An adaptive Bilateral Filter for Predicting Color Image Difference

Zhaohui Wang, Jon Yngve Hardeberg; Gjøvik University College; Gjøvik, Norway

Abstract

Color image difference metrics are of great importance in the field of color image reproduction. In this study, we introduce an adaptive bilateral filter for predicting color image difference. This filter is simple, employing two Gaussian smoothing filters in different domains, which avoids the loss of edge information when smoothing the image. However, the challenge is to select appropriate parameters to result in a better performance when applying for color image deference prediction. We propose a method to optimize the parameters, which are designed to be adaptive to the corresponding viewing conditions, and the quantity and homogeneity of information contained in an image. We have conducted psychophysical experiments to evaluate the performance of our approach. The experimental sample images are reproduced with variations in six image attributes: Lightness, Chroma, Hue, Compression, Noise, and Sharpness. The Pearson's correlation value between the predicted difference and the z-score of visual judgments was employed to evaluate the performance and compare it with that of s-CIELAB and iCAM.

Background

Theories of spatial characterization of the human visual system are of much current interest in the development of image difference metrics [1-4]. They all involve the conception that the human visual system is optimally designed to process the spatial information in images or complex scenes. The study [5] of the human visual system has shown that the human visual system is composed of spatial frequency channels. The light sensors of the human visual system, cones and rods, are sensitive to the spatial changes of stimuli. Both contrast sensitivity and color appearance vary as a function of the spatial pattern [6, 7]. Attempts to computationally assess color image difference have typically created models of human perception suitable for determining the discriminations introduced by spatial alteration, such as image compression, halftone reproduction, etc. On the other hand, the successful applications of color difference formulae, such as CIELAB 1976 color difference, CIE94, and CIEDE2000, have encouraged researchers to apply them also to image difference evaluation.

An important motivation of our work is the development of an image difference metric on various image reproduction tasks. Image difference may originate due to different image reproduction methods, such as the discriminations from chromatic and spatial modifications. Several studies [8-12] have measured the discriminations introduced by chromatic changes of the images alone. In this work, we study the general statistics over both spatial and chromatic image reproductions.

Spatial filtering was introduced into the color difference formula for measuring image reproduction errors, and later, replaced with the simulator of the human contrast sensitivity functions (CSFs) [2, 13, 14]. There are many models developed for simulating the CSFs. The model developed by Movshon and Kiorpes [15] was suggested [2] and also adopted by the CIE TC8-02 [16]. Generally, the spatial filters (or CSF models) are applied in the opponent color space to deduct the high frequency components in an image. The decrease in sensitivity at higher frequencies has been attributed to blurring because of the optical limitation of the eye and spatial summation in the human visual system [17]. Thus, a blurrier image is the output, in which the imperceptible information is attenuated, including, inevitably, high frequency edges. There is a broad consensus, however, that the human visual system is particularly sensitive to the edges in an image. Edge detection is believed to be necessary to distinguish objects from their background, and establish their shape and position. It has been proved to be a crucial early step in the process of scene analysis by the human visual system. To overcome the undesirable loss of edges whilst using the spatial filter, recent studies [3, 18] employed edge enhancement in the workflow for spatial localization.

Many image processing methods have been developed to smooth the image and keep the edges. Recently, Tomasi and Manduchi [19] described an alternative bilateral filter which extended the concept of Gaussian smoothing by weighting the filter coefficients with their corresponding relative pixel intensities. Two Gaussian filters are applied at a localized pixel neighborhood, one in the spatial domain (domain filter) and the other in the intensity domain (range filter). The result is a blurrier image than the original while preserving edges. However, the behavior of this filter is governed by a number of parameters which need to be selected with care for color image difference evaluation.

In this paper, we propose an adaptive bilateral filter for color image difference evaluation and design the parameters based on the spatial frequency and the quantity and the homogeneity of the information contained in a certain image. We describe a psychophysical experiment to validate its performance and compare it with other two models, sCIELAB and iCAM, which are both recognized as the human visual system based models. The testing images are reproduced in terms of both spatial and chromatic attributes. The evaluation is based on the Pearson's correlation value between the visual psychophysical judgments and the predicted difference.

