(\lambda)}}{S_{(\lambda)}}\right)} \quad (4)$$

where $K_{(\lambda)}$ is the absorption coefficient. The K and S spectra can be fitted to reflectance spectra as a function of wavelength over a series of thickness values using the KM equation, or by printing the colorants on multiple substrates of known reflectance. A method for calculating the K and S spectra over a series of thickness values can be found in the work of Dalal and Hoffman.[8,9]

Although the use of the KM model is relatively straightforward, it requires apriori knowledge of many of the parameters for the model to work. This work assumes that a minimal amount of data is available and that some parameters do not change from substrate to substrate.

Multilayer Kubelka-Munk Model

The basic KM model only predicts the reflectance of a single colorant layer. An extension of this model was derived by Dalal and Hoffman (1997)[8,9] that predicts the reflectance of multiple layer images by successively treating each predicted layer reflectance as the substrate reflectance of the subsequent layer (Fig. 2). That is, the modeled reflectance spectrum for the first toner layer is treated as the substrate reflectance spectrum, $Rp_{(\lambda)}$, for the second toner layer, and so on, ultimately resulting in the reflectance spectrum of the multi-layer sample.

The method assumes that the K and S spectra of each toner layer are invariant on each substrate. The thickness parameters for each layer were assumed known in our work and can be estimated using a set of measurements from a variety of substrates. It was also assumed that the R_∞ spectra or the individual K and S spectra are available for the toners in the xerographic printer. Note that the model used here assumes non-interspersed spatially uniform colorant layers, whereas in actual practice the colorant distributions are interspersed and do not have a uniform thickness spatially. If suitable statistics of the colorant layer distributions are available, the model can be extended to take these into account to obtain improved results. Details of this procedure can be found in Hoffman.[9]

The KM model only provides a mechanism for predicting the reflectance for solid colors and overprints and does not apply to halftone tints. Therefore it must be used in conjunction with a different model, such as Neugebauer[10,11] in order to predict the spectral response of halftone color prints.

Neugebauer Model

The Neugebauer model has been widely used to predict the colorimetric response of halftone color printers. The original model is essentially an extension of the Murray-Davies equation.[12]

The Neugebauer model is based on the assumption that the color of individual points on the printed substrate corresponds to one out of all possible *solid overprint colors* produced by overlaying one or more of the halftone colorants. These *solid overprint* colors are referred to as the Neugebauer primaries. For *CMYK* printers, there are 16 Neugebauer primaries, including white paper and all possible combinations of the four-color mixtures. In the process of viewing, the eye averages over a region, the color of a region is therefore predicted in the Neugebauer model as the weighted average of the colors for the Neugebauer primaries, where the weights are the respective fractions of area covered by the Neugebauer primaries. For a given digital input (C M Y K) with corresponding dot area coverages (*c, m, y, k*), the spectral Neugebauer equations[13] predict the average spectral reflectance of a printed patch using Eq. 5.

$$R_{CMYK(\lambda)} = \left[\sum_{i=1}^{16} w_i \cdot P_{i(\lambda)}^{1/n}\right]^n \quad (5)$$

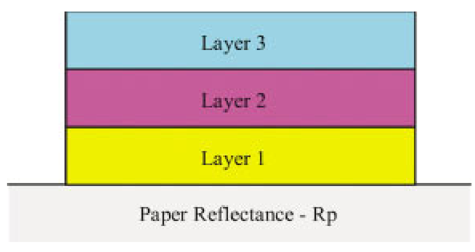

FIG. 2. Multi-layer Kubelka-Munk model.

where w_i represent the weight of each primary (i) calculated from the fractional area coverage of the individual colorants using the Demichel dot model equations. P_i is the reflectance of the (i)th solid color and R_{CMYK} is the predicted patch reflectance. A Yule-Nielsen[14] correction factor n is included to account for light scattering within the substrate. In this work, the fractional area coverages and Yule-Nielsen factor were derived using an iterative procedure that optimizes the model's prediction of single-separation ramps of *C, M, Y, K* (see ref. 10 for details).

The sixteen Neugebauer primaries can either be printed and measured, or modeled using the aforementioned KM model. Both approaches are explored in this article.

EMPIRICAL REGRESSION BASED ON PRINCIPAL COMPONENT ANALYSIS

An alternative to using model-based approaches, is to adopt an empirical technique that treats the problem of predicting the reflectance on the test substrate as a data-fitting problem. In this case, the reflectance measurements corresponding to a known set of device control values for a new substrate are inferred from the reflectance measurements corresponding to the same control values on the reference substrate by means of a pre-determined transform. The transform is determined from measurements of a set of training patches on the reference substrate and on the new substrate, both corresponding to the same set of device control values. Typically, the dimensionality of the transform directly impacts the amount of training data needed with larger amounts of data required as the dimensionality of the transform increases. Since it is desirable to keep the number of measurements on the new substrate to a minimum, it is beneficial to reduce the dimensionality of the problem as far as possible. We exploit principal component analysis (PCA) for this purpose. The basic idea is to transform the reflectance data into PCA coordinates and apply a transformation in these coordinates in order to account for a change in substrate. The transformation is estimated from the available set of corresponding training data on the reference and test substrates.

PCA is a mathematical technique for analysis of a set of multivariate data. The technique provides a representation of the multivariate data as a linear combination of orthonormal basis vectors, wherein each successive basis vector accounts for as much of the variation in the original data as possible. For typical print reflectance data, a very accurate approximation in a least squares sense can be obtained by using only the first few basis vectors or principal components. This approximation reduces the dimensionality of the original data from the number of variables in the multivariate data set to the number of principal components selected. Using a larger number of principal components (i.e. basis vectors) provides improved approximation accuracy at the cost of increased dimensionality. In an alternate view, the PCA analysis can be interpreted as successively searching for the dimensions along which the observations are maximally separated or spread out.[15] The principal components of the multivariate set of data correspond to orthonormal eigenvectors of the sample covariance matrix ordered according to decreasing eigenvalues. The principal components can be computed in a numerically stable and robust fashion using the Singular Value Decomposition.[16]

In a geometric sense, the best r ($<n$) dimensional representation of the original n-dimensional data is obtained by projecting the n-dimensional data into the r-dimensional subspace[17] defined by the first r principal components $\mathbf{v}_1$, $\mathbf{v}_2, \ldots, \mathbf{v}_r$. The approximation to an original n-dimensional vector $\mathbf{z}$ using the r principal components is given by a weighted combination of the principal components

$$\hat{\mathbf{z}} = \sum_{i=1}^{r} w_i \mathbf{v}_i \tag{6}$$

where the weights w_i are given by

$$w_i = \mathbf{v}_i^T \mathbf{z} \tag{7}$$

Here the superscript T represents the transpose operation.

For our application, we are dealing with spectral data, where typically a 10nm sampling interval is used over the range of wavelengths from 380–730nm. This corresponds to a 36-dimensional space. Since typical reflectance spectra of natural objects are smooth, PCA can provide an accurate approximation with as few as 5-10 principal components,[18] thus offering significant reduction in dimensionality with little loss in accuracy. Each reflectance spectrum $R(\lambda)$ is thus represented by the 5–10 weighting factors w_i corresponding to the chosen principal components $\mathbf{v}_i$.

The reduced dimensionality provided by PCA can be exploited in an empirical scheme for compensating for change in the substrate. With reference to Fig. 3, the first step is to determine the appropriate dimensionality r of the PCA representation of printed reflectance spectra. To this end, a large number of patches corresponding to different device control values (*CMYK*) are printed ❶ and measured using a spectrophotometer to obtain spectral reflectance values. In our work, 289 patches were used. The measured reflectance spectra are then normalized using a Yule-Nielsen modified form of the Beer's law assumption ❷,

$$R_{normalized(\lambda)} = \left(\frac{R_{measured(\lambda)}}{R_{paper(\lambda)}}\right)^{1/n} \tag{8}$$

where $R_{measured}$ is the reflectance factor of each patch, R_{paper} is the reflectance factor of the substrate white, and n is the Yule-Nielsen factor. Based on experiments with the different substrates, a value of $n = 5.0$ was chosen for all substrates. Equation [8] is analogous to pre-conditioning the data using the Beer's law model of Eq. [1] to reduce the variability that the subsequent regression needs to account for. The resulting normalized spectral data is analyzed using PCA to obtain a set of r basis vectors ❸. In our work, the 289 patches were printed on four different substrates and measured spectrally. For each value of r ranging between 1 and 20, the PCA approximations (given by Eq. [6]) were compared with the original measurements. In order to eval-

