

Gjøvik University College

HiGIA Gjøvik University College Institutional Archive

Wang, Z. & Hardeberg, J. Y. (2012) Development of an adaptive bilateral filter for evaluating color image difference. In: *Journal of Electronic Imaging (JEI)*, 21(2)

Internet address: http://dx.doi.org/10.1117/1.JEI.21.2.023021

Please notice: This is the journal's pdf version.

© Reprinted with permission from Society of Photo Optical Instrumentation Engineers

One print or electronic copy may be made for personal use only. Systematic electronic or print reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

Electronic Imaging

SPIEDigitalLibrary.org/jei

Development of an adaptive bilateral filter for evaluating color image difference

Zhaohui Wang Jon Yngve Hardeberg



Development of an adaptive bilateral filter for evaluating color image difference

Zhaohui Wang Jon Yngve Hardeberg Gjøvik University College The Norwegian Color Research Laboratory P.O. Box 191, Teknologiveien 22 N-2802 Gjøvik, Norway E-mail: Zhaohui.wang@live.cn

Abstract. Spatial filtering, which aims to mimic the contrast sensitivity function (CSF) of the human visual system (HVS), has previously been combined with color difference formulae for measuring color image reproduction errors. These spatial filters attenuate imperceptible information in images, unfortunately including high frequency edges, which are believed to be crucial in the process of scene analysis by the HVS. The adaptive bilateral filter represents a novel approach, which avoids the undesirable loss of edge information introduced by CSF-based filtering. The bilateral filter employs two Gaussian smoothing filters in different domains, i.e., spatial domain and intensity domain. We propose a method to decide the parameters, which are designed to be adaptive to the corresponding viewing conditions, and the quantity and homogeneity of information contained in an image. Experiments and discussions are given to support the proposal. A series of perceptual experiments were conducted to evaluate the performance of our approach. The experimental sample images were reproduced with variations in six image attributes: lightness, chroma, hue, compression, noise, and sharpness/blurriness. The Pearson's correlation values between the model-predicted image difference and the observed difference were employed to evaluate the performance, and compare it with that of spatial CIELAB and image appearance model. © 2012 SPIE and IS&T. [DOI: 10.1117/1.JEI.21.2.023021]

1 Introduction

The objective of a successful image difference model is to have a good agreement with the image difference perceived by observers. To achieve this, many of the characteristics of the human visual system (HVS) have to be considered in the process. Spatial characteristics are among the most important and of much current interest in the development of image difference metrics.

It is a common experience that an image reveals more details when we approach it and lose some details when we move away. Much of our understanding of this visual process is based on the studies of spatial characteristics of the HVS, which show that it is composed of spatial frequency channels.¹ According to Campbell and Robson,² the HVS contains a number of neural channels each one selectively sensitive to a different range of spatial

frequencies, and detecting the outputs of different channels independently.³ These studies typically employ sinusoidal gratings of different spatial frequency to which the sensitivity is measured and turns out the contrast sensitivity functions (CSFs). The functions are typically measured in opponent color space: luminance, red-green and yellow-blue. The shape of the function altering with the spatial frequency accordingly allows strong inferences to be made for current viewing process. The luminance CSF is essentially a band pass function. The fall-off at low spatial frequencies is usually for lateral inhibition,⁴ which plays a critical role in contrast (edge) enhancement.⁵ The decrease in sensitivity at high frequencies can be referred to as blurring because of the optical limitation of the eye and the spatial summation in the HVS.⁶ Both red-green and yellow-blue CSFs have lowpass characteristics with no low frequency attenuation.

The behavior of CSFs and its application to elucidate the visual processes of image discrimination has been extensively studied for half a century, and several researchers have developed models to predict the visible differences; e.g., models by Charman and Olin,⁷ Carlson and Cohen,³ and Barten.9 A detailed scheme proposed by Dalv10 described how CSFs can be normalized and integrated into a workflow to predict difference for a range of viewing distances. A further study conducted by Peli¹¹ indicated the relevant CSFs for a discrimination task had to be measured over a range of observation distances accordingly. A remarkable aspect of the above discussion is that models are expressed by only luminance CSF. Content in the image that falls below the observer's luminance contrast threshold is attenuated as imperceptible information. The application of both chromatic and achromatic CSFs has been introduced to color difference formulae to estimate the color image difference by Zhang and Wandell.¹² Later, a model based on contrast sensitivity measurement by Movshon and Kiorpes¹³ was suggested¹⁴ and recommended¹⁵ for a standard workflow instead of the application of CSF measured individually, in which, the band-pass luminance CSF becomes low-pass by normalizing the mean luminance; the image is filtered separately by three channels CSFs, and CIELAB formulae are applied to the filtered image to obtain a map of differences.

The performance of CSF filters has typically been characterized in terms of their effect on images. High frequency

Paper 11173 received Jul. 6, 2011; revised manuscript received Mar. 20, 2012; accepted for publication May 14, 2012; published online Jun. 22, 2012.

