User-guided White Balance for Mixed Lighting Conditions

Ivaylo Boyadzhiev
Cornell University

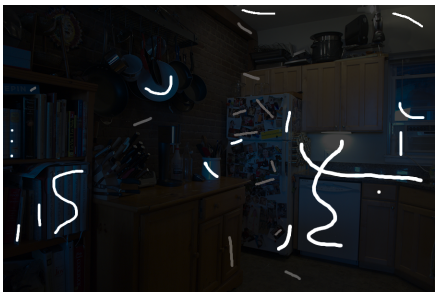
Kavita Bala
Cornell University

Sylvain Paris
Adobe

Frédo Durand
MIT CSAIL

(a) input with mixed lighting (daylight + neon under the cabinets + low-energy bulbs on ceiling) exhibits unsightly color casts everywhere

(b) naive single-light white balance makes the ceiling white, but other color casts remain



(c) user indicates regions that are neutral (white strokes) or correct after the single-light white balance (grey strokes)

(d) the image is improved, but color variations can still be observed, e.g., on the wooden cabinet

(e) user adds marks to specify uniform color, e.g., the cabinet and the wall

(f) our final output with no color casts

Figure 1: *In this photo, the ambient lighting, the cabinet light, and the ceiling lights all have different colors, which produces unpleasant color casts (a). In such situations, the single-light white balance tool provided in all photo editing software only improves a portion of the image, but the result is not satisfying (b). We address this issue by letting users make annotations on the photo. First, they mark objects of neutral color (i.e., white or gray), and regions that look fine after the standard white balance (c). This improves the result, but undesirable color variations are still visible, e.g., on the cabinetry and on the wall (d). Users can indicate that these elements should have a constant color (e), which yields a result free of color cast (f).*

Abstract

Proper white balance is essential in photographs to eliminate color casts due to illumination. The single-light case is hard to solve automatically but relatively easy for humans. Unfortunately, many scenes contain multiple light sources such as an indoor scene with a window, or when a flash is used in a tungsten-lit room. The light color can then vary on a per-pixel basis and the problem becomes challenging at best, even with advanced image editing tools.

We propose a solution to the ill-posed mixed light white balance problem, based on user guidance. Users scribble on a few regions that should have the same color, indicate one or more regions of neutral color , and select regions where the current color looks correct. We first expand the provided scribble groups to more regions using pixel similarity and a robust voting scheme. We formulate the spatially varying white balance problem as a sparse data interpolation problem in which the user scribbles and their extensions form constraints. We demonstrate that our approach can produce satisfying results on a variety of scenes with intuitive scribbles and without any knowledge about the lights.

Keywords: white balance, mixed lighting

Links: DL PDF

1 Introduction

White balance correction is a critical photography step, where the goal is ”to compensate for different colour temperatures of scene illuminants” [Jacobson 2000]. For example, tungsten lights cause images to have a yellowish cast. Proper white balance compensates for this color cast and yields photos where objects have their natural colors, as if taken under a neutral light [Hedgecoe 2009]. When all the lights have the same color, this problem is easy to solve for

a photographer who often indicates a white or gray object in the image, from which it is straightforward to recover the illuminant color. Unfortunately, many scenes exhibit a combination of illuminants such as artificially-lit indoor scenes with additional light from a window (Fig. 1) or from a flash. Adjusting the white balance is then a challenging task, even for skilled users. Each point can be lit by the mixture of several light sources, depending on their relative distances and orientations. Worse, modern low-consumption fluorescent and LED lights vary widely in their color temperature, and rooms with multiple bulbs exhibit a plethora of color casts (Fig. 1).

A few automatic techniques have been proposed, but the severely ill-posed nature of the problem restricts them to specific scenarios. For instance, Ebner [2004; 2009] produces perceptual renderings that are often not ideal from a photography perspective. Hsu et al. [2008] can handle only two light colors, need to know their exact values a priori, and cannot treat scenes with "a strong foreground-background separation." Riess et al. [2011] assume that photos can be decomposed into regions where a single illuminant dominates.

We introduce a user-guided approach to produce high-quality white balanced images for a broad range of photos with multiple light sources. We carefully designed a set of scribbles that are easy for humans to specify, such as, neutral-color objects, regions of constant color, and places where the color looks correct. From our experience, it is difficult for a human to estimate quantities related to illumination. Therefore, as a general guideline, our scribbles are related to reflectance properties, rather than illumination. For instance, the local light color on a textured material, such as fur or fabric, varies in nontrivial ways that cannot be easily understood by a human observer and described with scribbles. In comparison, the same observer can easily recognize and indicate regions where the fur has the same color. Similarly, human observers have no problem recognizing that a wall lit by a complex mixture of lights has a constant color. These situations are common in photographs and easy to identify. Our surroundings are full of constant-color objects such as walls and man-made objects – and our method is robust enough to also leverage textured materials, such as fur and fabric.

We formulate our method as an optimization problem that seeks to retrieve the color of the light mixture at each pixel under the constraints provided by the scribbles. We show that for canonical light-reflectance configurations, the white balance solution lies in the null space of the Matting Laplacian [Levin et al. 2006] , which motivates our use of this energy for regularization. We further identify points with similar patterns in the scene and constrain them to have the same reflectance. This scribble extension strategy allows users to achieve satisfying results with only a small number of scribbles. We demonstrate that our approach yields good results on a wide range of scenes; it can handle two or more light sources; and it does not require knowledge of the absolute values of lights or reflectances, since such knowledge is typically not easy to obtain. In practice, we show that it also copes with light mixtures and materials, beyond the theoretically studied base cases.