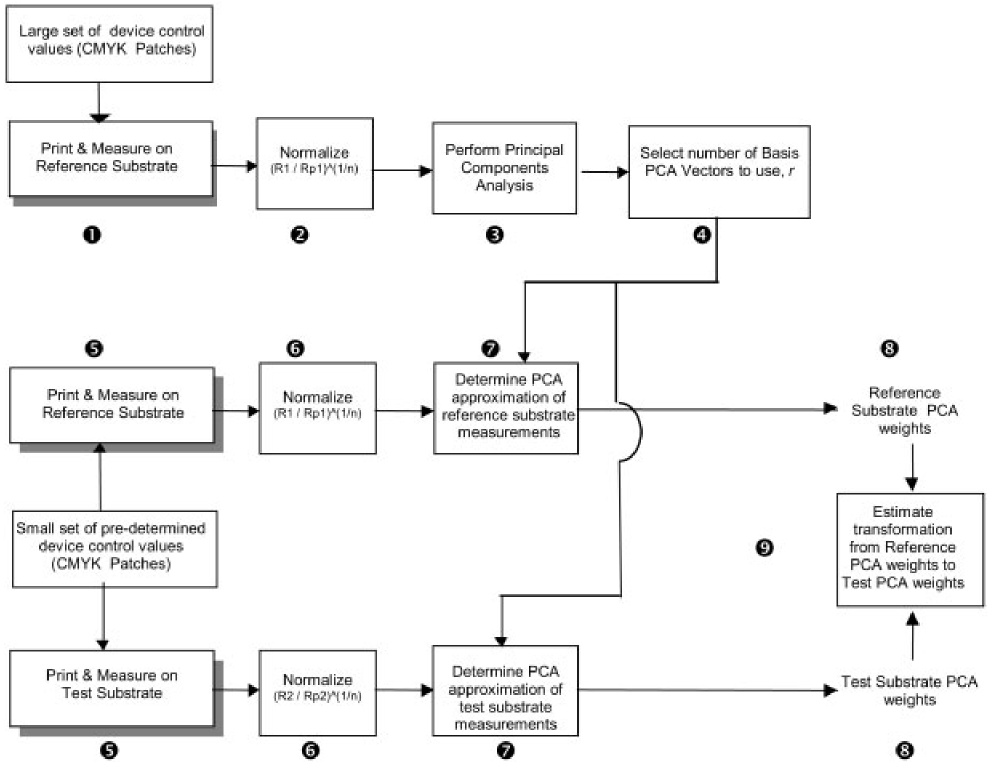

FIG. 3. Set-up procedure for relating spectra on reference and test substrates.

uate the PCA approximation with a perceptually meaningful metric, both the approximated and measured data were converted from their spectral representation to CIELAB and compared using the ΔE^*_{94} color difference metric.[19] This is given by:

$$\Delta E^*_{94} = \sqrt{\left(\frac{\Delta L^*}{k_L S_L}\right)^2 + \left(\frac{\Delta C^*}{k_C S_C}\right)^2 + \left(\frac{\Delta H^*}{k_H S_H}\right)^2}, \qquad (9)$$

where ΔL^*, ΔC^*, and ΔH^* are differences in lightness, chroma, and hue respectively; $k_L = k_C = k_H = 1$; $S_L = 1$, $S_C = 1 + 0.045C^*$, and $S_H = 1+0.015C^*$. The average ΔE^*_{94} color difference over the patches as a function of the number of PCA basis vectors is illustrated in Fig. 4 in tabular form and as a graph. As might be expected, the data in Fig. 4 show that the average color error in the reconstruction decreases monotonically with an increase in the number of basis vectors. Significant jumps in reconstruction accuracy can be seen when 3, 5, 10 or 17 basis vectors are used. Although increasing the number of basis vectors r improves the reconstruction, it is necessary to consider the fact that this also requires in an increased number of measurements on the test substrate for accurate regression. Furthermore,

# Vectors	Paper 1	Paper 3	Paper 7	Paper 11
1	19.6831	20.0577	20.4572	20.4411
2	14.0097	13.9687	14.4645	14.4148
3	4.5525	2.6744	3.1983	3.0807
4	3.8469	1.6457	2.2156	2.1114
5	0.5938	0.4471	0.4779	0.4914
6	0.5558	0.2508	0.3069	0.3247
7	0.5667	0.1815	0.2329	0.2629
8	0.2598	0.1417	0.1506	0.1534
9	0.2314	0.0816	0.1227	0.1016
10	0.0572	0.0491	0.0548	0.0638
11	0.0428	0.0459	0.0565	0.0628
12	0.0462	0.0323	0.0473	0.0436
13	0.05	0.043	0.0458	0.0465
14	0.0361	0.0449	0.0199	0.015
15	0.0319	0.042	0.0106	0.0139
16	0.0297	0.0084	0.0144	0.0106
17	0.0055	0.0014	0.0046	0.0025
18	0.003	0.0008	0.0013	0.0029
19	0.0029	0.0008	0.0011	0.0019
20	0.0014	0.0006	0.0011	0.0019

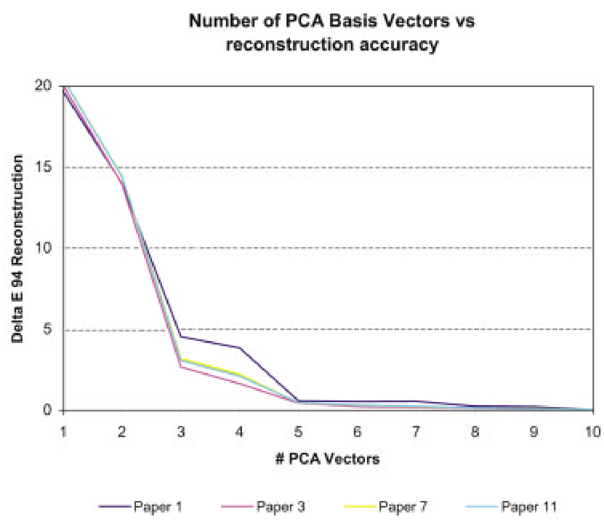

FIG. 4. Average ΔE^*_{94} error vs number of PCA basis vectors.

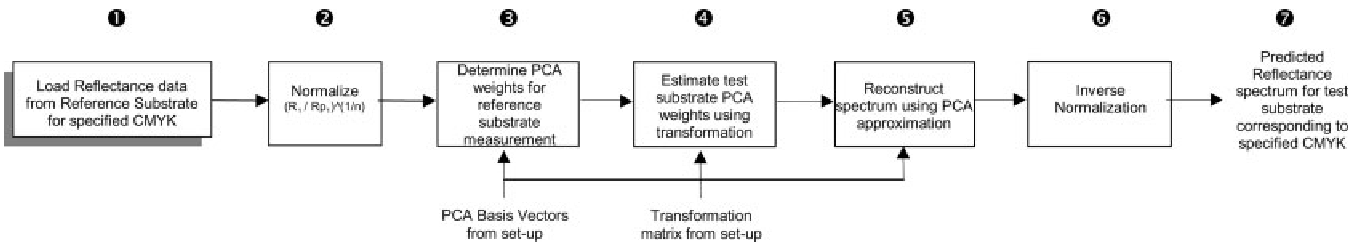

FIG. 5. Process for prediction of reflectance spectra on test spectra.

too many basis vectors may start modeling the noise within the spectral reflectance measurements and potentially cause increased error when applying the procedure to a new substrate. Based on these trade-offs, the number of basis vectors was empirically selected to be $r = 10$. ❹ For the four substrates compared in Fig. 4, this ensures an average color error under 0.1 ΔE^*_{94} units in the PCA based spectral approximations. Note that the more common technique for determining the number of PCA basis vectors is to determine the percent contribution to the variance from each successive PCA component, instead of the ΔE^*_{94} based evaluation employed here. While both techniques have similar qualitative characteristics, with successive vectors offering smaller and smaller gains, the technique employed here allows us to determine the contribution of the PCA approximation to the final color error and establish that the dimensionality of the PCA approximation is not the bottleneck in the empirical process.

