Predicting Perceived Colorfulness, Contrast, Naturalness and Quality for Color Images Reproduced on a Large Display

Seo Young Choi[1,2], M. Ronnier Luo[1], Michael R. Pointer[1] and Peter A. Rhodes[1]
[1]*Department of Color Science, University of Leeds, Leeds, UK;* [2]*Samsung SDI, Cheonan, Republic of Korea*

Abstract

Assessment of display quality is important for display manufacturers. This should be evaluated with two aspects – measuring physical performance and quantifying human perception of image appearance attributes. The present work particularly focuses on the development of an image quality model for the latter case. A developed model included three components (colorfulness, contrast and naturalness) that were all modeled using parameters derived from CAM02-UCS color-appearance attributes. In the combination of the three components, psychophysical relationships between each of them and image quality were exploited. The results exhibited that colorfulness and naturalness were critical factors in the judgment of image quality. A computational procedure was also given from a pair of color images to predict the image quality.

Introduction

The assessment of image quality is vital to the development of new displays. For example, the performance of a new model needs to be assessed against an existing model, or the performances of different types of display, e.g. LCD and OLED, need to be compared. Physical display characteristics such as color gamut can be evaluated by direct measurement, whereas image appearance attributes (image quality, naturalness, etc) are typically compared by a small number of experienced engineers. To overcome the subjective nature of these latter assessments, there has been a demand for imaging scientists to develop models capable of quantifying image appearance. Such models can provide quantitative results that can guide those working in display industries. The work described here has attempted to develop such a model.

For different applications, independent approaches have been made to develop computational metrics to quantify those attributes thought to be most relevant to the perception of image appearance. One approach has been to derive accurate functions describing the impairment between two versions of an image, i.e. the 'original' version against its 'manipulated' version. By taking into account the effect of human contrast sensitivity on the spatial structure of an image, various image difference metrics were constructed using color-difference formula based on CIELAB [1,2], CIECAM02 and CAM02-UCS [3,4], and CIEDE2000 [5]. These metrics have been used to evaluate image-compression tools, however what is calculated is actually the visual image-difference rather than the difference in perceptual image appearance. To overcome this limitation, Keelan [6] and Topfler [7] expressed changes in image quality and image preference in terms of a JND (just noticeable difference) in image quality, rather than image difference.

Another approach has been to establish subjective rules that are applied to the perception of image appearance, and then use them to model image-appearance attributes. Attributes which include sharpness, contrast and quality were assessed by a panel of observers for image stimuli that had been altered in terms of lightness, chroma, noise-spread, etc. This approach is time consuming, but might provide useful insight to improving image quality. A multi-attribute image-quality model was devised, amongst others, by Bartleson [8] who modeled image quality as a Minkowski sum of sharpness and the complement of graininess. Also, Engeldrum [9] used Gaussian and logistic functions to predict the non-linear variations in image-quality caused by changes in image colorfulness.

There is considerable more research on the determination of image-appearance attributes than those introduced here. It is, however, somewhat difficult to apply those results directly to the evaluation of the appearance of images seen on displays because the previous studies tended to consider other applications such as printed images. The present study therefore introduces four image-appearance functions that predict the ratios of perceived colorfulness, contrast, naturalness and quality of color images presented on two displays. The models include the typical applications to compare the appearance of color images viewed on two displays.

Psychophysical Experimental Setting

Eight test images (given in Appendix I) – including five natural scenes, one fruit and two portraits – were displayed on a 42-inch Samsung Plasma Display Panel (PDP) having 1024 × 768 pixel resolution in a dark room. The *RGB* values of each pixel for each test image were transformed to *XYZ* tristimulus values via the earlier developed characterization model for the PDP [10]. The *XYZ* values were then converted into CIECAM02 lightness (J) and chroma (C) appearance values. These attributes were manipulated in 17 different ways to produce the test images. Five transformations were made to the J spatial frequency channel to alter sharpness in the test images [11]. Table 1 outlines the rendering effects on the test images arising from all 22 methods. The eight (original) test images and their 22 derived manipulations were rated by 12 – 14 observers using a 9-point qualitative category scale in terms of contrast, sharpness, colorfulness, visual information content, naturalness and image quality. The category-scaling data in the psychophysical experiment were transformed to equal-interval scale values using Case V of Thurstone's Law. Overall, 79,200 estimations were made.

