

QUANTIFYING COLOR IMAGE DISTORTIONS BASED ON ADAPTIVE SPATIO-CHROMATIC SIGNAL DECOMPOSITIONS

Umesh Rajashekar¹, Zhou Wang², and Eero P. Simoncelli¹

¹ Ctr. for Neural Science, and Courant Inst. of Mathematical Sciences, New York University, USA

² Dept. of Electrical and Computer Engineering, University of Waterloo, Canada

Email: umesh@cns.nyu.edu, zhouwang@ieee.org, eero.simoncelli@nyu.edu

ABSTRACT

We describe a framework for quantifying color image distortion based on an adaptive signal decomposition. Specifically, local blocks of the image error are decomposed using a set of spatio-chromatic basis functions that are adapted to the spatial and color structure of the original image. The adaptive functions are chosen to isolate specific distortions such as luminance, hue, and saturation changes. These adaptive basis functions are used to augment a generic orthonormal basis, and the overall distortion is computed from the weighted sum of the coefficients of the resulting overcomplete decomposition, with smaller weights chosen for the adaptive terms. A set of preliminary experiments show that the proposed distortion measure is consistent with human perception of color images subjected to a variety of different common distortions. The framework may be easily extended to include any form of continuous spatio-chromatic distortion.

Index Terms— color image quality assessment, adaptive signal decomposition

1. INTRODUCTION

A perceptually accurate measure of color image distortion would be useful in many real-world applications. Examples include evaluating the quality of color prints and the effects of gamut mapping in printers, and assessing the performance of algorithms for color image restoration, compression, transmission, white balance correction, and demosaicing.

The simplest and most widely used distortion measure is the mean squared error (MSE). But the average of MSE values computed independently across three color channels does not correlate well with human color perception [1]. This is partly due to the fact that the RGB space used by most imaging devices is not well matched to the color sensitivities of the human visual system. There is an extensive literature on the perception of color, from both scientific and engineering perspectives, that describes color spaces that are more compatible with human perception than RGB. At present, the most widely used of these is the CIELAB space, which is optimized for quantifying perceived color differences of large uniform targets [2].

A distortion measure based on MSE in CIELAB space can perform better than in the original RGB space, but still offers a poor description of human perception, primarily because it does not take into account human sensitivities to spatial structure. The spatial CIELAB (S-CIELAB) metric [3] corrects this by incorporating spatial contrast sensitivity weighting into each of the CIELAB color bands, and provides good predictions of human performance in detecting distortions in halftone color images [1, 4]. There have also

been attempts to extend other spatial-domain error measures by applying them to individual color channels and combining the results. For example, the structural similarity (SSIM) approach [5] has been shown to produce reasonable results when averaged over the channels in transformed color spaces [6, 7].

Despite the successes of these recent attempts, it seems clear that there is room for improvement. In particular, these methods treat errors in the three color channels independently, and despite evidence that humans exhibit such separability when tested with square-wave stimuli [8] there are situations in which distortions can simultaneously affect several color channels. An example is a reduction in saturation, in which all three color channels are affected simultaneously, and in different ways. In this paper, we propose a novel approach for quantifying color image distortion based on an adaptive local spatio-chromatic signal decomposition. The methodology is based on a previously-developed adaptive linear system framework [9], in which signal-adaptive basis functions are used to separate “non-structural” image distortions (those that do not affect the perception of the structures of the objects in a visual scene) from the remaining structural distortions. We develop spatio-chromatic adaptive basis functions to capture changes in luminance, saturation, and hue, and show a set of simple examples demonstrating that a distortion measure based on these provides a better match to human perception than S-CIELAB or color-extended SSIM.

2. ADAPTIVE BASIS FRAMEWORK

We begin with a brief overview of the adaptive basis framework for grayscale image quality assessment developed in [9]. It is perhaps easier to appreciate the intuition behind this framework by considering the distortion of an original image \mathbf{x}_1 in two dimensions as shown in Fig. 1. The two axes could be, for example, the grayscale values of two neighboring pixels. An iso-quality contour around the reference vector \mathbf{x}_1 is a set of points in this space that represent distorted versions of \mathbf{x}_1 with the same visual quality. For example, if human vision were well described by mean squared error, the iso-quality contours would be circles, as illustrated in the Fig. 1.