Once the dimensionality and corresponding PCA basis vectors have been determined, the transformation from the reference substrate to the test substrate can be determined as follows. A small set of patches corresponding to a given set of device control values are printed on both the reference and test substrates ❺ and reflectance spectra for each of these patches are measured. Both sets of spectra are normalized using the Yule-Nielsen modified form of the Beer's law assumption ❻. For each normalized spectrum on the reference or test substrate, the weights corresponding to the 10 PCA basis vectors are computed ❼ (using Eq. [6]). This yields a set of PCA weights for the reference and test substrate corresponding to the same set of device control values ❽. Linear regression is then used to compute the optimal linear transformation $\mathbf{T}$ that maps the reference substrate PCA weights to the test substrate PCA weights. ❾ Specifically, if $\mathbf{x}$ is the $r \times 1$ vector of PCA basis weights for a patch printed on the reference substrate with a given set of device control values, the corresponding $r \times 1$ vector $\mathbf{y}$ of PCA basis weights for a patch printed on the test substrate with the same set of device control values is estimated as

$$\hat{\mathbf{y}} = \mathbf{T}\mathbf{x} \tag{9}$$

where $\mathbf{T}$ is the $r \times r$ matrix representing the linear transformation. If the vectors of PCA basis weights for the set of training patches printed on the reference substrate are $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, . . . $\mathbf{x}_K$ and corresponding vectors of PCA basis weights for the test substrate are $\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$, . . . $\mathbf{y}_K$, respectively, the optimal linear transformation $\mathbf{T}$ is given by:

$$\mathbf{T} = \mathbf{Y}\mathbf{X}^{\mathbf{T}}(\mathbf{X}\mathbf{X}^{\mathbf{T}})^{-1} \tag{10}$$

where

$$\begin{aligned} \mathbf{X} &= [\mathbf{x}_1|\mathbf{x}_2|\mathbf{x}_3|\cdots \mathbf{x}_K] \\ \mathbf{Y} &= [\mathbf{y}_1|\mathbf{y}_2|\mathbf{y}_3|\cdots \mathbf{y}_K] \end{aligned} \tag{11}$$

and $\lfloor \mathbf{X}^{\mathbf{T}}(\mathbf{X}\cdot\mathbf{X}^{\mathbf{T}})^{-1} \rfloor$ represents the pseudo-inverse of the $K \times r$ matrix $\mathbf{X}^{\mathbf{T}}$. In our work, the pseudo inverse was computed using the singular value decomposition.[16] Applying the computed transformation $\mathbf{T}$ to the PCA basis weights for a reflectance measured on the reference substrate provides an estimate of the PCA basis weights for the corresponding (same *CMYK* values) reflectance on the test substrate.[a]

A critical factor in the aforementioned derivation is the choice of *CMYK* values in the small sample set. The following are desirable guidelines in selecting this set:

i) The *CMYK* values should be a subset of the original characterization data set so that no new measurements are required on the reference substrate.
ii) In order to reduce the burden to the user, the size of this set should be significantly smaller than that of the original characterization data set.
iii) The patches should be chosen to ensure that $\mathbf{T}$ has sufficient rank. This basically means that each of the *C, M, Y, K* colorants should be sufficiently represented.

To this end, a set of 26 *CMYK* values was selected comprising the 16 Neugebauer primaries (i.e. all combinations of 0 and 100% area coverage of *C, M, Y, K*) and two levels of pure *C, M, Y* and *K*, and has two levels along the $C = M = Y$ axis. Once the transformation $\mathbf{T}$ is computed, the set-up procedure is complete. One can now estimate the spectral reflectance of a patch printed on the test substrate. Figure 5 illustrates this procedure. In order to estimate the

[a] It is conceptually feasible to eliminate the PCA data reduction and still employ the methodology presented here: In which case, the $\mathbf{x}_i$*'s* would correspond to individual sampled spectra and the matrix $\mathbf{T}$ would be a 36×36 matrix. This, however, poses two challenges, firstly a minimum of 36 corresponding measurements on both substrates are required and secondly, the 36×36 matrix ($\mathbf{X}\ \mathbf{X}^{\mathrm{T}}$) is highly ill-conditioned, making the solution in Eq. 10 extremely sensitive to the chosen sample set. This approach was therefore not pursued in this research.

TABLE I. Substrate names, coating characteristics weights, and optimum Yule-Nielsen value.

	Substrate	Coated	Weight	Yule-Nielsen value (n)
Paper 01	3R54	No	75 gsm	5.5
Paper 02	3R3874	No	90 gsm	5.0
Paper 03	Mead Moistrite Matte	Yes	80 gsm	1.0
Paper 04	Potltach Vintage Velvet Crème	Yes	80 gsm	2.0
Paper 05	S.D.Warren Strobe	Yes	80 gsm	2.0
Paper 06	Xerox Ultraspec Gloss	Yes	80 gsm	4.5
Paper 07	S.D.Warren Lustro Gloss Text lot sourced from Economy Paper	Yes	80 gsm	2.0
Paper 08	S.D.Warren Lustro Gloss Text Alling & Corey Lot	Yes	80 gsm	4.0
Paper 09	Consolidated Centure Gloss	Yes	80 gsm	3.0
Paper 10	Mead Vision	Yes	80 gsm	2.0
Paper 11	Westvaco Sterling Web Gloss	Yes	80 gsm	4.0
Paper 12	Mead Richgloss	Yes	80 gsm	1.0

reflectance value for a patch printed on the test substrate having specified *CMYK* values, the method begins with the reflectance spectrum for the patch on the reference substrate with the same CMYK values ❶. The spectrum reflectance is normalized ❷ as described in Eq. 8 and PCA weights required for approximating this normalized spectrum are determined using the basis vectors selected in the set-up phase ❸. The transformation matrix **T** computed in the set-up stage is then used to estimate the corresponding PCA weights for the test substrate ❹ as in Eq. 9. The estimated PCA weights for the test medium are used for reconstructing the PCA approximation to the normalized spectrum under for the test substrate ❺ using Eq. 6. The normalization procedure is inverted ❻ to obtain a reflectance estimate ❼ for the given *CMYK* patch printed on the test substrate. The inversion normalization can be expressed mathematically as:

$$R_{predicted(\lambda)} = (R_{normalized(\lambda)})^n \cdot R_{paper(\lambda)} \tag{12}$$

Since measurements for a complete printer characterization target are available for the reference substrate, estimates for the corresponding (having identical *CMYK* values) characterization target generated on the test substrate can be obtained using the procedure outlined above. The estimated values can then be used to characterize the printer for the new test substrate. Note also that the measurements of the complete characterization target on the reference medium can be used for the determination of the PCA representation in the set-up phase and a subset of these patches can be printed on the test medium and measured for estimating the transformation **T.** This process ensures that the additional prints and measurements required are only the small subset for the test substrate.

EXPERIMENTAL RESULTS AND ANALYSIS

All experiments were conducted on a four colorant *CMYK* Xerographic printer. Twelve different substrates varying in coating characteristics, weight, and paper color were chosen for the experimental investigation. Table I lists the substrates grouping them by coated and uncoated types and indicating their specified weight in grams per square meter (gsm).

For each of the substrates included in the study, a printer characterization target comprising 289 patches was printed and spectral measurements were made using a GretagMacbeth Spetralino device. The latter reports spectral data in the range 380–730nm in 10nm increments The spectral data was used to derive the parameters for the various models in this study.

The optimal Yule-Neilsen factor for the Neugebauer model is included in Table I for each substrate. It should be noted that this factor is derived from a mathematical optimization, and is only approximately related to the optical scattering phenomenon. The different values in Table I are roughly indicative of differences in optical scattering properties of the various substrates.

The spectral reflectance values corresponding to the characterization target on the test substrate were predicted from the spectral reflectance values on the reference target using each of the aforementioned techniques. The color error in these spectral estimates in comparison to the measured values was used as a figure of merit for comparing the different schemes. An exhaustive test was conducted in which each of the substrates was designated the reference substrate in turn and used to predict the measurements corresponding to the characterization target for each of the remaining 11 substrates. The results for this exhaustive pair-wise test are included in the appendix for each of the methods used for adjustment of the characterization for a test substrate. Only a small subset of representative results are presented here to illustrate the observed trends.

The techniques compared are (roughly in order of increasing effort): a) no adjustment, i.e., using the characterization for the reference substrate with no alteration for the test substrate; b) adjustment based on Beer's law using measured test substrate reflectance; c) adjustment based on the KM Model for predicting the Neugebauer primaries and a Neugebauer model for estimating the spectra of halftone tints, where the substrate reflectance and colorant *K* and *S* spectra are assumed known; d) adjustment based on the Neugebauer model for halftones with actual measuments of the 16 Neugebauer primaries; and e) adjustment based on

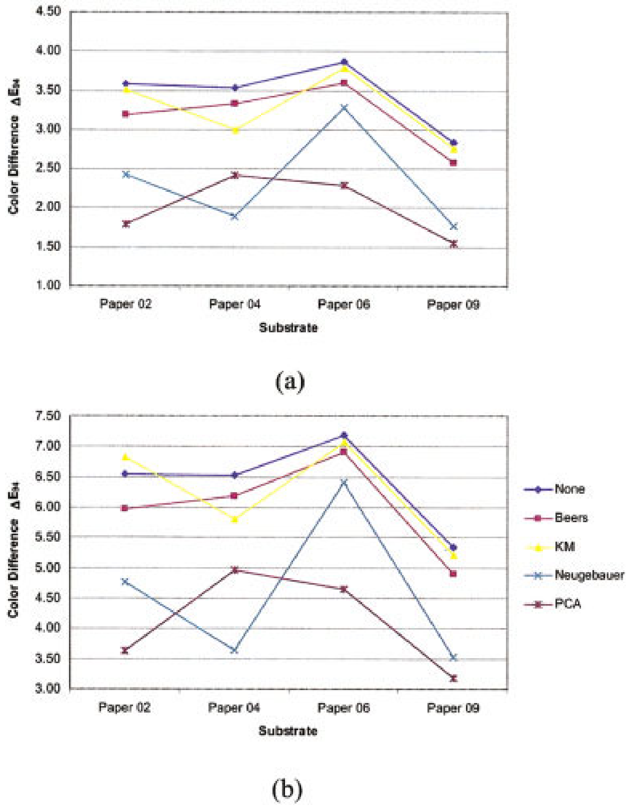

FIG. 6. (a) Average and (b) 95 percentile color error (ΔE^*_{94}) for predicted test substrate reflectances.

the PCA based empirical regression technique, where the transformation **T** was determined using a subset of 26 patches from the printer characterization target, printed on the respective substrates.