In summary, we introduce the following contributions:
▷ A practical algorithm to white balance scenes illuminated by a mixture of light sources, with no assumption on their number, color or configuration.
▷ Our user-assisted scribbles only deal with what humans identify most easily, reflectance properties.
▷ Our automatic extension of constraints reduces the number of user scribbles needed.
▷ We show that the null space of the Matting Laplacian spans the white-balance solution for canonical local configurations.

Table 1: *Table of Notation*

C_c^I	$\triangleq$	Input chromaticity, $I_c \;/\; (I_\mathrm{r} + I_\mathrm{g} + I_\mathrm{b})$
C_c^O	$\triangleq$	Output chromaticity, $O_c \;/\; (O_\mathrm{r} + O_\mathrm{g} + O_\mathrm{b})$
C_c^R	$\triangleq$	Reflectance chromaticity, $R_c \;/\; (R_\mathrm{r} + R_\mathrm{g} + R_\mathrm{b})$
W_c	$\triangleq$	Correction factors, $I_c \;/\; O_c$

$c \in \{\mathrm{r}, \mathrm{g}, \mathrm{b}\}$

2 Related work

The single-light case is well addressed with automatic methods, e.g. [Finlayson et al. 2006; Gehler et al. 2008], and photo editing tools. When used in cases where the lighting is mixed, they have to make compromises and residual color casts remain.

A few techniques deal with light mixtures. Ebner [2009], Riess et al. [2011], Bleier et al. [2011], and Gijsenij et al. [2011] assume that, locally, a single light dominates. This might approximate human perception, but from a photography perspective, the results have faded colors and retain local color casts Ebner [2009].

Hsu et al. [2008] focus on the two-light scenario and further assume that the light colors are known. In contrast, we do not seek a fully automatic technique and strive for an approach that handles an as-wide-as-possible range of scenes. Furthermore, our experience with their system shows that their voting stage is sensitive to the initial choice of the lights' RGB values, which is not always trivial to specify for scenes taken outside a controlled environment. Finally, as their paper mentions in Section 8, "scenes that exhibit a strong foreground-background separation may also cause problems." This is because they need to observe a given reflectance under a number of different mixtures.

Lischinski et al. [2006] describe a scribble interface to perform local edits. This approach can be used to correct a color cast that is well localized in the scene, but it can be tedious in scenes where the light mixture occurs everywhere. Further, this tool requires the user to specify the absolute correction to be applied, which is nontrivial in our case. Determining the local color of the illumination in a complex colorful scene is challenging even for a human observer. In comparison, we make sure that our scribbles only deal with relative characteristics of the scene reflectance, which is significantly easier than determining the absolute color of the local illumination.

Carroll et al. [2011] use a scribble interface to edit the color of inter-reflections. While related, our approach deals with effects that are more global and affect large portions of the image, whereas inter-reflections have a limited spatial extent. Further, their technique relies on scribbles that describe properties of the illumination, such as the fact that a light source does not affect a designated area. Since in our context light sources have a global impact, such information would be particularly challenging.

Bousseau et al. [2009] and Shen et al. [2011] compute intrinsic images, i.e. they separate an object's reflectance from the illumination reaching it. While related to white balance, intrinsic images are also significantly different because they often assume a monochromatic or nearly monochromatic illumination; the main challenge being the estimation of the light intensity at each point. In comparison, we focus on colored illumination and, as we shall see, seek to leave lighting intensity untouched, without estimating it. Further, the technique of Bousseau et al. [2009] requires scribbles about absolute and relative properties of the illumination such as "this point is fully lit" or "the illumination in this region is smooth." Since we are dealing with complex multi-source illumination, we cannot expect that users will be able to specify this, and argue that such information is hard to specify for novices and experts alike.

3 Mixed lighting white balance

Our approach starts with the standard workflow used for single-light white balance, but then introduces a set of scribbles to address the much more challenging problem of mixed white balance. First, the user globally adjusts the image to get an approximate result. This step is the same as the standard single-light white balance, and can be achieved with existing tools such as clicking on a white patch or adjusting the global color temperature. Then, we provide three brushes to annotate the image. *Neutral color* scribbles indicate objects that are white or gray. *Same color* scribbles show regions of constant reflectance or texture, but where the actual reflectance need not be specified. Finally, *Correct color* scribbles indicate areas that look correct in the current view. We designed these scribbles such that they are related to reflectance only, and do not require users to specify absolute values. We avoided scribbles related to lighting, because, in our experience, precise illumination properties are elusive for humans, especially on complex materials. After the user has marked the image with these scribbles, we solve a linear optimization problem that estimates the spatially varying RGB gain factors that explain the formation of the observed image with mixed lighting. We render the white-balanced image by applying the inverses of these factors. If needed, users can iterate and add more scribbles to refine the result. Figure 1 illustrates this process.

We first present our image formation model. We explain how we formulate the white-balance problem as a least-squares optimization based on the Matting Laplacian, and motivate this approach by analyzing the null space of this energy. Finally, we express scribbles as constraints in the optimization and show how we can extend these scribbles to find additional constraints.

