Modelling Vividness Perception for Colour Laser Printer Evaluation

Youn Jin Kim, Yousun Bang and Heui-Keun Choi; Digital Printing Division, Samsung Electronics; Suwon, South Korea

Abstract

The present study proposes metrics predicting level of vividness and preferred-vividness that are one of the important image-quality attributes for colour laser printers. Vividness is a term representing chromaticness of colours and also has been adopted as one of the colour adjectives in ISCC-NBS colour naming and PCCS systems. It is conceptually similar to chroma. According to ISO 20462-2, triplet comparison method*, which is a new psychophysical method that involves the simultaneous scaling of three test stimuli, was performed. As a result, an interval scale for vividness was established and it was modelled as a function of mean* C^*_{ab} *and* L^* *of primary colours, e.g. CMYRGB. Pearson correlation between the metric prediction and corresponding subjective data was about 0.96. Both preferred-vividness and vividness metrics were based upon chroma and lightness but contribution of lightness is a lot higher for the former (~ 40%) than for the latter (~ 10%). Consequently, our vividness metric confirms earlier findings published by Nayatani in 2005.*

Introduction

Vividness is a term representing chromaticness of colours and also has been adopted as one of the colour adjectives in ISCC-NBS[10] colour designation and PCCS[15] system. It has been frequently used for evaluating business graphic print quality in printing industry. Nayatani[1-3] reported that the concept of degree of vividness is similar to the definition of "chroma" in CIE International Lighting Vocabulary.[4] It can be used to estimate chromatic intensity of colours using interval or ratio scales. For example, achromatic colours have zero vividness and highly saturated colours would show a higher value of it. Nayatani (2005) [3] proposed an empirical model predicting degree of vividness (DV) as a function of Munsell chroma (C) and whiteness and blackness ([W-Bk]) based on the observations of NCS colour chart as shown in Equation 1. The compound characteristics of vividness affected by chroma and lightness channels could also be observed in ISCC-NBS system as well.[10] Significantly large differences in Munsell Values and Chromas of the central colours were revealed for different hues.

$$DV = C\{1 + 0.10[W\text{-}Bk]\} \qquad (1)$$

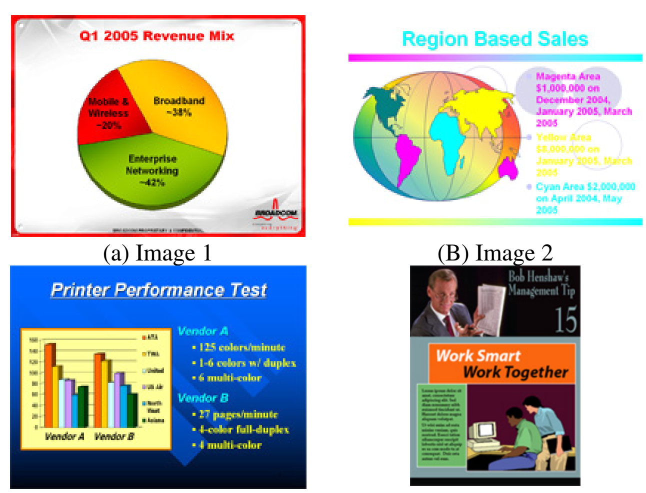
In the field of image quality, colour vividness has been understood as the degree of colourfulness[8-9] and there have been a number of efforts to predict colourfulness of images and its effects on the image quality.[11-13] Colourfulness, which is one of the perceptual attributes in colour appearance modelling, is another very similar concept to vividness. It is defined as an attribute of a visual sensation according to which an area appears to exhibit more or less light.[5, 14] The colourfulness of a given colour stimulus increases with luminance which is referred to as Hunt effect.[14] This effect describes the perceptual difference caused by large differences in illumination and the corresponding state of adaptation.[11]

This study aims to verify Nayatani's empirical vividness model and develop a metric that accurately predicts level of vividness and observer preference for evaluating colour laser printer. Since CIELAB colour space has been often used in imaging industry, metrics based upon CIELAB will be developed and compared with Nayatani's model. Four sets of psychophysical experiment have been done in this study as subsequently discussed in detail.

