ELSEVIER

A new universal colour image fidelity metric

Alexander Toet*, Marcel P. Lucassen

TNO Human Factors, Vision Group, Kampweg 5, 3769 DE Soesterberg, The Netherlands

Abstract

We extend a recently introduced universal grayscale image quality index to a newly developed perceptually decorrelated colour space. The resulting colour image fidelity metric quantifies the distortion of a processed colour image relative to its original version. We evaluated the new colour image fidelity metric through observer experiments in which subjects ranked images according to perceived distortion. The metric correlates strongly with human perception and can therefore be used to assess the performance of colour image coding and compression schemes, colour image enhancement algorithms, synthetic colour image generators, and colour image fusion schemes.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Image quality; Colour fidelity

1. Introduction

In many areas of research and development that are concerned with digital imagery there is a real need for a digital metric that quantifies how distorted a processed colour image appears relative to the original version of the same image, as perceived by a human observer. For instance, in computer graphics, we may wish to compare the degree of photorealism of two different rendering methods, or to evaluate visible errors between a synthetic and a real scene. In image coding, we may need to evaluate the results of two different compression methods. In image processing, we may need to assess the performance of a new colour image enhancement technique. In false-colour multispectral image fusion, we may need to quantify the degree of photorealism of fused imagery. Unfortunately, common metrics like the root mean square error (RMSE) are not viable for these tasks because the human visual system does not compare images this way [14]. Psychophysical evaluation of different image processing techniques is tedious, expensive, and difficult to automate. Hence there is a great need for a validated computational image fidelity metric that correlates with human perception and that can be used for automatic optimization of

the parameters involved in different image processing and rendering techniques.

Over the years a large number of objective metrics have been proposed to assess image and video fidelity [1–4,7–9, 11,19–26,28,30,32,33,35,37,39,42,43,45–47]. For an overview of different metrics see [3,49]. Ironically, the metrics that are most widely applied today, like the RMSE or the peak signal-to-noise ratio (PSNR), are also the ones that correlate least with human perception [12,13,15–18,29,31, 34,40,51]. Human visual system models are more successful, but they are not widely used, since they generally are computationally complex.

Most of the digital image distortion metrics in the literature apply to grayscale images. Only a few studies address colour image fidelity metrics [4,25,27,28,42,48,48, 50]. Grayscale image fidelity metrics can in principle be generalized to colour image fidelity metrics by applying them to the three (*RGB*) colour channels individually, and then weighing and combining the errors in the different channels together. However, this direct approach does not relate to human perception. The *RGB* representation is based on primary relative colour. This space does not represent colour as perceived and analyzed by the human visual system. The human visual system uses three paths to analyze colour images: one for achromatic information and two for chromatic contrast signals. As a result, the individual channels of an *RGB* colour image are perceptually highly correlated. Hence, the *RGB* image should first

---

* Corresponding author. Tel.: +31-346-356237; fax: +31-346-353977.
  E-mail address: toet@tm.tno.nl (A. Toet).

be transformed into a perceptually uncorrelated colour space before further analysis is performed.

Most of the algorithms used the literature to quantify the distortion of a processed grayscale image relative to its original version are designed for some special types of distortions, such as blocking artifacts [24,35,37]. Wang and Bovik [44] recently introduced a general grayscale image fidelity metric that quantifies a whole range of local image distortions. The quality index is computationally simple and correlates with the subjective evaluations of human observers for a wide variety of distortions. Here we show that this metric can be extended to colour imagery, by applying it to the individual channels of a colour image in the newly introduced perceptually decorrelated $l\alpha\beta$ colour space [36], and by combining the results from the individual channels into a weighted vector mean.

2. The grayscale image fidelity metric

Wang and Bovik [44] recently introduced a universal image fidelity metric Q that quantifies the distortion of a processed image relative to its original version. The distortion metric is defined as a combination of three factors: loss of correlation, luminance distortion, and contrast distortion. Let $x = \{x_i | i = 1, 2, ..., N\}$ and $y = \{y_i | i = 1, 2, ..., N\}$ be the original and processed image signals, respectively. The fidelity index Q is then given by:

$$Q = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \frac{2\bar{x}\bar{y}}{(\bar{x})^2 + (\bar{y})^2} \frac{2\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2} \tag{1}$$

where

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i, \qquad \bar{y} = \frac{1}{N} \sum_{i=1}^{N} y_i$$

$$\sigma_x^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2, \qquad \sigma_y^2 = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{y})^2$$

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})$$

The dynamic range of Q is $[-1,1]$. The maximal value 1 only occurs when both images are identical, i.e. $y_i = x_i$ for all $i = 1, 2, ..., N$. The minimal value -1 occurs when $y_i = 2\bar{x} - x_i$ for all $i = 1, 2, ..., N$. The first component in Eq. (1) is the correlation coefficient between x and y, which measures the degree of linear correlation between both images, and has a dynamic range of $[-1,1]$. The maximal value 1 is obtained when $y_i = ax_i + b$ for all $i = 1, 2, ..., N$, where a and b are constants and $a > 0$. Even if x and y are linearly related, there may still occur relative distortions between them. These are evaluated in the second and third components. The second component measures how close the mean luminance both images are, and ranges between

$[0,1]$. It equals 1 when $\bar{x} = \bar{y}$. Since σ_x and σ_y can be regarded as estimates of the contrast of x and y, the third component measures how similar the contrasts of both images are. It also ranges from $[0,1]$. The highest value 1 is obtained if and only if $\sigma_x = \sigma_y$.

