Novel spectral characterization method for color printer based on the cellular Neugebauer model

Jinyi Guo (郭晋一)[1,2], Haisong Xu (徐海松)[1*], and M. Ronnier Luo[2]

[1]*State Key Laboratory of Modern Optical Instrumentation, Zhejiang University, Hangzhou 310027, China*

[2]*Department of Colour Science, University of Leeds, Leeds LS2 9JT, UK*

[*]*E-mail: chsxu@zju.edu.cn*

Received March 29, 2010

The cellular Yule-Nielson spectral Neugebauer (CYNSN) model to characterize a typical CMYK (i.e., cyan, magenta, yellow, and black) 4-ink color printer is presented. By reconstructing spectral reflectance, models with high accuracy of spectral and colorimetric predictions are built. A novel cell-searching algorithm is proposed and used together with the iteration method to invert the cellular Neugebauer model efficiently. Large numbers of high quality hardcopy samples are produced to evaluate model performance and prove the feasibility of the algorithm. The spectral-based model performs better compared with the usual model based on CIE1931 *XYZ* tristimulus values.

OCIS codes: 330.1690, 330.1715, 330.1730, 300.6550.

doi: 10.3788/COL20100811.1106.

In the printing industry, color separation has always been an important issue in predicting ink combination for rendering a given color. Identifying printer characteristics is very important. Additionally, for cross-media color reproduction, image reproduction is required from computer displays to printers[1]. Reconstructing spectral reflectance of a color target for reproduction offers many advantages, such as higher accuracy and without any problems associated with metamerism[2].

In this letter, the cellular Yule-Nielson spectral Neugebauer (CYNSN) model was used to characterize a typical CMYK (i.e., cyan, magenta, yellow, and black) 4-ink color printer. A novel cell-searching algorithm was proposed together with the iteration method to invert the model efficiently. Large numbers of high quality hardcopy samples were produced. Corresponding spectral reflectances data were measured by a spectrophotometer in order to evaluate model performance.

For a CMYK 4-ink printer, the classical Yule-Nielson Neugebauer model[3,4] can be described as

$$R(\lambda) = \Big\{ \sum_{i=1}^{16} w_i [R_i(\lambda)]^{1/n} \Big\}^n, \tag{1}$$

where i denotes the ith of the 16 primary colors composed of all the possible combinations of the C, M, Y, K colors; $R(\lambda)$ and $R_i(\lambda)$ are the spectral reflectances of the desired sample and the ith primary color at wavelength λ, respectively; w_i is defined by the so-called Demichel equation[5], dependent on the normalized ink amount values of the color target; n is the coefficient of the power function, known as the Yule-Nielson n-factor. Yule-Nielson modification[4] is an efficient way to correct the problem caused by nonlinearity between the measured and predicted reflectances; it uses the power function. Another approach to improve model performance is to divide the classical Neugebauer model into small cells.

Equation (1) was applied on each cell. Prior to this, in each cell, the amount values of the colorant were renormalized to the range of 0–1. Generally, the more cellular levels used, the higher the accuracy obtained; however, in this present work, more samples need to be measured. Normally, two improved methods are used together as the CYNSN model.

In recent years, the inversion of the spectral Neugebauer model has been conducted through the linear regression iteration (LRI) method[6−8]. This method is an optimization process, minimizing the reflectance root mean square (RRMS) error to find the best ink amount values to match the spectral reflectances of the desired color. RRMS could show the matching grade between the predicted and the measured spectral reflectances, and could be calculated by

$$\mathrm{RRMS} = \sqrt{\frac{(\mathbf{R}_1 - \mathbf{R}_2) \times (\mathbf{R}_1 - \mathbf{R}_2)^{\mathrm{T}}}{N}}, \tag{2}$$

where $\mathbf{R}_1$ and $\mathbf{R}_2$ are the predicted and the measured spectral reflectances in a vector row form, respectively; T is the matrix transpose operation; N is the number of elements in a set of spectral reflectances.

Li *et al.* further improved the computation stage by introducing "QR" decomposition based on the matrix operation[9]. In addition, a different criterion for the optimization process was chosen to achieve spectral match, colorimetric match, or the combination of the two[5].