3.1 Image formation model and problem statement

The input of the algorithm is an image $(I_\mathrm{r}, I_\mathrm{g}, I_\mathrm{b})$ that represents a scene with reflectance $(R_\mathrm{r}, R_\mathrm{g}, R_\mathrm{b})$ lit by n_ℓ lights $\{(L_{i\mathrm{r}}, L_{i\mathrm{g}}, L_{i\mathrm{b}})\}$. We name λ_i the attenuation of the i-th light due to factors such as light travel and foreshortening. Further, we assume Lambertian materials and no inter-reflections. Using this notation, we model the observed image at a pixel p by:

$$\forall c \in \{\mathrm{r}, \mathrm{g}, \mathrm{b}\}, \quad I_c(p) = R_c(p) \left(\sum_{i=1}^{n_\ell} \lambda_i(p) L_{ic} \right) \tag{1}$$

In this equation, we only observe $(I_\mathrm{r}, I_\mathrm{g}, I_\mathrm{b})$, everything else is unknown. Our objective is to produce a white-balanced image $(O_\mathrm{r}, O_\mathrm{g}, O_\mathrm{b})$. Intuitively, we aim for rendering the scene as if each light had a neutral color, i.e., $(L_{i\mathrm{r}}, L_{i\mathrm{g}}, L_{i\mathrm{b}}) = (\ell_i, \ell_i, \ell_i)$ for some positive scalar ℓ_i. That is, we seek:

$$\forall c \in \{\mathrm{r}, \mathrm{g}, \mathrm{b}\}, \quad O_c(p) = R_c(p) \left(\sum_{i=1}^{n_\ell} \lambda_i(p) \ell_i \right) \tag{2}$$

However, this problem is severely under-constrained since we do not know the number of lights n_ℓ, their colors L_i, the corresponding spatially varying attenuation factors λ_i, and the spatially-varying scene reflectances R. Further, it is unclear what the ℓ_i values should be. In practice, this problem is intractable without additional hypotheses.

In this work, in addition to the "neutral light color" goal (Eq. 2), we also seek to preserve image intensities, i.e., we also aim for:

$$O_\mathrm{r} + O_\mathrm{g} + O_\mathrm{b} = I_\mathrm{r} + I_\mathrm{g} + I_\mathrm{b} \tag{3}$$

For the sake of clarity, we use a simple model of intensity represented by the sum of RGB channels. Optionally we could weight each channel according to its perceptual importance, which would not affect the rest of our formulation. Intensity preservation, as defined in Equation 3, has an intuitive interpretation in the photography context. Our approach alters only the chromaticity values. Everything related to intensities remains unchanged, e.g., our operator is orthogonal to tonal adjustments such as brightness and contrast. For the rest of the paper, we define the chromaticity of a pixel as its RGB channels divided by its intensity, that is, $C^I_c = I_c \,/\, (I_\mathrm{r} + I_\mathrm{g} + I_\mathrm{b})$ for $c \in \{\mathrm{r}, \mathrm{g}, \mathrm{b}\}$.

Our strategy to produce the white-balanced image $(O_\mathrm{r}, O_\mathrm{g}, O_\mathrm{b})$ is not to estimate all the unknown quantities in Equations 1 and 2. Instead, we seek $(W_\mathrm{r}, W_\mathrm{g}, W_\mathrm{b})$ factors such that at a pixel p:

$$\forall c \in \{\mathrm{r}, \mathrm{g}, \mathrm{b}\}, \quad I_c(p) = W_c(p)\, O_c(p) \tag{4}$$

This formulation drastically reduces the number of unknowns. Moreover, we show in the next section that in a number of cases, the W factors can be expressed as an affine combination of the observed chromaticity values $(C^I_\mathrm{r}, C^I_\mathrm{g}, C^I_\mathrm{b})$, which is the key of our approach based on the Matting Laplacian.

As we shall see later, it is useful to express W as a function of C^I. First, we use Equation 2 to get: $\sum_c O_c = (\sum_c R_c)(\sum_i \lambda_i \ell_i)$. Dividing Equation 2 by this result, we obtain $C^O = C^R$ since the light term $\sum_i \lambda_i \ell_i$ cancels out. Then dividing Equation 4 by Equation 3, we get

$$C^I_c = W_c\, C^O_c = W_c\, C^R_c \tag{5}$$

3.2 Standard scenarios and the Matting Laplacian

We study a number of standard cases under our model. We show that in all these cases, the W factors are an affine combination of the input chromaticities. We then use this result to adapt the Matting Laplacian introduced by Levin et al. [2006] in the context of image matting to the problem of white balance under mixed lighting.

3.2.1 Case studies

We now study a few standard cases in increasing order of complexity. We start with the single-light and single-reflectance scenarios, and then discuss the more complex case with two reflectance values lit by a 2D illumination.

Single-color illumination. Using the von Kries hypothesis [Chong et al. 2007], the effect of a single-color illumination can be modeled by globally scaling the RGB channels. That is, the W factors are constant over the image, which can be seen as a special case of an affine combination with zero coefficients affecting the chromaticity channels.

Monochromatic scenes. In the case where the reflectance R is constant over the scene, Equation 5 gives $W_c = C^I_c \,/\, C^R_c$. Since C^R is constant, it means that W is proportional to C^I, which is a special case of affine combination.

