A Simple Color Prediction Model based on Multiple Dot Gain Curves

Yuan Yuan Qu[*a] and Sasan Gooran[a]

[a]Department of Science and Technology, Linköping University, Norrköping Campus, Sweden

ABSTRACT

Most of the color prediction models use single dot gain curve, few assume that dot gain changes when ink superposition happens, but still, use single dot gain curve for each ink to compensate the effective ink coverage. Considering the fact that optical dot gain is the effect of light scattering in paper, it is reasonable that light with different wavelength might produce different optical dot gain for each ink. In this study, for each primary ink we utilized three different curves obtained by CIEX, Y and Z, which approximately stand for three special wavelength bands, to calculate color coordinates. In addition, we noticed that dot gain curves obtained from the print samples with single ink printed on paper do not work well for the prints where ink is printed on another, or others. Therefore, dot gain curves for different ink superposition situations are optimized by matching the calculated tri-stimulus values of training patches to their measurement counterpoints. For each ink, dividing the dot gain into several dot gain actions responding to different ink superposition situations, we got the final dot gain as a group of multiple curves that takes into account all possible 'dot gain actions' with certain probability coefficients.

Keywords: Color Prediction, Dot Gain, Color Difference

1. INTRODUCTION

It is common sense that dot gain composes of physical dot gain which refers to ink spreading and optical dot gain or Yule-Nielsen effect which is due to light scattering[1]. Murray-Davies equation or Neugebauer's equation were mostly used to compensate physical dot gain. Alternative approaches are also available, like modeling the physical dot gain with a transmission function defined on a blurred version of the halftone image[2].

Yule-Nielsen model with an n-factor was mostly used to numerically deduce the optical dot gain. Since the spectral reflectance measured by physical instruments technically includes the effects of both optical and physical dot gain[3], the n-factor practically is just a fitting factor and sometimes it is impossible to reach a 'best' one that approximately matches the measurement data. Of course, many complex methods were studied aiming to model the optical dot gain, amongst which the popular ones are point spread function (PSF) or probability functions fixed according to experiments[4,5], as well as Clapper-Yule which takes into account both halftone patterns and internal reflections occurring at the interface between the paper and air[6]. Last but not least, the Kubelka-Munk model, which counts the light reflection by two light fluxes using light absorption and scattering coefficients[7].

Based on those models, in addition with different assumptions and novel algorithms, there came out many research works on compensating dot gain for inks when certain halftone pattern and printing devices are used. Most of those works use single dot gain curve for each ink, few papers pointed out that dot gain might change in different ink superposition situations, but never synchronously consider that dot gain curves might also be wavelength dependent. Think about the fact that optical dot gain is caused by light scattering in paper, therefore it is reasonable that light with different wavelength produce different optical dot gain. A number of measurements and researches on dot gain also show that dot gain actually varies when the illumination wavelength is changed.

In the present paper, we utilized three different curves for each ink using CIEX, Y and Z, which approximately stand for three special wavelength bands, to calculate color coordinates. In addition, we noticed that dot gain curves obtained from the samples with single ink printed on paper don't work well to predict the color of the random prints with ink printed on another or others. We optimized dot gain curves of certain ink under different ink superposition situations by matching calculation of training patches to their measurement tri-stimulus values, and combined them by certain probabilities. The following sections are arranged to illustrate our approach in detail and present the results of our color prediction method using different print devices and papers.

*E-mail: qu.yuanyuan@liu.se;   phone (+46)11 363258;   fax   (+46)11 363270

2. THREE DOT GAIN CURVES FOR EACH INK ACCORDING TO CIEXYZ

The previous work of the present paper was introduced in reference 3, in which the tri-stimulus values of the print color were calculated simply by using three dot gain curves for each ink. Those curves are obtained by CIEX, CIEY and CIEZ according to Neugebauer's equation, as $dgc\_X(c_{ref})$, $dgc\_Y(c_{ref})$ and $dgc\_Z(c_{ref})$ denoted in Equation 1, in which cyan was taken as an example. Dot gain curves for magenta and yellow are obtained similarly.

