



A Perception-based Color Space for Illumination-invariant Image Processing

Citation

Chong, Hamilton, Steven J. Gortler, and Todd Zickler. 2008. A perception-based color space for illumination-invariant image processing. 35th annual conference on computer graphics and interactive techniques (SIGGRAPH 2008), August 11-15, Los Angeles, Calif. ACM Transactions on Graphics 27 (3): article 61, 1-7.

Published Version

<http://dx.doi.org/10.1145/1360612.1360660>

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:2640577>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

A Perception-based Color Space for Illumination-invariant Image Processing

Hamilton Y. Chong*
Harvard University

Steven J. Gortler†
Harvard University

Todd Zickler‡
Harvard University

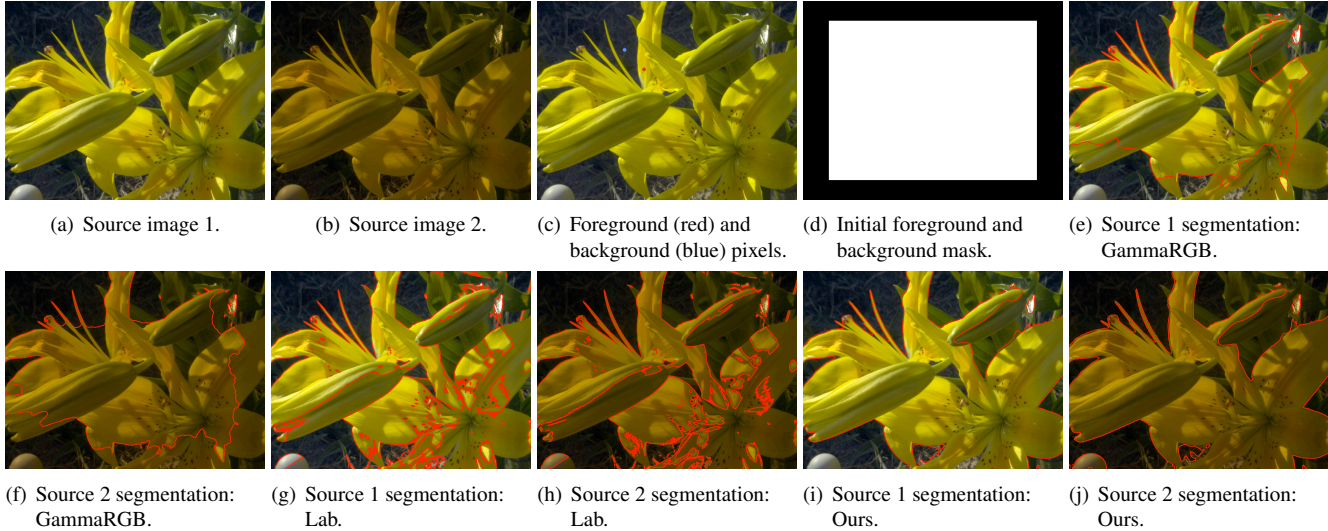


Figure 1: Segmentations of the same flower scene under two different illuminants. For (nonlinear) RGB and Lab color spaces, segmentation parameters for the two images were tweaked individually (optimal parameters for the image pairs were separated by 2 and 5 orders of magnitude respectively); for our color space, the same parameters were used for segmenting both images.

Abstract

Motivated by perceptual principles, we derive a new color space in which the associated metric approximates perceived distances and color displacements capture relationships that are robust to spectral changes in illumination. The resulting color space can be used with existing image processing algorithms with little or no change to the methods.

CR Categories: I.3.m [Computer Graphics]: Misc.—Perception

Keywords: perception, image processing, color space

1 Introduction

Given the importance of color processing in computer graphics, color spaces abound. While existing color spaces address a range of needs, none of them simultaneously capture two notable properties

*e-mail: hchong@post.harvard.edu

†e-mail: sjg@cs.harvard.edu

‡e-mail: zickler@seas.harvard.edu

required by a large class of applications that includes segmentation and Poisson image editing [Pérez et al. 2003]. In this work, we present a new color space designed specifically to address this deficiency.

We propose the following two color space desiderata for image processing:

- (1) difference vectors between color pixels are unchanged by re-illumination;
- (2) the ℓ_2 norm of a difference vector matches the perceptual distance between the two colors.

The first objective restricts our attention to three-dimensional color space parameterizations in which color displacements, or gradients, can be computed simply as component-wise subtractions. Furthermore, it expresses the desire for these color displacements—the most common relational quantities between pixels in image processing—to be invariant to changes in the spectrum of the scene illuminant. Illumination invariance is useful for applications where processing is intended to operate on intrinsic scene properties instead of intensities observed under one particular illuminant. For example, Figure 1 shows that when segmenting an image using usual color spaces, images that differ only in the scene’s illumination during capture can require nontrivial parameter tweaking before the resulting segmentations are clean and consistent. And even then, the segmentations are not necessarily reliable. Figure 4 illustrates the extreme sensitivity a Poisson image editing algorithm exhibits when the illuminant of the foreground object does not match the background’s illumination.

The second condition implies that the standard computational method of measuring error and distance in color space should match

the perceptual metric used by human viewers.

These desiderata have direct correspondence to widely-studied perceptual notions that possess some experimental support. Desideratum (1) corresponds to subtractive mechanisms in color image processing and to human color constancy. Desideratum (2) relates to the approximate “flatness” of perceptual space.

