Total Least Square Techniques in Color Printer Characterization

Minghui Xia[1], Eli Saber[1,2], Gaurav Sharma[2], A. Murat Tekalp[1] and Ronald Sosinski[2]

[1]Department of Electrical Engineering and Center for Electronic Imaging Systems
Hopeman 204, University of Rochester, Rochester, NY 14627-0126
{xia,saber,tekalp}@ee.rochester.edu
[2]Xerox Corporation, 800 Phillips Road, Webster, NY 14580
eli_saber@xn.xerox.com, sharma@wrc.xerox.com, ronald_sosinski@xn.xerox.com

Abstract

Neugebauer model is a powerful tool in obtaining end-to-end device characterization profiles for halftone color printer calibration. In this paper, we propose total least square (TLS) regression methods to estimate the parameters of Neugebauer models. Compared to the traditional least squares (LS) based methods, the TLS approach is a physically more appropriate procedure, because it accounts for errors in the measured reflectance of both the selected primaries and the modeled reflectance. The proposed TLS techniques are tested on a Xerox color printer with rotated halftone screen, and the results are compared with the LS based algorithms. Our experiments indicate that the TLS methods yield a significant improvement over the LS based techniques for model parameter estimation.

1 Color printer calibration and the Neugebauer model

In color printer calibration, a device characterization function, which maps from device control values to *device independent* (DVI) color space, is first determined. This characterization function is then inverted, so that the digital control values required to produce a given color (specified in a DVI color space) may be computed. This paper focuses on the characterization of halftone color printers, for which the Neugebauer model and its variants [1, 2, 3] offer an attractive characterization method. The model parameters can often be determined from a small number of measurements, as opposed to the large number of measurements required in an empirical scheme.

According to the Neugebauer model, for halftone color printers using Cyan (C), Magenta (M), Yellow (Y) and Black (K) colorants, up to $2^4 = 16$ different colored regions or primaries are produced on paper through subtractive overlap of none, one, two, three or four colorants. The color tristimulus (see [4] for clear definitions) of a halftone print can be expressed as the weighted average of the tristimuli of these 16 overlapping combinations, referred to as Neugebauer primaries. Recently, considerable success has been obtained with a spectral Neugebauer model [5, 3], where the spectral micro-reflectance (instead of tristimuli) of a halftoned region is expressed as the weighted average of the micro-reflectance of the individual Neugebauer primaries. Thus, it can be written as

$$r^{1/n}(\lambda;\mathbf{w}) = \sum_{i=1}^{P} w_i r_i{}^{1/n}(\lambda), \tag{1}$$

where P is the number of primaries defined in the model (e.g. 16), λ denotes the wavelength of light, $r(\lambda)$ is the predicted spectral reflectance corresponding to a halftone print, $r_i(\lambda)$ is the reflectance of the i^{th} primary, n represents the empirically determined *Yule-Nielsen* (YN) correction factor, which accounts for the penetration and scattering of light in paper, known as the *Yule-Nielsen effect* [6, 7], w_i denotes the fractional areas of the the i^{th} Neugebauer primary. Assuming that the distribution of the colorants on paper is independent, w_i are given by the following Demichel equations [8]:

$$\begin{aligned} w_i \in \{ & (1-c)(1-m)(1-y)(1-k), \\ & (1-c)(1-m)(1-y)k, \\ & (1-c)(1-m)y(1-k), \\ & (1-c)(1-m)yk, (1-c)m(1-y)(1-k), \\ & (1-c)m(1-y)k, (1-c)my(1-k), \\ & (1-c)myk, c(1-m)(1-y)(1-k), \\ & c(1-m)(1-y)k, c(1-m)y(1-k), \\ & c(1-m)yk, cm(1-y)(1-k), \\ & cm(1-y)k, cmy(1-k), cmyk \} \end{aligned} \tag{2}$$

0-8186-8821-1/98 $10.00

where c, m, y, and k represent the fractional areas covered by the C, M, Y, and K colorants respectively in their individual separations. The relationship between the actual dot areas c, m, y, k and the digital control values C, M, Y, K is usually nonlinear, and is often referred to as the dot area function.