The scale values of the eight test images fell within 95% confidence intervals from the mean of the eight test images for most image manipulations and for each of the six attributes. Hence, the six attributes could be scaled in an image-independent manner. Through multiple regression and factor analyses, colorfulness, contrast and naturalness were determined to be significant factors affecting image quality. A strong linear relationship was found between image quality and naturalness, whereas non-linear relationships were revealed between image quality and each of colorfulness and contrast. In the following sections, the concepts and computational procedures applied to model colorfulness, contrast and naturalness in this work are introduced. By considering the relationships that exist between image quality and each of these three attributes, an image-quality model is then constructed.

The Concept behind Modeling

As explained above, when a new or improved type of display is developed, the quality of its images must be evaluated by comparison with the image quality of existing displays. The main question to be asked is: 'Are images on the new display perceived to have a higher quality than those on the existing display?' Also, the difference in image quality needs to be quantified. The new display could, for example, reproduce images that appear to be more blurred, lighter or more colorful. To simulate these diverse but real scenarios, relative scale values obtained from the psychophysical experiment were calculated. Table 2 summarizes the seven sets of denominators (for images seen on the existing display) and numerators (for images seen on the new display) used to compute the ratios of scale values. The image-appearance comparisons between the existing and new displays that can be made by these seven sets are also described in the rightmost column. Image characteristics were computed using optimized parameters derived from lightness (J'), colorfulness (M') and pixel-based color difference ($\Delta E_{CAM02\text{-}UCS}$) so as to account for these properties [14]. These image characteristics were also calculated in a relative sense according to the suggested seven sets of

Table 1. Rendering effects on test images arising from the 22 different methods.

	Rendering effects on test images	Functions used to manipulate images
Lightness (J)	Decreasing lightness linearly	4 × linear functions
	Increasing lightness contrast	3 × sigmoid functions
	Decreasing lightness contrast	3 × inverse-sigmoid functions
	Dark pixel into lighter and light pixel into darker (Lightness inversion)	1 × local colour-correction method [12]
Chroma (C)	Decreasing chroma linearly	4 × linear functions
	Increasing chroma contrast	1 × sigmoid function
	Decreasing chroma contrast	1 × inverse-sigmoid function
Sharpness (spatial frequency of J)	Increasing sharpness	4 × high-frequency emphasis filters 1 × using information from Barten's contrast sensitivity function [13]

Table 2. Summary of numerators and denominators used to calculate the ratios of all image characteristics and scale values.

***Ratio = Numerator / Denominator* for all image characteristics and scale values**		
Denominator (images on the existing display)	***Numerator*** (images on the new display)	How images on the new display appear compared to those on the existing display
(1) Original Image	22 manipulated images	darker, having different contrast, sharper, less/more colorful
(2) and (3) Images darkened by 10% and 20% from the original	Original image and 22 manipulated images excluding the denominator used by images (2) and (3)	darker, lighter, having different contrast, less colorful, lighter and more colorful, darker and less colorful, lighter and sharper
(4) and (5) Images having reduced chroma by 10% and 20% from the original	Original image and 22 manipulated images excluding the denominator used by images (4) and (5)	having different contrast and colorfulness, sharper
(6) and (7) Images sharpened from the original	Original image and 22 manipulated images excluding the denominator used by images (6) and (7)	having different contrast and colorfulness, sharper, blurred

combinations in Table 2.

The models predicting colorfulness, contrast, naturalness and image quality were developed so that the following requirements could be satisfied: they were independent of image content and they were capable of predicting the inter-display ratios of perceived colorfulness, contrast, naturalness and quality of images. From the eight test images (see Appendix I), a subset of six images (Harbor, Pier, Seashore, Fruits, Kids and Adults) was used to derive the models by considering all seven sets of denominators and numerators mentioned in Table 2. The remaining two images (Park and Sheep) were used to evaluate the performance of the new models.