Subtractive mechanisms refer to the notion that humans perform spatial color comparisons by employing independent processing per channel [Kimmel et al. 2002; McCann 2005] and that such per-channel comparisons take a subtractive form. Physiological evidence for subtraction comes from experiments such as those revealing lateral inhibition in the retina [Fairchild 2005] and the existence of double opponent cells in the visual cortex (where each type provides a spatially opponent mechanism for comparing a select chromatic channel) [Palmer 1999].

Color constancy refers to the fact that the perceived color of an object in a scene tends to remain constant even when the illumination (and thus the tristimulus data recorded at the retina) is changed drastically. Color constancy is a surprising feature of the visual system because the light seen reflected from objects depends on the full spectrum of the illuminant and the per-wavelength attenuation of the material. Given a scene with arbitrary illuminant spectra and material reflectances, there is no reason that constancy should be even close to possible. Indeed, humans do not have perfect color constancy, but our color percepts are nonetheless surprisingly stable [Palmer 1999]. (Sometimes the term “chromatic adaptation” is used instead to emphasize the inability to achieve perfect constancy [Fairchild 2005].)

The approximate flatness of perceptual space refers to the relative empirical success—as evidenced by color spaces such as CIE $L^*a^*b^*$ and CIE $L^*u^*v^*$ —in approximating perceptual distance with Euclidean distance when the Euclidean computation is preceded by an appropriate nonlinear reparameterization of color space¹. There is ample evidence that perceptual space is unlikely to be exactly flat [Wyszecki and Stiles 1982], but there is also a fair amount of evidence that a large portion of it may be usefully treated as flat for many applications [Wyszecki and Stiles 1982; Judd and Wyszecki 1975].

Motivated by graphics applications, we explore the ramifications of enforcing these desiderata. The fact that both of our desiderata relate to human perceptual principles for low-level vision suggests that there is value in addressing each via an early-stage model, such as the choice of color space as opposed to the inclusion of additional structures at later stages of visual processing. We point out, however, that although the connections between our model and human perception are certainly suggestive, the utility of our approach does not depend on these assumptions regarding human visual processing being true.

To make our approach tractable, we make one further assumption.

- (3) materials and illuminants in our scenes are such that the effect of relighting is well-approximated by multiplying each tristimulus value (in an appropriate basis) by a scale factor that does not depend on the materials observed.

This assumption is commonly used in color correction algorithms [Finlayson et al. 1993]. Its veracity depends only on the spectral properties of the illuminants, the spectral reflectance of the materials, and the spectral sensitivities of the sensors (e.g., our cones). Evidence giving some quantitative support for this assumption can

¹This property is called flatness since such compatibility with a Euclidean distance metric implies a lack of Riemannian curvature in the underlying manifold.

be found in [Chong 2008; Chong et al. 2007; Finlayson et al. 1993]. We emphasize that this assumption does not require any additional assumptions about human color processing.

Assumption (3) also relates to the (generalized) von Kries coefficient rule, which asserts that the brain is able to discount illumination changes by applying the appropriate inverse scale factors [Fairchild 2005]. Color matching experiments evaluating the generalized von Kries hypothesis are therefore additional (indirect) tests of this assumption’s veracity [Wyszecki and Stiles 1982].

In this paper, we show that conditions (1)-(3) almost completely determine the analytical form of the three-dimensional color space parameterization. There are only a handful of degrees of freedom in this model, and these can be “fit” using measurements provided by perceptual difference experiments.

Once our model is fit to perceptual data, we find that Euclidian distances in our color space better match a standard perceptual difference dataset than the widely used CIE $L^*a^*b^*$ and CIE $L^*u^*v^*$ spaces². Beyond its matching of perceptual data, the utility of our approach is supported by its apparent utility for image processing applications (Figure 1 and Section 5). In particular, we show that both segmentation and Poisson editing can be made robust to illumination changes simply by re-expressing colors in our color space prior to processing.

2 Previous work

There has been extensive research on formulating “uniform color spaces,” or color spaces in which perceptual distances correspond to Euclidean distances [Judd and Wyszecki 1975; Wyszecki and Stiles 1982]. The most widely used spaces of this type are the CIE $L^*a^*b^*$ and CIE $L^*u^*v^*$ spaces. Some more recent attempts to predict perceptual distance modify CIE $L^*a^*b^*$ ’s Euclidean formula, but yield only difference equations, not parameterizations [Sharma et al. 2005; Luo 2006]. A disjoint effort was carried out by researchers trying to determine a “line element,” or Riemannian metric, for color space [Wyszecki and Stiles 1982; Vos 2006]. Despite leveraging the more general Riemannian formalism, a couple of the line elements nonetheless correspond to Euclidean spaces (e.g., Helmholtz, Stiles, Bouman-Walraven) [Vos 2006]. The Riemannian metrics do not seem to be in wide use.

A large body of work also surrounds the human visual perception issues we touch upon. For more on the theories and their experimental assessments, see [Wyszecki and Stiles 1982; Palmer 1999] and the references therein. Significant research has also focused on computational aspects such as spatial methods for computing per-channel gain factors and estimating the illuminant [Kimmel et al. 2002; Finlayson et al. 1993; McCann 2005]. Our approach is agnostic with regards to these methods and largely enables avoiding such computations altogether. Considerable work has also gone into investigating the conditions under which models such as the von Kries coefficient rule can work perfectly [Chong et al. 2007; Finlayson et al. 1993; West and Brill 1982]. We build upon [Chong et al. 2007] in section 3.1.1.