The fundamental philosophy of the adaptive basis framework is that there are certain directions that are less sensitive to distortions than others, and that *these directions generally depend on the original image*. As a simple example, consider scaling the vector \mathbf{x} such that the distorted vector is $\mathbf{y} = a\mathbf{x}$. If \mathbf{x} corresponds to an image, this might arise from an increase or decrease in the ambient illumination. The human visual system discounts these small changes in illumination (even when they are visible, they are far less objectionable than changes that distort the spatial pattern of intensities, such as additive white noise). Thus we can modify the iso-quality circle

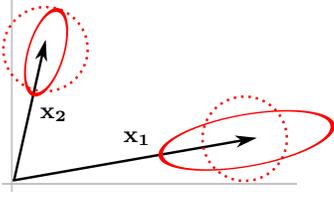


Fig. 1. The dotted circles represents a contour of equal MSE around a vector \mathbf{x} . An adaptive basis that is insensitive to modest changes in illumination strength can generate iso-quality contours that are elongated in the direction of the reference vector. The size and shape of the resulting ellipse can also change depending on \mathbf{x} . See text.

around \mathbf{x}_1 to allow more distortion along the direction of \mathbf{x} than in a direction perpendicular to \mathbf{x} , as shown by the ellipse in Fig. 1.

Modifying the shape of the iso-quality circle by asserting that certain axes in the image space are visually more relevant is embedded in ideas such as the use of the contrast sensitivity function in several image compression algorithms. However, what makes the adaptive basis framework different is that the direction and the shape of the iso-quality ellipse changes depending on the direction of the reference vector, as can be seen by comparing the iso-quality contours for the points \mathbf{x}_1 and \mathbf{x}_2 . The art of the adaptive basis framework lies in defining these special directions and the relative scaling of the iso-quality surface along them. By doing so, we define an adaptive subspace in which the errors are visually less objectionable. Signal-dependent modification of the iso-quality ellipse can also be found in other quality metrics [5, 10].

We can quantify the adaptive basis distortion measure using standard linear algebraic methods [9]. Given a reference signal, \mathbf{x} and a distorted version, \mathbf{y} , we write the error, $\mathbf{y} - \mathbf{x}$, as a weighted sum over a set of basis functions:

$$\mathbf{y} - \mathbf{x} = \mathbf{L}(\mathbf{x})\mathbf{c} = c_1\mathbf{l}_1 + c_2\mathbf{l}_2 + \dots \quad (1)$$

where $\mathbf{L}(\mathbf{x})$ represents a matrix whose columns contain the basis functions \mathbf{l}_k . The matrix is divided into two sub-matrices $\mathbf{L}(\mathbf{x}) = [\mathbf{A}(\mathbf{x}) \mid \mathbf{B}]$, where $\mathbf{A}(\mathbf{x})$ represents the collection of adaptive basis vectors that depend on \mathbf{x} . For example, the adaptive basis that discounts for the illumination change in Fig. 1 is simply the vector \mathbf{x} . The matrix \mathbf{B} contains a set of fixed orthogonal basis vectors spanning the space of the input vector. The distortion measure is now defined as a function of the coefficients, \mathbf{c} . In particular,

$$D(\mathbf{x}, \mathbf{y}) = \min_{\mathbf{c}: \mathbf{L}(\mathbf{x})\mathbf{c} = \mathbf{y} - \mathbf{x}} \|\mathbf{W}\mathbf{c}\|^2 \quad (2)$$

where, \mathbf{W} is a diagonal weight matrix, which can be used to define the shape of the iso-quality ellipse. Since $\mathbf{L}(\mathbf{x})$ is overcomplete, solving for the coefficients, \mathbf{c} , is a least squares optimization problem. The total distortion can be shown to be

$$D(\mathbf{x}, \mathbf{y}) = \left\| \mathbf{W}^{-1}\mathbf{L}(\mathbf{x})^T \left(\mathbf{L}(\mathbf{x})\mathbf{W}^{-2}\mathbf{L}(\mathbf{x})^T \right)^{-1} (\mathbf{y} - \mathbf{x}) \right\|^2 \quad (3)$$

Efficient implementations of (3) are discussed in [9].

