Color Constancy Beyond Bags of Pixels

Ayan Chakrabarti   Keigo Hirakawa   Todd Zickler

Harvard School of Engineering and Applied Sciences

ayanc@eecs.harvard.edu   hirakawa@stat.harvard.edu   zickler@seas.harvard.edu

Abstract

Estimating the color of a scene illuminant often plays a central role in computational color constancy. While this problem has received significant attention, the methods that exist do not maximally leverage spatial dependencies between pixels. Indeed, most methods treat the observed color (or its spatial derivative) at each pixel independently of its neighbors. We propose an alternative approach to illuminant estimation—one that employs an explicit statistical model to capture the spatial dependencies between pixels induced by the surfaces they observe. The parameters of this model are estimated from a training set of natural images captured under canonical illumination, and for a new image, an appropriate transform is found such that the corrected image best fits our model.

1. Introduction

Color is useful for characterizing objects only if we have a representation that is unaffected by changes in scene illumination. As the spectral content of an illuminant changes, so does the spectral radiance emitted by surfaces in a scene, and so do the spectral observations collected by a tri-chromatic sensor. For color to be of practical value, we require the ability to compute color descriptors that are invariant to these changes.

As a first step, we often consider the case in which the spectrum of the illumination is uniform across a scene. Here, the task is to compute a mapping from an input color image $y(n)$ to an illuminant-invariant representation $x(n)$. What makes the task difficult is that we do not know the input illuminant *a priori*.

The task of computing invariant color representations has received significant attention under a variety of titles, including color constancy, illuminant estimation, chromatic adaptation, and white balance. Many methods exist, and almost all of them leverage the assumed independence of each pixel. According to this paradigm, spatial information is discarded, and each pixel in a natural image is modeled as an independent draw. The well-known grey world hypoth-
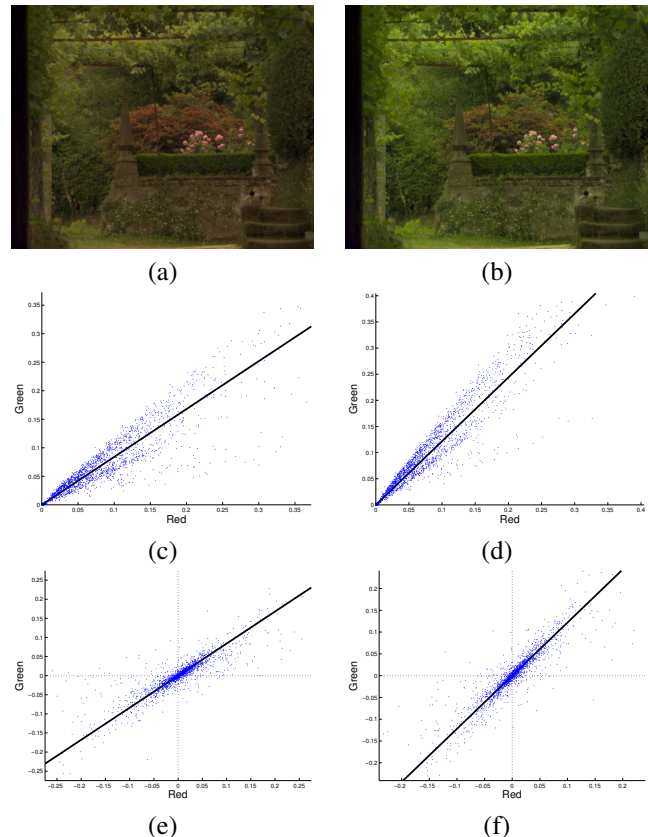

Figure 1. Color distributions under changing illumination. Images (a,b) were generated synthetically from a hyper-spectral reflectance image[12], standard color filters and two different illuminant spectra. Above are scatter plots for the red and green values of (c,d) individual pixels; and (e,f) 8×8 image patches projected onto a particular spatial basis vector. Black lines in (c-f) correspond to the illuminant direction. The distribution of individual pixels does not disambiguate between dominant colors in the image and the color of the illuminant.

esis is a good example; it simply states that the expected reflectance in an image is achromatic [14]. A wide variety of more sophisticated techniques take this approach as well. Methods based on the dichromatic model [10], gamut mapping [8, 11], color by correlation [9], Bayesian inference [1], neural networks [3], and the grey edge hypothe-

sis [18] are distinct in terms of the computational techniques they employ, but they all discard spatial information and effectively treat images as "bags of pixels."

Bag-of-pixels methods depend on the statistical distributions of individual pixels and ignore their spatial contexts. Such distributions convey only meager illuminant information, however, because the expected behavior of the models is counterbalanced by the strong dependencies between nearby pixels. This is demonstrated in Figure 1(c,d), for example, where it is clearly difficult to infer the illuminant direction with high precision.

