Image Colour-Quality Modelling for Mobile LCDs

Youn Jin Kim[1], Jangjin Yoo[1], M. Ronnier Luo[1], Peter Rhodes[1], Vien Cheung[2], Stephen Westland[2], Wonhee Choe[3], Seongdeok Lee[3] and Changyeong Kim[3]

[1]Colour & Imaging Group, University of Leeds, Leeds, UK

[2]Centre for Colour Design Technology, University of Leeds, Leeds, UK

[3]Samsung Advanced Institute of Technology, Yongin, South Korea

Abstract

This paper describes the development of an image colour-quality model based on individual physical image statistical measures for mobile liquid crystal displays (LCD). Five natural images were colour-rendered in terms of lightness, chroma and hue. Each of the images was displayed on a 2-inch QVGA mobile LCD and assessed by a panel of 10 observers in terms of image quality using a categorical judgment method. Only colour attribute modeling was carried out in this paper. Image statistical measures were established to quantify the image quality of natural colour images varying in colour. Those were memory colour reproduction ratio (MCRR), mean chroma and 95th percentile lightness. The Pearson correlation between the model predictions and their corresponding psychophysical data was 0.88 and the coefficient of variance was 18. The model outperformed observer accuracy in terms of those two measures. It is also significant that the subjective scale of image quality was bridged with objective metrics such as the image statistical measures.

Introduction

Image quality has been recognised as one of the top considerations in the display manufacturing industry, where there is a perpetual trade-off between quality and cost.[19] Hence, an image quality metric, which can accurately quantify the quality of an image according to human visual perception, is strongly desired. Objective evaluation involves physical measurement of images but generally fails to consider human visual characteristics. Therefore, psychophysical experimental results are required for developing metrics. Subjective image quality research can be divided into two major approaches: external (or relative image quality) and internal reference (or absolute image quality). The former assumes that the image quality of reproductions corresponds to perceptible visual difference from its original.[3] A number of these metrics have been suggested and widely used such as CIELAB colour difference equation[4] and CIEDE2000.[5, 6] S-CIELAB[7, 18] was developed in 1996 as an image difference metric accounting for image spatial properties.

Internal reference image quality is defined as the image quality of a given image corresponding to perceptible visual difference from its memory prototype. The category judgment method is appropriate for this approach, in which observers assess a single image by perceptual comparison with a cognitively represented reference, for which the original is not presented.[8, 9] There has been some effort to appraise an image without an original based on information theory[8] and the similarity to the memory colours of sky, grass, and Caucasian skin.[9] It was found that the appearance of particular memory colours are remembered slightly differently from the colour measurement from the real world.[10]

Previous Study

In the previous study[1], affective attributes in image quality modelling were investigated. These included naturalness and clearness. The experiment was originally designed to develop two types of image quality models: perceptual (or high-level) and physical (or low-level). The former was studied in the previous study and represents an image quality model that involves some perceptual attributes as input values such as naturalness, clearness or sharpness. Five natural test images were rendered in terms of seven physical parameters: two types of lightness rendering functions, chroma, hue, peak-white luminance, resolution, bit depth and correlated colour temperature (CCT). Observers used 9 categories (1 to 9) to appraise each image, according to 7 perceptual attributes: naturalness, clearness, sharpness, contrast, colourfulness, image quality and preference. The attributes were inter-compared using Pearson correlation.[28] Naturalness and image quality showed a very high correlation (0.96). Clearness seemed to be highly associated with sharpness (0.97). Image quality and preference were judged to be the same attribute (0.99). Image quality was modelled by the 5 perceptual attributes through a stepwise regression method[28], and it was found that naturalness and clearness are the principal affective attributes in the image quality of mobile displays.

The second image quality model is based on the physical attributes which include image statistical measures in the colour, spatial, or temporal domains. In the current paper, only colour attributes were considered and the accumulated mean opinion score (MOS) values of image quality from the previous study were used to develop an image colour-quality model based on image statistical measures such as memory colour reproduction, mean chroma and 95th percentile lightness. The spatial attributes were left for future research.

Experimental

Setup

Test stimuli were displayed on a Samsung SCH-S250 mobile phone.[15] It uses a 2-inch QVGA and its colour gamut is similar to sRGB as shown in the CIE 1931 xy chromaticity diagram.[21] A Minolta CS-1000 tele-spectroradiometer was used for measurement. The display was characterised using the PLCC (Piecewise Linear Interpolation Assuming Constant Chromaticity) method.[11] A 9 equal step greyscale was measured for training a characterization model. With the combinations of 0, 64, 128, 192 and 255, another 125 colours were selected to test the characterisation model.