2 Total least squares technique for estimating the Neugebauer model parameters

In order to utilize the Neugebauer model to predict the printer behavior, the model parameters (e.g. dot area function, YN correction) need to be estimated. The correctness of the Neugebauer model relies heavily on the accuracy of the estimated parameters. The YN correction parameter n is estimated by iterating through a set of candidate values within physically meaningful boundaries. The values leading to the smallest prediction error is selected as the optimum. Previously, the dot area function is estimated by a global search technique to minimize a predefined model prediction error [3]. Although the search provides satisfactory results, it is computationally expensive. An alternative approach is to perform the optimization by least squares (LS) estimation. The LS approach is based on the observation that the measurement and the model prediction of the reflectance $r(\lambda)$ is prone to error. Therefore, the representation of the Neugebauer model would be

$$r^{1/n}(\lambda;\mathbf{w}) = \sum_{i=1}^{P} w_i r_i^{1/n}(\lambda) + e(\lambda) \tag{3}$$

where $e(\lambda)$ represents the measurement and model error (in the YN corrected spectral space). Least-squares regression can then be used to determine the areas of the primaries w_i so as to minimize the mean squared error. However, it should be noted that the measurement of the primary reflectance $r_i(\lambda)$ is also subject to error. Hence, a more accurate model is the one that allows errors in all measured quantities, which is given by

$$r^{1/n}(\lambda;\mathbf{w}) + e(\lambda) = \sum_{i=1}^{P} w_i[r_i^{1/n}(\lambda) + e_i(\lambda)] \tag{4}$$

where $e(\lambda)$ denotes the measurement and model errors in the YN-corrected spectral space, and $e_i(\lambda)$ is the error in the YN-corrected measured reflectance of the i^{th} primary.

In order to apply the TLS method to the printer modeling problem, the unity sum constraint on the fractional areas of the primaries, e.g., $\sum_{i=1}^{P} w_i = 1$, must be incorporated into (4). To this effect, we assume without loss of generality that the first Neugebauer primary corresponds to paper white with reflectance $r_1(\lambda)$. Then, subtracting $r_1^{1/n}(\lambda)$ from both sides, Eq. (4) can be rewritten as

$$r'(\lambda;\mathbf{w}') + e'(\lambda) = \sum_{i=2}^{P} w_i(r_i'(\lambda) + e_i(\lambda)) \tag{5}$$

where

$$\begin{aligned}
\mathbf{w}' &= [w_2, w_3, \ldots w_p]^T, \\
r'(\lambda;\mathbf{w}') &= r^{1/n}(\lambda;\mathbf{w}) - r_1^{1/n}(\lambda), \\
r_i'(\lambda) &= r_i^{1/n}(\lambda) - r_1^{1/n}(\lambda), \\
e'(\lambda) &= e(\lambda) - w_1 e_1(\lambda).
\end{aligned}$$

Typically the color spectra are sampled at discrete wavelengths, $[\lambda_1, \lambda_2, \ldots \lambda_N]$, so that (5) can be written in matrix-vector notation as

$$\mathbf{r}'(\mathbf{w}') + \mathbf{e}' = \sum_{i=2}^{P} w_i(\mathbf{r}_i' + \mathbf{e}_i) = (\mathbf{R}_p' + \mathbf{E})\mathbf{w}' \tag{6}$$

where

$$\begin{aligned}
\mathbf{r}' &= [r'(\lambda_1;\mathbf{w}'), \ldots r'(\lambda_N;\mathbf{w}')]^T, \\
\mathbf{r}_i' &= [r_i'(\lambda_1), \ldots r_i'(\lambda_N)]^T, \\
\mathbf{R}_p' &= [\mathbf{r}_2', \mathbf{r}_3' \ldots \mathbf{r}_P'], \\
\mathbf{e}' &= [e'(\lambda_1), \ldots e'(\lambda_N)]^T, \\
\mathbf{e}_i &= [e_i(\lambda_1), \ldots e_i(\lambda_N)]^T, \\
\mathbf{E} &= [\mathbf{e}_2, \mathbf{e}_3 \ldots \mathbf{e}_P].
\end{aligned}$$