In this paper, we break from the bag-of-pixels paradigm by building an explicit statistical model of the spatial dependencies between nearby image points. These image dependencies echo those of the spatially-varying reflectance[17] of an observed scene, and we show that they can be exploited to distinguish the illuminant from the natural variability of the scene (Figure 1(e,f)).

We describe an efficient method for inferring scene illumination by examining the statistics of natural color images in the spatio-spectral sense. These statistics are learned from images collected under a known illuminant. Then, given an input image captured under a unknown illuminant, we can map it to its invariant (canonical) representation by fitting it to the learned model. Our results suggest that exploiting spatial information in this way can significantly improve our ability to achieve chromatic adaptation.

The rest of this paper is organized as follows. We begin with a brief review of a standard color image formation model in Section 2. A statistical model for a single color image patch is introduced in Section 3, and the optimal corrective transform for the illuminant is found via model-fitting in Section 4. The proposed model is empirically verified using available training data in Section 5.

2. Background: Image Formation

We assume a Lambertian model where $x : \mathbb{R} \times \mathbb{Z}^2 \to [0, 1]$ is the diffuse reflectivity of a surface corresponding to the image pixel location $n \in \mathbb{Z}^2$ as a function of the electromagnetic wavelength $\lambda \in \mathbb{R}$ in the visible range. The tri-stimulus value recorded by a color imaging device is

$$y(n) = \int f(\lambda)\ell(\lambda)x(\lambda, n)\ d\lambda, \qquad (1)$$

where $y(n) = [y^{\{1\}}(n), y^{\{2\}}(n), y^{\{3\}}(n)]^T$ is the tri-stimulus (*e.g.* RGB) value at pixel location n corresponding to the color matching functions $f(\lambda) = [f^{\{1\}}(\lambda), f^{\{2\}}(\lambda), f^{\{3\}}(\lambda)]^T$, $f^{\{1\}}, f^{\{2\}}, f^{\{3\}} : \mathbb{R} \to [0, 1]$, and $\ell : \mathbb{R} \to \mathbb{R}$ is the spectrum of the illuminant.

Our task is to map a color image $y(n)$ taken under an unknown illuminant to an illuminant-invariant representa-

tion $x(n)$[1]. In general, this computational chromatic adaptation problem is ill-posed. To make it tractable, we make the standard assumption that the mapping from y to x is algebraic/linear; and furthermore, that it is a diagonal transform (in RGB or some other linear color space[7]). This assumption effectively imposes joint restrictions on the color matching functions, the scene reflectivities, and the illuminant spectra[5, 19]. Under this assumption of (generalized) diagonal transforms, we can write:

$$y(n) = Lx(n), \qquad (2)$$

where $L = \text{diag}(\ell)$, $\ell \in \mathbb{R}^3$, $x(n) = [x^{\{1\}}(n)\ x^{\{2\}}(n)\ x^{\{3\}}(n)]^T \in [0, 1]^3$, and f is implicit in the algebraic constraints imposed.

3. Spatio-Spectral Analysis

Studies in photometry have established that the diffuse reflectivity for real-world materials as a function of λ are typically smooth and can be taken to live in a low-dimensional linear subspace[15]. That is, $x(\lambda, n) = \sum_{t=0}^{T-1} \phi_t(\lambda)c_t(n)$, where $\phi_t : \mathbb{R} \to \mathbb{R}$ is the basis and $c_t(n) \in \mathbb{R}$ are the corresponding coefficients that describes the reflectivity at location n. Empirically, we observe that the baseband reflectance ϕ_0 is constant across all λ ($\phi_0(\lambda) = \phi_0$) and the spatial variance along this dimension(*i.e.*, the variance in $c_0(n)$) is disproportionately larger than that along the rest.

The color image y can be written as a sum of the baseline and residual images:

$$y(n) = y_{\text{lum}}(n) + y_{\text{chr}}(n)$$

$$y_{\text{lum}}(n) = \int f(\lambda)\ell(\lambda)\phi_0 c_0(n)\ d\lambda = \ell\phi_0 c_0(n)$$

$$y_{\text{chr}}(n) = \sum_{t=1}^{T-1} \int f(\lambda)\ell(\lambda)\phi_t c_t(n)\ d\lambda. \qquad (3)$$

Here, the baseline "luminance" image contains the majority of energy in y and is proportional to the illuminant color $\ell \in \mathbb{R}^3$; we see from Figure 2 that y_{lum} marks the inter-object boundaries and intra-object textures. The residual "chrominance" image describes the "deviation" from the baseline intensity image, capturing the "color" variations in reflectance. Also, unlike the luminance image, it is largely void of high spatial frequency content.