Duochromatic scenes under 2D illumination. In the appendix, we show that under some reasonable assumptions, the previous result extends to scenes with two reflectances lit by an illumination that lies on 2D subspaces of the RGB cube. In this case, W can be expressed as an affine combination of $(C^I_\mathrm{r}, C^I_\mathrm{g}, C^I_\mathrm{b})$ in which all the coefficients are nonzero. This case illustrates that it is beneficial to consider all the channels at the same time and allow cross-talk, e.g., the red channel C^I_r is useful to estimate the blue factor W_b. Intuitively, the two reflectance values cannot induce arbitrary variations, and we build an affine combination that recovers the W factors while being insensitive to these variations.

Discussion. These results show the strong relationship between the observed values $(C^I_{\rm r}, C^I_{\rm g}, C^I_{\rm b})$ and the unknown $(W_{\rm r}, W_{\rm g}, W_{\rm b})$ factors that we seek. We use this link to guide the interpolation of a sparse set of user-specified constraints and obtain meaningful results over the entire image. In the next section, we explain how this affine relationship is related to the null space of the Matting Laplacian, and build upon this result to formulate our approach as a standard least-squares problem. Our least-squares approach is robust to other cases, as visible from our examples. Local windows that do not satisfy this model generate a higher residual but do not make the algorithm fail. As long as most windows satisfy these cases or are close to them, our approach produces satisfying outputs. Users can also add scribbles to constrain the solution.

3.2.2 Link with the null space of the Matting Laplacian

First, we summarize the properties and formulation of the Matting Laplacian, and then we explain how to adapt it to white balance.

Background on the Matting Laplacian. Levin et al. [2006] introduced the Matting Laplacian in the context of matting, i.e., to extract a foreground element from its surrounding background. They argued that the alpha values that represent the foreground-background mixture at each pixel should locally be an affine combination of the RGB channels. And they showed that this can be modeled with a least-squares functional based on a matrix $\mathbf{M}$ that they call the *Matting Laplacian*. $\mathbf{M}$ is a $n_{\rm p} \times n_{\rm p}$ matrix, with $n_{\rm p}$ the number of pixels in the image. Its $\mathbf{M}_{ij}$ coefficient is:

$$\sum_{\substack{k \text{ such that} \\ (i,j)\in w_k}} \left(\delta_{ij} - \frac{1}{|w_k|}\left(1 + (\mathbf{I}_i - \boldsymbol{\mu}_k)\left(\boldsymbol{\Sigma}_k + \frac{\epsilon}{|w_k|}\mathbf{Id}\right)^{-1}(\mathbf{I}_j - \boldsymbol{\mu}_k)\right)\right) \quad (6)$$

where i, j, and k refer to pixels, $\mathbf{I}_i$ is a vector containing the RGB components at pixel i, δ_{ij} is the Kronecker symbol , w_k is a window centered on pixel k, $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ are the mean vector and covariance matrix of the pixels within w_k, $\mathbf{Id}$ is the 3×3 identity matrix, and ϵ is a parameter controlling the smoothness of the result. Levin et al. showed that one can impose, in a least-squares sense, that the alpha values within each w_k are an affine combination of the RGB channels by minimizing the quadratic form $\mathbf{x}^{\mathsf{T}}\mathbf{M}\mathbf{x}$ where $\mathbf{x}$ is a $n_{\rm p}$-dimensional vector containing all the alpha values.

Null space of the Matting Laplacian. We have shown that for a number of standard cases, the W factors are an affine combination of the chromaticities $(C^I_{\rm r}, C^I_{\rm g}, C^I_{\rm b})$. By definition of the Matting Laplacian, this means that if we build $\mathbf{M}$ using the chromaticity values of the input image instead of the RGB channels, the W factors are in its null space. That is, $\mathbf{W}_c^{\mathsf{T}}\,\mathbf{M}\,\mathbf{W}_c = 0$ for all c in $\{{\rm r}, {\rm g}, {\rm b}\}$, where $\mathbf{W}_c$ is a large vector containing all the W_c factors of the image. This is a critical result for our task. While several other options exist to interpolate user scribbles, e.g. [Lischinski et al. 2006; Chen et al. 2007; An and Pellacini 2008], using the Matting Laplacian, constructed with the chromaticity values, ensures that we produce the correct result in the cases that we studied. In the next section, we build upon this result to design our energy function.

3.3 Mixed lighting white balance as optimization

We model white balance as a least-squares optimization. We build the energy, term by term, with the Matting Laplacian first, and then the user scribbles.

Affine combination of the chromaticities. We have shown that in a number of standard cases, the W factors are an affine combination of the chromaticity values, which can be expressed as $\mathbf{W}_c^{\mathsf{T}}\,\mathbf{M}\,\mathbf{W}_c = 0$ for all c in $\{{\rm r}, {\rm g}, {\rm b}\}$. However, on real-world images, there may be windows that do not fall in one of these standard scenarios. For instance, three or more different reflectances or lights can appear in some windows. We cope with these cases by modeling the affine-combination constraint in a least-squares sense. With the Matting Laplacian, this amounts to a quadratic term:

$$E_{\rm M} = \sum_{c\in\{{\rm r,g,b}\}} \mathbf{W}_c^{\mathsf{T}}\,\mathbf{M}\,\mathbf{W}_c \quad (7)$$

We use two settings to prevent the system from returning a trivial solution when the users have specified only neutral color scribbles. If the regularization is too weak, and only neutral colors have been indicated, the image can be interpreted as an uniformly white scene illuminated by many different light sources. To prevent this trivial solution, we use a strong regularization with $\epsilon = 10^{-2}$. When other scribbles are provided, the white-scene interpretation does not hold anymore and we relax the system with $\epsilon = 10^{-4}$ so that it better respects edges. Intuitively, the Matting Laplacian regularization factor controls the smoothness prior of our interpolation energy. In practice, we have found that values between $\epsilon = 10^{-4}$ and $\epsilon = 10^{-6}$ work fine for our application.