Experimental

Setup

In total, 12 colour laser printers produced by different manufacturers - such as Brother, Canon, Dell, HP, Minolta, Ricoh, Samsung and Xerox - were selected and compared each other in terms of vividness of the prints and observer preference. Those sampled printers show a wide range of printing performances from low-end for small business to high-end for industrial applications. For each printer, CIELAB coordinates of the maximum CMYRGB were measured using GretagMacbeth Spectroscan. Figure 1 provides test images used in this study. Since vividness is strongly related to purity of colours[1-3], business graphic images depicting highly saturated colours were chosen. For instance, test images 1 and 2 mainly include maximum cyan, magenta, yellow, red, green and blue colours and the others contain intermediate hue-levels of those primary colours.



(a) Image 1 (B) Image 2

(c) Image 3 (d) Image 4

***Figure 1**. Test stimuli*

Visual Assessment Methodology

According to ISO 20462-2 (Photography – Psychophysical Experimental Methods for Estimating Image Quality – Part2: Triplet Comparison Method)[6], two steps of psychophysical experiment were conducted. Triplet comparison is a new psychophysical method defined in ISO 20462 that involves the simultaneous scaling of three test stimuli with respect to image quality or an attribute thereof, in accordance with a set of instructions given to the observer.[6] It enables a large number of samples to be examined and provides precise scalability with a much lower observer stress than paired comparison method. Plus, the procedure is as simple as paired comparison and results are highly repeatable. For the all psychophysical experiments performed in this study, assessment distance between an observer and given stimuli was set to 25cm and the viewing condition was a typical office lighting (~ 800 lx). The viewing geometry was 45/0 to attenuate any glaring effects from the stimuli caused by specular reflection. The observers are all experts working in the field of colour imaging industry.

Step 1: Category step (Reducing the number of samples)

The first step of the triplet comparison method is a category step that aims to reduce the number of samples. All samples are categorised into three categories defined as '3: Favourable', '2: Acceptable' and '1: Unacceptable' and some of the samples can be excluded by clustering the data based upon the sample characteristics. (Note that equally perceived intervals were assumed between any two consecutive categories.) The number of reduced samples (*N*) should follow the following equations.

$$N = 6K + 1 \qquad \text{or} \qquad N = 6K + 3 \tag{2}$$

where *N* is the number of reduced samples and *K* is an integer number.

Four expert observers who have worked in the field of colour imaging industry participated in Step 1 and divided 48 images produced by the twelve sample printers (= 4 images × 12 printers) into the three categories in terms of their preference. Each observer assessed each print at a time under a typical office lighting environment. Sequence of the assessments was randomised and the collected subjective data were averaged for each printer. This is one of the common methods for analysing the category judgment data sets and it has been recommended by ITU-R BT.500-11.[7] The mean subjective score is often referred to as mean opinion score (MOS) that can be computed as

$$\bar{u}_{jk} = \frac{1}{n}\sum_{i=1}^{n} u_{ijk} \tag{3}$$

where u_{ijk} is a subjective score of observer *i* for test printer *j* and image *k* and the number of observers is *n*. The total number of observations is 192 (= 4 images × 12 printers × 4 observers). Through this category scaling, seven printers were selected to be used in the next step (triplet comparison).

Step 2: Triplet comparison

The second step is to derive a precise scaling based on an interval scale by comparing triplets of given samples. Specifically, three samples are compared simultaneously, thereby achieving high assessment accuracy while keeping the experimental scale realistic. Compared to paired comparison method, triplet comparison shortens assessment times so is expected to improve data accuracy and reproducibility. Following ISO 20462-2[6], Scheffe's method was applied for the statistical analysis to obtain an interval scale and it was converted into just noticeable difference (JND) values. The interval scale relies upon Thurston's law of comparison case V by computing cumulative frequency distribution matrix and probability. Because of the word limitation, mathematical derivations of the JND conversion are not given in this paper. (More detailed information can be found in Annex F of ISO 20462-2.) Thirteen expert observers including the four observers participated in the previous session (Step 1) assessed 364 triplets (= 4 images × 7 triplets × 13 observers) under the same office lighting environment. Each observer was asked to compare each triplet and rank the test stimuli in terms of vividness.