^{0091-3286/2012/\$25.00 © 2012} SPIE and IS&T

information is removed from the image and as a result much detail is lost. Edges in an image contain very high frequency information; loss of these frequencies blurs the image. There is a broad consensus, however, that the HVS is particularly sensitive to the edges.¹⁶ Edge detection is believed to be necessary to distinguish objects from their background, and establish their shape and position.¹⁷ Edge detection has been proved to be one of the first major steps in the process of scene analysis by the HVS¹⁶ and use both achromatic and chromatic information.^{18–20} A recent study by Bex et al.²¹ suggested that contrast sensitivity to natural scenes depends on edge as well as spatial frequency structure. To overcome the undesirable loss of edges whilst using the spatial CSF filters, edge enhancement techniques were added in the workflow for spatial localization, e.g. Refs. 22 and 23.

An important motivation of our work is to improve the performance of the spatial filters by using an appropriate edge preserving technique in an image difference workflow. In this study, we employ an adaptive bilateral filter for the purpose. The bilateral filter is adopted from Ref. 24, in which two Gaussian filters are applied at localized pixel neighborhood. The result is a blurrier image than the original while edges are preserved. Parameters have been investigated to modulate the bilateral filter for the purpose of image difference evaluation; parameters which are adaptive to the corresponding viewing conditions, and the quantity and homogeneity of information contained in an image.

The rest of this paper is organized as follows: first, the description of the proposed adaptive bilateral filter, then the investigation of viewing distance adaptation based on perceptual experiments, followed by the proposed adaptation to image content based on entropy analysis. Then we describe the experimental method used to evaluate our proposed model, followed by results and discussion, and finally, we conclude.

2 Proposed Adaptive Bilateral Filter

The investigation of CSF-based models in image difference metrics gave us the basic idea that the bilateral filter, as an edge preserving smoothing filter might be appropriate for evaluating color image difference.

The idea behind using a bilateral filter is to avoid the unwanted edge blur from Gaussian smoothing filter which averages the neighbor pixels across edges.^{25,26} A recent study by Tomasi and Manduchi²⁴ employed two Gaussian filters, one in the spatial domain (domain filter) and the other in the intensity domain (range filter). Pixels in the neighborhood which are geometrically closer and photometrically more similar to the filtering center will be weighted more, as illustrated in Fig. 1. Given a color image f(x), the bilateral filter²⁴ can be expressed as,

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

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 center x and a nearby point ξ :

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



Fig. 1 The principle of bilateral filter.

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

$$s(\xi, x) = e^{-\frac{[f(\xi) - f(x)]^2}{2\sigma_r^2}}.$$
(3)

The behavior of this filter is controlled by two parameters. The geometric spread σ_d in the spatial 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 and weighted. 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. We proposed these two parameters to be self-adaptive to the viewing condition and the image itself.

The domain spread σ_d in our proposal is determined by the viewing condition, 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, as portrayed in Fig. 2, the domain spread can be expressed as,

$$\sigma_d = \frac{\frac{n}{2}}{\frac{180}{\pi} \cdot \tan^{-1} \frac{l}{2m}},$$
(4)

which constructs a direct relationship between the smoothness of the image and the viewing condition. For example,



Fig. 2 Construction of domain spread in pixels per degree (ppd).

(2)

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. Entropy²⁷ is defined with the probability of occurrence of a certain pixel value,

$$E = -\sum_{i} p_i \log(p_i), \tag{5}$$

where p_i refers to the histogram of the pixel intensity values of an image. A high entropy value is associated with high variance in the pixel values of an image, and a 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

$$\sigma_r = \frac{K}{E},\tag{6}$$

where the constant *K* is used to rescale the image entropy into an optimized value and the 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. The image entropy is employed 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).

When applying CSF models on an image, blurry edges can be found. On the other hand, the three channels were filtered by CSF separately from one another in an opponent color space, which will increase the risk of disturbance of color balance. To avoid this problem, the adaptive bilateral filter operates 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 3 presents an example, which compares the results of an image processed by the CSF model¹³ and the adaptive bilateral filter. The results were converted to sRGB color space for display.

3 Viewing Distance Adaptation

The CSFs have been revealed to be spatial frequency and viewing distance related. Although many researchers have applied CSFs in prediction of image difference, relatively little is known about contrast sensitivity under different viewing distance. Here we investigated the variation of human perception under a set of distances when certain conditions hold.



Fig. 3 Comparison of images processed by the contrast sensitivity function (CSF) model¹³ and the adaptive bilateral filter (ABF): (a) original image; (b) image reproduced by CSF model; (c) image reproduced by ABF.

3.1 Image Manipulation

Image difference may originate due to different image reproduction methods, such as the discriminations from chromatic and spatial modifications. Attempts to computationally assess color image difference have typically applied models of CSF to determine the discriminations introduced by spatial alteration, such as image compression, halftone reproduction, etc. On the other hand, several studies^{28–31} 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.