Neutral color. Users can indicate pixels that have a neutral color. These are usually the first scribbles made by users. For these pixels, the RGB channels should be equal, and using Equation 3, we obtain $O_{\rm r} = O_{\rm g} = O_{\rm b} = \frac{1}{3}\sum I$. With Equation 4, this gives $\frac{1}{3}W_c = I_c \,/\, \sum I = C^I_c$, which we translate into a least-squares energy:

$$E_{\rm n} = \sum_{p\in\mathcal{S}_{\rm n}} \sum_{c\in\{{\rm r,g,b}\}} \left(\frac{1}{3}W_c(p) - C^I_c(p)\right)^2 \quad (8)$$

where, $\mathcal{S}_{\rm n}$ is the set of pixels covered by the scribbles indicating a neutral reflectance.

Correct color. Users can also specify that the chromaticity $\hat{C}^R$ currently visible in a region is correct: C^O should be equal to $\hat{C}^R$. Equation 5 leads to $\hat{C}^R_c W_c = C^I_c$ and the corresponding energy:

$$E_{\rm c} = \sum_{p\in\mathcal{S}_{\rm c}} \sum_{c\in\{{\rm r,g,b}\}} \left(\hat{C}^R_c\, W_c(p) - C^I_c(p)\right)^2 \quad (9)$$

Same color. Users can also mark regions that have the same chromaticity. They need not provide the common chromaticity, they only mark pixels that share it. The pixels p covering a scribble $\mathcal{S}_{\rm s}$ share the same $(\bar{C}^R_{\rm r}, \bar{C}^R_{\rm g}, \bar{C}^R_{\rm b})$ values. Using Equation 5 and summing over all the pixels, we get:

$$\forall c \in \{{\rm r, g, b}\},\ \ \bar{C}^R_c \left(\sum_{p\in\mathcal{S}_{\rm s}} W_c(p)\right) = \sum_{p\in\mathcal{S}_{\rm s}} C^I_c(p) \quad (10)$$

which leads to a linear relationship between $1/\bar{C}^R_c$ and the W_c factors under the scribble:

$$\frac{1}{\bar{C}^R_c} = \frac{\sum_{p\in\mathcal{S}_{\rm s}} W_c(p)}{\sum_{p\in\mathcal{S}_{\rm s}} C^I_c(p)} \quad (11)$$

Expressing $1/\bar{C}^R_c$ for a single pixel q gives: $W_c(q)/C^I_c(q)$. Since $\bar{C}^R_c$ is constant, we have:

$$\frac{W_c(q)}{C^I_c(q)} = \frac{\sum_{p\in\mathcal{S}_{\rm s}} W_c(p)}{\sum_{p\in\mathcal{S}_{\rm s}} C^I_c(p)} \quad (12a)$$

$$W_c(q) \sum_{p\in\mathcal{S}_{\rm s}} C^I_c(p) = C^I_c(q) \sum_{p\in\mathcal{S}_{\rm s}} W_c(p) \quad (12b)$$

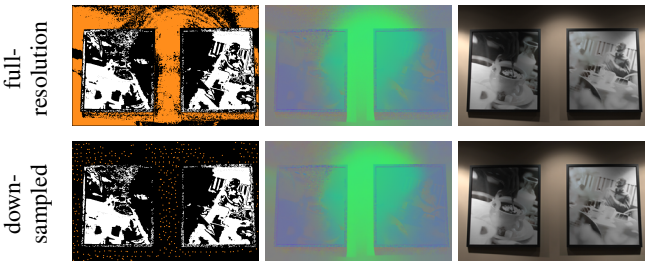
Equation (12b) avoids division and leads to a numerically more stable scheme. We turn it into a least-squares term $E_{\rm s}$:

$$\sum_{q\in\mathcal{S}_{\rm s}} \sum_{c\in\{{\rm r,g,b}\}} \left(W_c(q)\Big(\sum_{p\in\mathcal{S}_{\rm s}} C^I_c(p)\Big) - C^I_c(q)\Big(\sum_{p\in\mathcal{S}_{\rm s}} W_c(p)\Big)\right)^2 \quad (13)$$

We add one such term for each scribble made by users.

(a) input (b) user scribbles (c) output, no extensions (d) extended scribbles (e) output, with extensions

Figure 2: *Starting from an image under mixed lighting (a), using only the user-provided scribbles (b) does not fully remove the undesirable color cast (c). With the same user input and using our extension algorithm (d), we obtain a visually pleasing result with no color cast (e).*



(a) constraints (b) W factors (c) results

Figure 3: *We speed up our algorithm by subsampling the constraints defined by the scribbles. We produce similar results with full-resolution scribbles and with their downsampled counterparts. We use the same input as in Figure 2.*

Extending the scribbles. Our goal is to minimize the amount of user's input that will lead to a satisfying output. We use a similar strategy to Shen et al. [2008] and detect points that are likely to have the same reflectance as the pixels covered by a scribble and aggregate these points to this scribble. This automatically adds constraints into our system. This helps propagate scribble information more efficiently, and makes it possible to create satisfying results with only a small number of scribbles as shown in Figure 2.