$c_{eff}\_X(c_{ref})$, $c_{eff}\_Y(c_{ref})$ and $c_{eff}\_Z(c_{ref})$ are the effective dot coverage of cyan corresponding to the reference coverage value c_{ref}. Xp, Yp and Zp are CIEXYZ tri-stimulus value of paper while Xc, Yc and Zc are CIEXYZ tri-stimulus value of full tone cyan. The CIEXYZ measurements $Xc_{mea}(c_{ref})$, $Yc_{mea}(c_{ref})$ and $Zc_{mea}(c_{ref})$ of a group of single cyan print samples together with their c_{ref} range from 0 to 100% enable us to get three dot gain curves for cyan, as well as effective dot coverage curves. Then, the effective dot coverage for random reference coverage could be calculated by interpolation along those curves.

$$c_{eff}\_X(c_{ref}) = \frac{Xc_{mea}(c_{ref}) - Xp}{Xc - Xp} = dgc\_X(c_{ref}) + c_{ref}$$

$$c_{eff}\_Y(c_{ref}) = \frac{Yc_{mea}(c_{ref}) - Yp}{Yc - Yp} = dgc\_Y(c_{ref}) + c_{ref} \qquad \text{Equation 1}$$

$$c_{eff}\_Z(c_{ref}) = \frac{Zc_{mea}(c_{ref}) - Zp}{Zc - Zp} = dgc\_Z(c_{ref}) + c_{ref}$$

Figure 1 shows the calculated dot gain curves for cyan printed at 300 dpi by a laser printer (Xerox, Phaser 6180), using CIEX, Y and Z, respectively. The 'best dot gain' is an optimized single dot gain curve which is obtained by setting $dgc\_X(c_{ref}) = dgc\_Y(c_{ref}) = dgc\_Z(c_{ref}) = dgc\_best(c_{ref})$ in Equation 1, then apply the least square method to fix $dgc\_best(c_{ref})$, which is the one that gives the minimum color difference from the measurement data of the samples.
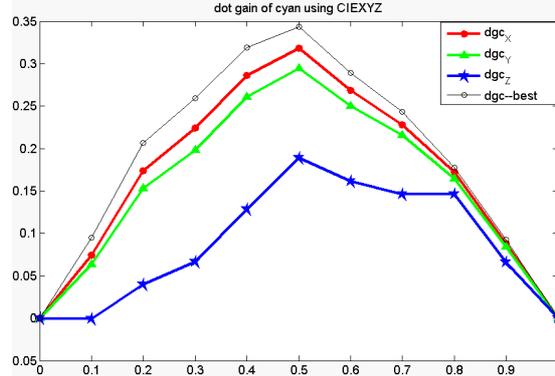

Figure 1. Dot gain curves using CIEX, CIEY and CIEZ for cyan printed by a laser printer at 300 dpi.

Since CIEXYZ tri-stimulus values have a linear relationship with the spectral reflectance, if the cyan ink and the paper were both ideal then it would only be possible to find the effective dot coverage for cyan by using the reflectance spectra in longer wavelength, which corresponds to CIEX values. Similarly, it would only be possible to find the effective dot coverage for magenta by using CIEY values and for yellow by using CIEZ values[3]. Figure 1 clearly shows that the effective cyan dot coverage using measured CIEX, CIEY, and CIEZ respectively differs from each other. The similar situation goes for magenta and yellow. Therefore, it is not completely correct to define only one dot gain curve for each ink.

On the other hand, there is actually optical dot gain included in the measurement data. An advantage proposal for color prediction is to use all the three dot gain curves based on CIEXYZ instead of plotting single curve for each ink just as most other prediction models do.

If the prints are single ink involved, those three dot gain curves indexed by *X*, *Y* and *Z* are used to calculate the corresponding three groups of effective coverage of the present ink; finally we could get the exact tri-stimulus values of the prints with very small color difference whatever halftone pattern is used.