The median colour difference between the model prediction of the test colours and their corresponding measurement was 4.0 CIELAB colour difference units. This discrepancy represents the typical colour characterisation accuracy for the display.[22] A colour reproduction exercise was carried out between the mobile and a

15-inch twisted nematic (TN) LCD and the performance of colour reproduction was tested through a visual examination. There was a reasonable match between the images reproduced by the two media. The characterisation model can reasonably express the relationship between the input signal and display output performance, despite having prediction error of 4.0 CIELAB colour difference units.

Test Stimuli

Figure 1 shows five test stimuli used in this study. It is common for images of facial and natural (sky, grass) scenes captured under outdoor daylight to be viewed. Hence, the images contained facial skin (Caucasian, Black, and Oriental), blue-sky, green-grass, water, and fruit colours.

Figure 1. *Test Stimuli*

Those images were rendered in terms of lightness, chroma and hue. For lightness and chroma, each pixel value was linearly scaled. Hue was altered by means of adding a scaling factor for each pixel in an image. In total, there were 95 images rendered. The levels, a rendering function and the total number of images of each data set are listed in Table 1. Hue, chroma and lightness data sets are identified as Data 1 through 3, respectively.

Table 1. Levels and rendering functions

ID	Parameter	Level (k)	Function
Data 1	Hue (°)	-60, -45, -30, -15, 0, 15, 30, 45, 60	Out = k + In
Data 2	Chroma	1.0, 0.8, 0.6, 0.4, 0.2	Out = k X In
Data 3	Lightness	1.0, 0.9, 0.8, 0.7, 0.6	Out = k X In

Table 2. The definition of Bartleson categories[13]

Category	Definition
1	Least imaginable “ness”
2	Very little “ness”
3	Mildly “ness”
4	Moderately “ness”
5	“Ness”
6	Moderately highly “ness”
7	Mildly highly “ness”
8	Very highly “ness”
9	Highest imaginable “ness”

Procedure

Ten observers with normal colour vision took part in this experiment. They were asked to rate the image quality of each of the displayed images on the mobile LCD at a distance of 25cm in a dark room using a 9-point scale. All categories were defined by a symmetrical design of quantitative adjectives originally suggested by Bartleson[13] and listed in Table 2. Equal-perception intervals were assumed between two consecutive categories.

Modelling Image Colour-Quality

The developed image colour-quality model is comprised of three parts: quantifying memory colour, chroma, and lightness. It is capable of predicting the quality of individual images in respect of colour variation. Each of the attributes affecting image quality was modelled separately and all three were combined into a single image colour-quality model.

Memory Colour Affecting Image Quality

Successfully reproducing the correct colour of a certain image content is vital to achieve high image quality. Correct colour can be defined in many ways. If the original can be immediately accessed, the correct colour of the image content would be the same as the original. However, it is not usually possible to compare side by side the reproduction with the original.[10] More often, the reproduction is seen at different places and different times under different viewing conditions (e.g. overcast or bright sunlight). This demonstrates the important role of colours stored in human brain (memory colour prototype).[10] The correct colour of a particular content in image quality can be defined as the similarity to the memory prototype.[9]

The concept of region of interest (ROI) was adopted at this point. Basically, it is assumed that when the ratio of reproduced colours in an ROI that are similar to its memory prototype is higher, the image should exhibit higher image quality. A method was developed in this study for quantifying the memory colour reproduction ratio (MCRR) and its conceptual process is illustrated in Figure 2.

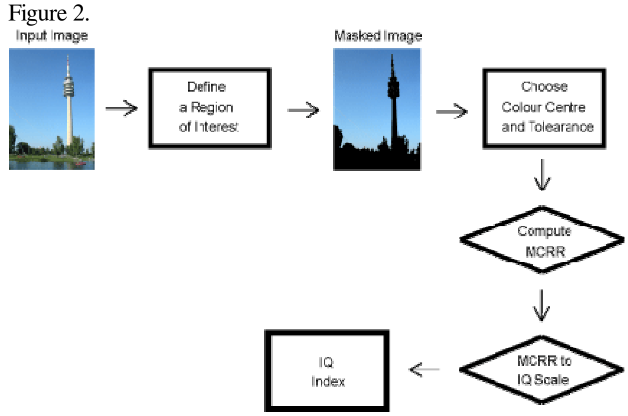

Figure 2. *A flowchart of the conceptual process of predicting image quality using memory colour reproduction ratio (MCRR) computing algorithm*