Specific Approach

Image characteristics responsible for variations in colorfulness, contrast and naturalness due to changes in lightness, chroma and sharpness of images were determined. Functions were then derived to relate these characteristics to the scale values obtained from the psychophysical experiment. It was found that as the images looked lighter or more chromatic, perceptual image colorfulness grew. Increased colorfulness caused by either higher brightness or chroma can be better predicted by CAM02-UCS colorfulness (M') rather than by chroma (C'). This is because C' is able to explain perceptual colorfulness change arising from just chroma variations. Thus, CAM02-UCS M' was chosen to compute a correlate of image colorfulness.

Image lightness, sharpness and chroma all influenced perceived image contrast. A significant increase in contrast was seen in sharpened images and those where dark areas were darker and light areas were lighter than in the original images. Conversely, blurred images and those having decreased chroma and lightness-contrast appeared to have noticeably less contrast. To address these visual effects, a correlate of image contrast was devised based on pixel-based color differences ($\Delta E_{CAM02\text{-}UCS}$) at three different image resolutions: 1024 × 768 (equivalent to 32 cpd [cycles per degree] of visual angle viewed at distance of 2 m), 512 × 384 (16 cpd) and 256 × 192 (8 cpd). The different resolutions were obtained by sub-sampling the original image and then displaying them at the same physical size as the original.

The visual factors found to be key to the perception of image naturalness in the psychophysical experiment were: image colorfulness, image sharpness, the reproduction of shadow-detail and a lack of washed-out appearance due to over-decreased lightness-contrast. These were modeled using the parameters derived from lightness (J'), colorfulness (M') and pixel-based color difference ($\Delta E_{CAM02\text{-}UCS}$) at a size of 128 × 96 pixels (4 cpd).

Image Quality Modeling

Image quality Models

An image-quality model was made by combining the three image-appearance attributes (colorfulness, contrast and naturalness) determined using the three developed functions. These attributes were considered to be independent variables that could explain variations in the dependent variable, image quality. Colorfulness and contrast were shown to have non-linear relationships with image quality, whereas naturalness was linearly related to image quality. To reflect the psychophysical relationships between image quality and its constituent attributes, two types of image quality model were considered. The first type was a first-order model. The second type was second-order no-interaction model where the expected change in image quality due to the change of one independent variable is non-linear. The term 'no-interaction' means that the expected change in image quality for a unit increase in one independent variable does not depend on any other independent variable. The reason for selecting no-interaction terms rather than cross-interaction terms was that there was only limited correlation between any two of colorfulness, contrast and naturalness.

The following series of equations, (1) – (5), describes the models developed to determine the relationships between perceived quality, contrast, colorfulness and naturalness for images presented on two displays. It is assumed that Display 1 was an existing model, and that a new contrast-enhancement algorithm was applied to Display 2. The goal is to predict the change in perceived colorfulness, contrast, naturalness and quality of images viewed on Display 2 compared with those on Display 1. The resolution of both displays was 1024 × 768 pixels and the viewing distance was 2 m with a dark surround. The same test image is reproduced on both displays: the image on Display 1 is named 'Image 1' and the same test image on Display 2 is called 'Image 2'. This assumption will be also applied to Figure 1 in the following section (*Computational Procedures*).