When more than one primary ink are involved, Demichel's equation are used to get the effective coverage with CIEX, CIEY and CIEZ as index for each primary and sub-primary colors. Equation 2 shows the effective coverage c_X, c_Y and c_Z for cyan when only cyan and magenta are involved. For other primary and sub-primary colors, the calculation is similar with that shown in Equation 2 for cyan.

$$c\_X = c_{eff}\_X \cdot (1 - m_{eff}\_X)$$
$$c\_Y = c_{eff}\_Y \cdot (1 - m_{eff}\_Y) \hspace{4cm} \text{Equation 2}$$
$$c\_Z = c_{eff}\_Z \cdot (1 - m_{eff}\_Z)$$

Finally, we calculate the tri-stimulus values (X_{cal}, Y_{cal}, Z_{cal}) of the prints by Equation 3 (take cyan companied with magenta as example). *Xc*, *Yc* and *Zc* are the CIEXYZ values of full tone cyan, the indices *c*, *m*, *b*, *p* refer to cyan, magenta, blue and paper respectively.

$$X_{cal} = c\_X \cdot Xc + m\_X \cdot Xm + b\_X \cdot Xb + p\_X \cdot Xp$$
$$Y_{cal} = c\_Y \cdot Yc + m\_Y \cdot Ym + b\_Y \cdot Yb + p\_Y \cdot Yp \hspace{2cm} \text{Equation 3}$$
$$Z_{cal} = c\_Z \cdot Zc + m\_Z \cdot Zm + b\_Z \cdot Zb + p\_Z \cdot Zp$$

As a comparison, we used the Yule-Nielsen model to predict the color. As for the dot gain curve of each ink in Yule-Nielsen model here, we tried to find the 'best dot gain curve' (introduced above) for each possible n-factor. It was carried out in the following steps:

1) Begin with n-factor equal to 1 and find the 'best dot gain curve';

2) Use this curve to estimate the effective dot coverage of each ink according to their reference coverage, and then the effective coverage of each primary and sub-primary color; do the prediction and record the ΔE_{94} value;

3) Increase n by 0.1 and iterate the above steps.

The best n-factor and dot gain curve are the ones that give the minimum average color difference ΔE_{94} between prediction and the measurement.

In the past work, 180 different color patches using only cyan and magenta, 120 different color patches using cyan, magenta and yellow were printed at two different resolutions, 150 dpi and 300 dpi, using a laser printer (Xerox, Phaser 6180) with uncoated office paper. In the case of two colors, our approach gave an average ΔE_{94} equals to 1.83 while an average ΔE_{94} equals to 2.96 was carried out based on Yule-Nielsen model introduced in the above paragraphs. For three colors, our approach gave an average ΔE_{94} of 4.96, not good enough but still better than the value 5.21, which was the result of Yule-Nielsen model.

To continue our work, we tried to figure out the behavior of our approach for different color prints. We applied Amplitude Modulated (AM) halftone prints at 150 lpi, first generation Frequency Modulated (FM1st), second generation Frequency Modulated with big dots, medium dots and small dots (FM2-b, FM2-m, FM2-s) prints at 1200 dpi by an offset press on coated paper. Those prints were measured by a Spectrophotometer (BARBIERI electronic Spectro LFP RT) using d50 light source for a 2° observer. The color differences between the predicted and measured values were calculated using color difference ΔE_{94} and are shown in table 1. As usual, we compared the results with that from Yule-Nielsen model.

Comparing the results that we got in previous experiments dealing with prints on uncoated paper using laser printers, our simple color prediction model works much better for prints on coated paper using offset press, for almost all halftone patterns we used, especially AM and Second generation FM with small dots (FM2-s).

Table 1. The Color difference ΔE94 between measurement and calculated values (offset press, coated paper).

Only Cyan& Magenta	Yule-Nielsen model				our approach	
	ΔE_{94}		n factor		ΔE_{94}	
	Max	Mean	n_c	n_m	Max	Mean
AM-150 lpi	3.06	1.66	1.9	1.2	3.19	1.72
FM1st-1200 dpi	9.16	4.67	1.5	1.8	3.88	1.93
FM2-B-1200 dpi	3.93	2.14	1.9	1.2	4.27	2.45
FM2-M-1200 dpi	4.28	2.21	2.3	1.3	4.33	2.35
FM2-S-1200 dpi	5.53	2.81	1.8	1.3	3.87	1.72