Defining a region of interest (ROI)

Defining an ROI depends on the judgment criteria of observers. When assessing an image, most observers focused on a particular object and its area is a ROI. Figure 3 shows masked images with the important area. The masking process can discard some wrongly included colours that show similarity to the memory prototype but which do not belong to the region. In addition, eyes, eyebrows and lips of the facial scene (Ladies) were masked.

***Figure 3.** Masked Stimuli*

Choosing a colour centre and tolerance

As mentioned earlier, an input image is compared with an internal memory prototype in image quality judgment. The internal memory prototype can be defined in terms of a colour centre and a tolerance. The colour centre is the mean colour coordinates of a certain memory colour and the tolerance is a level of acceptable colour difference unit from this colour centre. The scene-dependency effect in image quality judgment can be compensated by those two factors. Those used for each test image are given in Table 3. The tolerance for all images was optimised and set to 25 ΔE^*_{ab} except for "fruits" (35 ΔE^*_{ab}).

Table 3. List of region of interests and colour centres (*: Approximation)

Stimuli	ROI	Colour Centre (CIELAB)
Skytower	Blue-sky	(65.7, -10.4, -28.9)[25]
Picnic	Sky with clouds	(65.7, -2.8, -34.5)[9]
Grass	Green-grass	(45.5, -25.2, 41.9)[9]
Ladies	Oriental-skin	(81.8, 35.0, 5.0)*
Fruits	Orange	(68.2, 40.0, 50.0)*

Previously published mean memory colour from Tarczali et al.[25] and Yendrikhovskij et al.[9] were used here for blue-sky (Skytower), sky with cloud (Picnic) and grass (Grass). For the oriental face (Ladies) and orange (Fruits) images, data from Tarczali et al.[25] did not successfully explain the colour centres and so the colour centres in Table 4 were chosen through an optimisation process from our data set until the best approximation could be made.

Computing MCRR

The model calculation of the memory colour reproduction ratio (MCRR) is defined as the ratio of reproduced colours in a particular ROI, of which colour difference from its colour centre is less than the given tolerance, as shown in Equation 1.

$$MCRR = \frac{1}{m}\sum_{x=0}^{X-1}\sum_{y=0}^{Y-1} cat(x,y) \tag{1}$$

where X and Y are the numbers of horizontal and vertical pixels in the image considered and $cat(x,y)$ is a binary number at each pixel in an input image, i.e. 1: within tolerance or 0: out of tolerance. The total number of pixels categorised into the *ROI* is m.

MCRR Performance

Figure 4 shows the relationship between the MCRR model predictions after taking log and their corresponding MOS of the all data sets in Table 1. The Pearson correlation coefficient was 0.74, which does not show a clear relationship between them. The explanation for this scattered graph can be seen in Figures 5 to 7.

Figures 5 through 7 depict the relationship between Ln(MCRR) and MOS values of the different data sets. As can be seen in the series of plots, the hue effect (r = 0.90) was very important for quantifying image quality by memory colour reproduction ratio. The chroma effect could also be important, but the Pearson correlation was relatively lower (r = 0.74) for this than for the hue effect (Figure 6). For the lightness data set (Figure 7) the correlation between the Ln(MCRR) model prediction and visual results (MOS) was very low (r = 0.20). As a result, it can be concluded that memory colour can be an important factor to predict image quality, and that hue plays a very important role in memory colour. However, for the change in chroma and lightness data sets, the performance of Ln(MCRR) was not sufficient. Some additional terms to describe chroma and lightness effects are required to accurately quantify image quality.