We compare pixels using the Euclidean distance in chromaticity space. Using this metric, our goal is to find unmarked pixels that unambiguously relate to a scribble. For each unmarked pixel p and each scribble $\mathcal{S}$, we robustly estimate how related they are by averaging the distance between p and its 10 nearest neighbors in $\mathcal{S}$. We assign a pixel to the most similar scribble $\mathcal{S}_1$ if two conditions are satisfied. First, the distance s_1 has to be below a threshold t_s. That is, a pixel is added to a scribble only if it is closely related to it. Second, we check that the ratio $s_2/(s_1 + s_2)$ is above a threshold t_r, where s_2 is the second best choice. This ensures that we make only unambiguous assignments for which the second best choice is significantly worse than the best one. In practice, we always use $t_\text{s} = 0.01$ and $t_\text{r} = 0.8$ (with color channel values between 0 and 1).

Subsampling the same-color scribbles. The same-color scribbles create off-diagonal terms in the linear system. If we use all the pixels covered by these scribbles to define our energy, the corresponding linear system would be dense and slow to solve. Instead, we overlay a grid on the image and select a single pixel for each grid cell. Although a random selection achieves satisfying outputs, we found that we can improve the results by selecting pixels in smooth areas. The rationale is that if the signal varies quickly near a pixel, it may be on an edge or a corner where the W factors may also be discontinuous. In practice, we estimate the local amount of variation as the variance of I in a 3×3 window centered on the pixel, and we use 10×10 grid cells. To further improve the robustness of our automatic scribbles extension, we do not pick a representative point in a grid cell, if the number of similar pixels in it, as defined in the previous section, is less than 30% of all the points in the cell. Figure 3 shows the effects of subsampling.

In addition, we also discard samples where the image is dark because the signal-to-noise ratio in these regions is poor and the signal is unreliable. In practice, we discard pixels for which $I < t_\text{d}$ with $t_\text{d} = 0.01$. Later on, we fill in missing data in those regions, using interpolation from the nearest image pixels, whose W factor is reliable, i.e., $I_q > t_\text{d}$.

Putting it together. We get the final result by minimizing a least-squares energy that comprises all the terms that we have defined, that is, $E_\mathbf{M}$ that seeks to represent the W as an affine combination of the chromaticity channels, E_n, E_c, and E_s that model the users' scribbles. We weight each term to get the energy:

$$E = w_\mathbf{M} E_\mathbf{M} + w_\text{n} E_\text{n} + w_\text{c} E_\text{c} + w_\text{s} E_\text{s} \qquad (14)$$

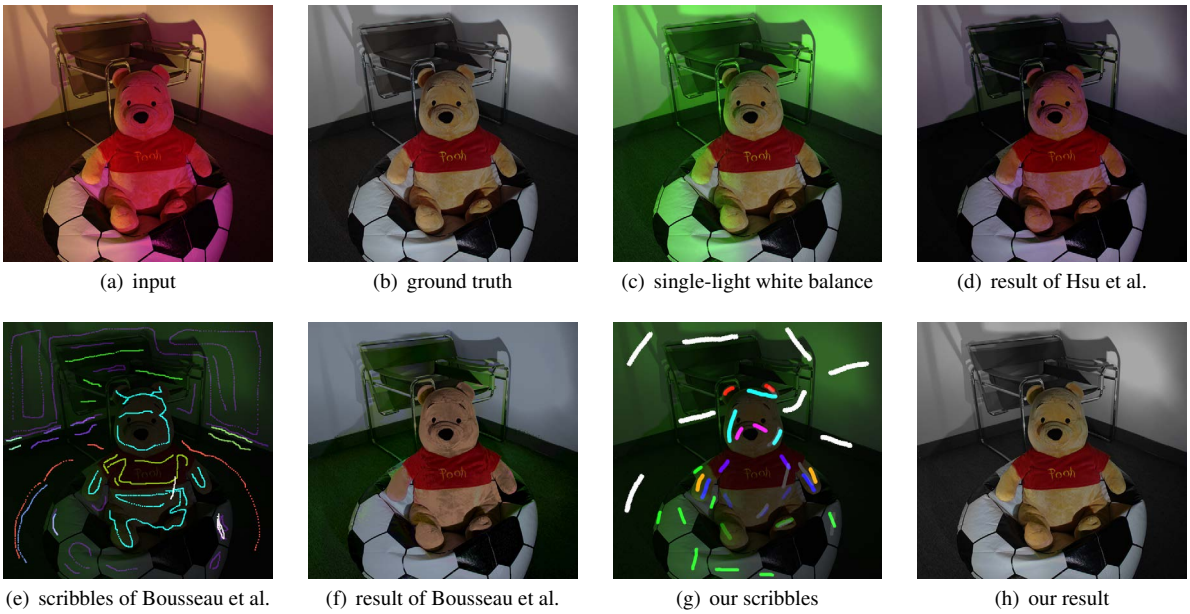
In practice, we use the following weights in all our results: $w_\mathbf{M} = 1$, $w_\text{n} = 10^3$, $w_\text{c} = 10^3$, and $w_\text{s} = 10^2$.