Existing literature in signal processing provides additional evidence that y_{chr} is generally a low-pass signal. For instance, Gunturk *et al.* [13] have shown that the Pearson product-moment correlation coefficient is typically above

---

[1]For convenience, we refer to $x(n)$ as the *reflectance image* and to ℓ as the *illuminant color*. In practice these may be, respectively, the image under a canonical illuminant and the entries of a diagonal "relighting transform". These interpretations are mathematically equivalent.
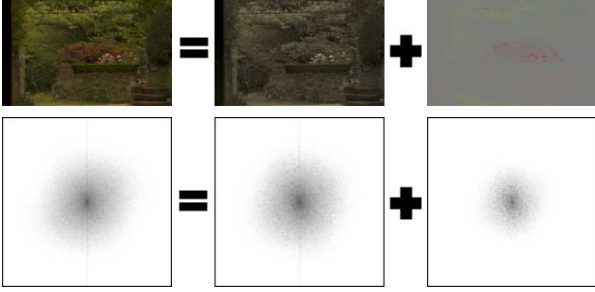
Figure 2. Decomposition of (left column) a color image $\boldsymbol{y}$ into (middle column) luminance $\boldsymbol{y}_{\text{lum}}$ and (right column) chrominance $\boldsymbol{y}_{\text{chr}}$ components. Log-magnitude of the Fourier coefficients in (bottom row) correspond to the images in (top row), respectively. Owing to the edge and texture information that comprise luminance image, luminance dominates chrominance in the high-pass components of $\boldsymbol{y}$.

0.9 for high-pass components of $y^{\{1\}}$, $y^{\{2\}}$, and $y^{\{3\}}$—suggesting that $\boldsymbol{y}_{\text{lum}}$ dominates high-pass components of $\boldsymbol{y}$. Figure 2 also illustrates Fourier support of a typical color image taken under a canonical illuminant, clearly confirming the band-limitedness of $\boldsymbol{y}_{\text{chr}}$. These observations are consistent with the contrast sensitivity function of human vision[14] (but see [16]) as well as the notion that the scene reflectivity $x(\lambda, \boldsymbol{n})$ is spatially coherent, with a high concentration of energy at low spatial frequencies.

All of this suggests that decomposing images by spatial frequency can aid in illuminant estimation. High-pass coefficients of an image $\boldsymbol{y}$ will be dominated by contributions from the luminance image $\boldsymbol{y}_{\text{lum}}$, and the contribution of $\boldsymbol{y}_{\text{chr}}$ (and thus the scene chrominance $\boldsymbol{x}_{\text{lum}}$) will be limited. Since the luminance image $\boldsymbol{y}_{\text{lum}}$ provides direct information about the illuminant color (equation (3)), so too will the high-pass image coefficients. This is demonstrated in Figure 1(e,f), which shows the color of 8×8 image patches projected onto a high-pass spatial basis function.

In subsequent sections, we develop a method to exploit the 'extra information' available in (high-pass coefficients of) spatial image patches.

3.1. Statistical Model

We seek to develop a statistical model for a $\sqrt{K} \times \sqrt{K}$ patch where $\boldsymbol{X}^{\{1\}}, \boldsymbol{X}^{\{2\}}$ and $\boldsymbol{X}^{\{3\}} \in \mathbb{R}^K$ are cropped from $x^{\{1\}}(\boldsymbol{n}), x^{\{2\}}(\boldsymbol{n})$ and $x^{\{3\}}(\boldsymbol{n})$ respectively. Rather than using a general model for patches of size $\sqrt{K} \times \sqrt{K} \times 3$, we employ a spatial decorrelating basis and represent such patches using a mutually independent collection of K three-vectors in terms of this basis. We use the discrete cosine transform(DCT) here, but the discrete wavelet transform(DWT), steerable pyramids, curvelets, *etc.* are other common transform domains that could also be used. This gives us a set of basis vectors $\{\boldsymbol{D}_k\}_{k=0\ldots(K-1)} \in \mathbb{R}^K$ where without loss of generality, $\boldsymbol{D}_0$ can be taken to cor-
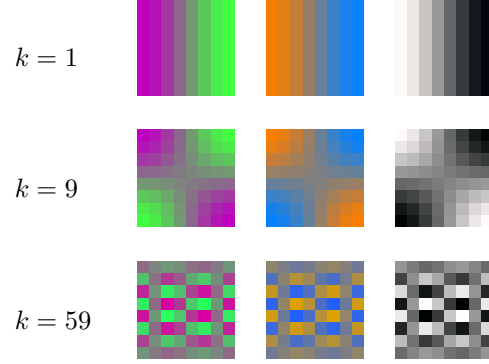

Figure 3. Eigen-vectors of the covariance matrices $\boldsymbol{\Lambda}_k$. The pattern in each patch corresponds to a basis vector used for spatial decorrelation (in this case a DCT filter) and the colors represent the eigen-vectors of the corresponding $\boldsymbol{\Lambda}_k$. The right-most column contains the most significant eigen-vectors that are found to be achromatic.

respond to the lowest frequency component or DC.