In terms of inverting the cellular Neugebauer model, the iteration process was performed inside each cell. This suggests that every cell could provide an optimized result by itself to best match the desired color. However, an optimal cell needs to be selected first. The straightforward approach is to obtain the prediction results from all cells by iteration, and then make a decision according to a certain criterion, such as the least RRMS difference or CIE color difference. However, this might result in too much computational cost, especially when using more cellular levels, as it greatly increases the number of times of iteration. Urban *et al.* proposed the cellular LRI (CLRI) method and expanded the iteration into multiple cells instead of from inside one cell alone[7]. They treated each colorant separately. The time required for

iteration from cell to cell (i.e., by changing only one colorant, such as cyan) in relation to all other colorants was fixed until the optimized value of the single colorant was obtained. The other colorants were then processed in series by using the same steps. This method could significantly reduce the number of times of iteration; thus, it could be implemented efficiently. However, when employed separately, the processing colorants might encounter some problems, especially in systems containing more than three colorants. Since it does not consider interactions between different colorants, the initial cell initializing the iteration, as well as the processing order of the colorants, could influence the final optimal results. This phenomenon is more apparent when the number of primary colorants is increased.

In this letter, a novel cell-searching algorithm is proposed. It can reduce cell-searching range through preselection while simultaneously considering changes in all colorants. It also requires fewer cells, and hence fast processing speed. The proposed inverse CYNSN model is described in detail by Fig. 1.

Step 1: Measure the reflectance spectra of all the cell corners and record them as training data.

Step 2: For each of the cells, calculate the reflectance spectrum $R_i(\lambda)$ of an imaginary cell center to represent that cell simply by averaging the spectral reflectances of all the corners of that cell at each wavelength.

Step 3: For the desired spectral reflectance R, calculate the differences ΔD_i between each R_i and R, according to the preselection criterion.

In this letter, the RRMS difference was used for the spectral match. We sorted the cells in ΔD_i ascending order. The rearranged indices of the cells were denoted as k_1, k_2,$\cdots$, k_t. The first t cells with the least ΔD_i were chosen as a subset in order to proceed to the next step. Steps 1–3 can be considered as the preselection process. By referring to the choice of the number t, which represents the number of cells chosen as a subset after the preselection, 2^m+1 was used for the m-colorant system. In an m-dimensional space, each cell has 2^m neighboring cells; thus, it is reasonable to assume that the optimal cell exists in the first 2^m+1 cells with the smallest differences from the target color. The supposition was proven feasible by experimental results in the following content. In this regard, the required searching cell number does not increase with the cellular level, which is superior to the CLRI method mentioned earlier.

Step 4: Predict the target reflectance spectrum in all of the t chosen cells by the iteration method.

In so doing, t sets of colorant values were obtained.

Step 5: Process all the t sets of colorant values by the forward CYNSN model to obtain t sets of predicted spectral reflectances.

This approach helps us decide which of the cells should be used. The chosen cell must have corresponding spectral reflectances manifesting the smallest difference with the desired color according to the decision criterion. RRMS was used so that the colorant values predicted in this cell could be treated as the final result of the inverse CYNSN model.

In the experiment stage, a typical CMYK 4-ink printer, the HP DesignJet 1055 CM, was employed and characterized. GretagMacbeth Spectrolino spectrophotometer was used to measure the spectral reflectance of the samples. The training samples contained 10000 grid color patches. Their normalized ink amount values were the full 10×10×10×10 combinations of the 10 steps for each of the C, M, Y, K color coordinates. The testing samples involved 256 other regular grid color patches, independent of the training samples, with four steps of 0, 0.3333, 0.6667, and 1 in each of the C, M, Y, K color coordinates. The amount of the ink was directly controlled by the "Adobe PostScript" page description language, which was quite helpful in obtaining accurate normalized ink amount values for the model.