Four color images were manipulated in six attributes: lightness (L), chroma (C), hue angle (H), compression (CO), noise (N), sharpness/blurriness (SB), according to the functions and parameters listed in Table 1. The purpose of the image manipulation is to generate images similar to the originals but under limited difference. The parameters of k in each rendering function are scaled as shown in Table 1.

3.2 Experimental Procedure

The experiment was conducted in a dark room. A 24-in. EIZO ColorEdge© CG241W LCD was used to display the image pairs. The screen resolution was 1920×1200 pixels and the refresh rate was 75 Hz. The display system was calibrated and characterized according to ISO 3664^{32} under illuminant D65. The image state was set to sRGB color space in a resolution of 800×600 pixels.

In each trial of the experiment, a pair of images, including an original and a manipulated image, was presented to the observer. The position of the presented images on the left/ right of the screen was randomized from trial to trial to minimize the effect of the non-uniformity of display. The methods of limits was employed, which is used mainly to obtain thresholds. Observers were presented with a series of manipulated images that are systematically increased (or decreased) according to the parameter of k defined in Table 1. The manipulated image (and its corresponding parameter k) at which the observer switches respond from "no" to "yes" is then taken as the threshold.

Ten observers participated in the experiment. All had normal color vision according to Ishihara test and their visual acuity reached 20/20 in all distances using the Snellen vision chart. The observers' task was to decide whether the two images on screen were identical or the difference was just noticeable. Observers were asked to finish the task by sitting at different distances to the display, which is demonstrated in Fig. 4. For each viewing distance, its corresponding viewing condition defined by Eq. (4) is presented in Fig. 4; e.g., when the observer is sitting at a viewing distance of 7 m, the corresponding viewing condition is 350 ppd.

3.3 Results

The experiment examined how the perceived image discrimination varies with the viewing distance in terms of different manipulation methods. Totally, a number of 360,080 (10 observers ×7 viewing distances ×5144 image pairs) visual judgments were collected for data analysis. The visual judgments were then averaged according to the viewing distance, which, in turn, presents the thresholds of parameter k, as shown in Table 2, when observers described answers as

Туре	Description	Scales of parameter k		
L	$L_1 = kL_0$	$\begin{cases} k_{i+1} = k_i + 0.0125 \\ k \in [0.5 \ 2] \end{cases}$		
С	$C_1 = kC_0$			
Н	$H_1 = kH_0$			
со	Matlab JPEG lossy model with quality variation of k , which the best quality is given by 100, and the scale of 0 produces the worst quality.	$\begin{cases} k_{i+1} = k_i + 1\\ k \in \begin{bmatrix} 0 & 100 \end{bmatrix} \end{cases}$		
Ν	Matlab Gaussian random noise with zero mean noise and variance of k , which 0 indicates no noise added and 0.05 means the largest noise level produced in this study.	$\begin{cases} k_{i+1} = k_i + 0.0001 \\ k \in [0 0.05] \end{cases}$		
В	Gaussian smooth function with standard deviation $k f(x, y) = \frac{1}{2\pi k^2} e^{-\frac{x^2 + y^2}{2k^2}}$	$\begin{cases} k_{i+1} = k_i + 0.05 \\ k \in [0.05 \ 11] \end{cases}$		
S	Unsharp masking $f_{\text{sharp}}(x, y) = f(x, y) + k \times [f(x, y) - f_{\text{smooth}}(x, y)] \begin{cases} k_{i+1} = k_i + 0.05 \\ k \in [0 5] \end{cases}$			

 Table 1
 Image manipulating functions and their corresponding parameters.



Fig. 4 Diagram of experimental setup (for each viewing distance, the corresponding ppd are presented).

just noticed difference (JND) or identical between images in each pair on screen.

Figure 5 presents the average scaling of all observers for different chromatic changes of the testing images. The horizontal axis displays spatial frequency in terms of pixels per degree which is corresponding to the different viewing distances, and the vertical axis shows the scales of parameter k defined in Table 1 at which observers agreed with the identical reproductions. The symbols of AS and DS in the figure denote the corresponding ascending and descending order of the presentation of reproduced images according to the parameter k, respectively. The error bars show the standard deviation of the average observations at each viewing distance. A statistical comparison of the average judgment

given in ascending and descending orders shows that the differences between the results are significantly different for lightness and chroma [p < 0.05, degree of freedom (df) = 14, Student's *t*-test, two-tailed] which obtained higher values of parameter *k* from the descending order and lower values from the ascending order as shown in Fig. 5(a) and 5(b). The visual judgments on hue reproduction are not statistically significantly different (p > 0.05, df = 14, Student's *t*-test, two-tailed). These differences between AS and DS are somehow inherited from the disadvantages of the experimental method—error of habituation and expectation³³—and can be minimized by averaging the results from both ascending and descending orders.