In practice we usually want to characterize an entire image using a single overall image fidelity measure. However, image fidelity is often spatially variant, meaning that different image regions may have different types of distortions. It is therefore more appropriate to measure statistical properties locally and combine them into a single measure. Following Wang [44] we therefore compute the image fidelity index Q over local image regions using a sliding window approach. Starting from the top-left corner of the image, a sliding window of size 8×8 moves pixel by pixel horizontally and vertically through all the rows and columns of the image until the bottom-right corner is reached. At the jth step in this procedure the local fidelity index Q_j is computed over the area of the 8×8 sliding window. If the total number of steps is equal to M, the overall image fidelity index is given by

$$Q = \frac{1}{M} \sum_{j=1}^{M} Q_j \tag{2}$$

3. The colour image fidelity metric

In this section we extend the grayscale fidelity metric Q to include colour by applying it to the individual dimensions of a perceptually decorrelated colour space, and combining the individual components in a (weighted) vector mean. The rationale for this approach is the fact that the human visual system processes the retinal image in three decorrelated visual channels: one luminance channel and two colour opponent channels. As a result, luminance and colour distortions will contribute independently to perceived image fidelity, and should therefore be calculated independently before combining them into a single overall perceived image fidelity metric.

The common *RGB* image representation is based on primary relative colour. This space does not represent colour as perceived and analyzed by the human visual system. In *RGB* space, there is a strong correlation between the individual image channels. For instance, most pixels will have large values for the red and green channels if the blue channel is large. The human visual system encodes the chromatic signals conveyed by the three types of retinal cone photoreceptors in an opponent fashion. This colour opponency is often interpreted as an attempt to remove correlations in the signals of different cone types that are introduced by the strong overlap of the cone spectral sensitivities [5]. Ruderman et al. [36] recently derived the perceptually decorrelated $l\alpha\beta$ colour space from a principal

component transform of a large ensemble of hyperspectral images that represents a good cross-section of natural scenes.

In the following sections we first discuss the *RGB* to *lαβ* transform. Then we construct the colour fidelity metric by applying the grayscale metric *Q* to each of the channels in the *lαβ* colour space.

3.1. The RGB to lαβ transform

First the *RGB* tristimulus values are converted to device independent *XYZ* tristimulus values. This conversion depends on the characteristics of the display on which the image was originally intended to be displayed. Because that information is rarely available, it is common practice to use a device-independent conversion that maps white in the chromaticity diagram to white in *RGB* space and vice versa [10].

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.5141 & 0.3239 & 0.1604 \\ 0.2651 & 0.6702 & 0.0641 \\ 0.0241 & 0.1228 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3)$$

The device independent *XYZ* values are then converted to *LMS* space by

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3897 & 0.6890 & -0.0787 \\ -0.2298 & 1.1834 & 0.0464 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (4)$$

Combination of Eqs. (3) and (4) results in

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3811 & 0.5783 & 0.0402 \\ 0.1967 & 0.7244 & 0.0782 \\ 0.0241 & 0.1288 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (5)$$

The data in this colour space shows a great deal of skew, which is largely eliminated by taking a logarithmic transform:

$$L = \log L \qquad M = \log M \qquad S = \log S \quad (6)$$

The inverse transform from *LMS* cone space back to *RGB* space is as follows. First, the *LMS* pixel values are raised to the power ten to go back to linear *LMS* space. Then, the data can be converted from *LMS* to *RGB* using the inverse transform of Eq. (5):

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 4.4679 & -3.5873 & 0.1193 \\ -1.2186 & 2.3809 & -0.1624 \\ 0.0497 & -0.2439 & 1.2045 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \quad (7)$$

Ruderman et al. [36] recently derived a colour space, called *lαβ*, which effectively minimises the correlation between the *LMS* axes. This result was derived from a principal component transform to the logarithmic *LMS* cone space representation of a large ensemble of

hyperspectral images that represented a good cross-section of natural scenes. The principal axes encode fluctuations along an achromatic direction (*l*), a yellow-blue opponent direction (*α*), and a red-green opponent direction (*β*). The resulting data representation is compact and symmetrical, and provides automatic decorrelation to higher than second order.

Ruderman et al. [36] presented the following simple transform to decorrelate the axes in the *LMS* space:

$$\begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \quad (8)$$

If we think of the *L* channel as red, the *M* as green, and *S* as blue, we see that this is a variant of a colour opponent model:

$$\text{Achromatic} \propto r + g + b$$
$$\text{Yellow-blue} \propto r + g - b \quad (9)$$
$$\text{Red-green} \propto r - g$$

The resulting data representation is compact and symmetrical, and provides automatic decorrelation to higher than second order.