three colors	Yule-Nielsen model					our approach	
	ΔE_{94}		n factor			ΔE_{94}	
	Max	Mean	n_c	n_m	n_y	Max	Mean
AM-150 lpi	4.47	2.00	1.9	1.2	1	3.96	1.64
FM1st-1200 dpi	10.21	5.34	1.5	1.8	1	7.48	3.07
FM2-B-1200 dpi	7.83	2.94	1.9	1.2	1.1	5.99	2.62
FM2-M-1200 dpi	5.24	2.94	2.3	1.3	1	4.67	2.49
FM2-S-1200 dpi	5.22	3.10	1.8	1.3	1	4.52	1.94
n_c is the n-factor value for cyan, n_m for magenta and n_y for yellow.							

We can also see that when three colors are involved, ΔE_{94} of AM prints is always small while the color difference of first generation FM (FM1st) is always worse than that of other halftone patterns. We plotted the 'best dot gain' curves of cyan under different halftone patterns and found out that the dot gain values are ordered as FM1st > FM2-S > FM2-M > FM2-B ≈ AM. It seems that color prediction models based on Neugebauer's equation might work better when halftone pattern with bigger dots was used.

3. MULTIPLE DOT GAIN CURVES

Since for laser printers with normal uncoated paper, our approach failed to give acceptable color difference, it is necessary to improve it although it gave less color difference than Yule-Nielsen model. According to experiences and references, dot gain is a complicated phenomenon that might be governed by ink dots superposition besides ink spreading and light wavelength. Consequently we tried to modify our simple approach in the following two ways.

Firstly, we divided the visible wavelength interval into more than three subintervals. In each subinterval, measured spectral data are used to plot one sub-dot gain curve for each ink; the predicted interval spectral reflectance was calculated by Neugebauer's equation using the effective ink coverage; then the full spectral reflectance is the sum of all the interval spectral reflectance. The color difference was counted by ΔE_{Lab}. It was unfortunate that this method did not work as well as we expected. The explanation is that one could predict the color spectrum better in this way, but it doesn't guarantee that the color difference ΔE_{Lab} will be less.

The second attempt was to figure out the complicated dot gain when ink superposition happens. We defined the dot gain under different ink superposition situations as different 'dot gain action'. Then there are several different dot gain actions act at the same time when ink drops are placed on the paper. All of those dot gain actions contribute to the total dot gain. Take cyan as an example, no matter which kind and how much there are the ink dots on the paper before cyan drops lay down, the possible dot gain actions are: dot gains along the edge between cyan and paper only; dot gains when cyan is mixed with magenta only; cyan is mixed with yellow only and cyan is mixed with both magenta and yellow.

To study dot gain when different ink combinations are present, we printed hundreds of training patches. Again, assume that only cyan and magenta are involved, the training patches include 4 bunches: one bunch is composed by 11 patches with increasing cyan coverage but constant magenta reference coverage at 25%. The other three bunches are similar with the first one, only difference between them is that the constant magenta reference coverage change to 50%; 75% and 100% respectively.