By using this decorrelating basis, modeling the distribution of color image patches $\boldsymbol{X}$ reduces to modeling the distribution of three-vectors $\boldsymbol{\mathcal{D}}_k^T \boldsymbol{X} \in \mathbb{R}^3$, $\forall k$, where $\boldsymbol{\mathcal{D}}_k$ computes the response of each of $\boldsymbol{X}^{\{1\}}, \boldsymbol{X}^{\{2\}}$ and $\boldsymbol{X}^{\{3\}}$ to $\boldsymbol{D}_k$ such that

$$\boldsymbol{\mathcal{D}}_k^T \boldsymbol{X} = \begin{bmatrix} \boldsymbol{D}_k^T & & \\ & \boldsymbol{D}_k^T & \\ & & \boldsymbol{D}_k^T \end{bmatrix} \begin{bmatrix} \boldsymbol{X}^{\{1\}} \\ \boldsymbol{X}^{\{2\}} \\ \boldsymbol{X}^{\{3\}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{D}_k^T \boldsymbol{X}^{\{1\}} \\ \boldsymbol{D}_k^T \boldsymbol{X}^{\{2\}} \\ \boldsymbol{D}_k^T \boldsymbol{X}^{\{3\}} \end{bmatrix}. \quad (4)$$

The DC component for natural images is known to have near uniform distributions[4]. The remaining components are modeled as Gaussian. Formally,

$$\boldsymbol{\mathcal{D}}_0^T \boldsymbol{X} \overset{\text{i.i.d.}}{\sim} \mathcal{U}(\boldsymbol{\nu}_{\text{min}} \times \boldsymbol{\nu}_{\text{max}})$$
$$\boldsymbol{\mathcal{D}}_k^T \boldsymbol{X} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \boldsymbol{\Lambda}_k), \qquad k > 0, \quad (5)$$

where $\boldsymbol{\Lambda}_k = E[\boldsymbol{\mathcal{D}}_k^T \boldsymbol{X} \boldsymbol{X}^T \boldsymbol{\mathcal{D}}_k]$, and $[\boldsymbol{\nu}_{\text{min}}, \boldsymbol{\nu}_{\text{max}}]$ is the range of the DC coefficients. The probability of the entire reflectance image patch is then given by

$$P(\boldsymbol{X}) \propto \prod_{k>0} \frac{1}{\det(\boldsymbol{\Lambda}_k)^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{\mathcal{D}}_k^T \boldsymbol{X})^T \boldsymbol{\Lambda}_k^{-1} \boldsymbol{\mathcal{D}}_k^T \boldsymbol{X}\right). \quad (6)$$

We can gain further insight from looking at the sample covariance matrices $\{\boldsymbol{\Lambda}_k\}$ computed from a set of natural images taken under a single (canonical) illuminant. The eigenvectors of $\boldsymbol{\Lambda}_k$ represent directions in tri-stimulus space, and Figure 3 visualizes these directions for three choices of k. For all $K > 0$ we find that the most significant eigenvector is achromatic, and that the corresponding eigenvalue is significantly larger than the other two. This is consistent with the scatter plots in Figure 1, where the distributions have a highly eccentric elliptical shape that is aligned with the illuminant direction.

4. Estimation Algorithm

In the previous section, a statistical model for a single color patch was proposed. The parameters of this model can be learned, for example, from a training set of natural images with a canonical illumination. In this section, we develop a method for color constancy that breaks an image into a "bag of patches" and then attempts to fit these patches to such a learned model.