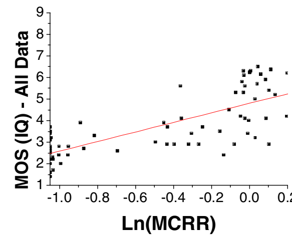

***Figure 4.** The relationship between Ln(MCRR) Prediction and MOS of all three data sets: r = 0.74*

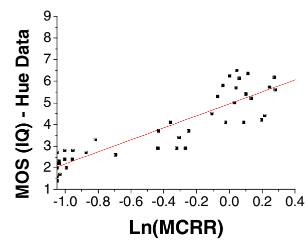

***Figure 5.** The relationship between Ln(MCRR) Prediction and MOS of data set 1 (Hue): r = 0.90*

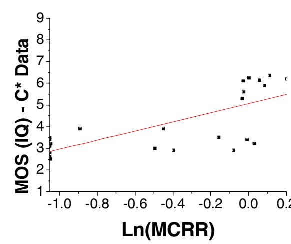

***Figure 6.** The relationship between Ln(MCRR) Prediction and MOS of data set 2 (Chroma): r = 0.74*

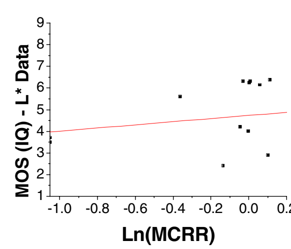

***Figure 7.** The relationship between Ln(MCRR) Prediction and MOS of data set 3 (Lightness): r = 0.20*

Chroma Affecting Image Quality

In order to quantify chroma effects in image quality, only data 2 (chroma-rendered images) were used in this section. Table 4 shows the Pearson correlation between a number of image chroma statistical measures and the MOS. In total, three different measures were implemented, i.e. mean C^*_{ab} (or $\overline{C^*_{ab}}$), mean ΔE_n (or $\overline{\Delta E_n}$) and C_Y. The former three are given in Equations 2 to 4. C_Y is the colourfulness model obtained by summation of the mean and standard deviation of saturation, as suggested by Yendrikhovskij et al.[9] The saturation term in CIELUV space was replaced by chroma in CIELAB, because the CIELAB space is more frequently used in the colour imaging industry than CIELUV.

$$\overline{C^*_{ab}} = \frac{1}{XY}\sum_{i=0}^{X}\sum_{j=0}^{Y}\sqrt{(a_{ij}^{*2} + b_{ij}^{*2})} \quad (2)$$

$$\overline{\Delta E_n} = \frac{1}{XY}\sum_{i=0}^{X}\sum_{j=0}^{Y}\sqrt{((L^*_{ij} - 50)^2 + a_{ij}^{*2} + b_{ij}^{*2})} \quad (3)$$

$$C_Y = \frac{1}{XY}\sum_{i=0}^{X}\sum_{j=0}^{Y}(\sqrt{(a_{ij}^{*2} + b_{ij}^{*2})}) + \sigma \quad (4)$$

where X and Y are the numbers of horizontal and vertical pixels in the image considered. L^*, a^* and b^* represent CIELAB coordinates for each pixel of the image, and σ is the standard deviation of C^*_{ab} for the pixels in an image.

Table 4. Performance of chroma statistical measures

Measure	Pearson Correlation	Description
$\overline{C^*_{ab}}$	0.87	Mean C^*_{ab}
$\overline{\Delta E_n}$	0.81	Mean distance from mid-grey (L^* of 50)
C_Y	0.84	Summation of mean and standard deviation of C^*_{ab}

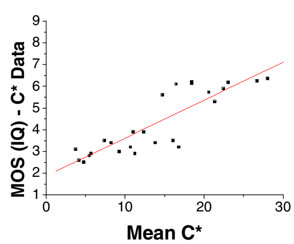

Figure 8. The relationship between mean C Prediction and MOS of data set 2 (Chroma): r = 0.87*

Mean C^*_{ab} showed the best correlation with MOS (r=0.87) in Table 4. Figure 8 plots mean C^*_{ab} against MOS for all the chroma-rendered image data (Data 2 shown in Table 1) and shows their linear fit.

Lightness Affecting Image Quality

Data 3 was the collected image quality MOS values of lightness rendering, and this data set was used to model lightness effect in image quality. Five different image lightness statistical measures were implemented: mean L^* ($\overline{L^*}$), maximum L^* (L^*_{Max}), 95 percentile L^* ($L^*_{95\%}$), *SIPk*[14] and *SQRI*.[16]

As shown in Table 5, mean L^* showed a very low correlation with MOS (r=0.62). However, high correlations with MOS could be obtained by Maximum L^* (r=0.93) and 95th percentile L^* (r=0.87). The mean lightness of each image was not a reliable measure due to its scene-dependency. Observers adapted to the overall lightness of each scene. Maximum and 95th percentile L^* can be more appropriate measures for predicting image quality change in lightness rendering. These measures can explain the range of lightness in an image. Boust et al.[17] also reported that image quality can be enhanced by increasing the available lightness gamut.