Large numbers of training data were obtained for building the cellular Neugebauer model, but not all of them were necessary. To find out how many cellular levels could achieve the best model performance with the fewest measured samples, some mathematical testing was conducted based on the CYNSN model with different cellular levels, ranging from 2 to 10. The 256 testing samples

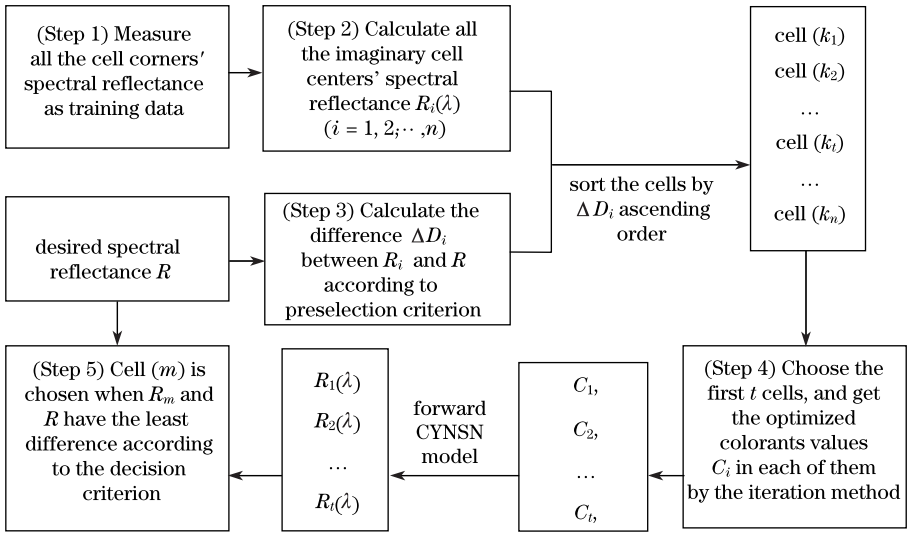

Fig. 1. Flow chart of the cell-searching algorithm for the inverse CYNSN model.

Table 1. Average Prediction Color Differences ΔE^*_{ab} ($D50/2^\circ$) of *XYZ*-Based Method and Spectral-Based Method

Cellular Level	2	3	4	5	6	8	10
XYZ-Based Method	4.98	2.33	1.33	0.85	0.70	0.69	0.68
Spectral-Based Method	3.93	1.61	1.08	0.75	0.69	0.68	0.68

were predicted by both the *XYZ*-based and spectral-based forward CYNSN models with different cellular levels. The *XYZ*-based models were obtained by replacing reflectance "*R*" in Eq. (2) with *XYZ* tristimulus values. Next, the average CIELAB color differences for all the predicted and measured spectral reflectances were calculated. The results are summarized in Table 1.

The more cellular levels used, the better the model performance became. Additionally, the fewer the cellular levels, the significantly better the accuracy of the spectral-based method was compared with the *XYZ*-based method. The spectral-based method (especially in the model using three cellular levels) could satisfy the image color difference threshold (i.e., 2.2 CIELAB units)[10], a possible threshold for human vision to judge image color difference between two images; the *XYZ*-based method could not perform as such. As a result, the spectral-based method was very effective in achieving higher model accuracy with fewer measured samples. When using more cellular levels (especially more than six), both the two methods obtained similar prediction accuracy (i.e., about a mean of 0.68 CIELAB units). Hence, model performance was mainly determined by cellular levels instead of the applied method. According to the above discussion, only three cellular levels were sufficiently accurate for the spectral model; six cellular levels achieved the best performance.