3.2. Construction of the colour fidelity metric

The colour fidelity metric Q_{colour} is defined as:

$$Q_{\text{color}} = \sqrt{w_l(Q_l)^2 + w_\alpha(Q_\alpha)^2 + w_\beta(Q_\beta)^2} \quad (10)$$

where Q_l, Q_α, Q_β represent, respectively, the fidelity factors given by Eq. (2), computed for each of the individual *lαβ* colour channels, and w_l, w_α, w_β are the corresponding weights attributed to the perceived distortions in each of these channels. The *Q* values corresponding to the image distortions used in this study were always positive. Eq. (10) is designed in analogy to most modern colour difference equations [6].

4. Observer experiments

To assess the agreement between our colour image fidelity metric and human visual perception we performed observer experiments in which subjects ranked images according to perceived distortion. The distortions were produced by quantizing the original images along each of the individual dimensions in the perceptually decorrelated *lαβ* colour space. Since there are many nonlinearities in the way the visual system responds to the retinal image, we

cannot expect a linear relation between the colour image fidelity metric and the perceived amount of distortion. However, since the metric does increase monotonically with the perceived amount of distortion, it induces a ranking according to the amount of distortion. This ranking should correlate with the ranking produced by the human observers.

4.1. Stimuli

The two original 24 bits colour reference images used in this study are shown in Fig. 1. These images were selected because they show a significant amount of spatial detail on different levels of resolution in combination with a large variation in colours. The 'Mandrill' image is 512×512 pixels in size, and the 'parrots' image is 384×256 pixels in size, each pixel being represented by three bytes (one for each of the R,G, and B channels). For each reference image, a set of degraded images was

Fig. 1. The original 24 bits *RGB* color images used in the experiments. The first image (a) represents a parrot, and the second image (b) shows the face of a Mandrill monkey.

constructed as follows. First, the original image was transformed into the perceptually decorrelated $l\alpha\beta$ representation. In this space, the individual channels of the reference image were progressively and uniformly quantized. Uniform quantization was performed by dividing the colour space range of the original images into a given number of equally large intervals. Coefficients inside an interval were attributed the value of the lower bound of the interval. The set of quantization intervals was successively chosen such that the fidelity index Q from Eq. (2) was approximately evenly distributed between 0.1 and 0.9. Finally, the resulting quantized images were transformed back to *RGB* space for display.

The effects of uniform quantization in the $l\alpha\beta$ colour space are illustrated in Table 1. This table shows the image of a colour wheel for a range of different quantization levels along the individual dimensions in the $l\alpha\beta$ colour space. The effect of quantization along the l dimension is a reduction of the mean luminance of the image. Quantization along the α and β dimensions results in an overall colour shift along, respectively, the yellow-blue and red-green opponent colour directions. The effects of $l\alpha\beta$ quantization applied to the two reference images from Fig. 1 are shown in Tables 2 and 3, respectively.

We used two different image sets. Images in the first set were quantized in a single channel only (l, α or β), and in seven progressive steps. Images in the second set were distorted in two channels simultaneously (either in l and α, l and β, or α and β), and again in seven steps. As a result, each of the two original reference images has 21 corresponding degraded versions in both sets (resulting from the seven quantization steps in, respectively, each of the three $l\alpha\beta$ channels in the first set, and for each of the three combinations of two individual channels in the second set).

We printed colour hardcopies of the reference images and their corresponding quantized representations on high fidelity glossy photographic paper, using a calibrated 600 dpi laser printer. The printed images were 7×7 cm^2 in size.

4.2. Subjects

Subjects were trained observers, men and women between the ages of 18 and 60. All had normal or corrected-to-normal vision, and no known colour deficiencies.

4.3. Experimental design

We collected rank ordering data from subjects on the sets of printed colour images. First, the subjects were given an original (unprocessed) image that served as a reference. Then, on each run, subjects were handed a randomly ordered set of images, corresponding to progressively quantized versions of the reference image. The subjects

Table 1
Illustration of the effects of uniform quantization in the $l\alpha\beta$ colour space