Results

Step 1: Category step

Table 1 gives mean opinion scores (MOS) of the 12 sample printers across the all test images and 13 observers obtained from Step 1 procedure. Printers A through E showed higher MOS values and MOS of printers J through L was under score of 1.5 which is much lower than the others. Therefore, the data could be clearly clustered by three groups. The first group (Group I) includes printers showing higher MOS larger than 2.5 (printers A through E) and the rest can also be divided into the middle (printers F through I: Group II) and lower (printers J through L: Group III) MOS groups. This clustered data distribution is graphically illustrated in Figure 2.

Table 1. Printer vs. MOS from Step 1

Printer		MOS
Group I	A	2.81
	B	2.63
	C	2.63
	D	2.56
	E	2.50
Group II	F	2.19
	G	2.06
	H	1.94
	I	1.81
Group III	J	1.25
	K	1.13
	L	1.00

As Equation 2 recommends, the number of samples should follow either *6K + 1* or *6K + 3*, (where *K* is an arbitrary integer) in triplet comparison method. Therefore, 5 of the 12 printers (B, C, G, I and L) were excluded in this study so the other 7 (A, D, E, F, H, J and K) were used only in a subsequent psychophysical experiment in Step 2. Printer L was taken out due to its unacceptably worse colour reproduction quality and B, C, G and I were also excluded so that that the remaining samples can show a wide range of printing performances from low-end for small business to high-end for industrial applications.

©2008 Society for Imaging Science and Technology

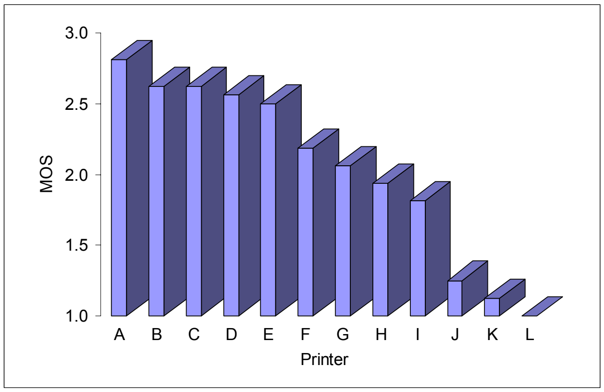


Figure 2. *Graphical illustration of each sample printer's mean opinion score (MOS) obtained from Step 1 procedure. Printers A through E showed higher MOS values and MOS of printers J through L was under score of 1.5 which is much lower than others. Therefore, those data can be clearly clustered into three groups (A-E / F-I / J-L).*

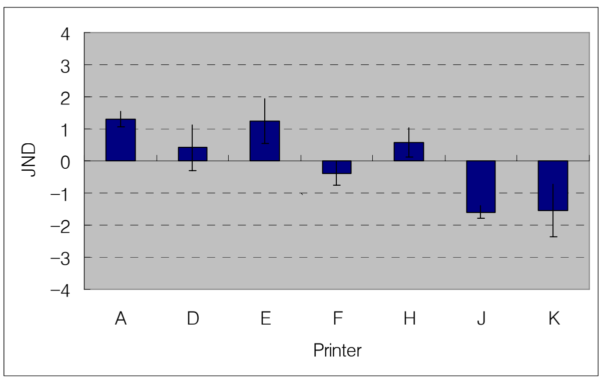


Figure 3 *Total JND computed obtained from Step 2 prcedure arcoss the all test images (Error bars show 95% confidence interval.)*