The forward model could accomplish very high model prediction accuracy. However, in practice, the inverse model provides more value. To investigate the choice of t number mentioned earlier in the cell-searching algorithm, for each testing sample, the iteration process was employed for all the 625 cells in a 6-level CYNSN model. The optimal cell was obtained by identifying the smallest RRMS difference from the target. The rearranged cell index of the optimal cell for each sample in the preselection was recorded. This index was listed as the sequence number of the optimal cell in the preselected cell array. As presented in this letter, 192 testing samples were used for testing. Their corresponding results are shown in Fig. 2. In our method, the recommended value for t is 17 for the CMYK 4-ink printer. In Fig. 2, for the 180 of 192 samples, the rearranged cell indices were under 17, which indicate that the optimal cell generally existed in the first 17 cells after preselection. For the other 12 samples, the results were different in terms of the iteration in all the 625 cells and in the first 17 cells. By observing further the other 12 samples, 10 of them were of nearly full black color because the cell centers were becoming denser with the addition of more black ink. The optimal result in the first 17 cells was highly accurate. In other words, the same color could be obtained in different mixes for the samples despite the high amount of black ink. In reality, however, increasing the amount of black ink seldom employed. The peak errors occurred in the low-black areas where the cell centers were dispersed. A wrong cell search would cause a huge color difference in those areas, similar to the remaining 2 of the 12 samples. Some restrained criterion was therefore added in the algorithm to solve the problem. When the optimal spectral reflectance in the 17-cell subset exhibited large differences in relation to the target, which could be justified by a predefined threshold (e.g., a RRMS of over 0.15 or ΔE^*_{ab} ($D50/2^\circ$) over 10 in this study), the searching range was enlarged by adding the next 16 cells onto the subset sporadically, but in accordance to the cell order in the preselection array. The optimal result in the new subset could then be obtained. The difference from the recalculated target and from the new operation manifested as a loop; consequently, the difference was below the threshold. The approach could effectively correct any accidental wrong cell search, and normally, it only required 1 or 2 times of enlargement without inducing much computational cost. This made the algorithm more robust and flexible.

Based on the iteration method and the proposed cell-searching algorithm, the measured spectral reflectance data of the 256 testing samples were used by the inverse 6-level CYNSN model to predict their CMYK values. The samples were further processed by the forward CYNSN model to predict corresponding spectral reflectance data, in which the error of the forward model was employed. The RRMS and ΔE^*_{ab} ($D50/2^\circ$) of the two sets of spectral reflectance data (Fig. 3) were

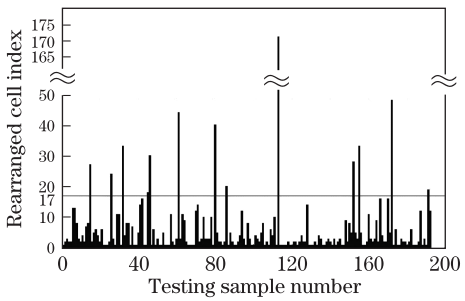

Fig. 2. Rearranged optimal cell indices of the 192 testing samples.

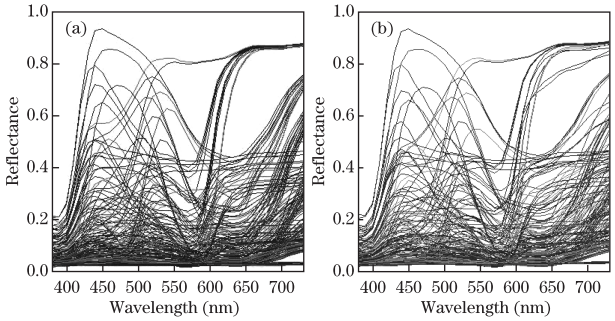

Fig. 3. (a) Measured and (b) predicted spectral reflectances of all the 256 testing samples.

Table 2. Inverse CYNSN Model Performance in Terms of RRMS Error and CIELAB Color Difference ΔE^*_{ab} ($D50/2^\circ$)

	Max	Mean	Median
ΔE^*_{ab}	13.6368	1.5929	0.8355
RRMS	0.1337	0.0104	0.0025

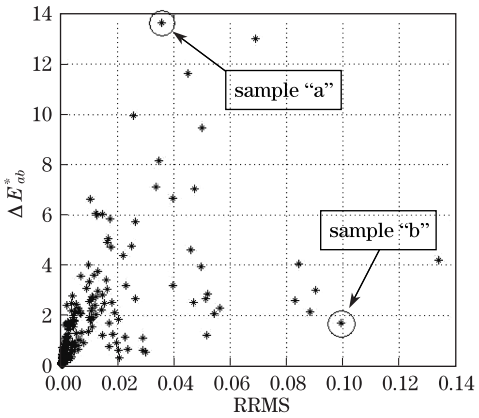

Fig. 4. Inverse CYNSN model prediction results of all the 256 testing samples in terms of RRMS error and CIELAB color difference ΔE^*_{ab} ($D50/2^\circ$).