4 Results

We demonstrate our approach on a variety of scenes, and compare to previous work, when possible. Our prototype is implemented in Matlab. We use the standard "backslash" function to minimize Equation 14. As long as no same-color strokes are specified (Section 3.3), the solver is interactive. When these strokes are present, they add off-diagonal terms and the solver can take up to 1 minute per channel for a 900×900 image. For larger resolutions, we first downsample the image, solve for the W factors, and upsample them using Joint Bilateral Upsampling [Kopf et al. 2007].

The kitchen in Figures 1 and 6 is a common case of an interior scene with multiple light chromaticities, which our technique can tackle. In contrast, existing tools cannot remove the color cast in many regions. Single-light white balance only improves part of the scene and produces strong color casts everywhere else (Fig. 1b). The method of Hsu et al. [2008] (Fig. 6c) shows its limitations on this photo because there are three light sources, all of a different color, and they only handle two. Even though it was not designed specifically for this task, we can use the chromaticity of the shading obtained by Bousseau et al. [2009] for white balance correction. However, in scenes such as the kitchen, this produces an overall desaturated result (Fig. 6b). This is probably because their method was designed for intrinsic images, not white balance, and their equations are derived for the case of monochromatic lights only. They explain that the use of the equations on a per-channel basis is only a heuristic. Furthermore, their method cannot handle well black-and-white reflectance variations such as the books in the lower-left of Fig. 6(b), which leads to blue color artifacts there.

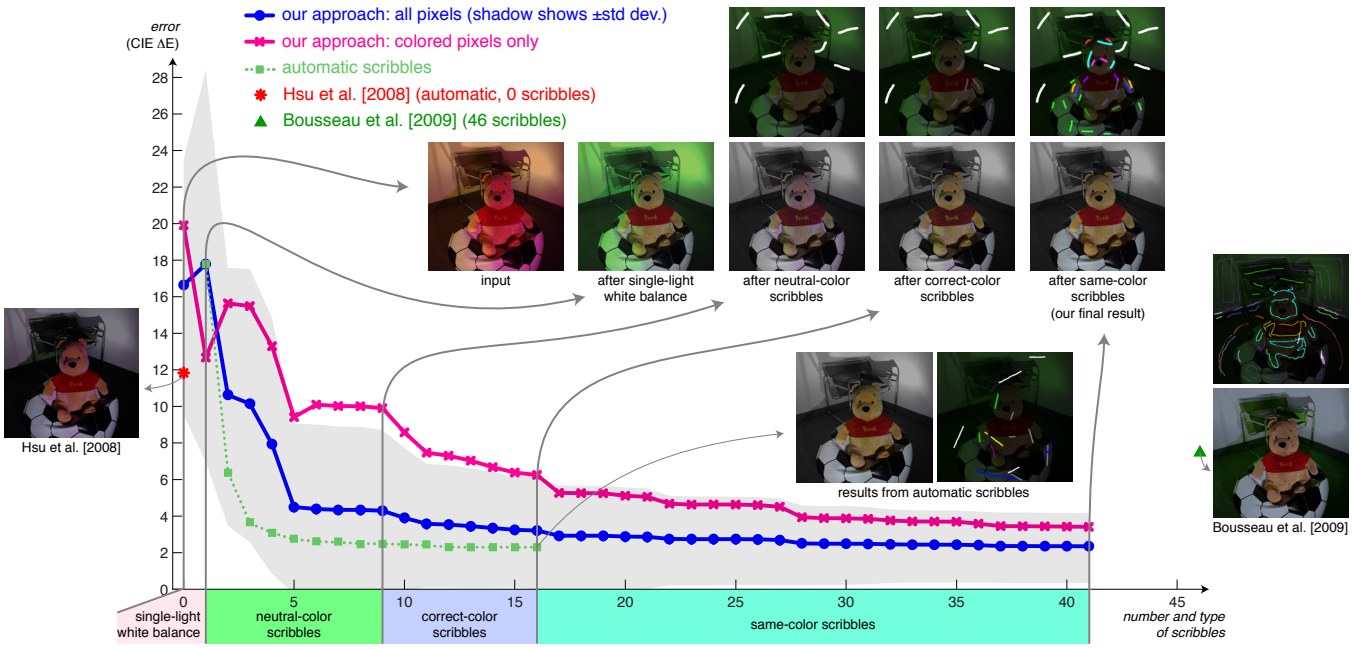
Figures 10 and 8 show that, in the two-light scenario, our approach performs as well as the method by Hsu et al. [2008], and better than that of Ebner [2009]. The main difference with the former is that our approach does not assume the light colors to be known a priori and relies on user input. Moreover, as previously discussed, and unlike the technique of Hsu et al., our approach can also cope with more than two lights.



(a) input (b) ground truth (c) single-light white balance (d) result of Hsu et al.