Let $\text{diag}(\boldsymbol{w}), \boldsymbol{w} = [1/\ell^{\{1\}} \ 1/\ell^{\{2\}} \ 1/\ell^{\{3\}}]$ represent the diagonal transform that maps the observed image to the reflectance image (or image under a canonical illumination). Dividing the observed image into a set of overlapping patches $\{\boldsymbol{Y}_j\}$, we wish to find the set of patches $\{\hat{\boldsymbol{X}}_j(\boldsymbol{w})\}$ that best fit the learned model from the previous section (in terms of log-likelihood) such that $\forall j$, $\hat{\boldsymbol{X}}_j$ is related to $\boldsymbol{Y}_j$ as

$$\hat{\boldsymbol{X}}_j(\boldsymbol{w}) = \left[w^{\{1\}}\boldsymbol{Y}_j^{\{1\}^T} \ w^{\{2\}}\boldsymbol{Y}_j^{\{2\}^T} \ w^{\{3\}}\boldsymbol{Y}_j^{\{3\}^T} \right]^T . \quad (7)$$

We choose to estimate $\boldsymbol{w}$ by model-fitting as follows:

$$\boldsymbol{w} = \arg\max_{\boldsymbol{w}'} \sum_j \log P\left(\hat{\boldsymbol{X}}_j\left(\boldsymbol{w}'\right)\right). \quad (8)$$

It is clear that (8) always admits the solution $\boldsymbol{w} = 0$. We therefore add the constraint that $\boldsymbol{w}^T\boldsymbol{w} = 3$ (so that $\boldsymbol{w} = [1\ 1\ 1]^T$ when $\boldsymbol{Y}$ is taken under canonical illumination).

This constrained optimization problem admits a closed form solution. To see this, let the eigen-vectors and eigen-values for $\boldsymbol{\Lambda}_k$ be given by $\{\mathcal{V}_{kh} = [\mathcal{V}_{kh}^{\{1\}} \ \mathcal{V}_{kh}^{\{2\}} \ \mathcal{V}_{kh}^{\{3\}}]\}_{h=\{1,2,3\}}$ and $\{\sigma_{kh}^2\}_{h=\{1,2,3\}}$ respectively. Then equation (8) simplifies as

$$\boldsymbol{w} = \arg\min_{\boldsymbol{w}'} \sum_{j,k>0,h} \frac{1}{2\sigma_{kh}^2} \left(w'^{\{1\}}\mathcal{V}_{kh}^{\{1\}}\boldsymbol{D}_k^T\boldsymbol{Y}_j^{\{1\}} \right.$$
$$\left. + w'^{\{2\}}\mathcal{V}_{kh}^{\{2\}}\boldsymbol{D}_k^T\boldsymbol{Y}_j^{\{2\}} + w'^{\{3\}}\mathcal{V}_{kh}^{\{3\}}\boldsymbol{D}_k^T\boldsymbol{Y}_j^{\{3\}} \right)^2$$
$$= \arg\min_{\boldsymbol{w}'} \sum_{j,k>0,h} \frac{1}{2\sigma_{kh}^2} \boldsymbol{w}'^T \boldsymbol{a}_{jkh}\boldsymbol{a}_{jkh}^T \boldsymbol{w}'$$
$$= \arg\min_{\boldsymbol{w}'} \boldsymbol{w}'^T \boldsymbol{A}\boldsymbol{w}', \quad (9)$$

subject to $\boldsymbol{w}^T\boldsymbol{w} = 3$, where

$$\boldsymbol{a}_{jkh} = \left[\mathcal{V}_{kh}^{\{1\}}\boldsymbol{D}_k^T\boldsymbol{Y}_j^{\{1\}} \ \mathcal{V}_{kh}^{\{2\}}\boldsymbol{D}_k^T\boldsymbol{Y}_j^{\{2\}} \ \mathcal{V}_{kh}^{\{3\}}\boldsymbol{D}_k^T\boldsymbol{Y}_j^{\{3\}} \right]^T$$

$$\boldsymbol{A} = \sum_{j,k>0,h} \frac{\boldsymbol{a}_{jkh}\boldsymbol{a}_{jkh}^T}{2\sigma_{kh}^2}. \quad (10)$$

The solution can now be found by an eigen-decomposition of $\boldsymbol{A}$. Note that the equivalue contours of $\boldsymbol{w}'^T\boldsymbol{A}\boldsymbol{w}'$ are ellipsoids of increasing size whose axes are given by the eigen-vectors of $\boldsymbol{A}$. Therefore, the point where the smallest
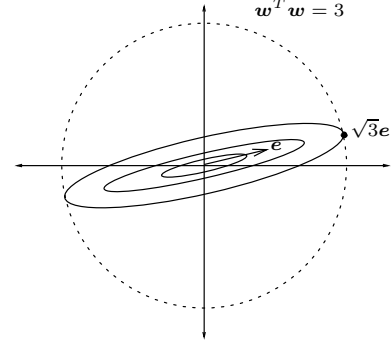

Figure 4. The concentric ellipses correspond to the equivalue contours of $\boldsymbol{w}'^T\boldsymbol{A}\boldsymbol{w}'$. The optimal point on the sphere $\boldsymbol{w}^T\boldsymbol{w} = 3$ therefore lies on the major axis of these ellipses.

ellipsoid touches the $\boldsymbol{w}^T\boldsymbol{w} = 3$ sphere is along the major axis, *i.e.* the eigen-vector $\boldsymbol{e}$ of $\boldsymbol{A}$ that corresponds to the minimum eigen-value. The solution to (8) is then given by $\sqrt{3}\boldsymbol{e}$. This is illustrated in Figure 4.