For other possible combinations, the training patches are created in the same way. Define *dgionj* as dot gain for ink *i* when *i* is mixed with *j*. Figure 2 shows the calculated dot gain for cyan when combined with magenta at different reference coverage, defined as *dgcon25m*, *dgcon50m*, *dgcon75m*, *dgcon100m*, and *dgconp* refers to the dot gain for cyan when neither magenta nor yellow on the paper, *dgconm* is the optimized dot gain according to *dgcon25m*, *dgcon50m*, *dgcon75m*, *dgcon100m*.
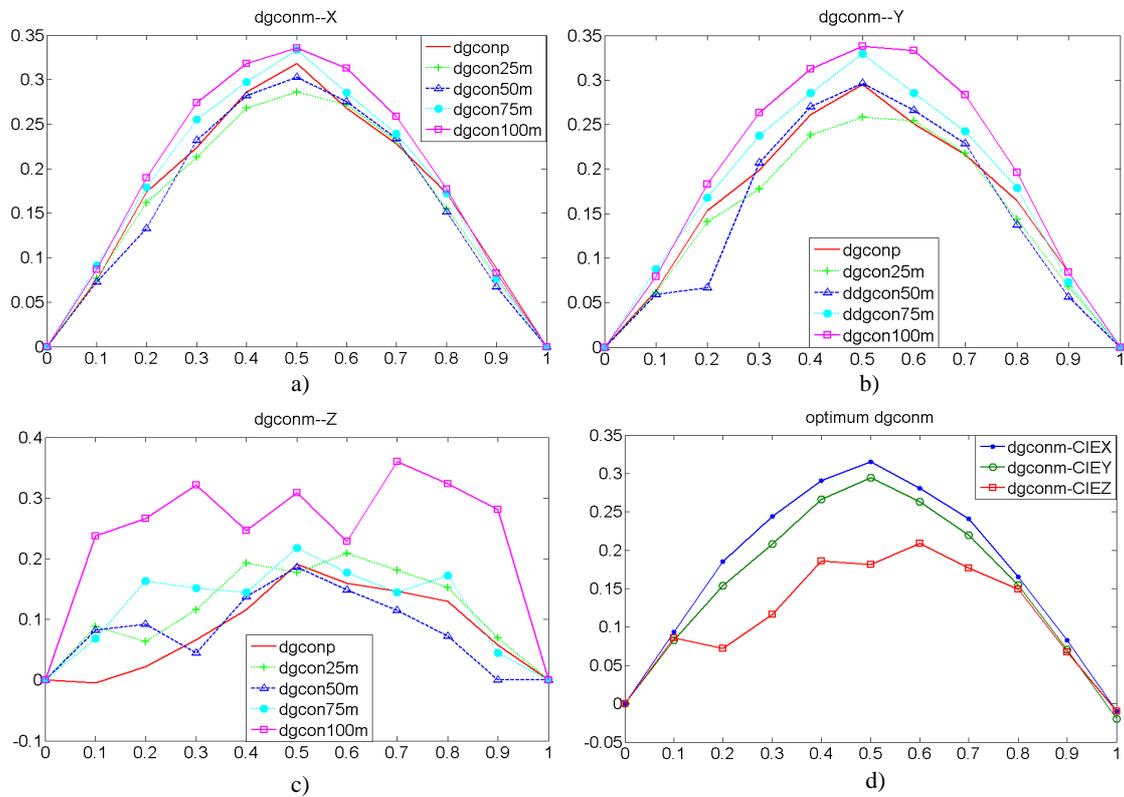

Figure 2. Dot gain curves for cyan when cyan dots mix with different amount of magenta dots, for a laser printer using FM at 300 dpi.

a). using CIEX   b). using CIEY   c). using CIEZ   d). optimum dot gain curves using CIEXYZ.

In this case, *dgcon25m* is calculated by Equation 4.

$$dgcon25m\_X = c_{eff}\_X - c_{ref}$$
$$dgcon25m\_Y = c_{eff}\_Y - c_{ref}$$
$$dgcon25m\_Z = c_{eff}\_Z - c_{ref}$$

Equation 4

Effective dot coverage for cyan c_{eff_X} (c_{eff_Y}, c_{eff_Z}) are calculated according to inverse forms of Equation 2 and Equation 3. The needed effective dot coverage for magenta m_{eff_X} (m_{eff_Y}, m_{eff_Z}) in Equation 2 are constant values here that could be calculated previously according to the measurement of the first patch with single reference coverage equals to 25% for magenta but zero for cyan. The needed X_{cal} (Y_{cal}, Z_{cal}) in the left side of Equation 3 are given by X_{mea} (Y_{mea}, Z_{mea}), which are measured CIEXYZ values of 11 patches with increasing cyan coverage from 0 to 100% but constant magenta reference coverage at 25%.

We can see from figure 2-a) that the dot gain for cyan differs when cyan was in different mixture with magenta. *To* split the difference and get the *dgconm_X*, optimization was done by finding the value amongst *dgcon25m_X, dgcon50m_X, dgcon75m_X* and *dgcon100m_X* to give the least color difference for all the print samples involved in this case. The optimized *dgconm_X (Y, Z)* were shown in figure 2-d).