Two different image-quality models having first-order terms (IQ 1) and second-order no-interaction terms (IQ 2) are described in eqns. (1) and (2). The derived models for contrast, colorfulness and naturalness are also expressed in eqns. (3), (4) and (5), respectively. Table 3 introduces the optimized weights used by the two models in eqns. (1) and (2) for the individual three attributes (independent variables). These attributes were all key factors affecting image quality; however all their possible combinations were also utilized in order to ascertain which factors are most critical in explaining variations in image quality. Naturalness had a linear relationship with image quality, so this alone was excluded from the non-linear image-quality model in Table 3.

$$\text{IQ 1} = w_1 \cdot \text{contrast} + w_2 \cdot \text{colorfulness} + w_3 \cdot \text{naturalness} + w_4 \qquad (1)$$

$$\text{IQ 2} = w_5 \cdot \text{contrast} + w_6 \cdot \text{colorfulness} + w_7 \cdot \text{naturalness} + w_8 \cdot \text{contrast}^2 + w_9 \cdot \text{colorfulness}^2 + w_{10} \cdot \text{naturalness}^2 + w_{11} \qquad (2)$$

$$\text{Contrast} = 0.24 \times \frac{PBCD_{image2\,(32cpd)}}{PBCD_{image1\,(32cpd)}} + 0.37 \times \frac{PBCD_{image2\,(16cpd)}}{PBCD_{image1\,(16cpd)}} + 0.37 \times \frac{PBCD_{image2\,(8cpd)}}{PBCD_{image1\,(8cpd)}} \qquad (3)$$

where $PBCD_{ij}$ is the pixel-based color difference calculated from image i at resolution j.

$$\text{Colorfulness} = 1.20 \times \left(\frac{M'_{image2}}{M'_{image1}} \right) - 0.20 \qquad (4)$$

where M'_i is the average colorfulness over all pixels in image i.

Table 3. Summary of weights for the independent variables in each of the two image-quality models (eqns. (1) and (2)).

Independent variables	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}	w_{11}
- contrast - colorfulness - naturalness	0.20	0.40	0.77	0.38	1.19	-0.46	-0.56	-0.44	0.40	0.81	0.05
- contrast - colorfulness	0.21	0.52		0.20	2.01	0.74		-0.81	-0.13		0.85
- colorfulness - naturalness		0.48	0.77	0.26		-0.32	-0.13		0.39	0.57	0.48
- contrast - naturalness	0.37		0.88	0.24	1.82		-0.74	-0.66		0.93	0.33
- contrast	0.44			0.50	2.90			-1.12			0.80
- colorfulness		0.61		0.32		0.91			-0.15		0.17
- naturalness			0.94	0.07							

$$\text{Naturalness} = 0.83 \times RSD + 0.99 \times IC + 0.34 \times IS - 1.18 \qquad (5)$$

RSD (Reproduction of Shadow-Detail) =

$$exp\left[2.17 - 2.16 \times \left(\frac{count(J' < 30)_{image2}}{count(J' < 30)_{image1}}\right) - 1.63 \times ln\left(\frac{count(J' < 30)_{image1}}{count(J' < 30)_{image2}}\right)\right]$$

IC (Image Colorfulness) =

$$exp\left[2.91 - 2.89 \times \left(\frac{M'_{image1}}{M'_{image2}}\right) - 2.68 \times ln\left(\frac{M'_{image2}}{M'_{image1}}\right)\right]$$

IS (Image Sharpness) =

$$exp\left[18.29 - 18.30 \times \left(\frac{PBCD_{image1(4cpd)}}{PBCD_{image2(4cpd)}}\right) - 17.70 \times ln\left(\frac{PBCD_{image2(4cpd)}}{PBCD_{image1(4cpd)}}\right)\right]$$

where $count(J'<30)_i$ is the number of pixels having J' less than 30 in image i, M'_i is the average colorfulness over all pixels in image i, and $PBCD_{ij}$ is the pixel-based color difference calculated from image i at resolution j.

Computational Procedures

Figure 1 describes the computational procedures for determining the changes in perceived colorfulness, contrast, naturalness and quality of images (e.g. Image 2) viewed on Display 2 compared with those (e.g. Image 1) on Display 1.

The following details the computation procedures for steps (1) – (13) in Figure 1.