Adaptive Bilateral Filter

The idea behind the bilateral filter is to combine domain and range filters together. Pixels in the neighborhood which are geometrically closer and photometrically more similar to the filtering centre will be weighted more. Given a color image *f(x)*, the bilateral filter [19] can be expressed as:

$$h(x) = k^{-1}(x)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\xi)c(\xi,x)s(f(\xi),f(x))d\xi ,$$

where $k(x)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}c(\xi,x)s(f(\xi),f(x))d\xi$,

and where the function $c(\xi,x)$ measures the geometric closeness between the neighborhood centre x and a nearby point ξ:

$$c(\xi,x)=e^{-\frac{(\xi-x)^2}{2\sigma_d^2}}$$

The function $s(\xi,x)$ measures the photometric similarity between the neighborhood centre x and a nearby point ξ:

$$s(\xi,x)=e^{-\frac{(f(\xi)-f(x))^2}{2\sigma_r^2}}$$

The behavior of this filter is controlled by two parameters. The geometric spread σ_d in the domain is determined by the desired amount of low-pass filtering. A large σ_d results in more blur effect since more neighbors are combined together. The photometric spread σ_r is used to achieve the desired amount of combination of similar pixel values. Pixels with values closer to each other than σ_r are mixed together.

In this paper, we propose that the domain spread σ_d is determined by the viewing conditions, which define the number of pixels per degree of viewing angle (ppd). Given an image whose width is n pixels corresponds to l meters of physical length. If the image is viewed from m meters away, the domain spread σ_d can be expressed as:

$$\sigma_d=\frac{n/2}{180/\pi\cdot\tan^{-1}(l/(2m))},$$

which constructs a direct relationship between the smoothness of the image and the viewing conditions. For example, when the viewing distance is kept constant, smaller images displaying on a certain screen will be blurred more and larger images on the same screen will be blurred less.

To determine the range spread σ_r, we propose to use the image entropy. Entropy [20] is defined with the probability of occurrence of a certain pixel value,

$$E=-\sum_i p_i\log(p_i),$$

where p_i refers to the histogram of the pixel intensity values of an image. High entropy value is associated with high variance in the pixel values of an image, and low entropy value indicates that the image is fairly uniform. Consequently, a uniform color patch will have an entropy value of zero. The range spread σ_r, is calculated using the image entropy by:

$$\sigma_r=K/E,$$

where constant K is used to rescale the image entropy into an optimized value and entropy E is larger than zero for images. It builds a direct relationship between the measurement of similarity (the function $s(\xi,x)$) and the "variance" of pixel values of an image, which, in turn, contributes to preserve edges perceptually.

Experimental Methods

Experiments were conducted in a dark room using a 21-inch LCD which was calibrated and characterized according to ISO 3664 [21].

Ten images were chosen which covered a wide range of natural scenes and artificial objects. The image state was set to sRGB color space in a resolution of 800x600 (96 pixels per inch) under D65. To understand the interrelations between spatial and chromatic effects in the color image discrimination, a set of reproduction methods were applied, including the manipulation in Lightness (L), Chroma (C), Hue (H), Compression (CO), Noise (N) and Sharpness (S). The transformations of Lightness, Chroma and Hue are using linear functions in CIELAB color space. The manipulations in Compression, Noise and Sharpness are using JPEG lossy compression, Gaussian random noise and unsharp-masking, respectively, in sRGB space.

Seven levels of transformation were applied to each manipulation method to prepare the testing images. The manipulations in Lightness, Chroma and Hue were in two directions, increase and decrease. The manipulation in Compression, Noise and Sharpness were in different ratios to produce images in seven levels. Image pairs were collected according to the color difference between the manipulated images and the original images, which are mainly in the range from the just noticeable difference [11, 16] to perceptible but acceptable [22] difference. Totally, 420 image pairs (10 images X 6 methods X 7 levels) were used in the experiment.