The vector $\mathbf{r}'$ represents the YN-corrected and paper "normalized" reflectance. With the measurement of multiple color patches, the above problem can be reduced to solving a set of linear equations $AX = b$ (one-dimensional case) or $AX = B$ (multi-dimensional case), depending on the type of color patches printed and measured. Considering the error associated with all the measurement terms, the solution can be obtained through a one-dimensional total least square (TLS) method, which seeks to

$$\text{minimize } \|[A;b] - [\hat{A};\hat{b}]\|_F, \text{subject to } \hat{A}x = \hat{b}, \tag{7}$$

or a multi-dimensional TLS method, which

$$\begin{aligned}
&\text{minimize } \|[A;B] - [\hat{A}\hat{B}]\|_F \\
&\text{subject to } \hat{A}X = \hat{B}; \quad [\hat{A};\hat{B}] \in \Re^{m\times(n+d)},
\end{aligned} \tag{8}$$

where 'F' denotes the Frobenius norm [9].

The TLS solution is computed through singular value decomposition (SVD). Compared with its LS

counterpart, TLS method is physically more appropriate because it accounts for errors in the measured reflectance of both the selected primaries and the modeled reflectance. The estimation of $\{w_i\}$ and the primary measurement error correction can be obtained at the same time. The correction to the primaries can then be utilized to update the primaries. A comparison of LS and TLS in the one-dimensional line-fitting problem is illustrated in Fig. 1. Both LS and TLS measures of goodness-of-fit are depicted. The LS method minimizes the squared sum of the vertical distances, whereas the TLS method minimizes the squared sum of the perpendicular distances.

2.1 TLS for single-colorant step-wedges

The TLS method can be utilized to obtain the dot area function from single-colorant prints. Usually, these prints are in a sequence with the (fractional) colorant coverage on the paper increasing monotonically from zero to one. The prints are generated by stepping through the digital values used for driving the printer, typically from 0 to 255, and are therefore referred to as step-wedges. Since there is only one colorant in this case, a simplified Neugebauer model with only 2 primaries (one colorant and paper white) is applicable. Thus, for a K-step cyan wedge, with digital values $0 \leq C_1 \leq C_2 \leq \ldots \leq C_K \leq 255$, the model of (6) reduces to

$$(\mathbf{r}'_{pc} + \mathbf{e}'_{pc})c_j = \mathbf{r}'_{c_j} + \mathbf{e}'_{c_j}, \qquad j = 1, 2, \ldots K \quad (9)$$

where c_j denotes the dot area corresponding to the digital step value C_j, $\mathbf{r}'_{pc}$ denotes the cyan primary reflectance, $\mathbf{r}'_{c_j}$ denotes the reflectance of the j^{th} step, $\mathbf{e}'_{c_j}$, and $\mathbf{e}'_{pc}$ denotes the measurement error in $\mathbf{r}'_{c_j}$ and $\mathbf{r}'_{pc}$ respectively. The K equations in (9) may be combined as

$$(\mathbf{r}'_{pc} + \mathbf{e}'_{pc})\mathbf{c}^T = \mathbf{R}'_c + \mathbf{E}'_c \quad (10)$$

where $\mathbf{c} = [c_1, c_2, \ldots c_K]^T$, $\mathbf{R}'_c = [\mathbf{r}'_{c_1}, \mathbf{r}'_{c_2}, \ldots \mathbf{r}'_{c_K}]$, and $\mathbf{E}'_c = [\mathbf{e}'_{c_1}, \mathbf{e}'_{c_2}, \ldots \mathbf{e}'_{c_K}]$.

Observing that the above equations represent a multi-dimensional TLS problem, we can "solve" them to obtain simultaneously the dot areas of cyan $\mathbf{c}$ and the correction $\mathbf{e}'_{pc}$ to the cyan primary reflectance.

2.2 TLS for multi-colorant step-wedges

The dot area functions estimated from single-colorant step-wedges can be iteratively refined, one colorant at a time, by using multi-colorant step-wedges in which the dot areas of other colorants are known and assumed fixed. Consider the process of refining the dot areas for the cyan colorant using a multi-colorant cyan step-wedge. Suppose the cyan digital values for this multi-colorant step-wedge are $0 \leq C_1 \leq C_2 \ldots \leq C_K \leq 255$ and the digital values (and therefore fractional areas) of other colorants are constant. In the resulting equations, on the right hand side of (6), we can group together the terms with the factor c (these correspond to primaries that include the cyan colorant) and other terms with the factor $(1-c)$ (these correspond to primaries that do not include the cyan colorant), to get