(1) Transform *RGB* values to *XYZ* values for each pixel.
(2) Transform *XYZ* values to CAM02-UCS, J', M', a'_M and b'_M values for each pixel.
(3) Count how many image pixels have J' less than 30.
(4) Calculate the ratio of the number of pixels having J' less than 30 in 'Image 2' to that in 'Image 1'.
(5) Calculate the average M' over all pixels in an image.
(6) Calculate the ratio of the average M' (over all pixels) of 'Image 2' to that of 'Image 1'.
(7) Create four images at four different resolutions: 1024×768 (32 cpd), 512×384 (16 cpd), 256×192 (8 cpd) and 128×96 (4 cpd). The images with 16 cpd, 8 cpd and 4 cpd resolutions are produced by averaging blocks of 2×2, 4×4 and 8×8 pixels from the original image to form new pixels.
(8) Calculate the pixel-based color difference (*PBCD*) by averaging the *color difference between the centre pixel and its surrounding pixels in each pixel* over all pixels in each of the four images at four different resolutions.
(9) Calculate the *PBCD* ratio by dividing the *PBCD* of 'Image 2' by that of 'Image 1'.
(10) Determine the perceived image-contrast ratio of 'Image 2' to 'Image 1' using the computation result from Step 9 and the image contrast model in Eq. (3).
(11) Determine the perceived image-colorfulness ratio of 'Image 2' to 'Image 1' using the computation result from Step 6 and the image colorfulness model in Eq. (4).
(12) Determine the perceived image-naturalness ratio of 'Image 2' to 'Image 1' using the computation result from steps 4, 6 and 9, and the image naturalness model in Eq. (5).
(13) Determine the perceived image-quality ratio of 'Image 2' to 'Image 1' using the computation result from steps 10 – 13 and the image quality models in eqns. (1) and (2).

Performance of Image Quality Models

The coefficient of variation (CV) was used as a measure of the agreement between the predicted and judged image-quality data. Table 4 lists CV values for the image-quality models that have different independent variables and use different types of function. Among the models having the same number of independent variables, the smallest CV values are written in bold type.

The main tendencies for both training and test images are summarized as follows.

- As the number of independent variables increases, CV values become smaller for all image quality models. This indicates better model performance.

- A mean CV value (across all observers) of 22 represented the inter-observer agreement of image quality. All the calculated CV values in Table 4 are less than this value, indicating that the typical observer uncertainty is larger than the predictive errors of the developed image-quality models.
- For the models derived using three attributes, variations in image quality can be best predicted by the second-order image-quality model. There are, however, negligible CV value differences between the two types of image quality model.
- For the models derived using two attributes, the image quality predicted using colorfulness and naturalness shows better agreement with judged image quality compared to those models using other combinations of two attributes. Additionally, there are almost no CV value differences between the two types of image quality model when they were developed using colorfulness and naturalness.
- For the models derived using a single attribute, naturalness is better when predicting the image quality of the training images, while colorfulness or contrast is better for the test images. On the other hand, the CV difference is largest between the test and training images when image quality was predicted using naturalness. In summary, changes in image quality can be equally explained by either colorfulness, contrast or naturalness.

In conclusion, either of the two types of model is both sufficient if three attributes (colorfulness, contrast and naturalness) or two attributes (colorfulness and naturalness) are used to determine image quality. Figures 2(a) and 2(b) show comparisons between the experimental quality data and the predictions made by the first-order model using the three attributes, for the testing and training images, respectively. The data points are shown using three different symbols for the three image-manipulation domains. A 45° line representing a perfect agreement between the experimental and predicted data is also plotted. As all data points are located close to the 45° line, it can be said that this model successfully predicts image quality variations arising either from changes in image lightness, chroma and sharpness.