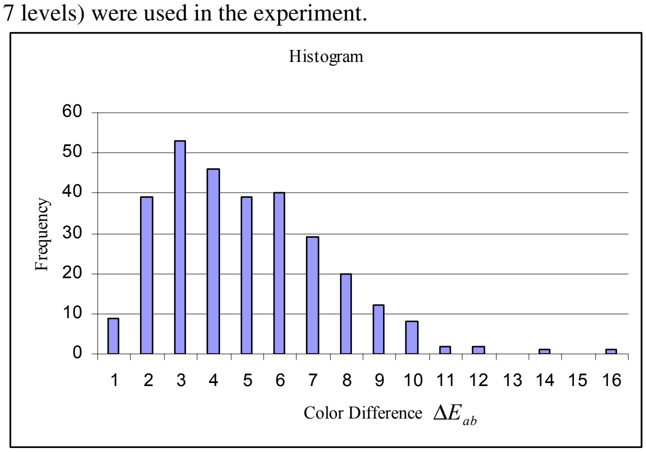

Figure 1 *The distribution of color difference of all testing images in terms of CIELAB*

Ten normal color vision observers (5 females and 5 males, all passed Ishihara test) participated in the experiments. Each observer was asked to evaluate the total image difference between an original image and a manipulated image using category judgment method. Seven categories (1 to 7) were used in the experiment as listed in Table 1. Each observer was given a training session. Experimental software was developed to present the image pairs in a random order.

Grade	Level of Difference
1	No Difference
2	Noticeable Difference
3	Moderate Difference
4	Acceptable Difference
5	Not Acceptable Difference
6	Very Large Difference
7	Extreme Difference

Table 1 Description of categories

Discussions and Results

Determination and extension of parameters

The investigation of CSF-based models in image difference metrics gave us the basic idea that the bilateral filter, as a smooth and edge preserving filter might be appropriate for evaluating the color image difference. The parameters are designed to be self-adaptive to the viewing condition and the image itself.

The smoothness degree, σ_d, is determined by the viewing condition which corresponds to the minimum noticeable amount of a change of the frequency component. We are certainly not arguing that the parameter does not reflect the spatial frequency response of chromatic channels. Over the general spatial range, the luminance CSF has band-pass characteristics and the spatial chromatic CSFs are low-pass. The chromatic CSFs are relative higher at low frequency and drop off earlier, as the frequency increasing, than luminance CSF. It is widely believed that the emphasis of luminance CSF is on the fine details and the chromatic CSFs give more information about large objects (or regions)[23] in images. The parameter σ_d used here is more emphasized on the perceptible details. The large objects are processed and weighted by the parameter σ_r using image entropy.

The image entropy is designed to determine the photometric parameter σ_r, which averages perceptually similar colors together. The constant K can be optimized by the experimental results (in this study a value of 100 is adopted). The image entropy is a useful measure to reflect the visual features of the image. Some studies [1, 24] argue that the image difference can be predicted from the large area or main objects in the image. The color appearance of single pixel can not be considered individually in an image, which is correlated with the neighbors. Thus, to consider the spatial response of similar color regions, the concept of image entropy may also be extended to the region based entropy and then weighted differently.

When applying CSF-based models in a workflow [4, 18], blurry edges and disturbance of color balance can be found if the three channels were filtered separately from one another in an opponent color space. To avoid this problem, the adaptive bilateral filter operated on the three channels, L*, a*, and b* of the CIELAB color space (another choice might be J, a, and b of the CIECAM02 color space) at once rather than filtering separately. Figure 2 presents an example, which compares the results of an image processed by the CSF model [15] and the adaptive bilateral filter. The results were converted to sRGB space for display.