$$\mathbf{r}^{\star\,\prime}_{c_j} + \mathbf{e}^{\star\,\prime}_{c_j} = c_j \sum_{k \in S_C} w^\star_k (\mathbf{r}'_{p_k} + \mathbf{e}'_{p_k}) + (1 - c_j) \sum_{l \in S_{NC}} w^\star_l (\mathbf{r}'_{p_l} + \mathbf{e}'_{p_l}), \quad j = 1, 2 \ldots K \quad (11)$$

where c_j denotes the fractional area of the cyan colorant corresponding to the digital value C_j; $\mathbf{r}^{\star\,\prime}_{c_j}$ is the reflectance of the multi-colorant step-wedge print with cyan digital value C_j; the sets S_C and S_{NC} represent a partition of the primary indices $\{2, 3, \ldots 15\}$, such that the elements of S_C correspond to the primaries having the cyan colorant as a constituent and elements of S_{NC} correspond to the primaries that do not have the cyan colorant as a constituent; and

$$w^\star_k = \begin{cases} \frac{w^j_k}{c_j} & k \in S_C \\ \frac{w^j_k}{1-c_j} & k \in S_{NC} \end{cases}$$

where w^j_k is the fractional area of the k^{th} primary for the j^{th} step print (as in (6)). Using the Demichel equation (2), we can see that the factors $w^\star_k, w^\star_l \in \{(1-m)(1-y)(1-k),\ (1-m)(1-y)k,\ (1-m)y(1-k),\ (1-m)yk,\ m(1-y)(1-k),\ m(1-y)k,\ my(1-k),\ myk\}$, where m, y, and k are the (fixed) dot areas for the magenta, yellow, and black colorants for the whole multi-colorant cyan step-wedge. These equations can be further rewritten as

$$(\mathbf{r}^\star_{pc} + \mathbf{e}^\star_{pc})c_j = \mathbf{r}^\star_{cw_j} + \mathbf{e}^\star_{cw_j}, \qquad j = 1, 2 \ldots K \quad (12)$$

where $\mathbf{r}^\star_{cw_j} = \mathbf{r}^{\star\,\prime}_{c_j} - \sum_{l \in S_{NC}} w^\star_l \mathbf{r}'_{p_l}$, $\mathbf{r}^\star_{pc} = \sum_{k \in S_C} w^\star_k \mathbf{r}'_{p_k} - \sum_{l \in S_{NC}} w^\star_l \mathbf{r}'_{p_l}$, and $\mathbf{e}^\star_{pc}$, $\mathbf{e}^\star_{cw_j}$ are the corresponding combined errors. The K equations in (12) may be combined as

$$(\mathbf{r}^\star_{pc} + \mathbf{e}^\star_{pc})\mathbf{c}^\mathbf{T} = \mathbf{R}^\star_{cw} + \mathbf{E}^\star_{cw} \quad (13)$$

where $\mathbf{R}^\star_{cw} = [\mathbf{r}^\star_{cw_1} \ldots \mathbf{r}^\star_{cw_N}]$, and $\mathbf{E}^\star_{cw} = [\mathbf{e}^\star_{cw_1} \cdots \mathbf{e}^\star_{cw_N}]$. If we constrain the digital control values of cyan in the multi-colorant step-wedge of cyan sweeps to be consistent with those in single-colorant cyan step-wedge, and combine single step-wedges (10) and multi-colorant step-wedges (13), we get

$$\left(\begin{bmatrix} \mathbf{r}'_{pc} \\ \mathbf{r}^\star_{pc} \end{bmatrix} + \begin{bmatrix} \mathbf{e}'_{pc} \\ \mathbf{e}^\star_{pc} \end{bmatrix} \right) \mathbf{c}^\mathbf{T} = \begin{bmatrix} \mathbf{R}'_c \\ \mathbf{R}^\star_{cw} \end{bmatrix} + \begin{bmatrix} \mathbf{E}'_c \\ \mathbf{E}^\star_{cw} \end{bmatrix} \quad (14)$$

where the terms are the same as defined in (10) and (13). The dot area function for the cyan colorant can now be obtained by "solving" the above system of equations as a multi-dimensional TLS problem. The same approach can be applied to the other colorants. Since the utilization of gray step-wedge is a special case of multi-colorant step-wedges, the equations follow a similar approach.