The variations of visual judgments are smaller in different viewing distance for lightness, chroma, and hue modifications. Comparing the results of lightness, chroma, and hue in Fig. 5, the tolerance of hue shift is smaller than that of lightness and chroma as shown by standard deviation. The results show that the visual judgment of image discrimination is viewing distance independent on the changes of lightness, chroma, and hue.

Figure 6 presents the average observations on the image spatial alterations. For each manipulation term, the average judgments are plotted with the corresponding standard deviation under each viewing distance. Obviously, the observers' visual judgments are related to viewing distance in these cases.

The results in Fig. 6(a) show that the tolerance on the quality of compressed reproductions is increased with the increasing viewing distance but remained relatively constant for larger viewing distance. The ability to discern small difference is easier when the distance is closer, higher quality parameter is demanded. When the viewing distance is increased, visual discrimination becomes more difficult and lower quality is tolerated. The differences between the presentation of images in ascending and descending orders are not significant (p > 0.05, df = 14, Student's *t*-test, two-tailed).

Table 2 Thresholds of parameter k under each viewing distance and manipulation methods.

	Lightness				Chroma				Hue			
Viewing Distance (ppd)	AS		DS		AS		DS		AS		DS	
	k	std	k	std	k	std	k	std	k	std	k	std
25	0.9859	0.0455	1.0406	0.0490	0.9500	0.0189	1.0563	0.0378	1.0016	0.0080	1.0031	0.0058
50	1.0203	0.0762	1.0609	0.0236	0.9828	0.0482	1.0281	0.0516	1.0000	0.0067	1.0000	0.0000
100	1.0203	0.0417	1.0625	0.0482	0.9875	0.0634	1.0266	0.0506	0.9906	0.0174	1.0047	0.0065
150	1.0000	0.0543	1.0766	0.0205	0.9750	0.0866	1.0203	0.0661	0.9922	0.0093	0.9984	0.0080
200	1.0000	0.0347	1.0750	0.0738	0.9141	0.0639	1.0594	0.0427	1.0016	0.0295	1.0000	0.0341
250	0.9797	0.0327	1.0859	0.0599	0.9734	0.0649	1.0438	0.1167	0.9984	0.0080	0.9984	0.0080
300	1.0031	0.0661	1.0547	0.0249	0.9266	0.0693	1.0625	0.0735	0.9969	0.0160	1.0094	0.0111
350	0.9781	0.0578	1.0734	0.0580	1.0047	0.1106	1.0797	0.0641	1.0000	0.0177	1.0016	0.0216
	Compression				Noiseness				Sharpness(DS)/Blurriness(AS)			
	AS		D	DS		AS		DS		AS		DS
Viewing Distance (ppd)	k	std	k	std	k	std	k	std	k	std	k	std
25	79	15	73	13	0.0003	0.0001	0.0001	0.0001	0.4175	0.1742	0.0525	0.2552
50	66	15	64	14	0.0005	0.0002	0.0003	0.0003	0.6525	0.4402	-0.0150	0.1961
100	47	19	52	17	0.0020	0.0013	0.0015	0.0008	0.9650	0.4467	-0.0300	0.2816
150	32	11	29	11	0.0052	0.0039	0.0067	0.0044	1.2800	0.6435	-0.0275	0.3432
200	21	8	25	11	0.0088	0.0050	0.0142	0.0060	1.8550	0.6145	-0.1125	0.3047
250	17	7	17	7	0.0133	0.0064	0.0165	0.0070	2.5500	0.9080	-0.1450	0.2417
300	14	6	12	5	0.0147	0.0039	0.0207	0.0077	2.7875	0.9601	-0.1800	0.3385
350	13	6	12	4	0.0221	0.0074	0.0258	0.0078	2.9275	0.6263	-0.3425	0.5425

Note: The symbol of "AS" and "DS" denote the corresponding ascending and descending order of the presentation of reproduced images according to the parameter *k*, respectively. The symbol of "std" represents the standard deviation.

Figure 6(b) presents the variation of observations in terms of noise addition. In the closer viewing distances (\leq 50 ppd), the judgment is remained in a quasi-constant small number amount of noise addition. Particularly in the smallest distance of 25 ppd, zero tolerance is obtained. When the viewing distance is increased, the tolerance of the amount of noise increased. The statistical test between the image presentation orders of ascending and descending shows that the differences of presenting order are not significantly different (p > 0.05, df = 14, Student's *t*-test, two-tailed).

In Fig. 6(c), the blurred reproductions are considered as the ascending order (+) of its reverse of the sharpened reproductions (-). The results of two presentation orders are found to be statistically significantly different (p < 0.05, df = 14, Student's *t*-test, two-tailed). As with the increase of viewing distance, the tolerance on the blurred degree is increased but remained relatively constant for larger distance. The variation of the tolerance on sharpness modification is far smaller than that of blurred reproductions with the viewing distance increase. The perceived identical reproductions fall on small sharpen scales. The average of the judgment is of -0.1 (the symbol of "–" means sharpened) with a standard deviation of 0.12, which shows the tolerance on sharpness is smaller than that on blurriness.