Figure 2 *Comparison of images processed by the adaptive bilateral filter and the CSF Model*

Psychophysical experiments and observer accuracy

Different psychophysical methods have been studied for objective image difference evaluation. Forced-choice pair comparison or rank order method is often employed for small difference or small number of image pairs. Grey scale method was used in a previous research [24], by which the objective judgment can be used directly to evaluate the performance of difference metrics. However, observers may be confused by the lightness difference between grey samples with total image difference. In the present study, the category judgment scaling method is employed according to Bartleson [25] and Miller [26]. The seven categories are also somewhat matched to the seven transformation levels of each manipulation method.

It is important to know the reliability in psychophysical experiments, which is often represented by observer accuracy. In this study, the observer accuracy is investigated in repeatability and variation.

A number of 190 image pairs were presented twice randomly to observers to test observer's repeatability. For each observer, the repeatability was computed in terms of coefficient of variation (CV). The results are summarized in Table 2. The observer's repeatability is higher than that in previous research[24], which arrived at average of 17 with 95% significant level of 20.

Observer	1	2	3	4	5	Average
CV	18	15	18	17	14	17
Observer	6	7	8	9	10	
CV	12	18	17	18	23	

Table 2 The performance of observer's repeatability

For observer's variation, the mean categorical judgment of each stimulus was calculated by averaging visual results from all observers. The CV value for each observer was then calculated by comparing individual results for all the stimuli with the mean category values of the corresponding stimuli. The results are shown in Table 3.

Observer	1	2	3	4	5	Average
CV	25	23	22	36	23	28
Observer	6	7	8	9	10	
CV	39	35	34	21	19	

Table 3 The performance of observer's variation

Results

Totally, 4200 (10 observers X 420 image pairs including repeated image pairs) visual judgments were collected. Togerson's Law of Categorical Judgment was applied to analyze the results. The raw data were transformed into an interval scale where scores are based on the relative position of stimuli with respect to category boundaries.

In the literature, a few models have been proposed based on the human visual system, such as, sCIELAB[4] and iCAM[18]. One difference has been mentioned in the previous section is the adaptive bilateral filter operating on three channels of CIELAB space together rather than separately in three channels of an opponent color space.

Experimental results were used in this study to compare the performances between adaptive bilateral filter, sCIELAB, and iCAM in terms of Pearson's correlation value. The Pearson's correlation value indicates the degree of linear relationship between two variables and ranges from -1 to 1. The Pearson's correlation values were calculated between the average scale values of each image pair and the predicted results by each model for each image pair.

Using the adaptive bilateral filter, the experimental image pairs were filtered in the CIELAB color space and then the average pixelwise differences were calculated using the CIELAB color difference formula. The iCAM workflow was implemented according to Fairchild *et al.*[18]. The sCIELAB workflow was carried out using the procedure provided by the CIE TC8-02 report[16] in which the identical CSFs as used in iCAM were recommended.

Figure 3 shows the performances on each manipulation method by three models. The closer Pearson's correlation value is to +/- 1, the higher the performance. The error bars indicate 95% confidence interval (CI) which is calculated by 95% $CI = 1.96/\sqrt{2N} = 0.02$, where N represents the number of overall observations. In most cases, the performances of the adaptive bilateral filter and sCIELAB are quite similar, except the Pearson correlation of sCIELAB is slightly higher than that of adaptive bilateral filter by manipulation method of chroma; however, this is not the case by manipulation method of noise. In the manipulation methods of compression, lightness and noise, the performances of iCAM are lower than that of other two models.

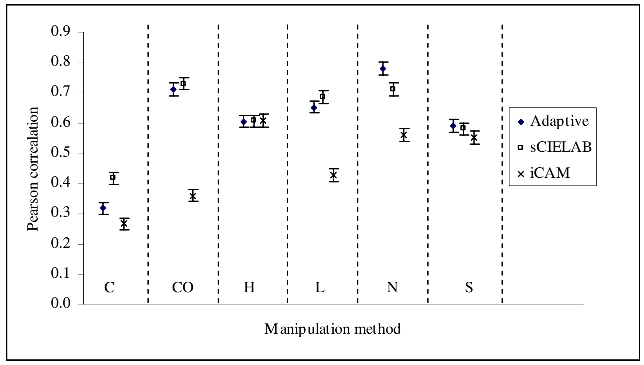

Figure 3 *The performances on different manipulation methods in terms of Pearson's correlation values*

Figure 4 shows the average color difference values of each model in different scale. It can be seen that the average values of the adaptive bilateral filter and iCAM result in high correlation coefficient. The correlation coefficient of sCIELAB is lower due to the higher average color difference value in scale 3 than that of scales 4 and 5.