In this work, we pursue a functional equations approach to developing a new color space. This more directly incorporates empirical data into performing model selection in addition to using data for parameter estimation [Castillo and Ruiz-Cobo 1992]. Perhaps the first to apply functional equations to psychophysics was Fechner,

²In this regard, our method does not outperform some color difference equations such as CIEDE2000 and CIE94. As discussed in Section 4, however, such equations do not satisfy either of our desiderata, and this limits their applicability for many graphics applications.

who proposed an explanation of Weber’s Law—a law relating physical magnitudes of stimuli to perceived intensities [Fechner 1877]. His analysis, however, applies only for scalar judgments such as “brightness,” not for the space of colors more generally [Luze 1993]. Mathematical psychologists have continued to refine the use of functional equations, but efforts seem mostly geared toward other areas of social science [Aczel et al. 2000]. One exception is the work of Resnikoff, who derived two candidates for the structure of color space—given a “homogeneity” assumption [Resnikoff 1975; Resnikoff 1989]. One candidate, Resnikoff points out, is unlikely to be a model for actual color vision; the other is the Stiles line element. We start with somewhat different perceptual principles that have exhibited some quantitative success.

3 A Perceptual Metric

Here we formalize the color space conditions presented in the introduction. Let us denote by \vec{x} the tristimulus values of a sensor represented in XYZ coordinates. Let us denote by F the 3D color space parameterization we wish to solve for.

According to our third assumption, when colors are written in some “appropriate” basis, the effect of relighting is simply a per-channel multiplication. We represent the fixed change to an appropriate basis by the matrix B . (In Section 3.1.1, we show how to select an “optimal” matrix B —and thus an optimal color basis—using a database of measured spectral distributions.) Under this assumption, the effect of relighting can be written as

$$\vec{x} \mapsto B^{-1}DB\vec{x}, \quad (1)$$

where D is a diagonal matrix depending only on the illuminants and not the materials. Under our reparameterization F , then, color displacements transform as

$$F(\vec{x}) - F(\vec{x}') \mapsto F(B^{-1}DB\vec{x}) - F(B^{-1}DB\vec{x}'). \quad (2)$$

Desideratum (1) states that displacements are unchanged by relighting. Thus, for all diagonal matrices D , and color vectors \vec{x} and \vec{x}' we must have

$$F(\vec{x}) - F(\vec{x}') = F(B^{-1}DB\vec{x}) - F(B^{-1}DB\vec{x}'). \quad (3)$$

Desideratum (2) states that for any two colors \vec{x} and \vec{x}' , the perceptual distance between them is computed as

$$d(\vec{x}, \vec{x}') = \|F(\vec{x}) - F(\vec{x}')\|, \quad (4)$$

where $\|\cdot\|$ denotes the usual ℓ_2 -norm.

Given these assumptions, Appendix A proves that the nonlinear function F of Equation (3) must take the form

$$F(\vec{x}) = A(\hat{\ln}(B\vec{x})), \quad (5)$$

where A and B are invertible 3×3 matrices and $\hat{\ln}$ denotes the component-wise natural logarithm. The matrix B transforms color coordinates to the basis in which relighting (best) corresponds to multiplication by a diagonal matrix, while the matrix A provides degrees of freedom that can be used to match perceptual distances.

3.1 Parameter Estimation

We first optimize for the matrix B using material and illuminant spectral data. Matrix A is then estimated using perceptual distance data.

3.1.1 Estimating the B matrix

To compute the B matrix, we used an approach adapted from [Chong et al. 2007], which we briefly review here.

Let us consider the relighting assumption of Equation (1). This is necessarily an approximation, since given scenes with arbitrarily complicated material spectral reflection functions and arbitrarily complicated illuminant spectral densities, such a simple model—where the D matrix does not depend on the spectral properties of the material and B is some pre-determined matrix mapping to a fixed preferred color basis—should not be possible. However, it can be shown that [Chong et al. 2007]:

- Such a relighting model can in fact work if the collection of illuminants and materials contained in scenes satisfies a certain “rank condition” described below.
- When the rank condition is satisfied by a collection of materials and lights, the B matrix is fixed (up to 3 scale factors).
- A large collection of naturally-occurring materials and lights can be well-approximated while satisfying the rank condition.
- Given such a dataset, one can numerically solve for a locally-optimal B matrix.

Both the “rank condition” and the description of the algorithm for estimating B require an auxiliary definition. Given a database of reflectance and illuminant spectra ($\{R_i\}$ and $\{L_j\}$ respectively), the XYZ color matching functions ($\{M_k\}_{k=1,2,3}$) can be used to define a *measurement tensor* (3D data block) indexed by reflectance, illuminant, and sensor (λ denotes wavelength):

$$T_{ijk} := \int R_i(\lambda)L_j(\lambda)M_k(\lambda)d\lambda. \quad (6)$$

A theorem of [Chong et al. 2007] then states that the relighting model of Equation (1) holds for the input reflectances and illuminants if and only if the measurement tensor is at most rank three—i.e., the tensor is expressible as the sum of three tensor outer-products (outer product denoted by \circ):

$$T = \sum_{n=1}^3 \vec{r}_n \circ \vec{l}_n \circ \vec{m}_n. \quad (7)$$

For notational simplicity, let the outer-product vectors $\{\vec{r}_n\}$, $\{\vec{l}_n\}$, and $\{\vec{m}_n\}$ correspond to the reflectance, illuminant, and sensor axes of the measurement tensor respectively. It turns out that such a decomposition of a rank three tensor into these outer-product vectors is almost always unique (modulo permutations and scalings).