Figure 6a depicts a plot of the average color error of these compensation techniques for four substrates chosen to represent the improvements obtained with the prediction process. The abscissa of the plot represents the test substrate for which reflectance values are estimated using each of the other substrates as reference substrates in turn. The ordinate represents the average color error in the estimates in CIE ΔE^*_{94} units across all the patches in the target and across the three choices of the reference substrate. The five individual curves in the plot represent the different adjustment techniques and are identified in the legend. Figure 6b is a corresponding plot with the 95th percentile values of the same set of color errors along the ordinate axis. From the curve corresponding to no adjustment (none), we can see that there are significant differences among the different substrates with average color difference around 3.5 ΔE^*_{94} units and a 95-percentile value for the color difference around 6.0 ΔE^*_{94} units. The curves for the minimal measurement model-based approaches (Beer's Law and KM) on these plots are quite close to the case of no adjustment, indicating that these approaches offer very marginal im-

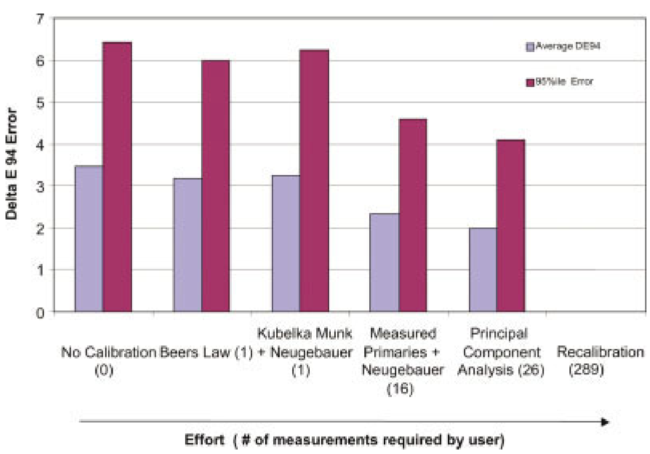

FIG. 7. Effort vs accuracy for prediction techniques.

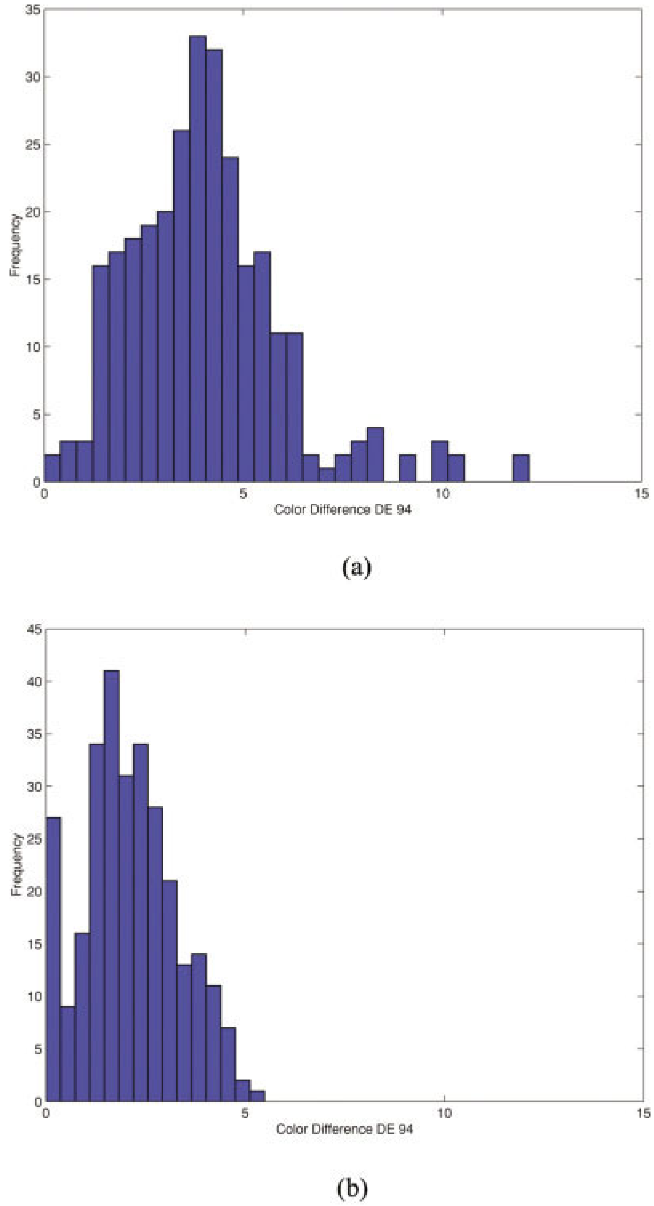

FIG. 8. (a) Histogram of ΔE^*_{94} errors between CIE Lab values measured on the reference (substrate 1) and test (substrate 4); (b) Histogram of ΔE^*_{94} errors between the CIE Lab values predicted by PCA and values measured on the text substrate.

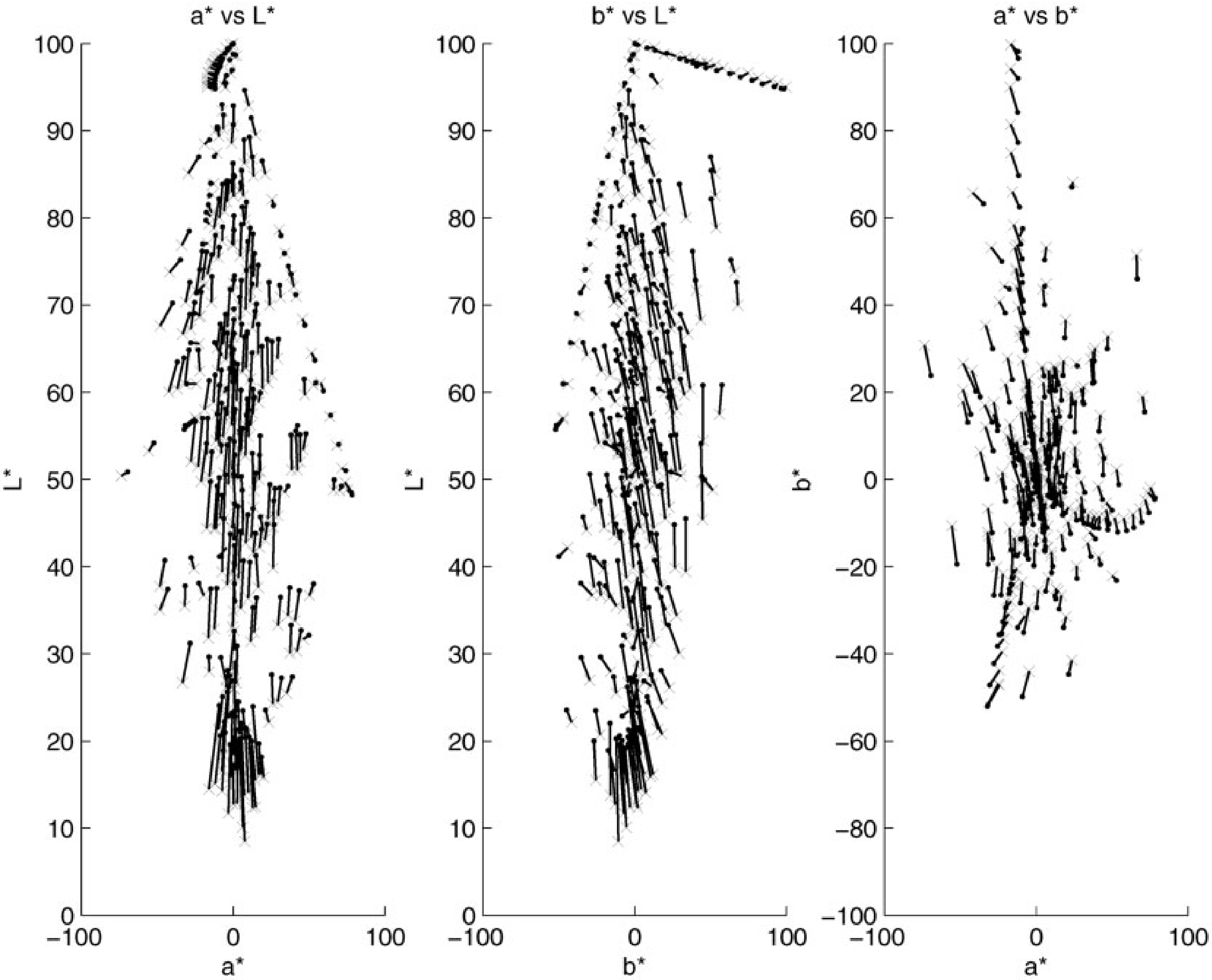

FIG. 9. CIE Lab values measured on a reference substrate (*x*) when compared to the CIE Lab values (•) measured on the test substrate.

provements at best. The approach based on the Neugebauer model incorporating actual measurements of the 16 Neugebauer primaries offers reasonable improvement for substrates 2, 4, and 9 but only a minor improvement for substrate 6. The empirical regression technique based on PCA consistently provides significantly lower error than the case of no adjustment, and among the techniques investigated, it is the most successful at predicting the colors for the characterization target on the test substrate.