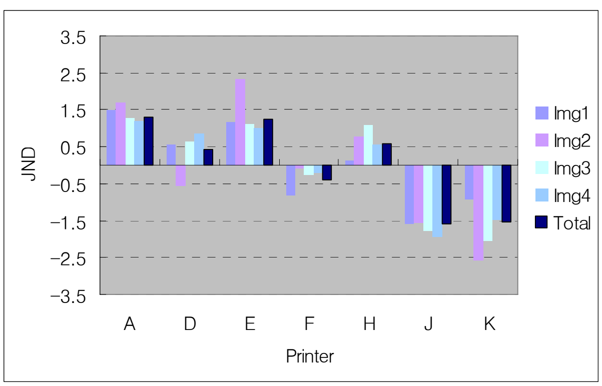


Figure 4. *Independency on different test images (Note that a lower JND in img2 for printer D is due to a huge banding artifact only shown in that case.)*

Step 2: Triplet Comparison

In Figure 3, total JND values across the all test images obtained from Step 2 procedure (triplet comparison) for those 7 selected sample printers are dipicted. The larger JND, the higher subjective vividness score. Printers A and E produced the highest vividness and H, D, F, J and K followed in order. Error bars show 95% confidence interval which can be computed as

$$\mu \pm t_{2.5\%}(n-1)\times SE \qquad (4)$$

where μ denotes the total JND value of each printer and *SE* denotes its standard error of mean which is standard deviation divided by square root of number of observations. (n-1) is the degree of freedom. The value that has 2.5% of t for the degree of freedom is given as $t_{2.5\%}$.

In Figure 4, JND values were separately computed for different test images and compared each other in a single plot. Different coloured bars represent different test images. Generally, a similar data trend for different test images can be seen. However, it should be noted that there is an exceptional case in image 2 for printer D. Its JND value was much smaller than the other images. Apparently, a considerable amount of banding artifact was observed in that specific image and resulted in the exceptionally lower vividness score.

Modeling Vividness

Memorisation

Following the structure of Nayatani's empirical vividness model[3], both chroma and achromatic intensity were selected as dependent variables in vividness modelling. Since CIELAB colour space has been often used in imaging industry, metrics based upon C^*_{ab} and L^* were developed. Precisely, an independent variable can be determined as a function of mean C^*_{ab} and L^* across printer primary colours, e.g. CMYRGB, as shown in Equation 5.

$$\psi = \frac{1}{n}(\omega_C \sum_i^n C^*_{abi} + \omega_L \sum_i^n L^*_i) \qquad (5)$$

where n denotes the number of primary colours to be used, e.g. n = 6 for CMYRGB. w_c and w_L represent weighting factors of C^*_{ab} and L^*, respectively. Those weighting factors were optimised using linear regression method so were determined as 0.91 (=w_c) and 0.09 (=w_L).This finding is similar to the results from an earlier vividness modelling[3] which predicts degree of vividness as a function of Munsell chroma (C) and whiteness-blackness ([W-Bk]) by Nayatani (2005). As shown in Equation 1, the coefficient of [W-Bk] is 0.10 while that of C is 1.0. Equation 6 shows a matrix form of Equation 5. The relation between vividness (**V**) and the dependent variables is determined by **T**.

$$\mathbf{V} = \mathbf{TS} \qquad (6)$$

where **S** is the corresponding mean C^*_{ab} and L^* of primary colours of a given colour laser printer, which constitute a 2 × 1 column matrix so the size of the transformation matrix **T** can be 1 × 2. This matrix **T** represents the relationship between the level of vividness achieved by a given colour laser printer and the printer's physical characteristics. In other words, this relation represents how to bridge the gap between them and can let us understand the observers' taste for vividness. Mathematically, least square was performed to minimise residual errors between known subjective vividness scores and their corresponding metric predictions. The solution for minimising the residual error is

$$\mathbf{T} = (\mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T\mathbf{V} \qquad (7)$$

where $\mathbf{S}^T$ denotes the transpose of **S**, and $\mathbf{S}^{-1}$ the inverse.