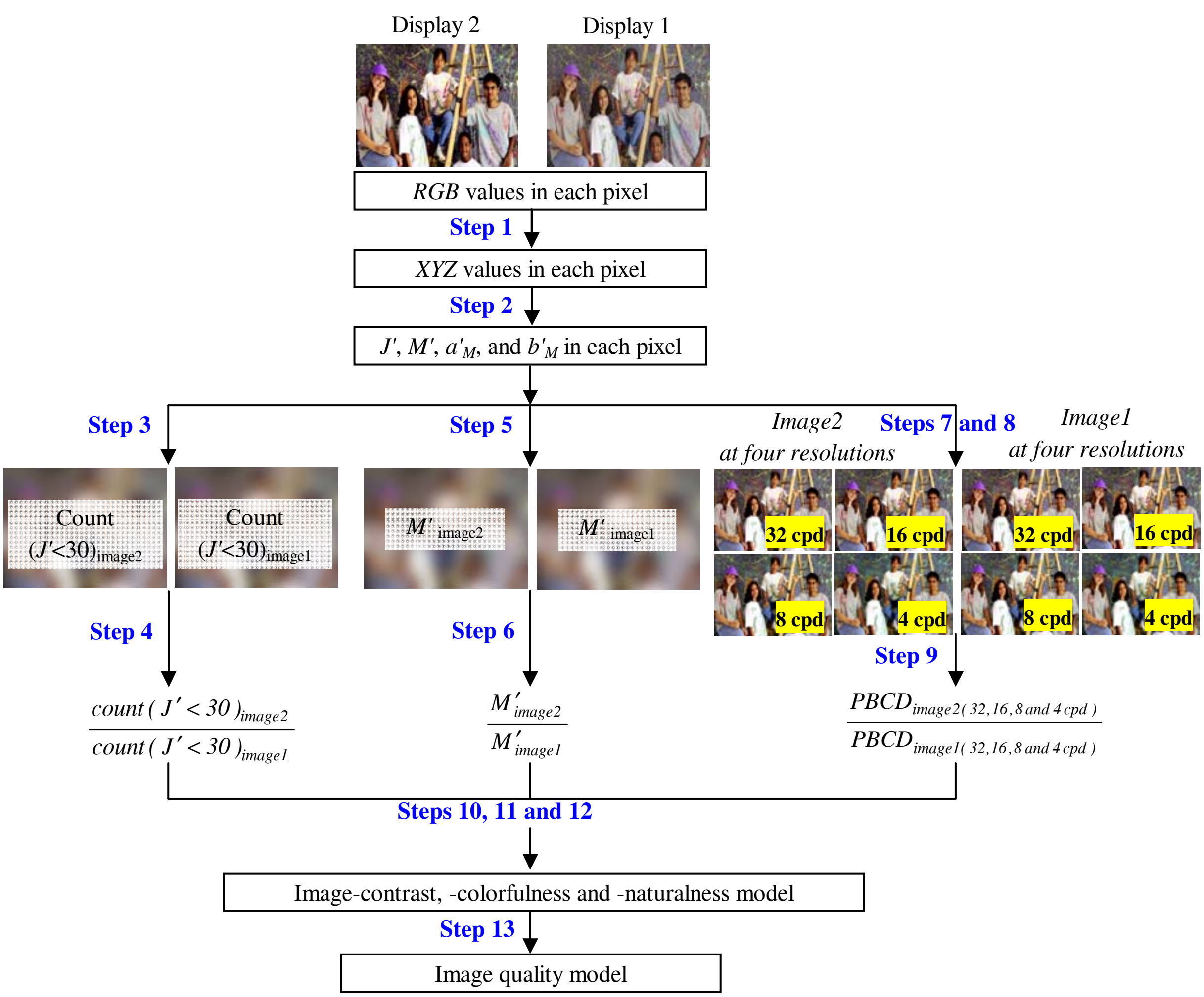
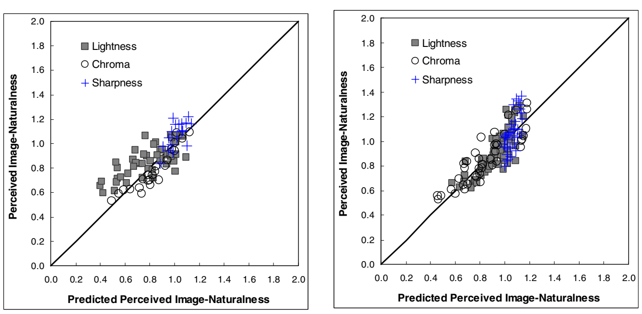


Figure 1. *Overall workflow to determine relationships of perceived colorfulness, contrast, naturalness and quality of images presented on two displays.*

Table 4. Summary of the calculated CV values between predicted and experimental image-quality data.