Figure 7 summarizes the performance of each of the techniques for prediction of test substrate reflectance based on a reference substrate in a bar graph. The abscissa of the bar graph is a rough depiction of the effort required in terms of additional measurements on the test substrate and the height of the bars indicate the average and 95-percentile values of the ΔE^*_{94} color error in the prediction. The short bar in each pair is the mean color error of all patches predicted using the model, and the tall bar the 95 percentile error. At the lowest end of the effort scale is the case of no adjustment, and at the highest end is the case of complete re-characterization on the test substrate by re-measuring the entire characterization target. Figure 7 shows that as the number of measurements increases, the accuracy of the spectral prediction increases.

Of the four techniques implemented, the best results were obtained with the empirical regression technique using PCA. The performance of the model based Beer's law and KM techniques is poor and offers only slight improvement over direct use of the reference substrate data. Several modeling assumptions and uncontrollable variables contribute to the poor performance of the model-based schemes. Beer's law assumes that the toners are transparent, exhibit zero scattering and have constant thickness across changes in substrates. All of these are unrealistic assumptions. Although the KM model eliminates some of Beer's law's deficiencies by compensating for absorption and scattering coefficients separately, it also makes several restrictive assumptions: e.g. a medium that is isotropic and homogeneous, no specular reflection, a planar non-interspersed configuration for the colorant layers, and thicknesses for the

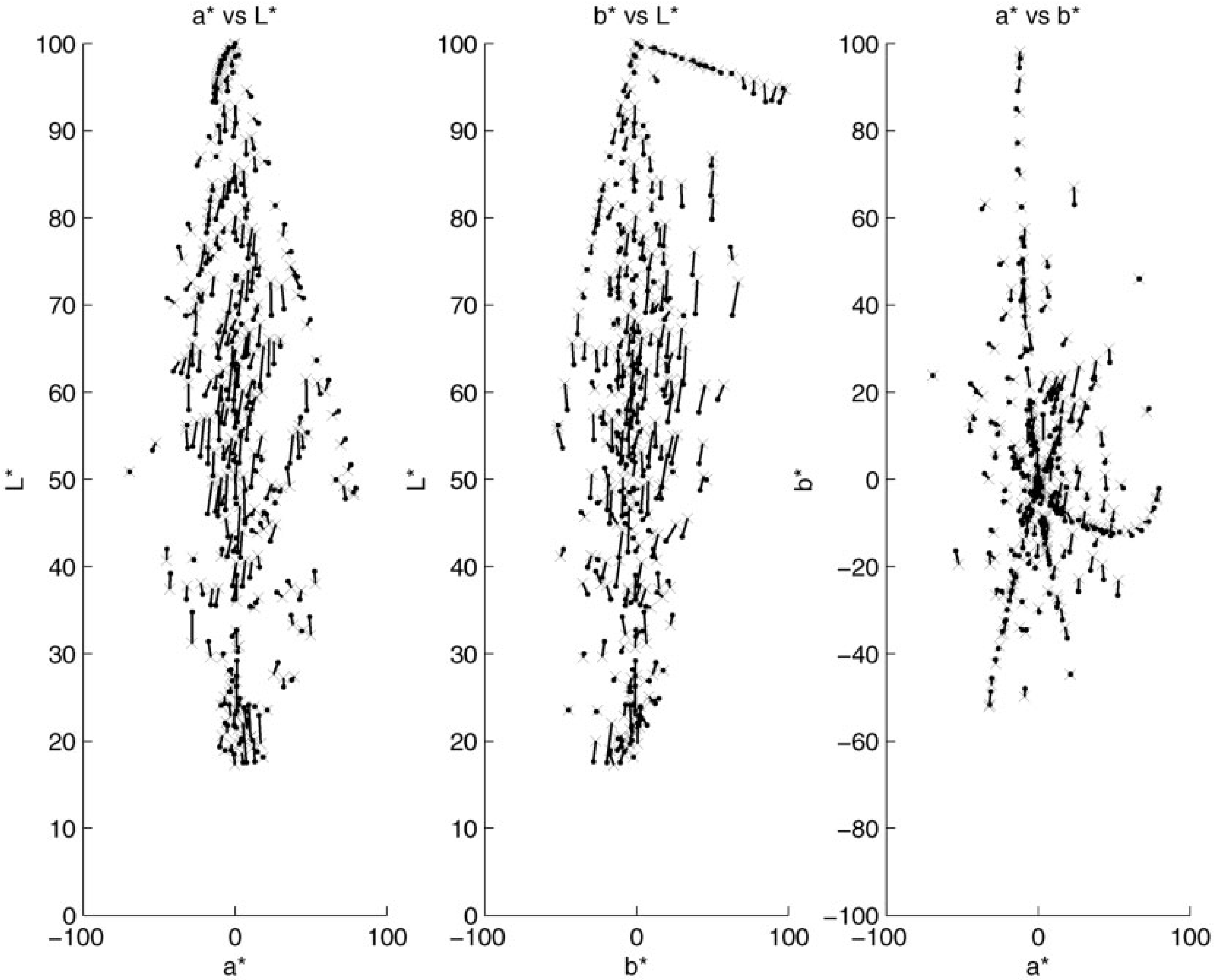

FIG. 10. CIE Lab values measured on a reference substrate (*x*) when compared to the reconstruction CIE Lab values (•) predicted using the model.

layers that are invariant under change in substrate and constant over the spatial extent of a patch. These assumptions are also unrealistic, but cannot be readily improved upon without additional measurement data for the test substrate. Note that if additional information is available the performance of the model based approaches can be improved. In particular, significant improvements have been demonstrated in situations where the statistical characteristics of the the mixing of the colorant layers and variations in thickness are known.[9]

Both the Neugebauer model with measured primaries, and PCA techniques improve prediction accuracy significantly. While these techniques do require additional measurements, the number of measurements (16 and 26, respectively) is much smaller than the 289 measurements required for complete characterization, and can significantly reduce the effort for re-characterizing the printer when a large number of substrates is involved.

Figure 8 shows histograms of the color errors for the case of no compensation and compensation based on the PCA technique. From the histograms, it is clear that the use of an incorrect substrate characterization will result in significant errors. The PCA model significantly reduces both the mean color difference and the standard deviation.

The distribution of errors in color space is further illustrated in Figs. 9 and 10. The errors for the individual patches on the printer characterization target for a sample reference and test substrate are shown in CIELAB space for the case of no compensation and for the PCA technique, respectively. In both figures the errors are depicted in plots of a^* vs L^*, b^* vs L^*, and a^* vs b^* so that the three dimensional distribution of the errors may be visualized. In the plots of Fig. 9, one can see that the dominant error caused by the direct use of the reference substrate characterization on the test substrate is a decrease in lightness. One can also see that the change in substrate causes a change in the black point. If not accounted for correctly, this will result in a reduction of the substrate's characterized dynamic range. In addition to the change in lightness, systematic trends in the color errors can also be seen in the a^* vs b^* plot. By using the

PCA based empirical regression technique to compensate for the change in substrate, it is possible to significantly improve the characterization for the new substrate. This can be seen in Fig. 10, where the errors are significantly smaller than those in Fig. 9 and only relatively minor trends can be seen. The algorithm compensates for the change in black point of the substrate, and significantly reduces the magnitude of the errors in hue and chroma.