Finally, we format the multiple dot gain curves using possibility coefficients. Imagine that when a cyan dot was impacted onto the paper, w1 part of it would be without any superposition or there is a possibility equals to w1 that dot gain acts as it does when only cyan dots on paper. With a similar reasoning, w2 part of the dot would be overlapped with magenta dots; w3 part of it might be superposition with yellow dots and w4 part of it might overlap with both magenta and yellow dots. We consider the coverage percentages *w1* to *w4* to be the possibilities of each 'dot gain action', and they are calculated according to the reference ink coverage based on Demichels' equation:

$$w1 = \frac{c_{ref} \cdot (1 - m_{ref}) \cdot (1 - y_{ref})}{c_{ref}} = (1 - m_{ref}) \cdot (1 - y_{ref})$$

$$w2 = \frac{c_{ref} \cdot m_{ref} \cdot (1 - y_{ref})}{c_{ref}} = m_{ref} \cdot (1 - y_{ref})$$

$$w3 = \frac{c_{ref} \cdot (1 - m_{ref}) \cdot y_{ref}}{c_{ref}} = (1 - m_{ref}) \cdot y_{ref}$$ Equation 5

$$w4 = \frac{c_{ref} \cdot m_{ref} \cdot y_{ref}}{c_{ref}} = m_{ref} \cdot y_{ref}$$

The final multiple dot gain curves of cyan with certain CMY reference coverage would be expressed by Equation 6:

$$dgc\_X = w1 \cdot dgconp\_X + w2 \cdot dgconm\_X + w3 \cdot dgcony\_X + w4 \cdot dgconmy\_X$$

$$dgc\_Y = w1 \cdot dgconp\_Y + w2 \cdot dgconm\_Y + w3 \cdot dgcony\_Y + w4 \cdot dgconmy\_Y$$ Equation 6

$$dgc\_Z = w1 \cdot dgconp\_Z + w2 \cdot dgconm\_Z + w3 \cdot dgcony\_Z + w4 \cdot dgconmy\_Z$$

As *dgconmy_X(Y, Z)* jumps irregularly, we consider the dot gain effect of cyan mixed with both magenta and yellow is the co-effect of *dgconm* and *dgcony* with coefficient α:

$$dgconmy = \alpha \cdot dgconm + (1 - \alpha) \cdot dgcony$$ Equation 7

For general calculation, we set α=0.5, therefore Equation 5 and Equation 6 become:

$$w1 = (1 - m_{ref}) \cdot (1 - y_{ref})$$

$$w2 = m_{ref} \cdot (1 - y_{ref}) + 0.5 \cdot m_{ref} \cdot y_{ref} = m_{ref} \cdot (1 - 0.5 \cdot y_{ref})$$ Equation 8

$$w3 = y_{ref} \cdot (1 - m_{ref}) + 0.5 \cdot m_{ref} \cdot y_{ref} = y_{ref} \cdot (1 - 0.5 \cdot m_{ref})$$

$$dgc\_X = w1 \cdot dgconp\_X + w2 \cdot dgconm\_X + w3 \cdot dgcony\_X$$

And $$dgc\_Y = w1 \cdot dgconp\_Y + w2 \cdot dgconm\_Y + w3 \cdot dgcony\_Y$$ Equation 9

$$dgc\_Z = w1 \cdot dgconp\_Z + w2 \cdot dgconm\_Z + w3 \cdot dgcony\_Z$$

Equation 9 expresses the multiple dot gain curves of cyan corresponding to certain reference coverage of CMY inks (c_{ref}, m_{ref} and y_{ref}), it means that not only the final effective coverage for cyan but also the dot gain for cyan will change even given the same c_{ref} but different m_{ref} and y_{ref}. Multiple dot gain curves for other primary inks could be fixed in the similar way.

4. RESULTS AND CONCLUSION

The proposed approach in this paper was tested by prints using a laser printer (Xerox, Phaser 6180) with normal office uncoated paper, 308 patches including 3 inks (cyan, magenta and yellow) and other 308 patches including only 2 inks (cyan, magenta) were created using FM1st at 300 dpi and AM at 600 dpi (100 lpi) respectively. Table 2 shows the prediction results by color difference ΔE_{94}, for both our approach and Yule-Nielsen model.

Table 2. The Color difference ΔE94 between measurement and calculated values (Laser printer, uncoated paper).