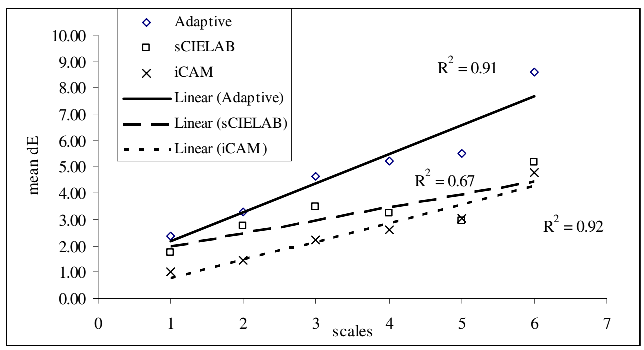

Figure 4 *The comparison of average in different scale values by different model*

Figure 5 compares the performance in terms of Pearson's correlation in each interval scale. It is also shown that the interval scale is not even. If comparing the Pearson values in Figure 3 and Figure 5, obviously, the Pearson correlation is lower when evaluating the performance of models in each scale, which might reveal that different manipulation methods can not be weighted equally.

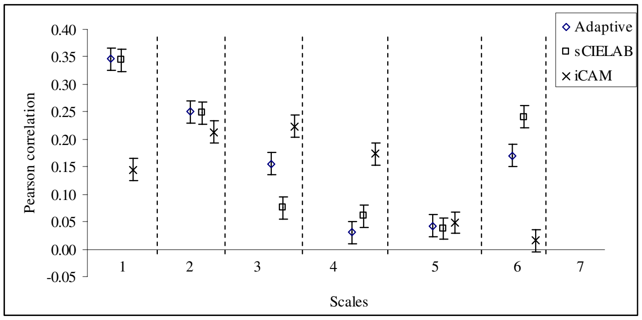

***Figure 5** The performances on different scale values in terms of Pearson's correlation values*

Summary

We have introduced a novel method for measuring the perceptual image difference. The optimized bilateral filter is adaptive to the corresponding viewing conditions and image entropy. A series of experiments were conducted using category judgment method. The experimental image pairs were manipulated in six image attributes, including both chromatic and spatial alterations. The Pearson's correlation values between the visual judgments and the predicted results by using the adaptive bilateral filter were analyzed. The performance was compared with that of sCIELAB and iCAM models. There are still some details need to be refined in future works, such as, using region based entropy instead of pixel entropy and understanding the differences between chromatic and spatial manipulation methods when evaluating the color image difference.

References

[1] G. Hong, M. R. Luo, and P. A. Rhodes, "A new algorithm for calculating perceived colour difference of images," *The Imaging Science Journal,* 54, pg. 86-91 (2006).

[2] G. M. Johnson and M. D. Fairchild, "On Contrast Sensitivity in an Image Difference Model," *IS&T PICS2002,* pg. 18-23 (2002).

[3] M. Orfanidou, S. Triantaphillidou, and E. Allen, "Predicting Image Quality Using a Modular Image Difference Model," *SPIE/IS&T Electronic Imaging 2008: Image Quality and System Performance,* 6808, pg. 0F1-0F12 (2008).

[4] X. Zhang and B. A. Wandell, "A Spatial Extension of CIELAB for digital Color-image Reproduction," *Proc. of the SID Symposiums,* pg. 731-734 (1996).

[5] F. W. Campbell, "The transmission of spatial information through the visual system," in *The Neurosciences Third Study Program*, F. Worden and F. O. Schmitt, Eds.(Cambridge, MA: MIT Press, 1973), pg. 95-103.