Table 5. Performance of lightness statistical measures

Measure	Pearson Correlation	Description
$\overline{L^*}$	0.62	Mean L^*
L^*_{Max}	0.93	Maximum L^*
L^*_{95th}	0.87	95th percentile L^*
SIPk	0.66	Ref 14
SQRI	0.63	Ref 16

Although the Pearson correlation of maximum L^* was higher than that of 95th percentile L^*, the former could be unreliable, when the results are dependent on outlier values. Hence, it was decided to adopt the latter as a lightness statistical measure. Figure 9 contains a plot between normalised 95th percentile L^* and MOS of the Data 3 and shows a linear fit.

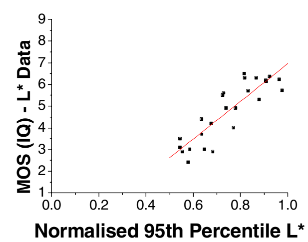

Figure 9. The relationship between normalized 95th percentile L Prediction and MOS of data set 3 (Lightness): r = 0.87*

Single image perceived (SIPk) contrast[14] is a function of image lightness standard deviation, chroma standard deviation and the standard deviation of high-frequency L* image information. The Pearson correlation between SIPk and MOS was 0.66. This low correlation was caused by the standard deviation terms being dependent on image contents.

The performance of square root integral (SQRI) was somewhat disappointing. SQRI is the square root integration of multiplication between display MTF[26,27] and the contrast sensitivity function (CSF).[16] Since mean luminance (cd/m^2) of an image, which is image-dependent as well as mean lightness, is used as an input to CSF, the output of SQRI should be also image dependent. This is why the performance of SQRI is similar to that of mean L^*.

Combined Image Colour-Quality

The three main effects (memory colour reproduction ratio in a region of interest, mean chroma, and 95th percentile lightness) in image colour-quality were modelled in the previous sections. The three models can be combined to form a single image colour-quality model (IQ_{CQ}). Only colour-related matters were considered in this model. It is possible using the model to make image colour-quality decisions about colour images and a comparison between the images reproduced on the same media. It is also important to note that this model can link subjective scales of image quality with objective measures such as image colour statistical measures. Equation 5 and Figure 10 show the image colour-quality model and its comparison with the corresponding MOS.

$$IQ_{CQ} = [a \quad b \quad c \quad d \quad e \quad f \quad g \quad h] \times \begin{bmatrix} M \\ C \\ L \\ M \times C \\ M \times L \\ C \times L \\ M \times C \times L \\ 1 \end{bmatrix} \tag{5}$$

where M is Ln(MCRR), C is mean C^*_{ab} and L is 95th percentile L^*. The coefficient values of a to h are listed in Table 6.

Table 6. Coefficients of the IQ_{CQ}

Coefficients	Value
a	8.85
b	-0.47
c	-9.37
d	-0.63
e	-10.04
f	0.62
g	0.84
h	11.57

The Pearson correlation between the model prediction and MOS values was 0.88 and the coefficient of variation (CV) was 18. Those values from the observer accuracy between the MOS and individual values can be directly compared with the model's predictions. The model accuracy outperformed the observer accuracy, which showed a Pearson correlation of 0.82 and a CV of 26. The larger Pearson correlation and the lower CV represent better performance.

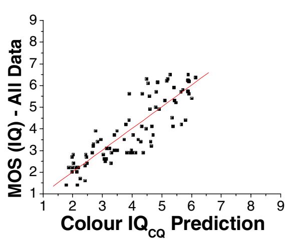

Figure 10. *The relationship between IQ_{CQ} Prediction and MOS of all three data sets: r = 0.88*

Summary

A single image colour-quality model was developed based on three components: memory colour reproduction ratio in a region of interest, mean chroma, and 95th percentile lightness of an image. It is capable of appraising a single image with a good correlation without the presence of an original image. It is also noteworthy that subjective image quality was linked to objective values such as image statistical measures.

References

[1] Y. J. Kim, M. R. Luo, P Rhodes, W. Choe, S. Lee, and C. Kim, Affective attributes in image quality of a mobile LCD (In Press), Proc. Third European Conference on Colour in Graphics, Imaging, and Vision, (2006).