	l			α			β	
#	Q_L	Image	#	Q_α	Image	#	Q_β	Image
128	0.92		128	0.96		128	0.94	
64	0.76		64	0.89		64	0.85	
32	0.53		32	0.73		32	0.65	
16	0.34		16	0.45		16	0.37	
8	0.24		8	0.18		8	0.14	
4	0.17		4	0.04		4	0.03	
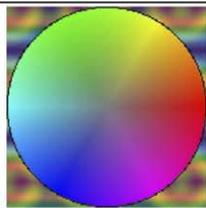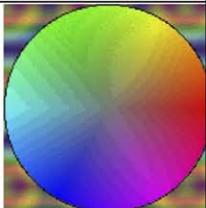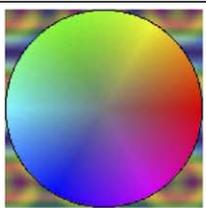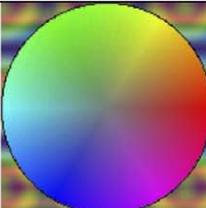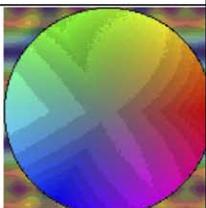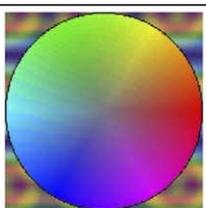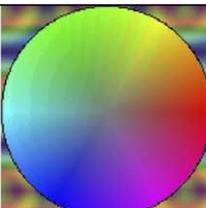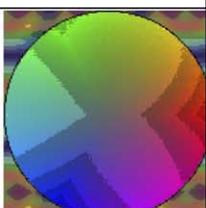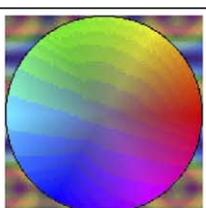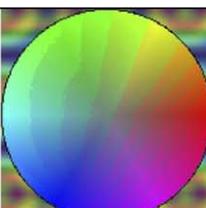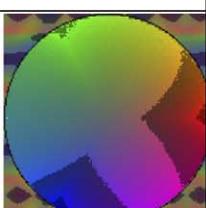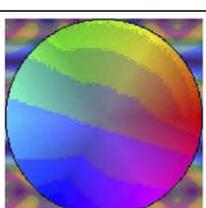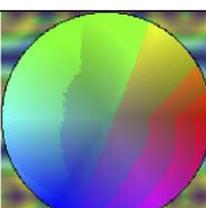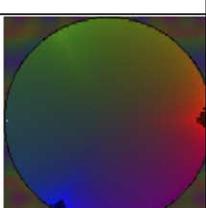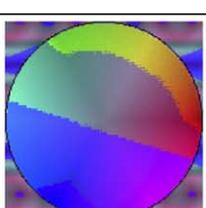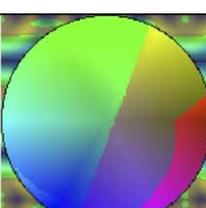
This table shows the number (#) of quantization levels and the associated fidelity factor for each of the three $l\alpha\beta$ channels of progressively quantized images of a.colour wheel. Note that quantization in the $l\alpha\beta$ space corresponds to a reduction of the available number of levels in, respectively, the luminance, yellow-blue and red-green colour opponent dimensions.

Table 2
The parrot images used in the observer fidelity ranking experiment

l			α			β		
#	Q_L	Image	#	Q_α	Image	#	Q_β	Image
64	0.88		64	0.86		128	0.81	
32	0.72		24	0.62		96	0.75	
24	0.64		16	0.50		64	0.62	
16	0.53		12	0.39		40	0.48	
12	0.43		8	0.27		24	0.32	
8	0.30		6	0.18		16	0.24	
4	0.16		5	0.11		8	0.12	

This table shows the number (#) of quantization levels and the associated fidelity factor for each of the three $l\alpha\beta$ channels.

were asked to rank each series of degraded (quantized) colour images by how similar each image was in comparison to the given reference image. The subjects had the ability to change, as required, the classification that they had already done by doing permutations between images until all images were ranked. They had no constraint of time to do this task. We asked the subjects to keep the images on a viewing stand as they sorted them and to keep their viewing distance fixed at approximately 30 cm. At this distance the visual angle of the images is about 4°.

Table 3
The Mandrill images used in the observer fidelity ranking experiment

l			α			β		
#	Q_L	Image	#	Q_α	Image	#	Q_β	Image
32	0.91		64	0.85		64	0.83	
16	0.79		32	0.69		45	0.75	
8	0.60		24	0.59		32	0.66	
7	0.55		16	0.42		16	0.41	
6	0.47		12	0.29		15	0.39	
5	0.36		10	0.20		8	0.17	
4	0.20		8	0.11		4	0.07	
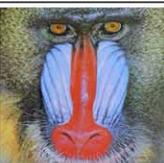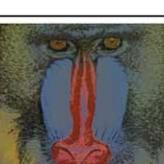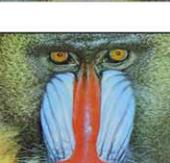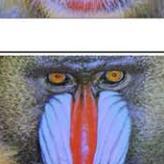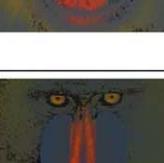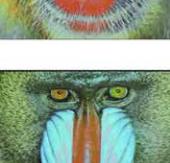
This table shows the number (#) of quantization levels and the associated fidelity factor for each of the three $l\alpha\beta$ channels.

The prints were presented to the observers in a Macbetch SpectraLight II booth, and thus viewed under homogeneous lighting. The luminance reflected from the print was maximally 450 cd/m^2 (white paper). The colour temperature of the illuminant is approximately 6430 K.

To measure how close the rankings produced by the human observers agree with the ranking induced by the fidelity metric we computed the coefficient T_c, which corresponds to the correlation between the set of observer rankings and the ranking induced by the metric [38]. T_c, is the average of the Kendall rank-order correlation coefficients between each ranker and the metric ranking.