One can compute the decomposition above for any rank three tensor, and once computed, the three vectors $\{m_n\}$ of the decomposition give us the columns of B^{-1} . So the desired matrix B is computable by collecting these vectors and inverting them.

When the measurement tensor is not rank three (which is more often the case), one can first compute the “closest” tensor that is rank three and then solve for an optimized B matrix via the decomposition. To jointly determine a close rank three tensor and its decomposition, [Chong et al. 2007] use an alternating least squares strategy in which the three outer-product vectors associated with each of the three tensor axes are resolved in turn. In each iteration, the vectors associated with two of the axes are fixed as constant while the free vectors are chosen so that Equation (7) best matches the given data tensor in a least squares sense. The alternating process is repeated until convergence.

While [Chong et al. 2007] measured error in the XYZ color space, we used the CIEDE2000 difference equation to measure perceptual

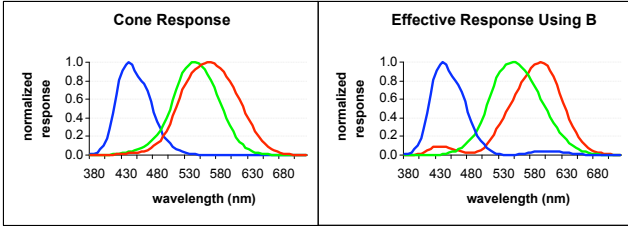


Figure 2: Color matching functions for cone sensors and effective sensors with B matrix applied.

error. CIEDE2000 is the latest CIE standard for computing perceptual distance and gives the best fit of any method to the perceptual distance datasets used by the CIE [Luo 2006]. We also constrained the solution to encompass the entire cone of human visible colors, so that the logarithm of Equation (5) is always well defined. In our case, our extra constraints (measuring error in CIEDE2000 sense and enforcing that the gamut include the entire human visible range) makes each step a nonlinear least squares optimization.

We used the SFU dataset for optimizing the B matrix [Barnard et al. 2002]. The database includes eight spectra simulating daylight (as in [Chong et al. 2007], we excluded fluorescent spectra since there is some indication that humans exhibit less color constancy under such lighting conditions). The database also includes 1,995 material reflectance spectra, including those of natural objects and paint chips. For color sensors, we used the CIE 1931 2-degree XYZ matching functions with Judd 1951 and Vos 1978 modifications [Wyszecki and Stiles 1982].

Our optimization procedure returned a rank three tensor approximation in which the colors had a root-mean-squared error of 1.1 CIEDE2000 units (a rough rule of thumb is that 1 CIEDE2000 unit is about 1 or 2 just noticeable differences). The mapping to our desired color basis for color constancy (assuming input coordinates in XYZ space) is given by

$$B = \begin{bmatrix} 9.465229 \times 10^{-1} & 2.946927 \times 10^{-1} & -1.313419 \times 10^{-1} \\ -1.179179 \times 10^{-1} & 9.929960 \times 10^{-1} & 7.371554 \times 10^{-3} \\ 9.230461 \times 10^{-2} & -4.645794 \times 10^{-2} & 9.946464 \times 10^{-1} \end{bmatrix}. \quad (8)$$

Figure 2 displays the matching functions of the effective sensors when the B matrix is used. The cone matching functions are shown for comparison. Note the “sharpening” of the curves.

3.1.2 Estimating the A matrix

We obtained two perceptual distance datasets for optimizing A . The RIT-Dupont dataset provides 312 pairs of colors that have been experimentally determined to be separated by a small and equal perceptual distance [Berns et al. 1991]. The Witt dataset provides both small and medium distance data for 418 color pairs [Witt 1999].

Instead of optimizing for A directly, we first optimized for the matrix $G := A^T A$. If for each color pair i , we define the displacement vector $\vec{v}_i := (\ln(B \vec{x}_i) - \ln(B \vec{x}'_i))$ and label its corresponding perceptual distance d_i , then we can rewrite Equation (4) as:

$$d_i(\vec{x}, \vec{x}') = \left(\vec{v}_i^T G \vec{v}_i \right)^{1/2}. \quad (9)$$

Since the \vec{v}_i and d_i are constants in the optimization, squaring Equation (9) leaves us with a set of constraints that are linear in the entries of G . Since G must be symmetric, we only have six variables to solve for. Such an overconstrained linear system is easily solved

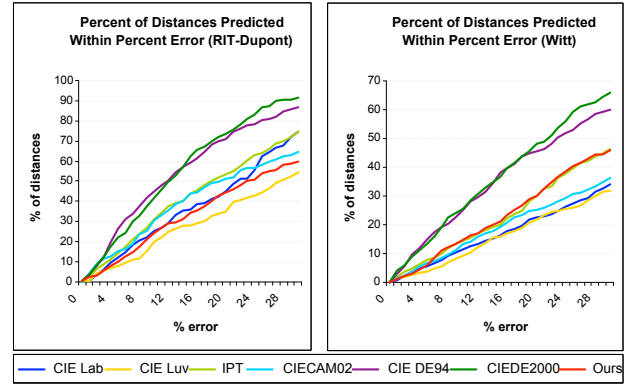


Figure 3: Perceptual error performance of our distance function against other color metrics. None of the other approaches satisfy all our desired requirements. For instance, CIEDE2000 and CIE94 do not provide 3D parameterizations, so it is not clear that they can be used in applications such as Poisson editing. Our metric was fit to the RIT-Dupont data.

in a least-squares sense. As an aside, we also desire that G be positive definite; fortunately, our optimization naturally produced such an optimum, so no additional constraint was necessary.