CONCLUSIONS

In this article several techniques were evaluated for estimating the printer characterization for a new substrate based on the available characterization for a reference substrate and a small number of measurements on the new substrate. This effort was motivated by the need to provide accurate color management for a large number of substrates in a typical printing environment with minimal additional effort. The proposed techniques included three model-based approaches: Beer's Law, Kubelka-Munk theory, and Neugebauer model; and an empirical regression technique that exploits principal components analysis (PCA) for reducing the dimensionality of the problem to a tractable level. The techniques were experimentally evaluated using a *CMYK* xerographic printer. While the only knowledge required for the Beer's law and Kubelka-Munk methods is the reflectance of the new substrate, these methods provided only small improvement over the direct use of the reference substrate characterization on the test substrate. The unrealistic modeling assumptions required in the absence of additional measurement data on the new substrate are the primary reason for the poor performance of these techniques. The Neugebauer model and the empirical regression technique based on PCA each utilize more measurements, 16 and 26, respectively. Both provided a significant improvement over the direct use of the reference substrate characterization, with PCA providing a more consistent and robust improvement than the Neugebauer model. While the techniques do require measurement of additional patches on the test substrate, the effort in measuring the additional patches is still significantly smaller than the effort required for measuring an entire printer characterization target on the test substrate.

Several advances of this research are conceivable. With the PCA algorithm, using more general mappings such and polynomials or neural-networks for the mapping from reference substrate PCA weights to test substrate PCA weights might achieve a better trade-off between effort and accuracy. Instead of least-squares, the objective function may also be changed to a more relevant metric with nonlinear optimization techniques such as simulated annealing or simplex methods in place of linear regression. Another important extension would be to model the fluorescence commonly encountered in substrates due to the presence of optical brighteners. Finally, modeling of other substrate properties such as surface roughness, gloss, etc. would likely yield even greater improvements in characterization accuracy, though this would also entail more sophisticated analysis of the substrates, which is often impractical.

1. Emmel P. Physical Models for Color Prediction. Chapter 3, *Digital Color Imaging Handbook*, G. Sharma (editor), CRC Press, Boca Raton, FL 2003 pp 173–238.
2. Rogers GL. A Generalized Clapper-Yule Model of Halftone Reflectance. Color Res Appl 2000;25:403–407.
3. Gustavson S. Color Gamut of Halftone Reproduction. J Imag Sci Tech 1997;41:283–290.
4. Arney JS, Wu T, Blehm C. Modeling the Yule-Nielsen Effect on Color Halftones. J Imag Sci Tech 1998;42:335–340.
5. Billmeyer FW, Saltzman M. Principles of Color Technology. Second Edition, John Wiley & Sons, New York, 1981.
6. Allen E. Colorant Formulation and Shading in Optical Radiation Measurement Vol. 2 Color Measurement, Gram F and Bartleson CJ (eds.), Academic Press, New York, 1980.
7. Judd DB, Wyszecki G. Color in Business, Science and Industry. John Wiley & Sons, Inc., New York, 1975.
8. Natale-Hoffman K, Dalal EN, Joseph R. Applications of the Kubelka-Munk Color Model to Xerographic Images. I: Planar Model. Internal Report, X97090406, Xerox Corporation, Webster, NY.
9. Hoffman K. Applications of the Kubelka-Munk Color Model to Xerographic Images. B S Thesis, Rochester Institute of Technology, Center for Imaging Science, 1998, available online at http://www.cis.rit.edu/research/theses.shtml
10. Balasubramanian R. Optimization of the Spectral Neugebauer Model for Printer Characterization. J Elec Imag 1999;8:156–166.
11. Xia M, Saber E, Sharma G, Tekalp AM. End-to-end Color Printer Calibration by Total Least Squares Regression, IEEE Trans Imag Proc 1999;8:700–716.
12. Yule JAC. Principles of Color Reproduction. John Wiley & Sons, New York, 1967.
13. Viggiano JAS. Modeling the Color of Multi-Colored Halftones. Proc TAGA, 1990, 44–62.
14. Yule JAC, Nielsen WJ. The penetration of light into paper and its effect on halftone reproduction. Proc. TAGA, 65–76, 1951.
15. Reachner AC. Methods of Multivariate Analysis. John Wiley & Sons, New York, 1995.
16. Golub GH, Van Loan CF. Matrix Computations. Second Ed., 1989, The Johns Hopkins University Press, Baltimore, MD.
17. Krzanowski WJ. Principles of Multivariate Analysis. Oxford university Press, New York, 1988.
18. Marimont D, Wandel BA. Linear models of surface and illuminant spectra. J Opt Soc A 1992;9:1905–1913.
19. Industrial color difference evaluation. CIE publication No. 116-1995, Central Bureau of the CIE, Vienna, 1995.

APPENDIX

Average and 95 Percentile Error Matrices with no re-calibration

Delta E 94 Statistics - Average Error

Reference Substrate	Paper 01	Paper 02	Paper 03	Paper 04	Paper 05	Paper 06	Paper 07	Paper 08	Paper 09	Paper 10	Paper 11	Paper 12	Average
	Predicted Substrate using Reference												
Paper 01	0.00	0.87	3.18	3.94	2.27	4.41	2.44	2.85	2.03	2.32	2.47	2.58	2.67
Paper 02	0.87	0.00	3.43	4.18	2.27	4.56	2.51	2.97	2.00	2.30	2.48	2.54	2.74
Paper 03	3.18	3.43	0.00	2.19	1.69	2.49	1.54	1.26	2.50	1.90	1.45	1.57	2.11
Paper 04	3.94	4.18	2.19	0.00	3.10	3.48	2.74	1.87	2.96	2.52	2.65	2.51	2.92
Paper 05	2.27	2.27	1.69	3.10	0.00	3.06	0.84	1.59	1.49	1.21	0.93	1.16	1.78
Paper 06	4.41	4.56	2.49	3.48	3.06	0.00	2.90	2.95	3.55	3.34	2.77	3.15	3.33
Paper 07	2.44	2.51	1.54	2.74	0.84	2.90	0.00	1.43	1.32	1.17	0.79	1.08	1.70
Paper 08	2.85	2.97	1.26	1.87	1.59	2.95	1.43	0.00	1.79	1.20	1.17	1.10	1.83
Paper 09	2.03	2.00	2.50	2.96	1.49	3.55	1.32	1.79	0.00	0.94	1.38	1.40	1.94
Paper 10	2.32	2.30	1.90	2.52	1.21	3.34	1.17	1.20	0.94	0.00	1.01	0.78	1.70
Paper 11	2.47	2.48	1.45	2.65	0.93	2.77	0.79	1.17	1.38	1.01	0.00	0.80	1.63
Paper 12	2.58	2.54	1.57	2.51	1.16	3.15	1.08	1.10	1.40	0.78	0.80	0.00	1.70
Average	2.67	2.74	2.11	2.92	1.78	3.33	1.70	1.83	1.94	1.70	1.63	1.70	

Delta E 94 Statistics - 95 Percentile Error

Reference Substrate	Paper 01	Paper 02	Paper 03	Paper 04	Paper 05	Paper 06	Paper 07	Paper 08	Paper 09	Paper 10	Paper 11	Paper 12	Average
	Predicted Substrate using Reference												
Paper 01	0.00	1.60	6.35	7.88	4.81	7.78	5.04	6.11	3.98	5.16	4.80	5.98	5.41
Paper 02	1.60	0.00	6.31	7.67	4.61	8.22	4.74	6.40	3.73	5.03	4.67	6.05	5.37
Paper 03	6.35	6.31	0.00	3.80	2.90	5.42	3.45	2.46	4.58	3.06	2.54	2.63	3.95
Paper 04	7.88	7.67	3.80	0.00	5.03	6.47	4.77	3.09	5.43	3.84	4.58	4.02	5.14
Paper 05	4.81	4.61	2.90	5.03	0.00	6.00	1.76	3.17	2.94	2.38	1.67	2.19	3.41
Paper 06	7.78	8.22	5.42	6.47	6.00	0.00	5.88	6.23	6.88	6.51	5.77	6.37	6.50
Paper 07	5.04	4.74	3.45	4.77	1.76	5.88	0.00	3.54	2.48	2.17	1.60	2.12	3.41
Paper 08	6.11	6.40	2.46	3.09	3.17	6.23	3.54	0.00	4.70	2.25	3.16	2.47	3.96
Paper 09	3.98	3.73	4.58	5.43	2.94	6.88	2.48	4.70	0.00	2.81	2.49	3.34	3.94
Paper 10	5.16	5.03	3.06	3.84	2.38	6.51	2.17	2.25	2.81	0.00	2.07	1.53	3.35
Paper 11	4.80	4.67	2.54	4.58	1.67	5.77	1.60	3.16	2.49	2.07	0.00	1.79	3.19
Paper 12	5.98	6.05	2.63	4.02	2.19	6.37	2.12	2.47	3.34	1.53	1.79	0.00	3.50
Average	5.41	5.37	3.95	5.14	3.41	6.50	3.41	3.96	3.94	3.35	3.19	3.50	