We performed three different experiments:

Experiment 1. In the first experiment, 16 observers ranked seven images that were distorted in a single colour channel only, for each of the colour channels separately and consecutively.

Experiment 2. In the second experiment, four subjects ranked a mixed subset of 12 of the images used in the first experiment, such that each set contained four progressively quantized versions of the original image for each of the three colour channels.

Experiment 3. In the third experiment, four observers ranked 21 images, corresponding to seven progressively quantized versions of the original image for each of the three colour channels. The images were selected such that the fidelity metric Q was well distributed for each of the individual colour channels.

4.4. Results

The results of Experiment 1 are shown in Fig. 2. In this experiment the images were distorted in a single channel (l, α or β) only, as described in Section 4.1. The subjects ranked the seven images for each of the three channels separately. Fig. 2 shows that a good correlation between the fidelity metric Q and the average ranking is obtained. Table 4 presents the correlation between the observer rankings and the ranking induced by the computed image fidelity metric, for, respectively, the l, α and β colour channels of the two test images. The correlation coefficient T_c, which can take on values between 0 (no agreement) and 1 (identical rankings), ranges between 0.83 and 0.99. This indicates significant agreement, at the 1% level, between the observer rankings and the ranking induced by the metric. To test the hypothesis that this observed agreement in rankings exceeds what one would expect if the rankings had been made randomly, we computed the z statistic. Since the probability of obtaining a z value greater than the computed values is $p \ll 0.00001$ we may conclude with a high degree of confidence that the raters as a group show strong agreement with the metric ranking. After establishing this fact, we decided to use a smaller number of subjects in the rest of the experiments.
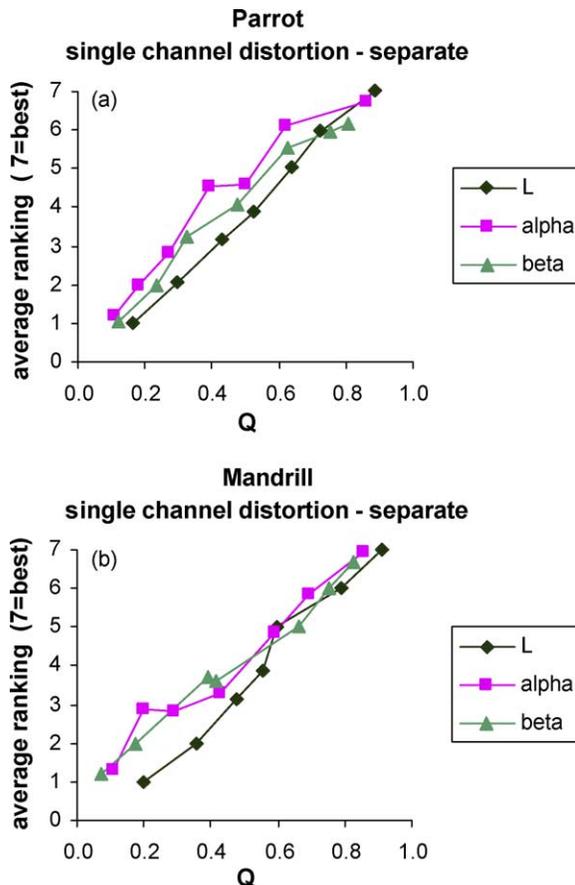
Fig. 2. Results of Experiment 1: the fidelity ranking experiment for quantization along each of the $l\alpha\beta$ colour channels individually. Results are shown for each of the two test images (Parrots and Mandrill). The average subjective ranking is plotted as a function of the image fidelity index Q, resulting from Eq. (2), computed for the individual and corresponding $l\alpha\beta$ colour channel.

The results of Experiment 2 are shown in Fig. 3. For this experiment we selected 12 images out of the overall set of 21 images deployed in Experiment 1, such that the fidelity metric in each channel was well distributed. The subjects ranked these 12 (mixed up) images with respect to their perceived fidelity. The data indicate that the fidelity levels of the three channels start to compete. At the same Q value, distortions in the l channel are attributed the lowest ranking, followed by the distortions in the α channel and the β channel, respectively. Or, in other words, distortions in

Table 4
Correlation between the observer rankings and the ranking induced by the image fidelity metric (as expressed by the correlation coefficient T_c; see text)