To recover A from G , we must specify three more degrees of freedom (corresponding to rotations). We chose A to align with an eigenbasis for G . Computing an SVD of G gives us a decomposition of the form

$$G = U^T \Lambda U, \quad (10)$$

where U is a rotation matrix and Λ is diagonal. If we denote the diagonal matrix whose elements are the square roots of Λ 's by $\sqrt{\Lambda}$, we have

$$A = \sqrt{\Lambda} \cdot U. \quad (11)$$

For our purposes, we trained on the RIT-Dupont dataset and tested on both RIT-Dupont and Witt. The resulting matrix is given by

$$A = \begin{bmatrix} 2.707439 \times 10^1 & -2.280783 \times 10^1 & -1.806681 \\ -5.646736 & -7.722125 & 1.286503 \times 10^1 \\ -4.163133 & -4.579428 & -4.576049 \end{bmatrix}. \quad (12)$$

In evaluating methods for computing perceptual distance, we were careful to not penalize uniform scalings of the distances. Rescaling a distance function does not change the color space geometry; it corresponds only to a different choice of units (e.g., although the numbers change depending on whether one uses inches or centimeters to measure spatial distance, such changes do not reflect *geometric* differences). For the RIT-Dupont dataset, in which all distances ought to be the same (without loss of generality, 1), we rescale computed distances such that the median distance is 1. For the Witt dataset, which has experimentally determined distances of varying magnitudes, we choose the rescaling that minimizes the ℓ_1 percentage error.

4 Metric Results

Figure 3 shows our metric’s performance on the RIT-Dupont and Witt datasets compared against other distance metrics. For RIT-Dupont’s small distance data, our metric performs comparably to CIE $L^*a^*b^*$ and CIE $L^*u^*v^*$. For Witt’s small and medium distance data, it outperforms both. This suggests that our mapping (F) is

more successful at mapping the geodesics of color space to straight lines.

The CIEDE2000 and CIE94 difference equations provide the best predictions of perceptual distance; however, these difference equations do not come with a 3D color space parameterization, so it is unclear whether they can be plugged into applications such as Poisson image editing. Even assuming these difference equations could be applied, the complexity of their formulas makes them less amenable to efficient computation, so many applications will likely benefit from simpler models that may sacrifice some of the accuracy. The work of [Sharma et al. 2005] also points out that CIEDE2000’s fit comes at the cost of other desirable features such as the continuity of color space.

5 Applications

By transforming the tristimulus values of an image according to Equation (5), one obtains color descriptors that are (approximately) invariant to illumination. In this section we show that we can take advantage of these (approximately) illumination-invariant color descriptors to produce algorithms that operate on the “intrinsic image” instead of working directly on RGB values. We focus in particular on segmentation and Poisson editing, although the same ideas can be applied to other visual tasks such as feature tracking and texture synthesis.

5.1 Segmentation

The objective of segmentation is often to separate a foreground object from background. If a segmentation algorithm is truly selecting “object,” then its segmentation should be invariant to changes in illuminant spectrum. If we assume that our world is (approximately) consistent with our relighting assumption (expressed in Equation (1)), then color differences between neighboring pixels are invariant to illumination changes *when computation is done in our coordinate system*. Thus any segmentation algorithm that is based on color differences (gradients) between neighboring pixels will be (approximately) unaffected by global illumination changes.

We tested such illumination invariance using iterations of a geometric active contour algorithm as implemented by GAC++ [Papandreou and Maragos 2007]. Given a “hyperspectral image” (one possessing captured reflectance spectra for each pixel) from the dataset provided by [Foster et al. 2004], and daylight illuminants provided by [Barnard et al. 2002], we rendered the hyperspectral image under two different illuminants to produce the input images shown in Figures 1(a) and 1(b). Since the segmentation algorithm required an estimated mean color for foreground and background, we chose arbitrary representative pixels for each (Figure 1(c)). We also provided the rough initial guess shown as a foreground/background mask (Figure 1(d)).

For each input image, we ran the segmentation algorithm in gamma-corrected RGB space (GammaRGB), CIE $L^*a^*b^*$ space (Lab), and our perceptual color space (Ours). The only free parameter to the segmentation algorithm is a constant controlling the discount factor for large gradients. For each run of the algorithm in gamma-corrected RGB space and CIE $L^*a^*b^*$ space, we manually tweaked the parameter to give the cleanest and most consistent segmentation between the light and dark image pairs. For the gamma-corrected RGB space, the optimal parameters were separated by two orders of magnitude. For the CIE $L^*a^*b^*$ space, the optimal parameters were separated by five orders of magnitude. When running the algorithm in our color space, we set the parameter once for one of the images, and ran the segmentation using the same parameter value for the other image (doing so for the other spaces

caused the algorithm to fail to even return an answer). The experiments showed that segmentation was relatively robust to illumination changes when run in our color space. Our hope is that such an approach, along with other methods, may help remove the need for manual tweaking of segmentation parameters.

5.2 Poisson Image Editing

Poisson image editing addresses the problem of deleting a portion of some image and filling it in using data from another source image [Pérez et al. 2003]. We would like the algorithm to be insensitive to the illumination used for the inserted object. Once again, in our coordinates, color gradients are (approximately) unaffected by re-illumination. Thus, a Poisson algorithm which only uses the color gradients from its insertion source and works in our coordinates will be approximately invariant to the inserted element’s illumination.