Figure 5 plots a linear relation between mean vividness metric predictions for the seven sample printers used in triplet comparison experiment and their corresponding subjective JND values. Pearson correlation (r) between the two data sets was found to be 0.972.

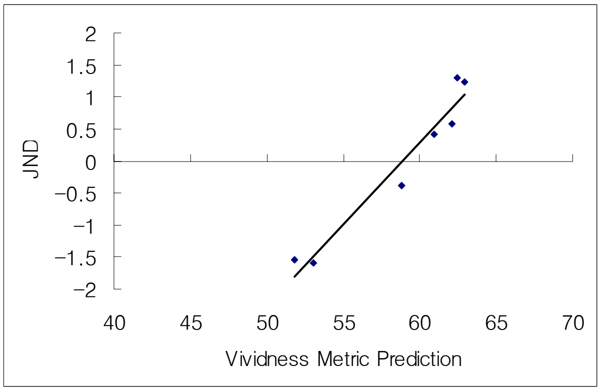


Figure 5. *Relation between subjective data (JND) and vividness metric prediction (r = 0.972)*

Generalisation

In order to verify merits and generality of our vividness metric, metric generalisation was performed through another set of category judgment method. Predicted values from our vividness metric were compared with their corresponding subjective data obtained from the following psychophysical procedure. A five-point scale, where all categories are defined by a symmetrical design of quantitative adjectives, was used. The categories were defined as '5: Highly Vivid', '4: Quite Vivid', '3: Vivid', '2: Quite Unvivid' and '1: Highly Unvivid'. Six expert observers rated printed test images using the five-point scale. The test images were printed by the 5 printers which were excluded in Step 1 procedure and were not used for developing the vividness metric in Step 2. The total number of observations is 120 (= 4 images × 5 printers × 6 observers).

In Figure 6, relation between mean vividness metric predictions for the 5 sample printers used in this metric generalisation procedure and their corresponding vividness scores (MOS). Pearson correlation (r) between the two data sets was found to be 0.964. (This quite high correlation may be due to the fact that visual assessment for vividness is relatively easier than that for other attributes.[3]) Consequently, the vividness metric developed in this study could very accurately predict subjective vividness of different colour laser printers that were not used in the metric's coefficient optimisation.

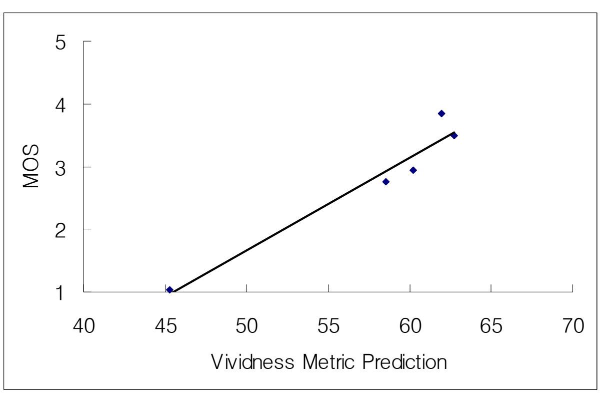
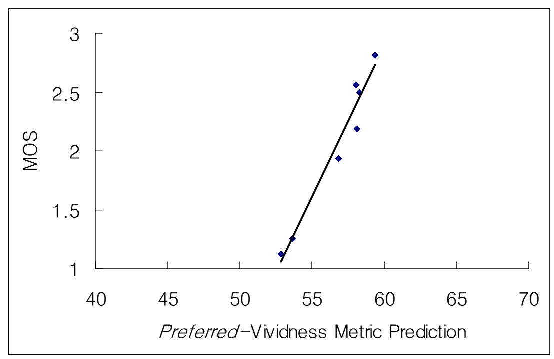