The smoothness degree, σ_d in Eq. (2), is determined by the viewing condition which reflects to the variation of minimum noticeable amount of a change of the frequency component, particularly in images' spatial alteration. To minimize the weight of σ_d on chromatic manipulations which show limited effect of viewing distance, the σ_d is operated on the luminance channel. 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



Fig. 5 Observers' visual judgment on lightness, chroma, and hue modifications under different viewing conditions in terms of ppd: (a) threshold of lightness corresponding to the average of all observers; (b) threshold of chroma corresponding to the average of all observers; (c) threshold of hue corresponding to the average of all observers.

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)³⁴ 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.

4 Image Content Adaptation

Previous research^{12,35,36} has demonstrated that color difference formulae, which were developed from large uniform color patches, cannot be applied directly for color image difference evaluation due to the great complexity of images. Further studies^{36–39} suggested that the image difference evaluation should be image dependent, for which global and local image features need to be considered.

Visual sensitivity is adjusted constantly by adaptation to the neighborhoods rather than in isolation. Light or



Fig. 6 The average of observers' visual judgment on compressions, noise and, blur/sharpness modifications under different viewing conditions in terms of ppd: (a) threshold of compression corresponding to the average of all observers; (b) threshold of noise corresponding to the average of all observers; (c) threshold of blur/sharpness corresponding to the average of all observers.

chromatic adaptation adjusts sensitivity according to the mean luminance and chromaticity averaged over some time and region of stimulus.⁴⁰ Thus, the color perception of a single pixel cannot be considered individually in an image, which is correlated with the neighbors. Much of the evidence comes from the well-known phenomena of Mach bands and simultaneous contrast that pixels interacting with one another will affect our perception accordingly. An investigation by Webster et al.⁴¹ suggested that natural images characterized by restricted color distribution may provide a potent stimulus for adaptation.

The Gestalt law^{42} of the perceptual organization states that the human perception is well organized according to the factors of proximity and similarity of the elements in a scene. The further study¹⁶ of perception suggested that the object perception is based on the local detection and tracking of edges. The HVS enhances local edges in order to better distinguish objects while looking at a natural scene. Pursuing this idea, Wang et al.³⁶ described an experimental method to predict color image difference from the large homogeneous area or main objects in an image. The experimental filters of each testing image were obtained from perceptual experiments which stand for the objects observed by subjects to tell the differences from image pairs; however, there were some difficulties with this method. The experimental filter was not realistic or applicable for practical applications.

We further extended the idea of the experimental filter³⁶ to get the measurement of homogeneity of the region of interest and consider the local adaptation to image content by using entropy. Entropy is the rigorous measure of disorder or system homogeneity. The color homogeneity measurements based on the intensities of the chroma channel provides more sensitivity, because the HVS is sensitive to large clusters of colors.⁴³ On the other hand, the entropy is useful to quantify the edges since the entropy of a color image is small when the change of lightness or hue is severe. Thus, an inverse proportion can be defined between the parameter of the range spread σ_r in Eq. (5) and the entropy, when the image is more homogeneous in color, the smaller range spread is. To consider the spatial response of similar color region, the concept of image entropy may also be extended to the region of interest (ROI) based entropy.

5 Performance Test

A series of psychophysical experiments were conducted to validate the performance of our proposed adaptive bilateral filter and compare it with two other models, spatial CIELAB (sCIELAB)¹² and image appearance model (iCAM).²² We gave our attention to these two algorithms because sCIELAB and iCAM are based on the spatial property of the HVS in which CSFs have been employed and derived for color image difference, which are also the motivation of our proposal.

Experiments were conducted in a dark room using a 21-inch LCD monitor which was calibrated and characterized according to ISO 3664.³² 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 800×600 (96 pixels per inch) under D65. A set of reproduction methods were applied, including the manipulation in lightness (*I*), chroma (*C*), hue (*H*), compression (CO), noise (*N*), and sharpness (*S*). The transformations have been defined in Table 1. However, to decrease the number of reproductions and the cost of perceptual experiments, only seven levels of transformation (including a level of zero which represents the original image) were applied to each manipulation method to prepare the testing images.

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^{15,30} to perceptible but acceptable⁴⁴ difference. Totally, 420 image pairs (10 images \times 6 methods \times 7 levels) were used in the experiment. Figure 7 shows the distribution of average color difference of all testing images in terms of CIELAB.



Fig. 7 The distribution of all testing images in terms of ΔE in CIELAB.

Ten normal color vision observers (all passed Ishihara test) participated in the experiments. In each observation, a pair of images, including original and reproduced images, was presented to the observer. The location of the original and reproduced images on the right or left was randomized from trial to trial. Observers were asked to evaluate the total image difference between an original image and a manipulated image using category judgment method. Each observer was given a training session. The categorical scale is employed according to Bartleson⁴⁵ and Miller.⁴⁶ The seven categories, which are listed in Table 3, are also somewhat matched to the seven transformation levels of each manipulation method.