		Training Images (1)	Training Images (2)	Test Images (1)	Test Images (2)
Three attributes		11	**10**	14	12
Two attributes	**contrast colorfulness**	15	15	15	**14**
	colorfulness naturalness	12	**11**	14	14
	contrast naturalness	13	**13**	18	14
One attribute	**contrast**	18	**17**	18	**17**
	colorfulness	18	18	**16**	**16**
	naturalness	**15**		20	

[Note: (1) is the image-quality model having first-order terms and (2) is the second-order image-quality model having no-interaction terms.]



(a) (b)

Figure 2. *Plot of the experimental image-quality data against the predicted image-quality data for (a) the two testing images and (b) the six training images.*

Conclusions

Four image-appearance functions were developed to predict the perceived colorfulness, contrast, naturalness and quality of color images. These functions are applicable to compare appearances of two displayed images in terms of ratios. The average CAM02-UCS M' values of all pixels in an image was chosen as a correlate of image colorfulness. Pixel-based color differences ($\Delta E_{CAM02\text{-}UCS}$) at different image resolutions were selected as a correlate of image contrast. Four visual factors were modeled using parameters derived from CAM02-UCS J', M' and $\Delta E_{CAM02\text{-}UCS}$ for image naturalness. Finally, two types of mage-quality model (1st-order and 2nd-order) were constructed by combining colorfulness, contrast and naturalness functions.

For the models derived using two attributes from the three, those involving colorfulness and naturalness exhibited somewhat better performance than those using other attribute pairs. It might be concluded that these two attributes govern the overall image quality: naturalness, based on an observer's memory of the real world, and colorfulness, evoking a pleasant feeling.

Disagreement between observers' judgments was larger than the predictive errors caused by all the developed models. The two types of model with different combinations of independent variables are, therefore, all suitable for use in determining image quality. Due to its simplicity, the linear regression model with two attributes (colorfulness and naturalness) is recommended for practical applications.

References

[1] X.M. Zhang and B.A. Wandell, "A Spatial Extension to CIELAB for Digital Color Image Reproduction," SID Symp. Tech. Digest, 27, 731 (1996).

[2] E.W. Jin, X.F. Feng and J. Newell, The Development of a Color Visual Difference Model (CVDM), Proc. PICS, pg. 154. (1998).

[3] D.G. Lee, M.R. Luo and G. Cui, Evaluation for Color Differences for Images, IS&T's 4th CGIV Conference, Terrassa, Spain, pg. 12. (2008).

[4] Z.H. Wang and M.R. Luo, Experimental Filters for Estimating IQ, IS&T's 4th CGIV Conference, Terrassa, Spain, pg. 112. (2008).

[5] G.M. Johnson and M.D. Fairchild, "A Top Down Description of S-CIELAB and CIEDE2000," Col. Res. Appl. 28, 425 (2003).

[6] B.W. Keelan, Predicting Multivariate Image Quality from Individual Perceptual Attributes, Proc. PICS, pg. 82. (2002).

[7] K. Topfler, B.W. Keelan, S. O'Dell and R. Cookingham, Preference in Image Quality Modeling, Proc. PICS, pg. 60. (2002). [8] C.J. Bartleson, "The Combined Influence of Sharpness and Graininess on the Quality of Color Prints," Jour. Photog. Science 30, 33 (1982).

[9] P.G. Engeldrum, Extending Image Quality Models, Proc. PICS, pg. 65 (2002).

[10] S.Y. Choi, M.R. Luo, P.A. Rhodes, E.G. Heo and I.S. Choi, "Colorimetric Characterization Model for Plasma Display Panel," Jour. Imaging Sci. Technol. 51, 337 (2007).