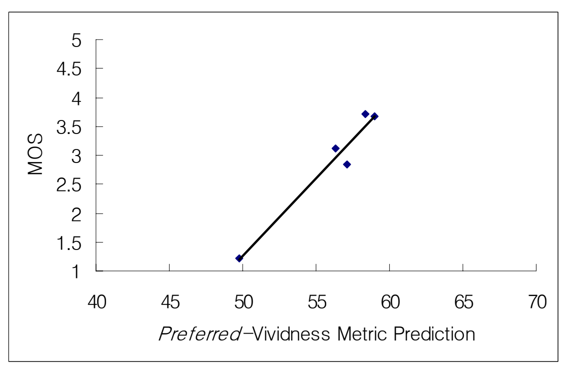
Figure 6. *Relation between subjective data (MOS) and vividness metric prediction (r = 0.964)*

Modeling Preferred-Vividness

The chroma and lightness weighting factors in Equation 4 were re-optimised using the data set from Step 1 which is based upon preference assessment. The original purpose of Step 1 procedure defined in ISO 20462-2 was to reduce the number of sample printers. However, it was also used as a training data set for *preferred*-vividness metric in this study. The same optimisation process previously discussed in Equations 5 and 6 was repeated and the weighting factors were determined to be 0.57 (=w_c) and 0.43 (=w_L). Merits and performance of the *preferred*-vividness metric were also tested via the previous metric generalisation procedure. The six observers participated in the previous experiment rated the test images printed by the 5 printers which were excluded in Step 1. The five categories used were defined as '5: Favourably Vivid', '4: Acceptably Vivid', '3: Just Acceptably Vivid', '2: Unacceptably Vivid' and '1: Poor'. The collected data were averaged across the observers and test images for each printer. Figures 7 (a) and (b) show results of the weight optimisation and metric generalisation. The abscissa represents prediction of preferred-vividness metric and the ordinate shows subjective data in MOS. Pearson correlation was 0.972 for the former and 0.978 for the latter. Therefore, it can be said that accuracy of the preferred-vividness is high enough to estimate subjective vividness preference of various colour laser printers.



(a) Memorisation (weight optimisation: r=0.972)



(b) Generalisation (r=0.978)

Figure 7. *Relation between subjective data (MOS) and preferred-vividness metric prediction (a) for metric memorisation (weight optimisation) and (b) for metric generalisation*

©2008 Society for Imaging Science and Technology

Table 2. Comparison of weights for vividness and *preferred*-vividness metrics

	w_c	w_L
Vividness	0.91	0.09
Preferred-Vividness	0.57	0.43

Table 2 lists and compares optimised weights between for vividness and *preferred*-vividness metrics. As can be seen, contribution of lightness is a lot higher for preferred-vividness (~ 40%) than for vividness (~ 10%). Apparently, a reasonably higher lightness level is also required as well as a higher chroma level to achieve a higher observer-preference. Performance of printers D and H can be good examples supporting this hypothesis. According to the triplet comparison data for vividness (Step 2), rankings of printers D and H were 4th and 3rd, respectively, as indicated in bold in Table 3. However, their order was reversed in preferred-vividness data (D for 2nd and H for 5th). It is due to the fact that the mean L^* of printer H is much lower than the others so its preference score was decreased despite its quite higher chroma (or vividness). It should be noted that the ratio (L^*/ C^*_{ab}) of H is much lower than that of D. In contrary, printer D obtained the 2nd highest preference score with the aid of its high chroma and lightness values.

Table 3. Comparison of Mean C^*_{ab} and L^* and their ratio for seven sample printers (The ratio of H is much lower than that of D.)

Printer	Mean C^*_{ab}	Mean L^*	Mean L^*/C^*_{ab}	Vividness Rank	Preferred-Vividness Rank
A	63.23	54.31	0.86	1	1
D	61.67	53.21	**0.86**	**4**	**2**
E	64.16	50.54	0.79	2	3
F	58.98	56.91	0.96	5	4
H	63.46	48.21	**0.76**	**3**	**5**
J	52.93	54.49	1.03	7	6
K	51.50	54.56	1.06	6	7