6 Results and Discussion

In this study, observer accuracy is investigated in repeatability and variation. The observers' accuracy is calculated for each stimulus in terms of coefficient of variation using equation

$$CV = \frac{100 \times \sqrt{\frac{\sum_{i=1}^{N} (x_i - \bar{x})^2}{N-1}}}{\bar{x}},$$
 (7)

where x_i represents each observation; \bar{x} is the average category value of all observers; N is the number of observers, i.e., 10 in this experiment. The value of CV ranges from 1 to 100. The smaller the CV is, the higher the observers' precision.

Table 3 Description of categories.

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				

Ohaamuan		•	•		-	A
Observer	I	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 5 Performance of observers' repeatability.

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	-

A number of 190 image pairs selected from the testing images were presented twice randomly to observers to test observer's repeatability. The results are summarized in Table 4, which arrived at average of 17 with 95% significant level of 20, which are considered to be relatively high. These CV values are also higher than those in previous research.^{29,36}

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 5.

Totally, 4200 (10 observers × 420 image pairs including repeated image pairs) visual judgments were collected. Torgerson'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. As a result, a *z*-score matrix, which presents the unit normal deviate corresponding to the proportion of times stimulus is sorted below category boundary. The values of *z*-score were considered to represent psychologically different stimuli, which are plotted against the prediction of ABF to investigate the performance.

To investigate the observation of image reproductions on spatial and chromatic properties, we average the results of all 4200 visual judgments according to original images, manipulation methods, and reproduction levels. For each manipulation method, the performances of adaptive bilateral filter are compared with z-scores. A higher z-score indicates a larger grade of visual judgment. The results for manipulation methods of L, C, and H are shown in Fig. 8(a). At the same value of z-score (observation result), the prediction of adaptive bilateral filter on lightness reproductions are slightly higher than that of reproductions on chroma and hue channel, which gave the lightness is the most tolerated. To improve the performance, the relationship between L and the other two is suggested to be set at a higher ratio of 3:1:1 which gives a Pearson value of 0.75 compared to 0.54 for the ratio of 1:1:1. Figure 8(b) presents the performance



Fig. 8 The performance of adaptive bilateral filter in terms of manipulation methods: (a) Performance of adaptive bilateral filter (dE-ABF) on manipulation methods of *L* (open triangle), *C* (filled diamond), and *H* (open circle); (b) Performance of adaptive bilateral filter (dE-ABF) on manipulation methods of CO (open diamond), *N* (filled circle), and SB (filled triangle).

of adaptive bilateral filter on the manipulation methods of compression (CO), noise (N), and sharpness/blurriness (SB). Comparing with the results of Fig. 8(a), the predictions of adaptive bilateral filter are mostly in a lower level at the same value of *z*-score, which suggests a higher ratio between the predictions on spatial and chromatics alterations by adaptive bilateral filter.

The performance of adaptive bilateral filter is analyzed in terms of Pearson's correlation value, which 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.

The performance of adaptive bilateral filter was compared with that of sCIELAB and iCAM. The difference between adaptive bilateral filter and other two methods is that adaptive bilateral filter operates on three channels of CIELAB together rather than separately in three channels of an opponent color space. Using adaptive bilateral filter, the experimental image pairs were filtered on L, C, and H channels in CIELAB color space and then the average pixel-wise differences were calculated using the CIELAB color difference formulae. The iCAM workflow was implemented according



Fig. 9 The performances of adaptive bilateral filter (filled circle), sCIE-LAB (open square), and iCAM (open diamond) on different manipulation methods in terms of Pearson's correlation value.

to Fairchild et al.²² The sCIELAB workflow was carried out using the procedure provided by the CIE TC8-02 report¹⁵ in which the identical CSFs as used in iCAM were recommended.

Figure 9 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. The differences of performances given by three models are statistically significant [repeated measures ANOVA, $F(2, 15) > F_{\text{critical}}, P < 0.05$]. In most cases, the adaptive bilateral filter gives relative higher Pearson's values except that the Pearson correlation of sCIELAB is slightly higher by manipulation method of compression (CO) and noise (NO).

7 Conclusion

An adaptive bilateral filter was proposed for color image difference evaluation. The filter requires two parameters to control the behavior. The bilateral filter is adaptive to the corresponding viewing conditions and the homogeneity of chromatic information contained in an image. The experiments of viewing distance adaptation showed that the visual judgments on chromatic and spatial alteration are different; the latter is affected stronger by the viewing distance. Thus, the filter's ability to smooth the image by the domain spread is adjusted by the viewing condition to attenuate the imperceptible information. Based on the analysis of the statistic and homogeneity of the image, the filter's ability of edge preserving is controlled by the image chromatic entropy.