	Parrots	Mandrill
l	0.96	0.99
α	0.90	0.83
β	0.83	0.85
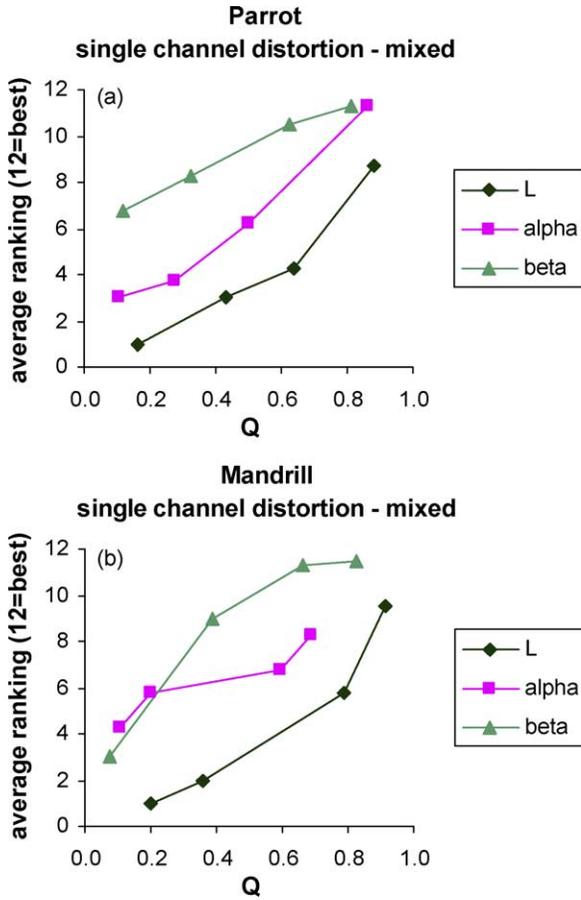
Fig. 3. Results of Experiment 2: the fidelity ranking experiment for a mixed subset of images that have been quantized along each of the $l\alpha\beta$ colour channels individually. Results are shown for each of the two test images (Parrot and Mandrill). The average subjective ranking is plotted as a function of the image fidelity index Q, resulting from Eq. (2), computed for the individual and corresponding $l\alpha\beta$ colour channel.

Fig. 4. Results of Experiment 3: the fidelity ranking experiment for images that were distorted in two channels simultaneously (either in l and α, l and β, or α and β). Results are shown for each of the two test images (Parrots and Mandrill). The average subjective ranking is plotted as a function of the computed objective image fidelity ranking induced by the image fidelity index given in Eq. (10).

the β channel are less visible than the same amount of distortion in the α channel and the l channel.

The results of Experiment 3 are shown in Fig. 4. In this experiment we used 21 images that were distorted in two channels simultaneously (either in l and α, l and β, or α and β). The subjects ranked these 21 (mixed up) images with respect to their perceived fidelity. Although this task was considerably more difficult than the one in Experiments 1 and 2, it could still be performed with satisfaction. We obtained the objective rankings as follows. First, we computed the colour fidelity metric Q_{colour} (Eq. (10)) for different values of the weighting factors w_l, w_α and w_β. The values at which the Spearman rank-order correlation coefficient (between objective and average subjective ranking) was at maximum are, respectively, $w_l = 3.3, w_\alpha = 1.3$ and $w_\beta = 0.9$ for the Parrot image, and $w_l = 2.8, w_\alpha = 0.9$ and $w_\beta = 0.8$ for the Mandrill image. Apparently, the fidelity metric for l dominates the colour fidelity metric.

5. Discussion

In this paper we extend a recently introduced universal grayscale image quality index to a newly developed perceptually decorrelated colour space. The resulting colour image fidelity metric quantifies the distortion of a processed colour image relative to its original version. The metric is computationally simple, which makes real-time implementation feasible.

We evaluated the new colour image fidelity metric through observer experiments in which subjects ranked images according to perceived distortion. The metric correlates strongly with human perception. Hence, it provides a meaningful objective measure of overall image fidelity, and can therefore be used to assess the performance of colour image coding and compression schemes, colour image enhancement algorithms, synthetic colour image generators, and colour image fusion

schemes. This minimises the need for time-consuming and intricate subjective tests in many digital colour image processing applications.

There are a number of issues in the evaluation of the new colour image fidelity metric that are worth further investigation. The relative image fidelity ranking technique used in the present study forces the observers to collapse the ensemble of local fidelity variations over the entire image plane into a single judgment. Observers may use different individual criteria to weigh the relative importance of image distortions. These criteria may depend on the type of distortion and on the composition of the local image region. We may gain insight into the way these factors affect the overall perceived image fidelity by giving the observer a clear instruction on how the judgments should be made. For example, the observer can be instructed to judge image fidelity based on the worst artifact. Or, alternatively, to only judge the fidelity of smooth image regions, or the fidelity of edge regions, etc. When artifacts are suprathreshold, such instructions will change the observer's ratings. Consequently, the instructions given to the observer should reflect the technique used in the fidelity metric to sum errors across space and assign a fidelity rating to the image. Another way of handling this issue is to divide the image into small subregions and have the user specify image fidelity for these different image regions [13]. For some observers the fidelity of the whole image is quite as important as the fidelity of each image element; whereas for others the colour appearance of the background elements is more relevant than the colour appearance of the object elements [41]. Within the scope of this study only a small number of distortions could be investigated. Future research will focus on the perceptual attributes that prevail in the perceptual fidelity judgement, and the effects of their spatial distribution (image composition).

References

[1] A.J. Ahumada, Computational image quality metrics: a review, in: J. Morreale (Ed.), Society for Information Display International Symposium Digest of Technical Papers, Society for Information Display, Playa del Rey, California, 1993, pp. 305–308.