3 Results

Both the LS and TLS algorithms discussed above were tested on a Xerox color printer with rotated screen. A training chart with four sets of single-colorant step-wedges, a gray step-wedge, and multi-colorant step-wedges was printed. Each single-colorant step-wedge has 17 steps evenly distributed between 0 and 255 (0 and 255 included). The gray step-wedge (C=M=Y) also had 17 steps between 0 and 255. The multi-colorant step-wedges had one colorant varying from 0 to 255 at digital values identical to the single-colorant step-wedges, while the other 3 colorants are kept fixed at a constant level. Another test chart with $5 \times 5 \times 5 = 125$ samples was generated separately to test the various algorithms. Both the training and the test chart were measured using a Gretag spectrophotometer. The measurements from the training chart were then used to estimate the dot area functions for the C, M, Y, K colorants and the YN parameter n. CIELAB values [4] were computed from the spectral measurement of the test chart and ΔE differences between the measured $L^*a^*b^*$ values and the Neugebauer model predictions were obtained. CIE Illuminant D50 was used to calculate the $L^*a^*b^*$ values. Furthermore, ΔE^*_{CMC} and ΔE^*_{94} [10] were computed due to their added consistency over ΔE^*_{ab} with human visual sensitivity. Four techniques were tested: 1) LS estimation, employing single-colorant step-wedges; 2) TLS estimation, employing single-colorant step-wedges; 3) TLS estimation, employing single-colorant and gray step-wedges; 4) TLS estimation, employing single-colorant, gray and multi-colorant step-wedges. Correction of the primary reflectance curves (only for R_C, R_M, R_Y, R_K) was performed in Techniques 2, 3, and 4. An example of cyan primary (R_c) correction is given in Fig. 2. Fig. 3 shows the average ΔE^*_{ab}, ΔE^*_{CMC}, and ΔE^*_{94} errors between the $L^*a^*b^*$ measurements of the test chart and the corresponding model predictions computed by utilizing the LS and TLS techniques respectively. It can be clearly seen that the TLS based technique produces smaller ΔE errors in all three cases than the LS based technique with single-colorant step-wedges. By utilizing gray and multi-colorant step-wedges, the parameter estimation accuracy were further improved.

4 Conclusions

Parameter estimation accuracy for the Neugebauer model directly affects the model performance in color printer calibration. In this paper, one-dimensional or multi-dimensional total least squares techniques were applied to the parameter estimation problem, depending on the selection of color patches. The proposed TLS techniques were shown to produce better results than the single-colorant based LS technique on a Xerox color printer with rotated screen. Finally, by utilizing gray and multi-colorant patches, the estimation accuracy was further improved.

Acknowledgments

The authors wish to thank the Invention Opportunity Program Committee and the Production Controller Software Development Team of Xerox Corporation for sponsoring this research and providing the appropriate facilities.

References

[1] H. E. J. Neugebauer, "Die theoretischen Grundlagen des Mehrfarbenbuchdrucks," *Zeitschrift für wissenschaftliche Photographie Photophysik und Photochemie*, vol. 36, no. 4, pp. 73–89, Apr. 1937, reprinted in [11, pp. 194-202].

[2] J. A. S. Viggiano, "The comparison of radiance ratio spectra: Assessing a model's "goodness of fit"," in *Adv. Printing of Conf. Summaries: SPSE's 43rd Annual Conference, Rochester, NY*, May 1990, pp. 222–225.

[3] R. Balasubramanian, "Colorimetric modeling of binary color printers," in *Proc. IEEE Intl. Conference on Image Proc.*, Nov. 1995, vol. II, pp. 327–330.

[4] CIE, "Colorimetry," CIE Publication No. 15.2, Central Bureau of the CIE, Vienna, 1986.

[5] J. A. S. Viggiano, "Modeling the color of multi-colored halftones," *TAGA Proc.*, pp. 44–62, 1990.

[6] J. A. C. Yule and W. J. Nielsen, "The penetration of light into paper and its effect on halftone reproduction," in *TAGA Proc.*, May 1951, pp. 65–76.

