Total Least Squares Regression in Neugebauer Model Parameter Estimation for Dot-on-Dot Halftone Screens

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

[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

¿

Abstract

Parameter estimation is crucial to the accuracy of Neugebauer model in characterizing digital halftone printers. In this paper, a novel total least square (TLS) regression technique is proposed for the model parameter estimation problem for dot-on-dot screen printers. Compared to the traditional least square (LS) approach, the TLS method is physically more appropriate because it accounts for errors in both the measured reflectance of the selected primaries and the modeled reflectance. At the same time, TLS provides increased mathematical flexibility in fitting the Neugebauer model to the experiment data. Both the TLS and LS techniques are tested on a Xerox color printer with a dot-on-dot halftone screen. Compared to the LS based techniques, the TLS methods yield a more accurate model for the printer.

1. Neugebauer model and its variation for dot-on-dot halftone screens

The Neugebauer model and its variants [1, 2, 3] offer an attractive characterization method for color printer calibration, where the model parameters can often be determined from a small number of measurements. 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. As a result, a halftone print can be expressed as the weighted average of the tristimuli of these 16 overlapping combinations, referred to as Neugebauer primaries. The well known spectral Neugebauer model can be written as:

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

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* [4, 5], w_i denotes the fractional areas of the the i^{th} Neugebauer primary. The fractional areas w_i, a function of the dot areas of each individual colorant, where the functional form is determined by the halftone geometry. However, since the relationship between the actual dot areas c, m, y, k and the digital control values C, M, Y, K is usually nonlinear (which is often referred to as the dot area function), sophisticated procedure needs to be performed to estimate the function.

For dot-on-dot screens, because the dots of individual colorant are on top of each other, at most 5 of the 16 primaries are active [6]. Fig. 1 shows an example of the arrangement of dots for a dot-on-dot screen. As a result, if we let $p_1,\ p_2,\ p_3,\ p_4$ denote the printer colorants in increasing dot area coverage, and $a_1,\ a_2,\ a_3,\ a_4$ the corresponding dot areas, Eq. (1) can be rewritten as [6]:

$$r^{1/n}(\lambda) \quad = \quad \sum_{i=1}^{5} w_i {r_i}^{1/n}(\lambda) \qquad (2)$$

$r_i \in \{r_{p_1p_2p_3p_4}(\lambda), r_{p_2p_3p_4}(\lambda), r_{p_3p_4}(\lambda), r_{p_4}(\lambda), r_w(\lambda)\}$ denotes the 5 primaries, and $w_i \in \{a_1, a_2 - a_1, a_3 - a_2, a_4 - a_3, 1 - a_4\}$ are the fractional areas.

In practice, the printing process is subject to noise and mis-registration effects. Therefore, a combination of the dot-on-dot (2) and the general model (1) (where the fractional areas w_i are computed by Demichel Equations [7]) is introduced in [6] to improve the prediction accuracy. The combined model represents the predicted reflectance as:

$$r(\lambda) = (1-\alpha) r_d(\lambda) + \alpha r_r(\lambda), \quad (3)$$

where $r_d(\lambda)$ is the reflectance predicted by the dot-on-dot model (2), $r_r(\lambda)$ is the reflectance predicted by the general model (1), and α is a "noise factor" (within the range of $(0,1)$) which determines the relative contributions of the two models to the mixing process.

2. Neugebauer model parameters estimation by total least square regression

In the Neugebauer model, dot area function and YN correction factor n need to be estimated, so is the noise factor α in the combined model(3). The YN correction parameter n and the noise factor α can be estimated by iterating through a set of candidate values within empirically established boundaries. The dot area function can be estimated by either least square (LS) or total least square (TLS) approach. 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

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

where $e(\lambda)$ represents the measurement and model error (in the YN corrected spectral space). However, it should be noted that the measurement of the primary reflectance $r_i(\lambda)$ is also subject to error. Therefore, a more accurate model is the one that allows errors in all measured quantities, which is given by

$$r^{1/n}(\lambda; \mathbf{w}) + e(\lambda) = \sum_{i=1}^{P} w_i [r_i^{1/n}(\lambda) + e_i(\lambda)] \quad (5)$$

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.