[2] N. Avadhanam, V.R. Algazi, Evaluation of a human-vision-system-based image fidelity metric for image compression, in: A.G. Tescher (Ed.), Applications of Digital Image Processing XXII, The International Society for Optical Engineering, Bellingham, WA, 1999, pp. 569–579.

[3] I. Avcibas, B. Sankur, K. Sayood, Statistical evaluation of image quality measures, Journal of Electronic Imaging 11 (2) (2002) 206–223.

[4] M.R. Bolin, G.W. Meyer, S.O. Aase, A visual difference metric for realistic image synthesis, in: B.E. Rogowitz, T.N. Pappas (Eds.), Human Vision and Electronic Imaging IV, The International Society for Optical Engineering, Bellingham, WA, 1999, pp. 106–120.

[5] G. Buchsbaum, A. Gottschalk, Trichromacy, opponent colours coding and optimum colour information transmission in the retina, Proceedings of the Royal Society of London B 220 (1218) (1983) 89–113.

[6] CIE, Industrial Colour Difference Evaluation (Report CIE 116), International Commission on Illumination CIE, Austria, 1995.

[7] S.J. Daly, The visible differences predictor: an algorithm for the assessment of image fidelity, in: A.B. Watson (Ed.), Digital Imaging and Human Vision, MIT Press, Cambridge, MA, 1993, pp. 179–206.

[8] S.J. Daly, Visible differences predictor: an algorithm for the assessment of image fidelity, in: B.E. Rogowitz (Ed.), Human Vision, Visual Processing, and Digital Display III, The International Society for Optical Engineering, Bellingham, WA, 2000, pp. 2–15.

[9] A.M. Eskicioglu, P.S. Fisher, Image quality measures and their performance, IEEE Transactions on Communications 43 (12) (1995) 2959–2965.

[10] M.D. Fairchild, Color Appearance Models, Addison Wesley, Reading, MA, 1998.

[11] J.E. Farrell, X. Zhang, C.J. van den Branden Lambrecht, Image Quality Metrics Based on Single and Multi-Channel Models of Visual Processing, Proceedings of COMPCON, 1997.

[12] X.R. Fdez-Vidal, A. Toet, J.A. Garcia, J. Fdez-Valdivia, Computing visual target distinctness through selective filtering, statistical features, and visual patterns, Optical Engineering 39 (1) (2000) 267–281.

[13] D.R. Fuhrmann, J.A. Baro, J.R. Cox, Experimental evaluation of psychophysical distortion metrics for JPEG-encoded images, Journal of Electronic Imaging 4 (4) (1995) 397–406.

[14] A. Gaddipatti, R. Machiraju, R. Yagel, Steering image generation with wavelet based perceptual metric, EUROGRAPHICS'97 16 (3) (1997) C241–C251.

[15] J.A. Garcia, J. Fdez-Valdivia, X.R. Fdez-Vidal, R. Rodriguez-Sánchez, Computational Models for Predicting Visual Target Distinctness, PIE Optical Engineering Press, Bellingham, Washington, 2001.

[16] J.A. Garcia, J. Fdez-Valdivia, X.R. Fdez-Vidal, R. Rodriguez-Sánchez, Information theoretic measure for visual target distinctness, IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI 23 (4) (2001) 362–383.

[17] N.C. Griswold, Perceptual coding in the cosine transform domain, Optical Engineering 19 (3) (1980) 306–311.

[18] A.K. Jain, Image data compression: a review, Proceedings of the IEEE 69 (3) (1981) 349–389.

[19] R. Janssen, Computational Image Quality, The International Society for Optical Engineering, Bellingham, WA, 2001.

[20] T.J.W.M. Janssen, F.J.J. Blommaert, A computational approach to image quality, Displays 21 (2000) 129–142.

[21] T.J.W.M. Janssen, F.J.J. Blommaert, Information processing approach to image quality, in: T.N. Pappas (Ed.), Human Vision and Electronic Imaging V, The International Society for Optical Engineering, Bellingham, WA, 2000, pp. 78–87.

[22] N. Jayant, J. Johnston, R. Safranek, Signal compression based on models of human perception, Proceedings of the IEEE 81 (10) (1993) 1385–1422.

[23] M. Jung, D. Léger, M. Gazalat, Univariant assessment of the quality of images, Journal of Electronic Imaging 11 (3) (2002) 354–364.

[24] S.A. Karunasekera, N.G. Kingsbury, A distortion measure for blocking artifacts in images based on human visual sensitivity, IEEE Transactions on Image Processing 4 (6) (1995) 713–724.

[25] Y.-K. Lai, J. Guo, C.-C.J. Kuo, Perceptual fidelity measure of digital color images, in: B.E. Rogowitz, T.N. Pappas (Eds.), Human Vision and Electronic Imaging III, The International Society for Optical Engineering, Bellingham, WA, 2002, pp. 221–231.