[7] F. R. Clapper and J. A. C. Yule, "Reproduction of color with halftone images," in *Proc. Seventh Annual Tech. Mtg. TAGA*, May 1955, pp. 1–14.

[8] E. Demichel, in *Procédé*, 1924, vol. 26, pp. 17–21,26–27.

[9] S. Van Huffel and J. Vandewalle, *The total least squares problem : computational aspects and analysis*, Society for Industrial and Applied Mathematics, Philadelphia, 1991.

[10] CIE, "Industrial color difference evaluation," CIE Publication No. 116-1995, Central Bureau of the CIE, Vienna, 1995.

[11] K. Sayangi, Ed., *Proc. SPIE: Neugebauer Memorial Seminar on Color Reproduction*, vol. 1184, SPIE, Bellingham, WA, Dec. 1989.

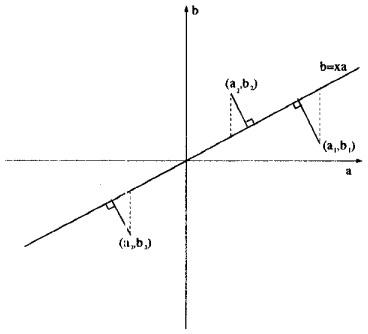

Figure 1: LS versus TLS. The dashed lines denote LS error and the solid lines perpendicular to line b=xa denote TLS error.

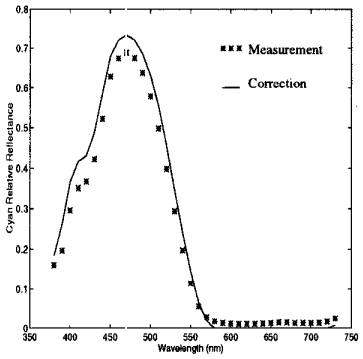

Figure 2: Primary cyan reflectance curve correction in TLS estimation (employing single-colorant step-wedges)

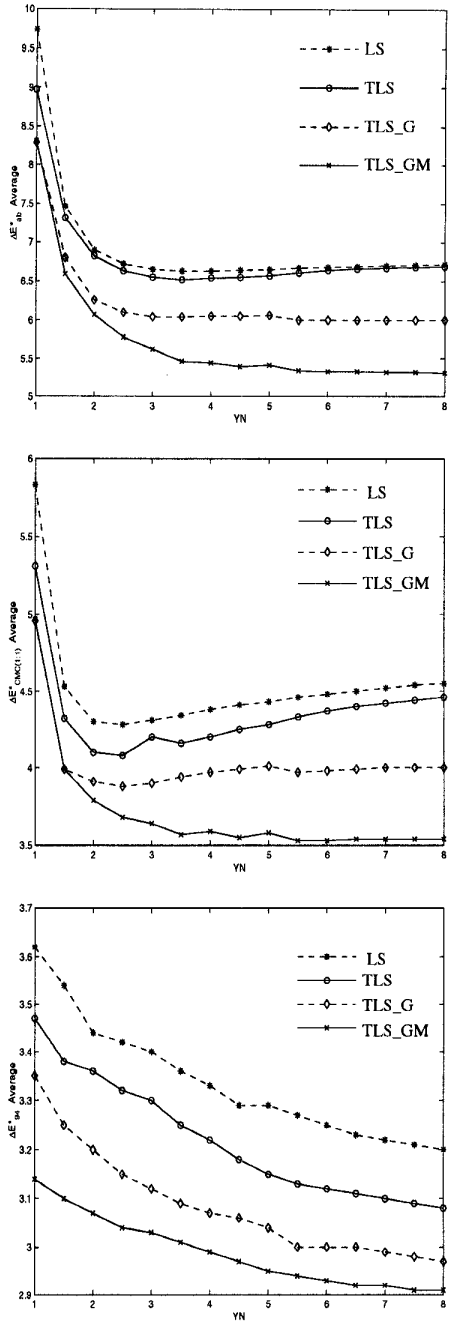

Figure 3: Average ΔE errors of the test chart for the random mixing model - (1) LS estimation, employing single-colorant step-wedges (LS); (2) TLS estimation, employing single-colorant step-wedges (TLS); (3) TLS estimation, employing single-colorant and gray step-wedges (TLS_G); (4) TLS estimation, employing single-colorant, gray and multi-colorant step-wedges (TLS_GM)