The solution to the above equation must incorporate the unity sum constraint on the fractional areas of the primaries, e.g., $\sum_{i=1}^{P} w_i = 1$. If we assume 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. (5) can be rewritten as

$$r'(\lambda; \mathbf{w}') + e'(\lambda) = \sum_{i=2}^{P} w_i (r_i'(\lambda) + e_i(\lambda)) \quad (6)$$

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

The above equation at all wavelengths $[\lambda_1, \ldots \lambda_N]$ can be combined as

$$\mathbf{r}'(\mathbf{w}') + \mathbf{e}' = \sum_{i=2}^{P} w_i (\mathbf{r}_i' + \mathbf{e}_i) = (\mathbf{R}_p' + \mathbf{E})\mathbf{w}' \quad (7)$$

where each vector represents the corresponding term in Eq. 6 at all sampling wavelengths.

The above problem is equivalent to "solving" a set of linear equations $AX \approx B$. Considering the error associated with all the measurement terms, it can be solved as a TLS problem, which seeks to

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

where 'F' denotes the Frobenius norm [8]. 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. At the same time, it provides more flexibility in fitting the model to the experiment data. $\{w_i\}$ and the primary measurement error correction can be estimated at the same time. The correction to the primaries can then be utilized to update the primaries. Fig. 2 illustrates graphically the difference between the one dimensional LS and TLS case. 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 applied to single-colorant step-wedges

We first give an example of applying TLS to single-colorant prints. These 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. Therefore, for a K-step cyan wedge, with digital values $0 \le C_1 \le C_2 \le \ldots \le C_K \le 255$, the model of (7) 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 for 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 $\mathbf{e}_{pc}'$ denote 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 $\mathbf{E}_c' = [\mathbf{e}_{c_1}', \mathbf{e}_{c_2}', \ldots \mathbf{e}_{c_K}']$. By solving these equations through the TLS approach discussed in the previous section, the dot areas of cyan $\mathbf{c}$ and the correction $\mathbf{e}'_{pc}$ to the cyan primary reflectance can be simultaneously obtained.

2.2. TLS applied to multi-colorant step-wedges

Multi-colorant prints can be further employed to enhance the accuracy of the estimation. If we follow the same notation for the dot-on-dot mixing model in Section 2, then from (2) and (7), we get:

$$(\mathbf{r}'_{p_1p_2p_3p_4} - \mathbf{r}'_{p_2p_3p_4})a_1^j = \mathbf{r}' - a_2\mathbf{r}'_{p_2p_3p_4} - (a_3 - a_2)\mathbf{r}'_{p_3p_4} - (a_4 - a_3)\mathbf{r}'_{p_4} \quad (11)$$

After incorporating the error terms, we can rewrite the equation as

$$(\mathbf{r}^\star_{pa_1} + \mathbf{e}^\star_{pa_1})a_1^j = \mathbf{r}^\star + \mathbf{e}^\star \quad (12)$$

where $\mathbf{r}^\star_{pa_1} = \mathbf{r}'_{p_1p_2p_3p_4} - \mathbf{r}'_{p_2p_3p_4}$, $\mathbf{r}^\star = \mathbf{r}' - a_2\mathbf{r}'_{p_2p_3p_4} - (a_3 - a_2)\mathbf{r}'_{p_3p_4} - (a_4 - a_3)\mathbf{r}'_{p_4}$, and $\mathbf{e}^\star_{pa_1}$, $\mathbf{e}^\star$ represent the corresponding combined errors. This equation set can again be solved by TLS approach.

In our experiment, the combined model (3) is utilized to account for the noise in printing process. In order to estimate the parameters by a linear approach, the combined model is modified to

$$r^{1/n}(\lambda) = (1-\alpha)r_d^{1/n}(\lambda) + \alpha r_r^{1/n}(\lambda) \quad (13)$$

where the combination is done in the YN-corrected reflectance space. Details of TLS solution to this combined model can be readily developed using [8].

3. Results

We tested the LS and TLS techniques on a Xerox color printer with dot-on-dot halftone screen. A training chart was first printed and measured with a Gretag spectrophotometer. The chart has four single-colorant and one gray step-wedges [1], each with 17 steps evenly distributed between 0 and 255 (0 and 255 included). After the model parameters were estimated from the measurement, a test chart was generated and measured to test the various algorithms. The test chart contains $5 \times 5 \times 5 = 125$ samples evenly distributed throughout the color space. CIELAB values [9] were computed (under the CIE viewing Illuminant D50) from the spectral measurement of the test chart and ΔE differences between the measured $L^\star a^\star b^\star$ values and the Neugebauer model predictions were calculated. Color differences between the measurements and the predictions were computed using three common color difference metrics: $\Delta E^\star_{ab}$ [9], $\Delta E^\star_{CMC}$ [10], and $\Delta E^\star_{94}$ [11]. Three 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. Correction of the primary reflectance was performed in Techniques 2) and 3). Fig. 3 shows the average ΔE errors of the three color difference metrics. It can be observed that the TLS based technique produces smaller ΔE errors in all three cases than the LS based technique with single-colorant step-wedges. By employing gray step-wedges, the TLS estimation accuracy were further improved.