[26] Y.-K. Lai, C.-C.J. Kuo, A Haar wavelet approach to compressed image quality measurement, Journal of Visual Communication and Image Representation 11 (1) (2000) 17–40.

[27] P. Le Callet, D. Barba, Perceptual color image quality metric using adequate error, in: B.E. Rogowitz, T.N. Pappas (Eds.), Human Vision and Electronic Imaging VII, The International Society for Optical Engineering, Bellingham, WA, 2002.

[28] M.-S. Lian, Image evaluation using a color visual difference predictor (CVDP), in: B.E. Rogowitz, T.N. Pappas (Eds.), Human Vision and Electronic Imaging VI, The International Society for Optical Engineering, Bellingham, WA, 2001, pp. 175–186.

[29] J.O. Limb, Distortion criteria of the human observer, IEEE Transactions on Systems, Man and Cybernetics SMC 9 (12) (1979) 778–793.

[30] J. Lubin, A visual discrimination model for imaging system design and evaluation, in: E. Peli (Ed.), Vision Models for Target Detection and Recognition, World Scientific, Singapore, 1995, pp. 245–283.

[31] F.X.J. Lukas, Z.L. Budrikis, Picture quality prediction based on a visual model, IEEE Transactions on Communications 30 (7) (1982) 1679–1692.

[32] J. Malo, A.M. Pons, J.M. Artigas, Subjective image fidelity metric based on bit allocation of the human visual system in the DCT domain, Image and Vision Computing 15 (7) (1997) 535–548.

[33] J.L. Mannos, D.J. Sakrison, The effects of a visual fidelity criterion on the encoding of images, IEEE Transactions on Information Theory 20 (4) (1974) 525–536.

[34] H. Marmolin, Subjective MSE measures, IEEE Transactions on Systems, Man and Cybernetics SMC 16 (1986) 486–489.

[35] N.B. Nill, A visual model weighted cosine transform for image compression and quality assessment, IEEE Transactions on Communications 33 (6) (1985) 551–557.

[36] D.L. Ruderman, T.W. Cronin, C.-C. Chiao, Statistics of cone responses to natural images: implications for visual coding, Journal of the Optical Society of America A 15 (8) (1998) 2036–2045.

[37] J.A. Saghri, Image quality measure based on a human visual system model, Optical Engineering 28 (7) (1989) 813–818.

[38] S. Siegel, N.J. Castellan, Nonparametric Statistics for the Behavioral Sciences, second ed., McGraw-Hill, Boston, MA, 1988.

[39] P.C. Teo, D.J. Heeger, Perceptual Image Distortion, Human Vision, Visual Processing, and Digital Display V, The International Society for Optical Engineering, Bellingham, WA, 1994, p. 127–141.

[40] A. Toet, Computing Visual Target Distinctness (Report TNO-TM 1997 A-039), TNO Human Factors, Soesterberg, The Netherlands, 1997.

[41] A. Trémeau, C. Charrier, Influence of chromatic changes on the perception of color image quality, Color Research and Application 25 (3) (2000) 200–213.

[42] C.J. van den Branden Lambrecht, Color Moving Pictures Quality Metric, Proceedings of the IEEE International Conference on Image Processing, 1996, p. 885–888.

[43] C.J. van den Branden Lambrecht, A Working Spatio-Temporal Model of the Human Visual System for Image Restoration and Quality Assessment Applications, Proceedings of ICASSP 96, 2000.

[44] Z. Wang, A.C. Bovik, A universal image quality index, IEEE Signal Processing Letters 9 (3) (2002) 81–84.

[45] A.B. Watson, Toward a Perceptual Video Quality Metric, Human Vision and Electronic Imaging III, The Society of Photo-Optical Instrumentation Engineers, Bellingham,WA, 1998, p. 139–147.

[46] A.B. Watson, J. Hu, J.F. McGowan, Digital video quality metric based on human vision, Journal of Electronic Imaging 10 (1) (2001) 20–29.

[47] A.B. Watson, J. Hu, J.F. McGowan, J.B. Mulligan, Design and Performance of a Digital Video Quality Metric, Human Vision, Visual Processing, and Digital Display IX, The Society of Photo-Optical Instrumentation Engineers, Bellingham, WA, 1999, p. 168–174.

[48] S. Winkler, Perceptual distortion metric for digital color video, in: B.E. Rogowitz, T.N. Pappas (Eds.), Human Vision and Electronic Imaging IV, Bellingham, WA, The International Society for Optical Engineering, 1999, p. 175–184.

[49] H.R. Wu, C.J. van den Branden Lambrecht, M. Yuen, B. Qin, Quantitative Quality and Impairment Metrics for Digital Coded Images and Image Sequences, Proceedings of the Australian Telecommunication Networks and Applications, 1996.

[50] X. Zhang, D.A. Silverstein, J.E. Farrell, B.A. Wandell, Color Image Quality Metric S-CIELAB and its Application on Halftone Texture, 2000.

[51] X. Zhang, B.A. Wandell, Colour image fidelity metrics evaluated using image distortion maps, Signal Processing 70 (3) (1998) 201–214.