This framework is related to methods that use a function $\mathbf{f}(\cdot)$ to transform the reference and distorted images to a perceptual domain, where distortion is measured as $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\|^2$ [11]. Using a Taylor series approximation, the Jacobian of $\mathbf{f}(\cdot)$ can be shown to produce signal-dependent basis vectors for analyzing the error image [10]. In these approaches, $\mathbf{f}(\cdot)$ is usually designed using physiological [11]

or psychophysical models of human vision [2, 10]. In contrast, our adaptive basis vectors are based on assumptions about the visual system's invariance to changes in imaging/viewing conditions.

3. COLOR ADAPTIVE BASIS

In this section, we describe our choice of the adaptive and fixed basis functions for measuring color distortions. We hypothesize that human judgments of image quality are relatively insensitive to small changes in the viewing or imaging conditions. For example, a slight decrease or increase in the illumination level, even if visible, does not cause a noticeable degradation in the appearance of the image. Similarly slight shifts in the spectral properties of the illuminant, or in the saturation of colors is likely to be discounted by the human visual system when evaluating the quality of color images. Finally, as suggested by the so-called MacAdam ellipses [12], minor hue changes are discounted as well. Accordingly, we have developed six adaptive basis functions to reflect these color features. Consider a reference image, \mathbf{x} , of size N pixels in the RGB space consisting of a red channel $\mathbf{r} = [r_1 \dots r_N]$, a green channel, $\mathbf{g} = [g_1 \dots g_N]$, and a blue channel, $\mathbf{b} = [b_1 \dots b_N]$. The six adaptive basis are:

- **Red, Green, & Blue:** The column vectors $\mathbf{a}_1 = [r_1 \ 0 \ 0 \ r_2 \ 0 \ 0 \ \dots \ r_N \ 0 \ 0]^T$, $\mathbf{a}_2 = [0 \ g_1 \ 0 \ 0 \ g_2 \ 0 \ \dots \ 0 \ g_N \ 0]^T$, and $\mathbf{a}_3 = [0 \ 0 \ b_1 \ 0 \ 0 \ b_2 \ \dots \ 0 \ 0 \ b_N]^T$, each of length $3N$ are used to account for changes in the white balance of an image or changes in the spectrum of the illuminant.
- **Luminance:** Define the luminance at pixel k to be $l_k = (r_k + g_k + b_k)/3$. The basis $\mathbf{a}_4 = [l_1 \ l_1 \ l_1 \ l_2 \ l_2 \ l_2 \ \dots \ l_N \ l_N \ l_N]^T$ is used to account for small changes to the luminance.
- **Chroma:** The basis $\mathbf{a}_5 = \mathbf{x} - \mathbf{a}_4$ captures changes in chroma. It is easy to visualize this vector in the HSV color space.
- **Hue:** In the HSV space, hue is measured as a circular variable in a plane that contains the chroma vector and is normal to the luminance vector. We use a linear approximation and define the hue adaptive basis, h_k , for a pixel location, k , as the cross product between, the chroma and the luminance of that pixel. The adaptive basis is defined as the vector $\mathbf{a}_6 = [h_1 \ h_2 \ \dots \ h_N]^T$.

The six column vectors are rescaled to have unit norm, and combined to form the matrix $\mathbf{A} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3 \ \mathbf{a}_4 \ \mathbf{a}_5 \ \mathbf{a}_6]$. We then used the identity basis of size $3N \times 3N$ for the fixed non-adaptive basis, \mathbf{B} . The weight matrices, \mathbf{W}_A and \mathbf{W}_B associated with adaptive and fixed basis vectors respectively, control the relative amount of energy captured by that basis. Here, we assumed a fixed weight of 0.1 for all the adaptive basis except the hue, \mathbf{a}_6 , which was assigned a weight of 0.5. This is based on our intuition that changes in hue are more noticeable than changes in the other adaptive basis directions. The weights for the fixed basis functions were set to unity.

For our analysis, we compute the distortion metrics locally. Given an original image, \mathbf{x} and a distorted image, \mathbf{y} , we extract a 3×3 window at each spatial location from the original image to create the basis functions, $\mathbf{L}(\mathbf{x})$. The error vector, $\mathbf{y} - \mathbf{x} = [\Delta r_1 \ \Delta g_1 \ \Delta b_1 \ \dots \ \Delta r_N \ \Delta g_N \ \Delta b_N]^T$ at the corresponding location is then analyzed to compute a local distortion value according to (3). This results in a distortion map across the entire image, which is then averaged to report an overall distortion value.