Average and 95 Percentile Error Matrices for Beers Law prediction

Delta E 94 Statistics - Average Error

Reference Substrate	Paper 01	Paper 02	Paper 03	Paper 04	Paper 05	Paper 06	Paper 07	Paper 08	Paper 09	Paper 10	Paper 11	Paper 12	Average
	Predicted Substrate using Reference												
Paper 01	0.00	0.84	3.03	3.31	2.19	4.18	2.34	2.65	1.87	2.15	2.34	2.43	2.48
Paper 02	0.82	0.00	3.28	3.46	2.21	4.32	2.42	2.76	1.80	2.10	2.36	2.39	2.54
Paper 03	2.97	3.31	0.00	1.76	1.66	2.36	1.49	1.18	2.38	1.79	1.37	1.48	1.98
Paper 04	3.63	4.01	1.93	0.00	2.92	3.20	2.55	1.67	2.76	2.33	2.42	2.32	2.70
Paper 05	2.11	2.17	1.61	2.44	0.00	2.90	0.79	1.44	1.36	1.08	0.87	1.07	1.62
Paper 06	4.10	4.35	2.37	3.06	2.96	0.00	2.78	2.77	3.37	3.14	2.62	2.95	3.13
Paper 07	2.28	2.42	1.47	2.22	0.80	2.76	0.00	1.34	1.26	1.09	0.75	1.02	1.58
Paper 08	2.64	2.86	1.20	1.48	1.51	2.80	1.36	0.00	1.69	1.12	1.10	1.04	1.71
Paper 09	1.85	1.86	2.37	2.51	1.36	3.36	1.25	1.67	0.00	0.88	1.26	1.28	1.79
Paper 10	2.15	2.18	1.80	2.07	1.11	3.17	1.11	1.12	0.89	0.00	0.93	0.71	1.57
Paper 11	2.30	2.38	1.38	2.09	0.89	2.63	0.76	1.08	1.27	0.92	0.00	0.74	1.50
Paper 12	2.41	2.44	1.50	2.00	1.10	2.98	1.04	1.02	1.29	0.71	0.75	0.00	1.57
Average	2.48	2.62	1.99	2.40	1.70	3.15	1.63	1.70	1.81	1.57	1.52	1.58	

Delta E 94 Statistics - 95 Percentile Error

Reference Substrate	Paper 01	Paper 02	Paper 03	Paper 04	Paper 05	Paper 06	Paper 07	Paper 08	Paper 09	Paper 10	Paper 11	Paper 12	Average
	Predicted Substrate using Reference												
Paper 01	0.00	1.56	6.07	6.77	4.62	7.45	4.82	5.65	3.83	4.86	4.54	5.69	5.08
Paper 02	1.52	0.00	6.05	6.58	4.47	7.90	4.49	5.86	3.41	4.75	4.40	5.61	5.00
Paper 03	5.89	6.06	0.00	3.11	2.81	5.14	3.33	2.33	4.35	2.87	2.40	2.46	3.70
Paper 04	7.31	7.34	3.53	0.00	4.78	6.00	4.55	2.75	5.20	3.49	4.19	3.74	4.81
Paper 05	4.44	4.38	2.73	4.03	0.00	5.82	1.66	2.93	2.76	2.15	1.57	2.11	3.15
Paper 06	7.28	7.85	5.15	6.40	5.85	0.00	5.66	5.89	6.47	6.13	5.46	6.00	6.19
Paper 07	4.74	4.45	3.32	3.96	1.65	5.63	0.00	3.29	2.37	2.01	1.50	1.99	3.17
Paper 08	5.81	6.15	2.30	2.46	3.10	5.96	3.40	0.00	4.45	2.11	2.98	2.34	3.73
Paper 09	3.78	3.52	4.34	4.82	2.76	6.36	2.34	4.40	0.00	2.62	2.26	3.10	3.66
Paper 10	4.81	4.84	2.89	3.19	2.12	6.17	2.08	2.11	2.66	0.00	1.94	1.44	3.11
Paper 11	4.45	4.43	2.41	3.55	1.60	5.51	1.52	2.92	2.27	1.93	0.00	1.68	2.93
Paper 12	5.60	5.89	2.48	3.27	2.20	6.08	2.03	2.29	3.15	1.43	1.69	0.00	3.28
Average	5.06	5.13	3.75	4.38	3.27	6.18	3.26	3.68	3.72	3.12	3.00	3.29	

Average and 95 Percentile Error Matrices for Neugebauer, using Kubelka Munk prediction for Primaries

Delta E 94 Statistics - Average Error

Predicted Substrate using Reference

Reference Substrate	*Paper 01*	*Paper 02*	*Paper 03*	*Paper 04*	*Paper 05*	*Paper 06*	*Paper 07*	*Paper 08*	*Paper 09*	*Paper 10*	*Paper 11*	*Paper 12*	*Average*
Paper 01	0.00	0.82	3.29	3.68	2.34	4.48	2.51	2.84	1.95	2.30	2.52	2.60	2.67
Paper 02	0.89	0.00	3.50	3.99	2.33	4.59	2.57	2.97	1.94	2.28	2.53	2.56	2.74
Paper 03	3.40	3.50	0.00	1.53	1.55	2.48	1.49	1.30	2.48	1.99	1.49	1.66	2.08
Paper 04	3.47	3.50	1.43	0.00	2.05	3.05	1.91	1.29	2.46	2.00	1.87	1.82	2.26
Paper 05	2.44	2.38	1.65	2.74	0.00	3.02	0.81	1.51	1.45	1.21	0.93	1.18	1.76
Paper 06	4.57	4.57	2.48	3.29	2.95	0.00	2.85	3.00	3.52	3.40	2.79	3.20	3.33
Paper 07	2.56	2.58	1.53	2.48	0.81	2.88	0.00	1.41	1.33	1.19	0.79	1.10	1.70
Paper 08	2.89	2.87	1.29	1.38	1.33	2.96	1.32	0.00	1.69	1.19	1.08	1.05	1.73
Paper 09	2.02	1.94	2.48	2.81	1.40	3.52	1.31	1.75	0.00	0.91	1.30	1.31	1.89
Paper 10	2.34	2.22	1.95	2.18	1.10	3.36	1.16	1.19	0.88	0.00	0.98	0.73	1.64
Paper 11	2.56	2.51	1.47	2.26	0.89	2.77	0.78	1.13	1.28	0.99	0.00	0.81	1.59
Paper 12	2.66	2.52	1.62	2.03	1.06	3.16	1.05	1.06	1.27	0.74	0.79	0.00	1.63
Average	2.71	2.67	2.06	2.58	1.62	3.30	1.61	1.77	1.84	1.65	1.55	1.64	

Delta E 94 Statistics - 95 Percentile Error

Predicted Substrate using Reference

Reference Substrate	*Paper 01*	*Paper 02*	*Paper 03*	*Paper 04*	*Paper 05*	*Paper 06*	*Paper 07*	*Paper 08*	*Paper 09*	*Paper 10*	*Paper 11*	*Paper 12*	*Average*
Paper 01	0.00	1.60	6.37	7.96	4.48	8.01	4.91	6.09	4.02	5.12	4.84	5.82	5.38
Paper 02	1.72	0.00	6.43	8.51	4.54	8.35	4.69	6.44	3.63	5.10	4.83	6.19	5.49
Paper 03	6.74	6.49	0.00	2.97	2.69	5.42	3.37	2.53	4.55	3.18	2.60	2.75	3.94
Paper 04	7.01	6.74	2.64	0.00	3.41	6.04	3.64	2.30	4.64	3.13	3.47	3.02	4.19
Paper 05	4.88	4.72	2.88	4.76	0.00	6.02	1.79	2.97	2.89	2.43	1.75	2.23	3.39
Paper 06	8.15	8.37	5.42	6.23	5.92	0.00	5.83	6.33	6.63	6.61	5.81	6.51	6.53
Paper 07	5.22	4.78	3.43	4.67	1.72	5.89	0.00	3.49	2.46	2.16	1.55	2.12	3.41
Paper 08	6.16	6.08	2.47	2.57	2.65	6.24	3.27	0.00	4.57	2.25	3.00	2.47	3.79
Paper 09	4.14	3.62	4.57	5.39	2.84	6.63	2.46	4.71	0.00	2.75	2.36	3.20	3.88
Paper 10	5.23	4.99	3.12	3.49	2.19	6.52	2.08	2.25	2.66	0.00	1.98	1.51	3.27
Paper 11	5.01	4.64	2.56	4.19	1.62	5.76	1.53	3.11	2.33	2.01	0.00	1.76	3.14
Paper 12	5.95	5.83	2.70	3.55	2.02	6.41	2.01	2.48	3.10	1.50	1.72	0.00	3.39
Average	5.47	5.26	3.87	4.94	3.10	6.48	3.23	3.88	3.77	3.29	3.08	3.42	

Average and 95 Percentile Error Matrices for Neugebauer, using measured primaries