[6] S. Anderson, K. Mullen, and R. Hess, "Human peripheral spatial resolution for achromatic and chromatic stimuli: Limits imposed by optical and retinal factors," *J. Physiol.,* 442, pg. 47-64 (1991).

[7] A. B. Poirson and B. A. Wandell, "The appearance of colored patterns: Pattern-color separability," *J. Opt. Soc. Am. A,* 10, pg. 2458-2470 (1993).

[8] D. G. Lee, M. R. Luo, G. Cui, and P. A. Rhodes, "Evaluating for Colour Differences for Images," in *CGIV2008* Terrassa, Spain: IS&T/SID (2008).

[9] J. Morovi and P.-L. Sun, "Predicting Image Differences in Colour Reproduction from Their Colorimetric Correlates," *Journal of Imaging Science and Technology* 47, pg. 509-516 (2003).

[10] T. Song, "Colour difference Evaluation for Complex Images." PhD Thesis Derby, England: University of Derby (2007).

[11] M. Stokes, "Colorimetric tolerances of digital images," MSc Thesis Rochester, NY: Rochester Institute of Technology (1991).

[12] J. Uroz, "Colour Difference Perceptibility for Large-Size Printed Images," MSc Thesis Derby, England: University of Derby (1999).

[13] G. M. Johnson and M. D. Fairchild, "A Top Down Description of S-CIELAB and CIEDE2000," *Color Research and Application,* 28, pg. 425-435 (2003).

[14] X. Zhang, J. E. Farrell, and B. A. Wandell, "Applications of a spatial extension to CIELAB," in *SPIE97* (1997).

[15] J. A. Movshon and L. Kiorpes, "Analysis of the development of spatial sensitivity in monkey and human infants," *J. Opt. Soc. Am. A,* 5, pg. 2166- (1988).

[16] CIE, "CIE TC8-02 Technical Report: Methods for Deriving Colour Differences in Images - Draft Version 0.12," Central Bureau of the CIE, Vienna (2007).

[17] F. W. Campbell and D. G. Green, "Optical and retinal factors affecting visual resolution," *J. Physiol. (London),* 181, pg. 576-593 (1965).

[18] M. D. Fairchild and G. M. Johnson, "The iCAM framework for image appearance, image difference, and image quality," *J. of Electronic Imaging,* pg. 126-138 (2004).

[19] C. Tomasi and R. Manduchi, "Bilateral Filtering for Gray and Color Images," in *IEEE International Conference on Computer Vision*, Bombay, India, pg. 839-846 (1998).

[20] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal,* 27, pg. 379-423, 623-656 (1948).

[21] ISO, "ISO 3664:2000, Viewing conditions --- Graphic technology and photography," (2000).

[22] J. Y. Hardeberg, "Acquistion and reproduction of colour images: colorimetric and multispectral approaches," in *Ecole National Supérieure des Télécommunications Département TSI*. PhD Thesis, Paris, France (1999).

[23] R. L. Devalois and K. k. Devalois, *Spatial Vision*: Oxford University Press (1991).

[24] Z. Wang and M. R. Luo, "Experimental Filters for Estimating Image Differences," in *CGIV08* Terrassa: IS&T/SID (2008).

[25] C. J. Bartleson, "Measuring Differences," in *Optical Radiation Measurement, Vol. 5 - Visual Measurements*, C. J. Bartleson and F. Grum, Eds., (Academic Express, Inc., 1984), pg. 441-489.

[26] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information," *Psychological Review,* 63, pg. 81-97 (1956).

Author Biography

Zhaohui Wang received his MSc in color imaging from the Color & Imaging Institute at the University of Derby (2004) and his PhD in color imaging from the Department of Color Science at the University of Leeds (2008). Currently, he is a postdoctoral researcher in the Norwegian Color Research Laboratory at Gjøvik University College. His work has focused on the development of a perceptual image difference metrics, and is sponsored by the Research Council of Norway.