[1] The gray step-wedge which has no black colorant and equal digital values for the cyan, magenta and yellow colorants, and is therefore a special case of multi-colorant prints.

4. Conclusions

In this paper, a total least square regression technique was proposed to estimate the parameters of the Neugebauer model for halftone printers with dot-on-dot screen. The TLS approach is better than LS due to its physical appropriateness and greater mathematical flexibility. The test results on a Xerox printer have shown that the TLS method yield a more accurate model than its LS counterpart if single-colorant step-wedges were employed. The incorporation of gray step-wedges into TLS estimation further improved the results.

Acknowledgments

The authors wish to thank the Invention Opportunity Program Committee and the Production Controller Software Development Team of Xerox 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 [12, 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] 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.

[5] 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.

[6] R. Balasubramanian, "A printer model for dot-on-dot halftone screens," in *Proc. SPIE: Color hard copy and graphic arts IV*, J. Bares, Ed., 1995, vol. 2413, pp. 356–364.

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

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

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

[10] F. J. J. Clarke, R. McDonald, and B. Rigg, "Modification to the JPC79 colour-difference formula," *J. Soc. Dyers Colourists*, vol. 100, pp. 128–132, 1984.

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

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

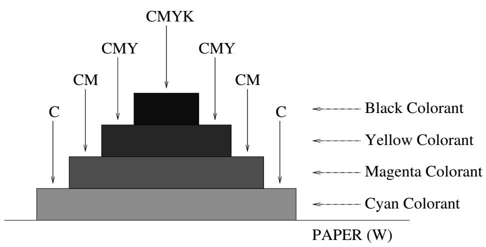

Figure 1: *Example of dot-on-dot screen configuration.*

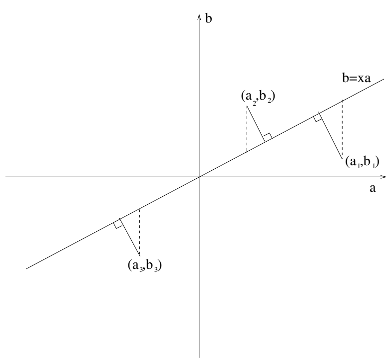

Figure 2: *LS versus TLS. The dashed lines denote LS error and the solid lines perpendicular to line b=xa denote TLS error.*

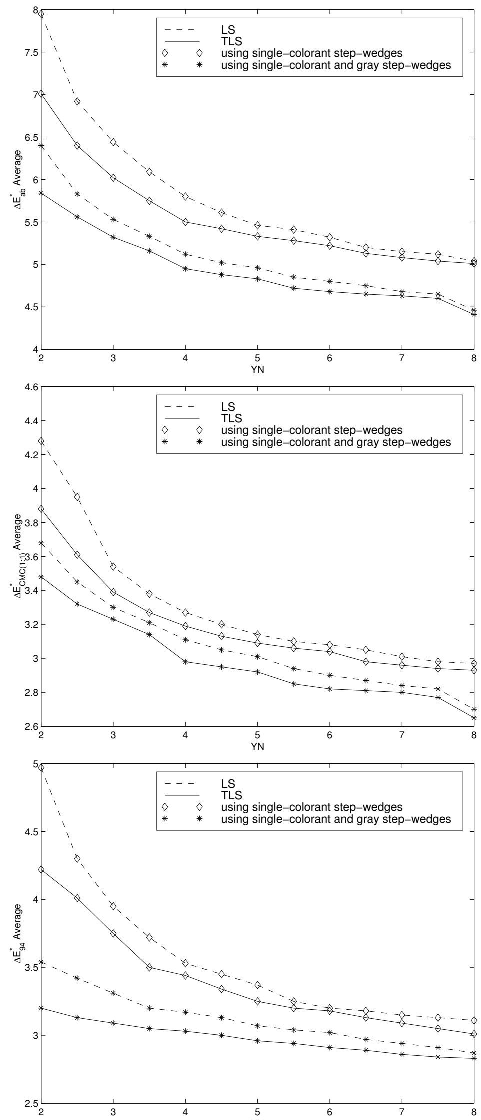

Figure 3: *Average* ΔE *errors for the test chart - (1) LS and TLS estimation, employing single-colorant step-wedges; (2) LS and TLS estimation, utilizing combined model (*$\alpha = 0.5$*), employing single-colorant and gray step-wedges.*