4. RESULTS

Although there exist color image databases such as the UTEXAS LIVE (<http://live.ece.utexas.edu/research/quality/>) and Cornell A57



Fig. 2. Examples of distortions to the ‘Peppers’ image. The left column shows distortions that do not interfere with local image structures. The right column shows distortions that can interfere with local image structures. Distortion values for various metrics are shown in Table. 1.

(<http://foulard.ece.cornell.edu/dmc27/vsnr/vsnr.html>), the distortion types are mainly of a spatial nature. As a result we used two specific types of color and spatio-chromatic distortions to demonstrate the performance of our method. First, we introduced global distortions along the adaptive basis directions as shown in the left column of Fig. 2. In particular, we simulated the following distortions: an incorrect white balance setting, where the red, green, and blue channels were each multiplied by a different scalar; an overall reduction in the brightness of the image; a slight decrease in the chroma of each color; and finally a shift in the hue of each pixel. Most of these distortions (except of the consistent hue shift) represent real world scenarios. The second set of distortions involved modifying the original image in directions that were not necessarily along the six adaptive basis directions. In particular, as shown in the right column of Fig. 2, the original image was subjected to corruption by white noise, blurring, JPEG/JPEG2000 compression. We also show a practical example of gamut mapping in which the sRGB gamut of the image was mapped to a printer’s smaller CMYK gamut.

Each of these distorted images were analyzed using four color quality measures. First, we compute the mean squared error between the original and the distorted image, averaged over all color channels. Second, we applied the grayscale SSIM metric [5] to each color channel and used the geometric mean as an overall distortion measure. The SSIM values are between 0 and 1, with larger values indicating better quality. Third, we used S-CIELAB [3], with parameters set to match our viewing scenario. The mean value of the resulting S-CIELAB map was used as the distortion metric. Finally, we computed the adaptive basis error as given by (3). We also report the amount of distortion captured by the adaptive basis, $D_A = \|\mathbf{W}_A \mathbf{c}_A\|^2$, and those by the fixed basis, $D_B = \|\mathbf{W}_B \mathbf{c}_B\|^2$. All distortions (except the gamut mapping and large offset in white balance) were adjusted so that they had approximately the same mean squared error. The results are shown in Table 1.

The “natural” distortions in the left column of Fig. 2 are, by design, less visually noticeable than the “unnatural” distortions introduced by white noise, JPEG and JPEG 2000 compression, and blurring. The values reported by our method are consistent with this observation for reasons that are easily understood by considering the relative values of the adaptive and non-adaptive components: the errors for the natural distortions are largely represented using the adaptive basis elements, which are then weighted by small values for computing the overall distortion value. The errors for the unnatural distortions are primarily represented using the fixed basis. For example, in the case of white noise corruption, the error is decorrelated with the image, and thus not well represented by the adaptive basis functions. The gamut mapping example involved clipping/scaling of the hue and saturation of the source gamut to simulate newspaper prints, resulting in activity in both the adaptive and non-adaptive basis. However, the distortion is not as noticeable as other unnatural distortions. On the other hand, the S-CIELAB values reported in Table 1 are not consistent with these basic perceptual observation. Specifically, the values are smallest for white noise, JPEG and JPEG 2000 artifacts, and blur. This is somewhat expected: as described in [1], the method is not able to produce accurate predictions of error visibility in JPEG images. The color-extended SSIM index correlates reasonably well with perception for most distortions, except for white noise. The large change in the white balance (Fig. 2(j)) is an interesting example where the distortions exist along the adaptive red, green, and blue directions, but the changes are significant enough to produce a large value even along the adaptive basis. Since SSIM is computed on individual color channels, and the distortion along each channel is smooth, the resulting SSIM value is large and

	MSE	SSIM	S-CIELAB	$D(D_A, D_B)$
Original	0	1.000	0.00	0.00 (0.00, 0.00)
Luminance	200	0.983	11.66	1.60 (1.59, 0.01)
White Balance	201	0.972	12.93	1.33 (1.31, 0.02)
Chroma	203	0.921	5.99	0.68 (0.55, 0.12)
Hue	211	0.892	8.84	3.88 (1.31, 2.57)
Gamut	289	0.879	13.01	4.56 (1.86, 2.70)
Large White Balance	1129	0.908	25.26	8.31 (8.24, 0.07)
White Noise	199	0.558	2.72	39.72 (0.42, 39.31)
JPEG2000	194	0.733	6.26	59.95 (2.21, 57.74)
JPEG	206	0.720	6.14	61.95 (1.97, 59.98)
Blur	200	0.768	4.49	68.27 (2.91, 65.37)

Table 1. Distortion values for the images in Fig. 2 using various image quality metrics. The rightmost column shows the error captured by our method with the error captured by the adaptive and fixed part shown in parenthesis. The rows are approximately ordered according to decreasing visual quality in Fig. 2.

thus predicts a good quality for this distorted image.