Let us denote the image being edited as I , and its damaged or masked out region by Ω . Let us denote the source image used to fill in Ω by J . In filling in Ω , our goal is to preserve the perceived color relations between the pixels in J as much as possible. This suggests that we solve the following problem [Pérez et al. 2003]:

$$\min_I \int_{\Omega} \|\nabla I - \nabla J\|^2 dA, \text{ subject to } I \text{ being fixed on } \partial\Omega. \quad (13)$$

The Euler-Lagrange equations give us the following conditions:

$$\Delta I_i = \Delta J_i, \quad i = 1, \dots, \# \text{ color components}. \quad (14)$$

Figure 4 demonstrates a typical Poisson edit in which three objects that have been captured under different environmental conditions are inserted into a single scene. We show comparisons with results derived from solving Equation (14) in various other color spaces: gamma-corrected RGB (GammaRGB), linear RGB (LinearRGB), log of linear RGB (LogRGB), IPT [Ebner and Fairchild 1998], CIECAM02 [Moroney et al. 2002; Chalmers et al. 2007], and CIE $L^*a^*b^*$ (Lab).

We also compare against the related method of [Georgiev 2006] (Ratio), which instead solves

$$\Delta \left(\frac{I_i}{J_i} \right) = 0, \quad i = 1, \dots, \# \text{ color components}, \quad (15)$$

where I and J are expressed in linear RGB coordinates.

To demonstrate performance in a case where ground truth is available, we also experiment with the (somewhat contrived) example in which the object to be inserted is actually coming from a version of the original image in its original location, the only difference being that this source image is captured under different illumination conditions.

The experiment is as follows: render a hyperspectral image under two different illuminants (as in Section 5.1); in one image (denoted by I above), excise data in some region (denoted by Ω); use the other image (denoted by J) along with I ’s excision mask to provide source data for filling in Ω . The output of each algorithm can then be compared against the original image I before the data excision. A perfect algorithm is one that produces exact matches.

We ran the experiment for a couple different hyperspectral images and different pairs of illuminant spectra. Figure 5 shows one of the runs. These images have been designed for display on a monitor calibrated with a gamma of 2.2 and D65 white point. Ratio and LogRGB were both able to handle mild forms of illumination change because they account for multiplicative scalings of linear RGB.

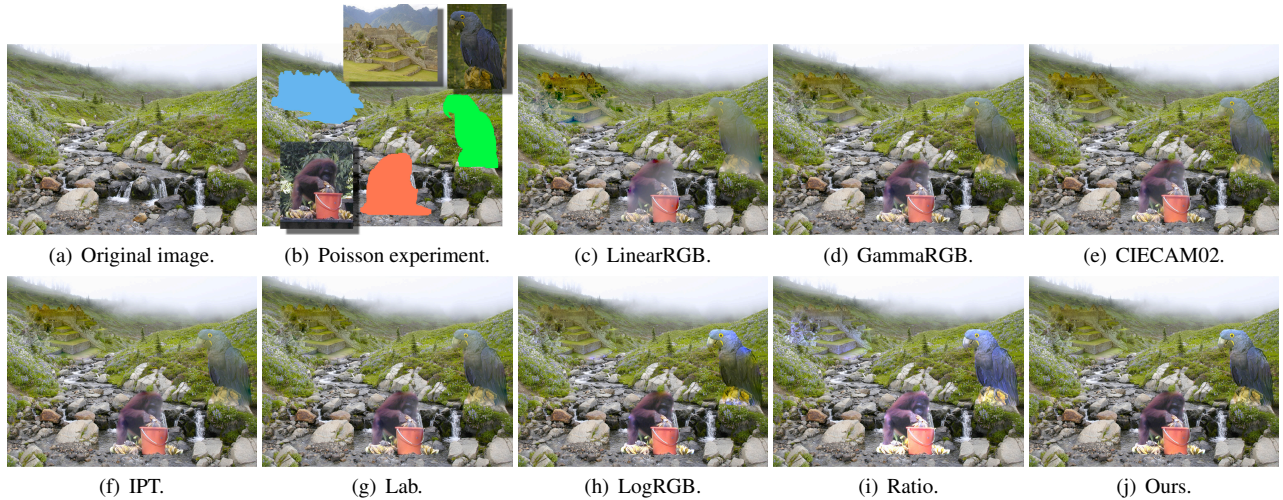


Figure 4: Three objects are inserted into a scene by running Poisson image editing in various color spaces.