[2] S. R. Fernandez, M. D. Fairchild, and K. Braun, Analysis of observer and cultural variability while generating "preferred" color reproductions of pictorial images. Journal of Imaging Science and Technology, 49(1), 96 (2005).

[3] M. D. Fairchild, Measuring and modelling image quality, Chester F. Carlson Center for Imaging Science industrial associates program, Rochester (1999).

[4] R. W. G. Hunt, Measuring colour, 3rd ed. Fountain Press, London, England (1998).

[5] M. R. Luo, G. Cui, and B. Rigg, The development of the CIE 2000 colour-difference formula: CIEDE2000, Color Research and Application, 26, 340 (2001).

[6] S. Westland and C. Ripamonti, Computational colour science, John Wiley & Sons, Ltd, Chichester, England (2004).

[7] X. Zhang and B. A. Wandell, A spatial extension of CIELAB for digital color image reproduction, SID 96 Digest (1996).

[8] T. J. W. M. Janssen, Computational image quality, PhD thesis, Eindhoven, IPO (1999).

[9] S. N. Yendrikhovskij, F. J. J. Blommaert, and H. de Ridder, Colour reproduction and the naturalness constraint. Color Research and Application, 24, 53 (1999).

[10] R. W. G. Hunt, The reproduction of colour, 6th ed., John Wiley & Sons Ltd, Chichester, England (2004).

[11] D. L. Post and C. S. Calhoun, An evaluation of methods for producing desired colors on CRT monitors, Color Research and Application. 14, 172 (1989).

[12] J. Morovic and Y. Wang, Influence of test image choice on experimental results, Proc. IS&T/SID Eleventh Color Imaging Conference, 143 (2003).

[13] C. J. Bartleson, Visual measurements. In Bartleson, C. J. & Grum, F. (Ed), Optical Radiation Measurements, 5, Orlando, FL: Academic Press (1984).

[14] A. J. Calabria and M. D. Fairchild, Perceived Image Contrast and Observer Preerence II. Empirical modelling of perceived image contrast and observer preference data, Journal of Imaging Science and Technology, 47(6), 479 (2003).

[15] http://www.samsung.com/PressCenter/PressRelease

[16] P. G. J. Barten, Contrast sensitivity of the human eye and its effects on image quality, SPIE-The international society for optical engineering, Washington (1999).

[17] C. Boust, F. Cittadini, M. B. Chouikha, H. Brettel, F. Vienot, S. Berche, and G. Alquie, Does an expert use memory colors to adjust images? Proc. IS&T/ SID Twelfth Color Imaging Conference, 347 (2004).

[18] M. D. Fairchild, Color appearance models 2nd ed., John Wiley & Sons Ltd, Chichester, England (2005).

[19] P. G. Engeldrum, Psychometric scaling: a toolkit for imaging systems development, Imcotek Press, Winchester (2000).

[20] B. W. Keelan, Handbook of image quality: characterization and prediction, Marcel Dekker, New York (1994).

[21] S. Susstrunk, R. Buckley, and S. Swen, Standard RGB color spaces, Proc. IS&T/SID Seventh Color Imaging Conference, 127 (1999).

[22] A. Sharma, Understanding color management, Thomson Delmar Learning, New York (2004).

[23] C. Sano, T. Song, and M. R. Luo, Colour differences for complex images, Proc. IS&T/SID Eleventh Color Imaging Conference, 121 (2003).
[24] B. Oicherman and A. Johnson, A method to derive local colour differences tolerances in complex images, Proc.Tenth Congress of the International Colour Association, 317 (2005).
[25] T. Tarczali, D, Park, P. Bodrogi, C. Y. Kim, Long-term memory colors of Korean and Hungarian observers, Color Research and Application, 31, 176 (2006).
[26] J. C. Danity and R. Shaw, Image science: principles, analysis and evaluation of photographic –type processes, Academic Press, London. (1974)
[27] S. Triantaphillidou and R. E. Jacobson, Measurements of the modulation transfer function of image displays, Journal of Imaging Science and Technology, 48(1), 58 (2004).
[28] N. R. Draper and H. Smith, Applied regression analysis, John Wiley & Sons, Inc., New York (1966).

Author Biography

Youn Jin Kim received his BSc degree in Physics from Daejin University (Korea) in 2003 and MSc in Imaging Science from University of Derby (UK) in 2004. He has worked as a PhD student at the Colour & Imaging Group at the University of Leeds (UK) since 2004. His research interests are in colour image quality and enhancement.