5. Experimental Results

In this section, we evaluate the performance of the proposed method on a database collected specifically for color constancy research[6]. While this database suffers from a variety of non-idealities—JPEG artifacts, demosaicking, non-linear effects such as gamma correction, *etc.*— it is frequently used in the literature to measure the performance and therefore provides a useful benchmark[6, 18]. The database contains a large number of images captured in different lighting conditions. Every image has a small grey sphere in the bottom right corner that provides the "ground truth". Since the sphere is known to be perfectly grey, its mean color (or rather, the mean color of the 5% brightest pixels to account for the sphere being partially in shadow) is taken to be the color of the illuminant.

Training was done on all overlapping patches in a set of 100 images that are color corrected based on the sphere, *i.e.* for each image the illuminant was estimated from the sphere and then every pixel was diagonally transformed by the inverse of the illuminant. The patch size was chosen to be 8×8 and the DCT was used for spatial decorrelation. For "relighting" images, we chose to apply diagonal transforms directly in RGB color space, and it is important to keep in mind that the results would likely improve (for all methods we consider) by first "sharpening" the color matching functions (*e.g.* [5, 7]).

The performance of the estimation algorithm was evaluated on 20 images from the same database. These images were chosen *a-priori* such that they did not represent any of the scenes used in training, and also such that the sphere was approximately in the same light as the rest of the scene. The proposed algorithm was compared with the *Grey-World*[2] and *Grey-Edge*[18] methods. An implementation provided
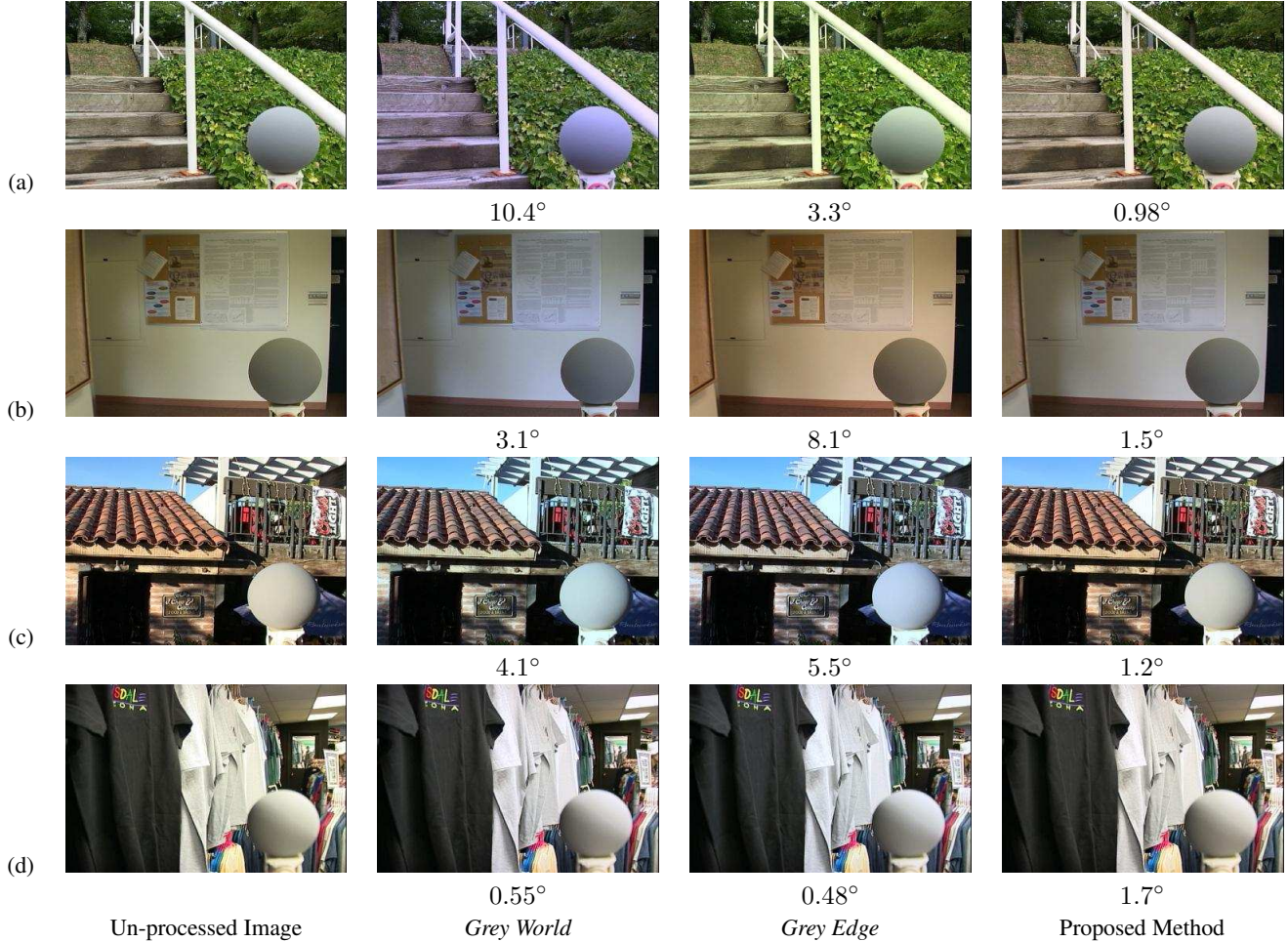
Figure 5. A selection of images from the test set corrected by different methods with the corresponding angular errors.