Perceptual experiments were conducted using category judgment method to evaluate the performance of the proposed methods and compare with the performances of sCIELAB and iCAM. 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 were employed to analyze the results. The adaptive bilateral filter shows a higher Pearson's values than other two models.

References

F. W. Campbell, "The transmission of spatial information through the visual system," in *The Neurosciences: Third Study Program*, 1. F. O. Schmitt and F. Worden, eds., pp. 95-103, MIT Press, Cambridge, MA (1974).

- F. W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings," *J. Physiol.* **197**(1), 551–556 (1968).
 M. B. Sachs, J. Nachmias, and J. G. Robson, "Spatial-frequency channels in human vision," *J. Opt. Soc. Am.* **61**(9), 1176–1186 (1971).
- T. Cornsweet, Visual Perception, Academic Press, New York and London (1970).
- 5. F. Ratliff, Studies on Excitation and Inhibition in the Retina: A collection of Papers from the Laboratories of H. Keffer. Hartline, Chapman and Hall, London (1974).
- F. W. Campbell and D. G. Green, "Optical and retinal factors affecting visual resolution," *J. Physiol.* 181(3), 576–593 (1965).
 W. N. Charman and A. Olin, "Image quality criteria for aerial camera 6.
- 7.
- systems," *Phot. Sci. Eng.* **9**(6), 385–397 (1965). C. R. Carlson and R. W. Cohen, "A simple psychophysical model for predicting the visibility of displayed information," *Proc. Soc. Inf. Displ.* 8 21, 229-246 (1980).
- P. G. J. Barten, "Evaluation of subjective image quality with the square-root integral method," *J. Opt. Soc. Am. A* 7(10), 2024–2031 (1990).
 S. Daly, "The visible differences predictor: an algorithm for the assess-
- ment of image fidelity," in Digital Images and Human Vision, A. B. Watson, ed., MIT Press, Cambridge, MA (1993).
- E. Peli, "Contrast sensitivity functions and image discrimination," J. Opt. Soc. Am. A 18(2), 282–393 (2001).
- X. Zhang and B. A. Wandell, "A spatial extension of CIELAB for digital color image reproduction," *Proc. Soc. Inf. Displ.* 27, 731–734 (1996).
- 13. J. A. Movshon and L. Kiorpes, "Analysis of the development of spatial contrast sensitivity in monkey and human infants," J. Opt. Soc. Am. A **5**(12), 2166–2172 (1988).
- 14. G. M. Johnson and M. D. Fairchild, "On contrast sensitivity in an image difference model," in IS&T PICS2002, Portland, USA, pp. 18-23 (2001)
- CIE, "CIE TC8-02 technical report: methods for deriving colour differ-ences in images," CIE TC8-02 technical report, Central Bureau of the CIE, Vienna, Austria (2007).
- 16. D. H. Hubel and T. N. Wiesel, "Receptive fields and functional architecture of monkey striate cortex," J. Physiol. 195(1), 215-243 (1968).
- 17. D. G. Lowe, Perceptual Organization and Visual Recognition, Kluwer Academic Publisher, Nowell, MA (1985).
- 18. I. Fine, D. I. A. MacLeod, and G. M. Boynton, "Surface segmentation based on the luminance and color statistics of natural scenes," J. Opt. oc. Am. A 20(7), 1283-1291 (2003).
- 19. T. Hansen and K. R. Gegenfurtner, "Independence of color and luminance edges in natural scenes," *Vis. Neurosci.* **26**(1), 35–49 (2009). 20. M. S. Livingstone and D. H. Hubel, "Segregation of form, color,
- movement, and depth: anatomy, physiology, and perception," *Science* **240**(4853), 740–749 (1988).
- 21. P. J. Bex, S. G. Solomon, and S. C. Dakin, "Contrast sensitivity in natural scenes depends on edge as well as spatial frequency structure,' J. Vis. 9(10), 1–19 (2009).
- M. D. Fairchild and G. M. Johnson, "The iCAM framework for image appearance, difference, and quality," *J. Electron. Imag.* 13(1), 126–138 (2004)
- M. Orfanidou and S. Triantaphillidou, "Predicting image quality using a modular image difference model," *Proc. SPIE* 6808, 68080F (2008).
- 24. C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in IEEE 6th Int. Conf. Comput. Vis., Bombay, India, pp. 839–846 (1998).
- 25 V. Aurich and J. Weule, "Non-linear Gaussian filters performing edge preserving diffusion," in 17th DAGM Symp., Beielefeld, Germany, pp. 538–545 (1995).
- S. M. Smith and J. M. Brady, "SUSAN—a new approach to low level image processing," *J. Comput. Vis.* 23(1), 45–78 (1997).
- 27. C. E. Shannon, "A mathematical theory of communication," Bell System Tech. J. 27(3), 379-423, 623-656 (1948)
- 28. J. Morovič and P. Sun, "Predicting image differences in colour reproduction from their colorimetric correlates," J. Imag. Sci. Tech. 47(6), 509-516 (2003).
- T. Song and M. R. Luo, "Testing colour-difference formulae on com-plex images using a CRT monitor," in *IS&T/SID 8th Color Imag. Conf.*, Scottsdale, AZ, pp. 44–48, IS&T, Springfield, VA (2000).
- 30. M. Stokes, Colorimetric Tolerance of Digital Images, Rochester Institute of Technology, Rochester, NY (1991).
- J. Uroz, M. R. Luo, and J. Morovič, "Colour difference perceptibility for large size printed images," in *Proc. Color Image Science 2000* Conf., pp. 138–151, University of Derby, Derby, England (2000).
- 32. ISO 3664: 2000(E), "Viewing conditions-graphic technology and
- R. L. D. Valois and K. K. D. Valois, *Spatial Vision*, pp. 147–176, Oxford University Press, NewYork, NY (1990). 34.
- 35. G. M. Johnson and M. D. Fairchild, "A top down description of S-CIE-LAB and CIEDE2000," Color Res. Appl. 28(6), 425-435 (2003).