[11] S.Y. Choi, M.R. Luo, M.R. Pointer and P.A. Rhodes, "Investigation of Large Display Color Image Appearance I: Important Factors Affecting Perceived Quality," Jour. Imaging Sci. Technol. 52, 040904-1 (2008).

[12] N. Moroney, Local Color Correction Using Non-Linear Masking, IS&T/SID's 8th Color Imaging Conference, Scottsdale, Arizona, pg. 108. (2000).

[13] P. Barten, "Evaluation of Subjective Image Quality with the Square-Root Integral Method," Jour. Opt. Soc. Am. 7, 2024 (1990).

[14] M.R. Luo, G. Cui and C.J. Li, "Uniform Color Spaces Based on CIECAM02 Color Appearance Models," Col. Res. Appl. 31, 320 (2006).

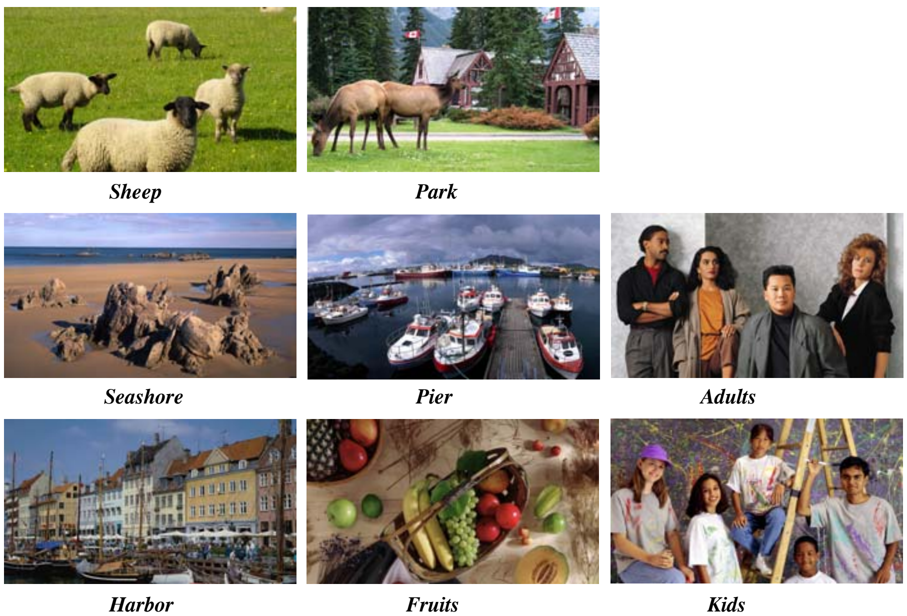
Author Biography

SEO YOUNG CHOI received her PhD degree in 2008 from the Dept. of Color Science at the University of Leeds. Her current interests include the development of new types of display and improving digital-signal workflow.

M. RONNIER LUO is Professor of Color and Imaging Science at the Department of Color Science at the University of Leeds. He has over 300 publications in the field of color and imaging science. He is actively involved with the Commission International de l'Éclairage (CIE) as the Director of Division 1: Vision & Color. He is also a Fellow of the Society of Dyers and Colorists and the Society for Imaging Science and Technology.

Dr. MICHAEL R. POINTER received his PhD from Imperial College, London and then worked in the Research Division of Kodak Limited on fundamental issues of color science applied to the photographic system. After a period at the National Physical Laboratory, he is now a Visiting Professor at the University of Leeds as well as working as a consultant scientist. He has authored over 70 scientific papers, is a Fellow of The Royal Photographic Society and the Institute of Physics, and Secretary of CIE Division 1 (Vision & Color).

Dr. PETER A. RHODES is a Senior Research Fellow at the University of Leeds, and was awarded a PhD for his thesis "Computer Mediated Color Fidelity and Communication" from Loughborough University of Technology. He has produced almost 40 academic publications, including book chapters on color management and color notation systems. In addition to contributing towards the Department's MSc programme in Color and Imaging Science, he is actively engaged in a number of commercially-oriented R&D projects.



Appendix I. *Eight test images: two testing images (top row) and six training images (bottom two rows).*