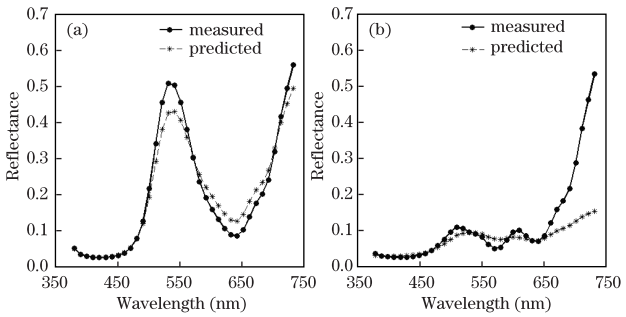

Fig. 5. Measured and predicted spectral reflectance diagrams of samples (a) "a" and (b) "b" .

calculated to estimate the model performance. The results are shown in Table 2 and Fig. 4.

The average prediction accuracy of the inverse method, as shown in Table 2, could achieve the color difference of 1.59 CIELAB units. This is rather outstanding compared with those presented by reports using other similar experiments[11]. In Fig. 4, most of the points are distributed near the point of origin; they have low RRMS and ΔE^*_{ab} values. In this area, the RRMS and ΔE^*_{ab} values showed consistent trends, indicating that excellent spectral and colorimetric match was accomplished between the original color and the reproduction. As for the samples being off the point of origin, they had poor prediction results and manifested poor consistency.

Two samples, denoted as sample "a" and sample "b", showing an extreme situation in Fig. 4, were explored. The ΔE^*_{ab} of sample "a" was 13.63, the highest value among all the 256 testing samples, but its RRMS value (0.035) was not correspondingly significant. Meanwhile, sample "b" was the complete opposite; it has a high RRMS value (0.099) but a low ΔE^*_{ab} (1.70). To investigate this phenomenon, the measured and predicted spectral reflectance diagrams of the two samples were illustrated (Fig. 5). The high ΔE^*_{ab} of sample "a" was due to its spectral differences being mainly distributed in the wavelength range of 480–680 nm, the light component of which greatly influenced the XYZ tristimulus values. On the other hand, sample "b" had significant spectral differences in the longer wavelength range (i.e., over 650 nm) at which the human eyes are not so sensitive to. A better match in the middle wavelengths was observed, which resulted in a relatively small colorimetric difference.

In conclusion, the CYNSN models with high accuracy is built with optimized parameters and confirmed by experiments. The methodology involved in inverting the CYNSN model is discussed. A novel cell-searching method with supplementary correcting method is proposed, and this presents advantages, especially for models with more cellular levels. The spectral-based method is compared with the XYZ-based method in terms of prediction accuracy. Experimental results indicate that with fewer cellular levels, the accuracy of the spectral method is significantly better compared with the XYZ method. As to the problem of the spectral model, whereby using reflectance data requires high computational cost, further studies should be conducted to improve the calculation algorithm and the processing speed.

This work was sponsored by Clariant Co., Ltd.

References

1. X. Zhang and H. Xu, Chin. Opt. Lett. **7,** 873 (2009).
2. Y. Wang and H. Xu, Chin. Opt. Lett. **3,** 725 (2005).
3. H. E. J. Neugebauer, D. Wyble, and A. Kraushaar, Col. Res. Appl. **30,** 322 (2005).
4. D. R. Wyble and R. S. Berns, Col. Res. Appl. **25,** 4 (2000).
5. J. Gerhardt and J. Y. Hardeberg, Col. Res. Appl. **33,** 494 (2008).
6. P. Urban and R.-R. Grigat, Col. Res. Appl. **31,** 229 (2006).
7. P. Urban, M. R. Rosen, and R. S. Berns, in *Proceedings of IS and T/SID Fifteenth Color Imaging* 178 (2007).
8. M. Mahy and P. Delabastita, Col. Res. Appl. **21,** 404 (1996).
9. C. Li and M. R. Luo, in *Proceedings of IS and T/SID Sixteenth Color Imaging* 84 (2008).
10. C. Sano, T. Song, and M. R. Luo, in *Proceedings of IS and T/SID Eleventh Color Imaging* 121 (2003).
11. R. Balasubramanian, J. Electron. Imaging **8,** 156 (1999).