5. CONCLUSION

We developed a new form of color image quality metric based on the adaptive basis framework of [9] and showed that the framework is effective at discounting distortions corresponding to naturally occurring changes in viewing or illumination conditions, while still being able to capture other distortions such as blurring, noise, and compression artifacts. The framework is quite flexible: the basis functions at each spatial location can be any function (linear or nonlinear) of the reference image at that location. But at the same time, once the basis has been computed from the reference image, the framework provides a quadratic error metric that allows for easy optimization. We expect this will prove invaluable when using this metric as an objective function for optimizing image processing applications such as denoising or compression.

The examples shown here serve as an initial demonstration of the feasibility of the method, but there is much room for refinement in the choice of basis functions and weights. Incorporating other adaptive basis directions is as easy as adding the appropriate signal dependent basis function to the matrix, $\mathbf{A}(\mathbf{x})$ in (1). We envision adding bases to account for spatio-chromatic distortions that occur naturally as a result of changes in lighting or viewpoint, under the assumption that the human visual system is designed to discount these changes [5]. The non-adaptive basis could be replaced by Fourier or wavelet type basis functions, and the associated weights could be set to reflect the contrast sensitivity behavior of the human visual system. By using fixed weights, \mathbf{W} , and unit norm vectors in $\mathbf{L}(\mathbf{x})$, the iso-quality ellipse is constrained to be a fixed size. We are currently developing a simple protocol for estimating the adaptive weights from perceptual measurements. We are also working to develop a multiscale framework for adaptive bases that allows for coarse-to-fine refinement in image distortion analysis.

6. REFERENCES

- [1] X. Zhang and B. A. Wandell, "Color image fidelity metrics evaluated using image distortion maps," *Signal Processing*, vol. 70, no. 3, pp. 201–214, Nov. 1998.
- [2] "Industrial color-difference evaluation, publication CIE 116-95," 1995.
- [3] X. Zhang and B. A. Wandell, "A spatial extension of cielab for digital color image reproduction," *SID Journal*, 1997.
- [4] X. Zhang, D.A. Silverstein, J.E. Farrell, and B.A. Wandell, "Color image quality metric S-CIELAB and its application on halftone texture visibility," in *Compton '97. Proceedings, IEEE*, 1997, pp. 44–48.
- [5] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600–612, April 2004.
- [6] A. Toet and M. P. Lucassen, "A new universal colour image fidelity metric," *Displays*, vol. 24, no. 4-5, pp. 197–207, Dec. 2003.
- [7] N. Bonnier, F. Schmitt, H. Brettel, and S. Berche, "Evaluation of spatial gamut mapping algorithms," in *Proc. of the 14th Color Imaging Conference*, Scottsdale Arizona, 2006.
- [8] A. B. Poirson and B. A. Wandell, "Appearance of colored patterns: pattern-color separability," *J Opt Soc Am A Opt Image Sci Vis*, vol. 10, no. 12, pp. 2458–2470, Dec 1993.
- [9] Z. Wang and E. P. Simoncelli, "An adaptive linear system framework for image distortion analysis," in *Proc. IEEE Int. Conf. on Image Processing ICIP 2005*, 11–14 Sept. 2005, vol. 3, pp. III–1160–3.
- [10] A. M. Pons, J. Malo, J. M. Artigas, and P. Capilla, "Image quality metric based on multidimensional contrast perception models," *Displays*, vol. 20, no. 2, pp. 93–110, Aug. 1999.
- [11] Javier Portilla, "A perceptual bayesian estimation framework and its application to image denoising," in *EUSIPCO*, 2006.
- [12] R. S. Berns, *Billmeyer and Saltzman's Principles of Color Technology, 3rd Edition*, Wiley-Interscience, 3 edition, Mar. 2000.