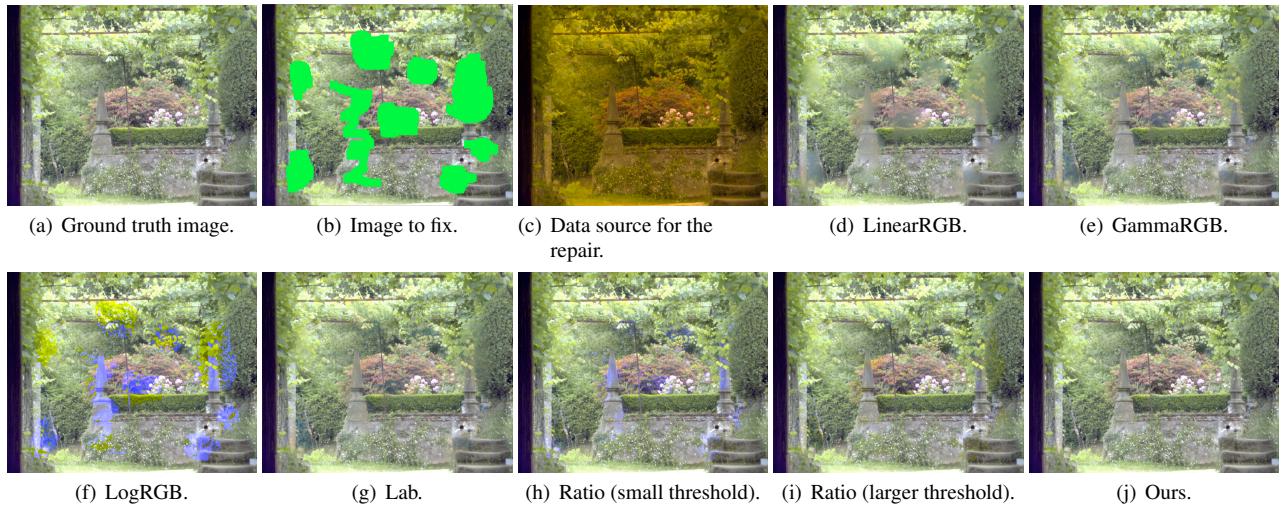


Figure 5: Filling in a damaged image using data from an image with different illumination conditions.

However, their implicit use of the RGB basis for von Kries mappings and failure to use perceptual weightings on the RGB channels led to some noticeable artifacts for more extreme illumination changes. In Figure 5, we see both suffer from taking the logarithm or dividing by a small value. Some of the problems in Ratio were ameliorated by raising the min-value threshold (for algorithm details, see [Georgiev 2006]). The threshold lower bounds the darks in J , and when set too high causes noticeable chromaticity shifts (in particular, note the discontinuous green chromaticity shift on the bush on the right in Figure 5(i) and the excessive yellowness added to the flowers on left). In Figure 5, the threshold for the two Ratio images correspond to 10 gamma-corrected RGB units and 20 units respectively. Lab generally outperformed GammaRGB. LinearRGB always performed worst. IPT and CIECAM02 performed very similarly to CIE $L^*a^*b^*$ in all our experiments, so we omit showing their outputs in Figure 5.

6 Conclusions and Future Work

We have presented a new color space that endows image processing algorithms with some robustness to illumination change. The associated metric also matches the Witt perceptual difference dataset better than CIE $L^*a^*b^*$ and CIE $L^*u^*v^*$ (and performs comparably

on RIT-Dupont), but is equally easy to use. We have also drawn some connections between our desiderata and various human perceptual principles. It would be interesting to look for further perceptual correspondences. For instance, one wonders to what extent the deviation from perfect color constancy in humans can be predicted by the use of our color space.

With regards to applications, continued acquisition of color constancy and perceptual difference datasets (perhaps with more saturated colors as well), and continued search for even better optima based on these datasets, may further improve our metric’s predictive ability. We note that if only illumination invariance is desired, much of the invariance comes from the combination of using logarithms and choosing a B matrix such that the implied color gamut tightly includes the full range of human visible colors (larger than typical RGB working space gamuts but smaller than the implied cone gamut).

References

- ACZEL, J., FALMAGNE, J.-C., AND LUCE, R. D. 2000. Functional equations in the behavioral sciences. *Math. Japonica* 52, 3, 469–512.

BARNARD, K., MARTIN, L., FUNT, B., AND COATH, A. 2002. A data set for colour research. *Color Research and Application* 27, 3, 147–151.

BERNS, R. S., ALMAN, D. H., RENIFF, L., SNYDER, G. D., AND BALONON-ROSEN, M. R. 1991. Visual determination of suprathreshold color-difference tolerances using probit analysis. *Color Research and Applications* 16, 5, 297–316.

CASTILLO, E., AND RUIZ-COBO, M. R. 1992. *Functional Equations and Modelling in Science and Engineering*. CRC.

CHALMERS, A. N., SOLTIC, S., AND JAMMALAMADAKA, R. 2007. A case for a ciecam02 colour appearance space. *Manukau Dept. ECE Conf. Papers*.

CHONG, H. Y., GORTLER, S. J., AND ZICKLER, T. 2007. The von kries hypothesis and a basis for color constancy. *ICCV*.

CHONG, H. Y. 2008. Geometric methods in perceptual image processing. *Harvard University PhD Thesis* (May).

EBNER, F., AND FAIRCHILD, M. D. 1998. Development and testing of a color space (ipt) with improved hue uniformity. *Proc. 6th Color Imaging Conf.*, 8–13.

FAIRCHILD, M. D. 2005. *Color Appearance Models*, 2nd ed. Wiley-IS&T.

FECHNER, G. T. 1877. *In Sachen der Psychophysik*. Leipzig.

FINLAYSON, G., DREW, M., AND FUNT, B. 1993. Diagonal transforms suffice for color constancy. In *ICCV93*, 164–171.

FOSTER, D. H., NASCIMENTO, S. M. C., AND AMANO, K. 2004. Information limits on neural identification of coloured surfaces in natural scenes. *Visual Neuroscience* 21, 331–336.

GEORGIEV, T. 2006. Covariant derivatives and vision. In *ECCV '06*, Springer-Verlag, Part IV, 56–69.

JUDD, D. B., AND WYSZECKI, G. 1975. *Color In Business, Science, and Industry*. John Wiley and Sons.

KIMMEL, R., ELAD, M., SHAKED, D., KESHET, R., AND SOBEL, K., 2002. A variational framework for retinex.

LUCE, R. D. 1993. Let's not promulgate either fechner's erroneous algorithm or his unidimensional approach. *Behavior and Brain Sciences* 16, 155–156.