- Z. Wang and M. R. Luo, "Experimental filters for estimating image 36. differences," in IS&T 4th Euro. Conf. Colour Graph., Imag. Vis. (CGIV), pp. 112–115, IS&T, Terrassa, Spain (2008).
- 37. G. Hong and M. R. Luo, "Perceptually based colour difference for complex images," in 9th Congress Int. Colour Assoc., pp. 618–621, AIC, Rochester, NY (2001).
- G. Hong, M. R. Luo, and P. A. Rhodes, "A new algorithm for calculat-ing perceived colour differece of images," *Imag. Sci. J.* 54(2), 86–91 (2006).
- (2006).
 Z. Wang and J. Y. Hardeberg, "An adaptive bialetral filter for predicting color image difference," in *IS&T/SID 17th Color Imag. Conf. (CIC)*, Albuquerque, NM, pp. 27–31, IS&T, Springfield, VA (2009).
 M. A. Webster, "Human colour perception and its adaptation," *Network: Computation Neural System* 7, 587–634 (1996).
 M. A. Webster and O. H. MacLin, "Influence of adaptation on the perception of distortions in patron improve *I Elevana*, *Inspin*, 10(1).
- ception of distortions in natural images," J. Electron. Imaging 10(1), 100–109 (1998). V. Bruce, M. A. Georgeson, and P. R. Green, *Visual Perception–Phy-*
- 42 siology, Psychology and Ecology, 4th ed., Psychology Press, East Sussex, UK (2003).
- A. M. Treisman and G. Gelade, "A feature-integration theory of atten-tion," *Cog. Psychol.* **12**(1), 97–136 (1980).
- J. Y. Hardeberg, "Acquistion and reproduction of color images: colorimetric and multispectral approaches," Ecole Nationale Supérieure des Télécommunications, Paris, France (1999).
 C. J. Bartleson, "Measuring differences," in *Optical Radiation Measurements*, Vol. 5, C. J. Bartleson and F. Grum, Eds., pp. 441–489,
- Academic Express, Inc., Orlando, FL (1984).
- G. A. Miller, "The magical number seven, plus or minus two: some limits on our capacity for processing information," *Psychol. Rev.* 63 (2), 81–97 (1956). 46



Zhaohui Wang received his MSc degree in color and imaging science from the Color & Imaging Institute, University of Derby, United Kingdom in 2004, and his PhD from Department of Color Science, Faculty of Mathematics and Physical Sciences, University of Leeds, United Kingdom, in 2008. He then joined staff at the Norwegian Color Research Laboratory as a postdoctoral researcher to begin working on "SHP-Perceptual Image Difference Metrics: A unifying approach to

image representation and reproduction," funded by the Research Council of Norway. The project was completed at the end of 2011. His general research interests are in color and imaging science; his specific research interests include color science, color perception, color image difference, color image appearance, gamut mapping, image guality, psychophysics, and imaging device technology.



Jon Yngve Hardeberg received his MSc degree in signal processing from the Norwegian Institute of Technology, Trondheim, Norway, in 1995, and his PhD from Ecole Nationale Supérieure des Télécommunications, Paris, France, in 1999. After a short but extremely valuable industry career near Seattle, Washington, where he designed, implemented, and evaluated color imaging system solutions for multifunction peripherals and other imaging devices and systems, he

joined Gjøvik University College (GUC) in Gjøvik, Norway, in 2001. He is currently a professor of color imaging at GUC's Faculty of Computer Science and Media Technology, and director of the Norwegian Color Research Laboratory. His current research interests include multispectral color imaging, print and image quality, colorimetric device characterization, and color management, and he has coauthored more than 150 publications within the field. His professional memberships include IS&T, SPIE, and ISCC, he is GUC's representative in IARIGAI, and the Norwegian delegate to CIE Division 8.