#	Grey-World[2]	Grey-Edge[18]	Proposed
1	7.4°	5.9°	**2.5°**
2	4.1°	5.5°	**1.2°**
3	10.4°	3.3°	**0.98°**
4	3.1°	8.1°	**1.5°**
5	11.3°	1.8°	**0.31°**
6	4.3°	1.8°	**0.90°**
7	2.2°	4.2°	**1.4°**
8	4.4°	1.9°	**1.2°**
9	3.3°	1.7°	**1.1°**
10	2.6°	0.91°	**0.42°**
11	4.4°	1.9°	**1.7°**
12	**2.5°**	3.6°	2.6°
13	2.4°	**2.1°**	2.6°
14	4.6°	**0.80°**	1.4°
15	14.7°	**6.8°**	7.7°
16	7.2°	**2.2°**	3.1°
17	13.7°	**0.96°**	1.9°
18	6.9°	**3.1°**	4.3°
19	0.55°	**0.48°**	1.7°
20	3.9°	**0.08°**	2.2°
Mean	5.7°	2.9°	**2.0°**

Table 1. Angular errors for different color constancy algorithms.

by the authors of [18] was used for both these methods, and for *Grey-Edge* the parameters that were described in [18] to perform best were chosen (*i.e.* second-order edges, a Minkowski norm of 7 and a smoothing standard devia-

tion of 5). For all algorithms, the right portion of the image was masked out so that the sphere would not be included in the estimation process. The angular deviation of the sphere color in the corrected image from $[1\ 1\ 1]^T$ was chosen as the error metric.

Table 1 shows the angular errors for each of the three algorithms for all images . The proposed method does better than *Grey-World* in 17 and better than *Grey-Edge* in 12 of the 20 images. Some of the actual color corrected images are shown in Figure 5. In Figure 5(a-c), the proposed method outperforms both *Grey-World* and *Grey-Edge*. In the first case, we see that since the image has green as a very dominant color, *Grey-World* performs poorly and infers the illuminant to be green. For images (b) and (c), there are many edges (*e.g.* the roof in (c)) with the same color distribution across them, and this causes the *Grey-Edge* method to perform poorly. In both cases, the proposed method benefits from spatial correlations and cues from complex image features. In Figure 5(d), both *Grey-World* and *Grey-Edge* do better than the proposed method. This is because most of
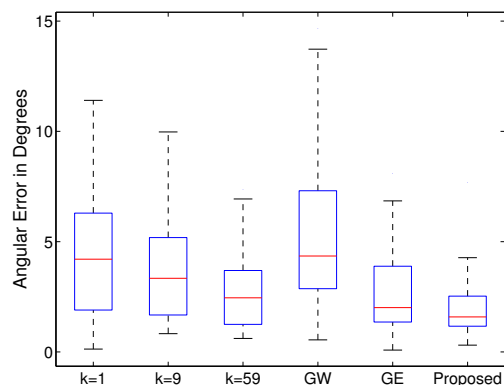
Figure 6. This *box-plot* summarizes the performance of three individual spatial components ($\mathcal{D}_k$), showing the median and quantiles of angular errors across the test set. These are also compared to the *Grey World*(GW) and *Grey Edge*(GE) algorithms, and the proposed method that combines cues from all spatial components. The proposed method performs best—having the lowest average error as well as the smallest variance.

the objects in the scene are truly achromatic (*i.e.* their true color is grey/white/black) and therefore the image closely satisfies the underlying hypothesis for those algorithms.