LUO, M. R. 2006. Colour difference formulae: Past, present and future. *CIE Expert Symposium*.

MCCANN, J. J. 2005. Do humans discount the illuminant? *Proc. SPIE 5666*, 9–16.

MORONEY, N., FAIRCHILD, M. D., HUNT, R. W. G., LI, C., LUO, M. R., AND NEWMAN, T. 2002. The ciecam02 color appearance model. In *Color Imaging Conference*, 23–27.

PALMER, S. E. 1999. *Vision Science: Photons to Phenomenology*. MIT Press.

PAPANDREOU, G., AND MARAGOS, P. 2007. Multigrid geometric active contour models. *IEEE Trans. on Image Processing* 16, 1 (Jan.), 229–240.

PÉREZ, P., GANGNET, M., AND BLAKE, A. 2003. Poisson image editing. *SIGGRAPH* 22, 3, 313–318.

RESNIKOFF, H. L. 1975. Differential geometry and color perception. *Journal of Mathematical Biology* 2, 4, 97–131.

RESNIKOFF, H. L. 1989. *The Illusion of Reality*. Springer-Verlag.

SHARMA, G., WU, W., AND DALAL, E. N. 2005. The ciede2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research and Application* 30, 1, 21–30.

VOS, J. 2006. From lower to higher colour metrics: a historical account. *Clinical & Experimental Optometry* 86, 6, 348–360.

WEST, G., AND BRILL, M. H. 1982. Necessary and sufficient conditions for von kries chromatic adaptation to give color constancy. *Journal of Mathematical Biology* 15, 2, 249–258.

WITT, K. 1999. Geometric relations between scales of small colour differences. *Color Research and Applications* 24, 2, 78–92.

WYSZECKI, G., AND STILES, W. 1982. *Color Science: concepts and methods, quantitative data and formulae*. Wiley.

A Deriving the Functional Form

In the notation of Equation (3), define the displacement vector $V(\vec{x}, \vec{x}') := F(\vec{x}) - F(\vec{x}')$. V satisfies the following property (reflecting the Euclidean nature of our space): $\forall \vec{x}, \vec{x}', \vec{x}''$,

$$\begin{aligned} V(\vec{x}'', \vec{x}') + V(\vec{x}', \vec{x}) &= F(\vec{x}'') - F(\vec{x}') + F(\vec{x}') - F(\vec{x}) \quad (16) \\ &= V(\vec{x}'', \vec{x}). \quad (17) \end{aligned}$$

Meanwhile, Equation (3) states that $V(\vec{x}, \vec{x}')$ is unchanged if both arguments are transformed by $B^{-1}DB$. Define $\vec{c} := B\vec{x}$ and $\vec{c}' := B\vec{x}'$. Then since equation (3) holds for all diagonal matrices D , let us choose the diagonal matrix whose diagonal elements are the reciprocal of the components of \vec{c}' . This gives us: $\forall \vec{x}, \vec{x}'$,

$$V(\vec{x}, \vec{x}') = V(B^{-1}DB[x_1, x_2, x_3]^T, B^{-1}DB[x'_1, x'_2, x'_3]^T) \quad (18)$$

$$= V(B^{-1} \begin{bmatrix} c_1 & c_2 & c_3 \\ c'_1 & c'_2 & c'_3 \end{bmatrix}^T, B^{-1}[1, 1, 1]^T). \quad (19)$$

Since B^{-1} and $B^{-1}[1, 1, 1]^T$ are just fixed constants, define:

$$W\left(\frac{c_1}{c'_1}, \frac{c_2}{c'_2}, \frac{c_3}{c'_3}\right) := V(B^{-1} \begin{bmatrix} c_1 & c_2 & c_3 \\ c'_1 & c'_2 & c'_3 \end{bmatrix}^T, B^{-1}[1, 1, 1]^T). \quad (20)$$

The function $W(c_1/c'_1, c_2/c'_2, c_3/c'_3)$ inherits the additive property of Equations (16)-(17). In other words:

$$W\left(\frac{c''_1}{c'_1}, \frac{c''_2}{c'_2}, \frac{c''_3}{c'_3}\right) + W\left(\frac{c'_1}{c_1}, \frac{c'_2}{c_2}, \frac{c'_3}{c_3}\right) = W\left(\frac{c''_1}{c_1}, \frac{c''_2}{c_2}, \frac{c''_3}{c_3}\right). \quad (21)$$

If we make the variable substitutions $u_i = c''_i/c'_i$, $v_i = c'_i/c_i$, then the additive property becomes the set of functional equations:

$$W_i(u_1 v_1, u_2 v_2, u_3 v_3) = W_i(u_1, u_2, u_3) + W_i(v_1, v_2, v_3), \quad i = 1, 2, 3. \quad (22)$$

The only non-constant and continuous solutions (defined for positive u_i and v_i) are (see [Castillo and Ruiz-Cobo 1992]):

$$W_i(u_1, u_2, u_3) = \sum_{j=1}^3 a_{ij} \ln u_j, \quad a_{ij} \in \mathbb{R}. \quad (23)$$

Substituting back into Equation (20) yields the parameterization F :

$$F(\vec{x}) = A(\hat{\ln}(B\vec{x})) + \vec{k}, \quad (24)$$

where A and B are invertible 3×3 matrices and $\hat{\ln}$ denotes component-wise natural logarithm. The vector \vec{k} might as well be taken to be zero since the range of F is an affine space (only displacements, their magnitudes, and higher order descriptors have illumination-invariant meaning).