Conclusion

The present study proposes a metric predicting level of vividness that is one of the important image-quality attributes for colour laser printers. A number of psychophysical assessments were carried out following a procedure recommended by ISO 20462-2[6] and vividness and preferred-vividness were quantified as a function of mean C^*_{ab} and L^* of primary and secondary colours, e.g. CMYRGB. The merits and performance of those metrics were evaluated by means of comparing with corresponding subjective results (r > 0.96). The vividness metric proposed in this study is based upon chroma and lightness defined in CIELAB colour space but weight of lightness (~ 10%) is a lot higher than that of chroma (~ 90%). Contribution of lightness for vividness metric also agrees with earlier findings[3] by Nayatani (2005). For future studies, effective hues to the vividness perception will be studied. Currently, we have used mean chroma and lightness values of given printers' primary and secondary colours under an assumption that there is no significant impacts from hue. However, more experiments will be conducted to separate their effects based upon colour appearance and vision theories.

References

[1] Nayatani Y. Proposal of an opponent-colours system based on colour-appearance and colour-vision studies. Colour Res Appl 2004;29:135-150.

[2] Nayatani Y. Why two kinds of colour order systems are necessary? Colour Res Appl 2005;30:295-303.

[3] Nayatani Y. Relationships among chromatic tone, perceived lightness, and degree of vividness. Colour Res Appl 2005;30:221-234.

[4] CIE Publ. No. 17.4. International Lighting Vocabulary; 1987.

[5] Fairchild MD. Colour appearance models, 2nd ed., John Wiley & Sons, 2005.

[6] ISO 20462-2, Photography – Psychophysical experimental methods for estimating image quality. Part 2: Triplet comparison method, 2005.

[7] ITU-R Rec. BT. 500-10, Methodology for the subjective assessment of the quality of television pictures, 2002.

[8] Fedorovskaya EA., de Ridder H., Blommaert FJJ. Chroma variations and perceived quality of colour images of natural scenes. Colour Res Appl 1997;22:96-110.

[9] Tremeau A., Charrier C. Influence of chromatic changes on the perception of colour image quality. Colour Res Appl 2000;25:200-213.

[10] Nickerson D, Newhall. Central notations for ISCC-NBS colour names. J Opt Soc Am 1941;31:587-591.

[11] Cui C. Which colour is more colourful, the lighter one or the darker one? Colour Res Appl 2003;28:168-174.

[12] Yendrikhovskij SN, Blommaert FJJ, de Ridder H. Colour reproduction and the naturalness constraint, Colour Res Appl 1999;24:52-67.

[13] Kim YJ, Luo MR, Rhodes P, Cheung V, Westland S, Choe W, Lee S, Kim C. Image colour-quality modelling for mobile LCDs, *Proc. 14th Col. Img. Con.* 159-164 (2006).

[14] Hunt RWG. The reproduction of colour, 6th ed., John Wiley & Sons, 2004.

[15] Kawasaki H, Kodoma A. A study of practical colour coordinate system. *Proc. AIC Colour 97* Vol I 467-470 (1997).

Author Biography

Youn Jin Kim *received his B.S. degree in Physics from Daejin University, Korea, in 2003, M.S. degree in Imaging Science from the University of Derby, UK with Distinction, in 2004, and Ph.D. in Colour Chemistry from the University of Leeds, UK, in 2008. Since then he has worked in the Digital Printing Division at Samsung Electronics in Korea. His interest has focused on the development of image quality evaluation tools for colour laser printers and scanners.*

Yousun Bang *received her B.S. and M.S. degrees in mathematics from Ewha Womans University, Seoul, Korea, in 1994 and 1996, respectively. She received her Ph.D. degree in the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, in 2005. She has been working at Samsung Electronics Company since 2004. Her research interests include colour management system, and image rendering, and print quality.*

Heui-Keun Choh *received his B.S. degree in physics and M.S. degree in solid state physics from Korea university, Seoul, Korea, in 1984 and 1986, respectively. He has been with Samsung Electronics, Korea. His research interests include modeling of colour imaging devices, appearance colour modeling, cross-media colour reproduction for colour management, digital printing, and image processing for colour imagers.*