Finally, the performance of each individual spatial subband component was evaluated. That is, we observed how well the proposed method performed when estimating w using the statistics of each $\mathcal{D}_k Y^j$ alone, for every k. Figure 6 shows a box-plot summarizing the angular errors across the test set for three representative values of k and compares them with the *Grey-World* and *Grey-Edge* algorithm as well as the proposed method which combines all components. Each single component outperforms *Grey-World* and some are comparable to *Grey-Edge*. The proposed method, which uses a statistical model to weight and combine cues from all components, performs best.

6. Conclusion and Future Work

In this paper, we presented a novel solution to the computational chromatic adaptation task through an explicit statistical modeling of the spatial dependencies between pixels. Local image features are modeled using a combination of spatially decorrelating transforms and an evaluation of the spectral correlation in this transform domain. The experimental verifications suggest that this joint spatio-spectral modeling strategy is effective.

The ideas explored in this paper underscores the benefits to exploiting spatio-spectral statistics for color constancy. We expect further improvements from a model based on heavy-tailed probability distribution functions for the transform coefficients. Also, many bag-of-pixel approaches to color constancy can be adapted to use bags of patches instead, especially Bayesian methods [1] that fit naturally into our statistical framework. Examining spatially-varying illumination is also within the scope of our future work.

References

[1] D. Brainard and W. Freeman. Bayesian color constancy. *J. of the Optical Soc. of Am. A*, 14(7):1393–1411, 1993.

[2] G. Buchsbaum. A spatial processor model for object colour perception. *J. Franklin Inst.*, 310(1):1–26, 1980.

[3] V. Cardei, B. Funt, and K. Barnard. Estimating the scene illumination chromaticity using a neural network. *J. of the Optical Soc. of Am. A*, 19(12):2374–2386, 2002.

[4] A. Chakrabarti and K. Hirakawa. Effective Separation of Sparse and Non-Sparse Image Features for Denoising. In *Proc. ICASSP*, 2008.

[5] H. Chong, S. Gortler, and T. Zickler. The von Kries hypothesis and a basis for color constancy. In *Proc. ICCV*, 2007.

[6] F. Ciurea and B. Funt. A Large Image Database for Color Constancy Research. In *Proc. IS&T/SID Color Imaging Conf.*, pages 160–164, 2003.

[7] G. Finlayson, M. Drew, and B. Funt. Diagonal transforms suffice for color constancy. In *Proc. ICCV*, 1993.

[8] G. Finlayson and S. Hordley. Gamut constrained illumination estimation. *Intl. J. of Comp. Vis.*, 67(1):93–109, 2006.

[9] G. Finlayson, S. Hordley, and P. M. Hubel. Color by correlation: A simple, unifying framework for color constancy. *IEEE Trans. PAMI*, 23(11):1209–1221, 2001.

[10] G. Finlayson and G. Schaefer. Convex and non-convex illuminant constraints for dichromatic colour constancy. In *Proc. CVPR*, 2001.

[11] D. Forsyth. A novel algorithm for color constancy. *Intl. J. of Comp. Vis.*, 5(1), 1990.

[12] D. Foster, S. Nascimento, and K. Amano. Information limits on neural identification of colored surfaces in natural scenes. *Visual Neuroscience*, 21(03):331–336, 2005.

[13] B. K. Gunturk, Y. Altunbasak, and R. M. Mersereau. Color plane interpolation using alternating projections. *IEEE Trans. Image Processing*, 11(9):997–1013, 2002.

[14] E. Land. The retinex theory of colour vision. In *Proc. R. Instn. Gr. Br.*, volume 47, pages 23–58, 1974.

[15] H.-C. Lee. *Introduction to Color Imaging Science*. Camb. Univ. Press, 2005.

[16] C. Parraga, G. Brelstaff, T. Troscianko, and I. Moorehead. Color and luminance information in natural scenes. *J. of the Optical Soc. of Am. A*, 15(3):563–569, 1998.

[17] B. Singh, W. Freeman, and D. Brainard. Exploiting Spatial and Spectral Image Regularities for Color Constancy. *Workshop on Stat. and Comp. Theories of Vis.*, 2003.

[18] J. van de Weijer, T. Gevers, and A. Gijsenij. Edge-Based Color Constancy. *IEEE Trans. on Image Processing*, 16(9):2207–2214, 2007.

[19] G. West and M. H. Brill. Necessary and sufficient conditions for von kries chromatic adaptation to give color constancy. *J. of Math. Bio.*, 15(2):249–258, 1982.