Only Cyan& Magenta	Yule-Nielsen model		our former approach		Multiple curves	
	ΔE_{94}		ΔE_{94}		ΔE_{94}	
	Max	Mean	Max	Mean	Max	Mean
AM-600 dpi	8.00	3.87	8.33	2.81	6.02	2.02
FM1st-300 dpi	9.56	3.78	7.87	2.39	7.89	1.98
three colors	Yule-Nielsen model		our former approach		Multiple curves	
	ΔE_{94}		ΔE_{94}		ΔE_{94}	
	Max	Mean	Max	Mean	Max	Mean
AM-600 dpi	12.26	4.92	13.66	4.75	9.85	3.43
FM1st-300 dpi	13.14	4.81	12.44	4.48	11.68	3.77
For YN model n_c =10, n_m =3.5, n_y=1 for FM1st prints; n_c =1.8, n_m =3.2, n_y=1 for AM prints;						

As shown in table 2 we succeeded to decrease the average ΔE_{94} between predicted and measured data from 4.75 (using our previous model without dealing with ink superposition) to 3.43 for AM at 600 dpi and from 4.48 (using our previous model which without dealing with ink superposition) to 3.77 for FM at 300 dpi, which are also better than Yule-Nielsen model (AM 600 dpi /4.92; FM 300 dpi /4.81). We think the result is quite satisfactory considering we have used uncoated paper and office laser printer. Consider also the fact that the measured CIEXYZ values for paper at different positions over the same sheet sometimes differ by a ΔE_{94} around 1.

According to all the color differences shown in our study, prediction models based on dot gain compensation, like our model and Yule-Nielsen model, might work better for prints with smaller dot gain or bigger dots. Meanwhile, obvious difference between dot gain curves based on CIEX, CIEY and CIEZ stands for the thinking that only using one dot gain curve is not enough to correctly deduce the dot gain. By using CIEXYZ to plot three dot gain curves for each ink, the calculated tri-stimulus values are closer to the measured values compared to the results obtained by Yule-Nielsen model with a fitted n-factor.

Furthermore, we showed that the dot gain for each ink also differs when different inks are mixed. Therefore we tried to include various dot gain curves for different superposition situation by weighting them with certain probability coefficients according to reference ink coverage of the prints. The final dot gain curves for each ink change as long as any reference ink coverage varies.

As shown in table 2, the usage of multiple dot gain curves decreases the color difference considerably for prints on office uncoated paper using a laser printer. Although our former model worked well for offset prints on coated papers (results are in table 1), we will continue our work by beginning with testing the developed method on prints using coated paper and offset press.

REFERENCES

[1] Wyble D., Berns R., "A Critical Review of Spectral Models Applied to Binary Color Printing," Journal of Color Research and Application. Papers 25(1), 4-19 (2000).

[2] P. Emmel and R. D. Hersh, "Modeling Ink Spreading for Color Prediction," J. Imaging Sci. Technol., Papers 46(3), 237-46 (2002).

[3] S. Gooran, M. Namedanian and H. Hedman, "A New Approach to Calculate Color Values of Halftone Prints, " IARIGAI 36th Research Conference, Advances in Printing and Media Technology, Sweden (2009).

[4] L.Yang, S. Gooran, and B. Kruse, "Simulation of Optical Dot Gain in Multi-chromatic Tone Reproduction," Journal of Imaging Science and Technology, Papers 45(2), 198-204, (2001).

[5] S. Gustavason, "Dot Gain in Color Halftones," PHD thesis, Dept. of Electrical Engineering, Linkoping University, Sept. (1997).

[6] F.R. Clapper and J. A. C.Yule, "The Effect of Multiple Internal Reflections on the Densities of Halftone Prints on Paper," Journal of the Optical Society of America. Papers 43(7), 600-603 (1953).

[7] P. Emmel and Roger D. Hersh, "A Model for Color Prediction of Halftoned Samples Incorporating Light Scattering and Ink Spreading, " Proceeding of the IS&T/SID 7th Color Imaging Conference: Color science, Systems and Applications, Scottsdale, Arizona, USA, pp. 173-181 (1999).