Delta E 94 Statistics - Average Error

Predicted Substrate using Reference

Reference Substrate	*Paper 01*	*Paper 02*	*Paper 03*	*Paper 04*	*Paper 05*	*Paper 06*	*Paper 07*	*Paper 08*	*Paper 09*	*Paper 10*	*Paper 11*	*Paper 12*	*Average*
Paper 01	1.19	1.35	2.23	2.33	2.35	4.44	2.88	2.33	2.32	2.21	2.54	2.37	2.38
Paper 02	1.21	1.20	2.11	2.24	2.21	4.22	2.73	2.20	2.17	2.09	2.40	2.28	2.26
Paper 03	1.96	1.92	1.60	1.87	1.60	3.31	1.79	1.82	1.65	1.75	1.70	1.94	1.91
Paper 04	1.98	1.91	1.80	1.74	1.77	3.42	1.93	1.87	1.75	1.90	1.83	2.05	2.00
Paper 05	1.97	1.91	1.66	1.85	1.59	3.27	1.83	1.79	1.64	1.77	1.71	1.98	1.91
Paper 06	3.71	3.51	2.69	2.82	2.62	1.97	2.49	2.76	2.68	2.92	2.44	3.11	2.81
Paper 07	2.45	2.38	1.81	1.95	1.79	3.16	1.72	1.98	1.84	2.03	1.79	2.16	2.09
Paper 08	1.78	1.72	1.71	1.81	1.72	3.53	2.06	1.64	1.63	1.72	1.74	1.95	1.92
Paper 09	1.88	1.80	1.64	1.79	1.61	3.36	1.86	1.68	1.53	1.68	1.64	1.88	1.86
Paper 10	1.87	1.88	1.94	2.17	1.91	3.87	2.24	2.09	1.86	1.84	2.04	2.02	2.14
Paper 11	2.22	2.13	1.67	1.81	1.64	3.11	1.74	1.70	1.59	1.77	1.53	1.97	1.91
Paper 12	2.16	2.18	2.12	2.39	2.08	3.92	2.27	2.34	2.09	2.09	2.20	2.12	2.33
Average	2.03	1.99	1.92	2.06	1.91	3.47	2.13	2.02	1.90	1.98	1.96	2.15	

Delta E 94 Statistics - 95 Percentile Error

Predicted Substrate using Reference

Reference Substrate	*Paper 01*	*Paper 02*	*Paper 03*	*Paper 04*	*Paper 05*	*Paper 06*	*Paper 07*	*Paper 08*	*Paper 09*	*Paper 10*	*Paper 11*	*Paper 12*	*Average*
Paper 01	2.86	3.06	4.48	4.54	4.56	8.86	5.70	4.66	4.55	4.46	4.94	4.59	4.77
Paper 02	2.61	2.75	4.17	4.25	4.37	8.52	5.29	4.40	4.29	4.09	4.63	4.52	4.49
Paper 03	3.94	3.98	3.99	4.01	3.98	7.13	4.03	4.56	3.79	3.98	3.85	4.69	4.33
Paper 04	4.08	3.99	4.14	3.96	4.48	7.22	4.38	4.56	3.98	4.45	4.28	5.04	4.55
Paper 05	4.06	3.99	3.82	3.92	3.99	6.93	4.15	4.62	3.65	4.25	3.92	5.03	4.36
Paper 06	6.88	6.79	5.80	6.08	5.55	4.51	5.08	5.49	5.56	6.30	5.01	7.13	5.85
Paper 07	4.96	5.01	4.27	4.19	4.20	6.59	4.36	5.21	4.73	5.07	4.34	5.46	4.87
Paper 08	3.77	3.73	4.07	4.21	4.29	7.28	4.50	4.27	3.94	4.13	4.14	4.71	4.42
Paper 09	3.90	3.76	3.71	4.00	4.09	6.89	4.15	4.39	3.55	3.80	4.00	4.57	4.23
Paper 10	3.75	3.78	4.43	4.68	4.51	8.24	4.78	4.82	4.31	4.59	4.74	4.57	4.77
Paper 11	4.40	4.43	3.83	4.11	3.87	6.37	4.12	4.71	3.67	4.53	3.97	5.12	4.43
Paper 12	4.38	4.31	4.52	4.80	4.80	8.16	4.90	5.31	4.58	4.61	4.86	4.60	4.99
Average	4.13	4.13	4.27	4.40	4.39	7.23	4.62	4.75	4.22	4.52	4.39	5.00	

Average and 95 Percentile Error Matrices for PCA Analysis Calibration

Delta E 94 Statistics - Average Error

Predicted Substrate using Reference (columns); Reference Substrate (rows)

	Paper 01	Paper 02	Paper 03	Paper 04	Paper 05	Paper 06	Paper 07	Paper 08	Paper 09	Paper 10	Paper 11	Paper 12	Average
Paper 01	0.00	0.77	2.03	1.96	1.92	2.82	2.07	2.11	1.73	1.97	1.96	2.13	1.95
Paper 02	1.06	0.00	1.59	1.49	1.55	2.52	1.70	1.88	1.36	1.61	1.61	1.76	1.65
Paper 03	2.17	1.99	0.00	0.69	1.04	2.30	1.15	1.07	1.26	0.82	0.95	0.94	1.31
Paper 04	3.28	2.98	0.79	0.00	1.37	2.47	1.57	1.10	1.77	1.00	1.32	1.15	1.71
Paper 05	1.71	1.57	0.75	0.90	0.00	2.06	0.65	1.07	0.66	0.72	0.89	1.03	1.09
Paper 06	3.02	2.81	2.02	2.00	2.01	0.00	1.95	2.33	2.03	2.14	1.87	2.21	2.22
Paper 07	2.09	1.82	1.03	1.02	0.70	2.14	0.00	1.21	0.89	0.87	0.87	0.99	1.24
Paper 08	2.39	2.13	1.12	0.99	1.37	2.57	1.49	0.00	1.48	0.96	1.43	1.06	1.54
Paper 09	1.77	1.52	1.17	1.07	0.78	2.04	0.78	1.29	0.00	0.79	0.77	0.94	1.18
Paper 10	2.42	2.18	0.93	0.96	0.85	2.30	1.02	0.91	1.13	0.00	0.93	0.71	1.30
Paper 11	2.05	1.79	0.77	0.92	0.82	1.90	0.83	1.06	0.84	0.81	0.00	0.98	1.16
Paper 12	2.30	2.08	1.03	1.07	1.09	2.55	1.12	1.07	1.31	0.68	1.21	0.00	1.41
Average	2.21	1.97	1.20	1.19	1.23	2.33	1.30	1.37	1.31	1.12	1.26	1.26	

Delta E 94 Statistics - 95 Percentile Error

Predicted Substrate using Reference (columns); Reference Substrate (rows)

	Paper 01	Paper 02	Paper 03	Paper 04	Paper 05	Paper 06	Paper 07	Paper 08	Paper 09	Paper 10	Paper 11	Paper 12	Average
Paper 01	0.00	1.58	4.15	4.03	3.79	5.50	4.07	4.16	3.41	3.74	3.85	3.87	3.83
Paper 02	1.81	0.00	3.21	3.31	2.90	5.03	3.12	3.88	2.53	3.29	3.15	3.39	3.24
Paper 03	4.36	4.05	0.00	1.41	2.31	4.78	2.31	2.27	2.57	1.69	2.06	1.95	2.71
Paper 04	6.61	5.83	1.74	0.00	3.02	5.36	3.33	1.97	3.70	2.08	3.10	2.45	3.56
Paper 05	3.02	2.89	1.79	1.74	0.00	4.17	1.28	2.21	1.50	1.64	1.78	2.49	2.23
Paper 06	6.46	6.00	3.95	3.90	4.05	0.00	3.96	4.59	4.02	4.01	3.92	4.09	4.45
Paper 07	3.93	3.28	2.13	2.00	1.33	4.31	0.00	2.26	1.61	1.61	1.66	2.04	2.38
Paper 08	4.55	4.07	2.11	1.80	2.50	5.00	2.81	0.00	2.87	2.00	2.76	2.33	2.98
Paper 09	3.43	3.00	2.22	2.37	1.44	4.17	1.54	2.43	0.00	1.47	1.62	1.75	2.31
Paper 10	4.84	4.06	1.99	2.08	1.69	4.53	1.95	1.97	2.20	0.00	1.98	1.79	2.64
Paper 11	4.02	3.63	1.78	1.77	1.63	3.83	1.58	1.97	1.72	1.53	0.00	2.05	2.32
Paper 12	4.04	3.99	2.11	2.23	2.08	5.02	2.10	2.51	2.40	1.40	2.39	0.00	2.75
Average	4.28	3.85	2.47	2.42	2.43	4.70	2.55	2.75	2.59	2.22	2.57	2.56	