(e) scribbles of Bousseau et al. (f) result of Bousseau et al. (g) our scribbles (h) our result

Figure 4: *Ground truth comparison against existing approaches. We captured two photos with a white light at different positions. We applied different color filters to each image and combined them to get the input image (a). The ground-truth version is a direct combination of the photos taken under the white light (b). Compensating for the color of one of the lights only does not yield a satisfying result (c). Because the areas lit by each light are mostly disconnected, the Pooh for the red light and the wall for the yellow light, the automatic technique of Hsu et al. [2008] does not produce a good result (d). The technique by Bousseau et al. [2009] relies on scribbles made users (kindly provided by A. Bousseau) (e) and works better, but the colors are desaturated (f). In comparison, our approach uses a number of scribbles on the same order as Bousseau et al. and renders a satisfying result (h) close the ground truth.*



(a) Scribbles error evaluation

Figure 5: *We evaluated the effect of our scribbles on the semi-synthetic image (Fig. 4). The use of neutral scribbles alone fixes the gray floor and the white wall, but extrapolates wrong colors elsewhere. By adding correct-color scribbles to fix a few of the natural looking colors, produced by the single-light white balance stage, we get better looking regions. Finally, we introduce a couple of same-color scribbles to propagate correct color information to other parts of the image. The solution produced by our system is free of color casts, visually (Fig. 4h) and numerically closer to the ground truth, compared to the methods of [Bousseau et al. 2009] and [Hsu et al. 2008].*

Figure 11 shows that our approach can deal with complex materials, such as fur. A few, easy to specify, same-color scribbles are enough for our system to produce a result that is free of color casts (Fig. 11e). For comparison, the system of Hsu et al. [2008] struggles to estimate the two light mixtures because of the complex appearance of the fur (Fig. 11c).

(a) scribbles of Bousseau et al. (b) result of Bousseau et al.

(c) result of Hsu et al. (d) our result

Figure 6: *Comparison with other techniques on the same input as Figure 1. Bousseau et al. [2009] produce results with desaturated colors (b), e.g., the cabinetry. The two-light technique of Hsu et al. [2008] does not fully remove color casts on the left wall and on the dishwasher because there are three different lights (c). In contrast, our approach produces a satisfying output (d).*

4.1 Evaluation using ground-truth data

We use the ground-truth data provided by Hsu et al. [2008] to evaluate the performance of our approach by comparing it to their results (reproduced from their article), and to the results of Bousseau et al. [2009] (kindly provided by A. Bousseau). Our result (Fig. 9c) is numerically and visually closer to the ground-truth image (Fig. 9e). We provide more details in the supplemental material.

Energy evaluation. Figure 7 evaluates how well our energy models the W factors for various ground-truth mixed lighting conditions. We show comparisons to the all-purpose interpolation energy proposed by Lischinski et al. [2006], and to the KNN Matting energy [Chen et al. 2012], which improves the clustering Laplacian based on the nonlocal principle [Lee and Wu 2011]. We used 10 pairs of input/ground truth images, and for each one of them we randomly placed a small set of known-white-balance constraints through the entire image; approximately 0.001% of all pixels in an image were constrained (every 300×300 pixels on average). Using the same constraints, we optimized the three energies to interpolate the missing values. We repeated this step 10 times for each of the 10 data sets, and report the mean relative per pixel error for the W factors. The plot shows that for the problem of white balance our energy performs consistently better than both the all-purpose energy and the clustering Laplacian energy (visual results are in the supplemental material). We also tested the clustering Laplacian energy computed using the chrominance channels and observed only a minor improvement. On average, the Matting Laplacian still performed $2.5\times$ better. The general-purpose energy and clustering Laplacian rely on appearance similarity to interpolate the missing data. While this performs well for some applications, e.g., [Chen et al. 2012] achieves state-of-the-art matting results, our experiment suggests that the affine model of the Matting Laplacian is better suited to the white balance problem.

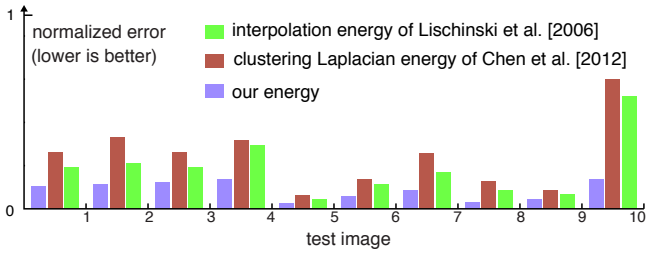


Figure 7: *We compare our energy to the all-purpose interpolation energy proposed by Lischinski et al. [2006], and to the clustering Laplacian energy by Chen et al. [2012]. With a sparse set of constraints, our energy interpolates missing values that are closer to the ground truth light mixtures. This corresponds to visually better white balanced results (see the supplemental material).*

Scribble impact. In Figure 5, we evaluate the impact of our scribbles on an image with ground truth (Fig. 4). The neutral-color scribbles quickly improve the overall result. Then, correct-color scribbles help identify a few regions that look good after the single-light white balance stage. Finally, the same-color scribbles help propagate this information to other areas through the scribbles extension mechanism. To isolate the effect of the white wall in the background, we also plot the energy restricted to the color pixels as determined by the ground-truth data. The initial gain from the neutral-color scribbles is lower, but the trend remains the same. We provide another such analysis in the supplemental material.

User-independent evaluation of our model. Scribble-based techniques are hard to evaluate because they depend on a user's decisions. It makes it difficult to get a sense of their convergence or the amount of user annotations that are fundamentally needed. We propose a new methodology to evaluate scribble-based approaches independent of human input and based on ground truth data. Ideally, we would like to know how a method performs with the best possible user input. To make things tractable, however, we propose an approximation based on a set of tentative scribbles obtained from an image segmentation, and a greedy strategy. By definition, this does not inform us on how easy it would be for a user to choose these scribbles, but it provides strong information about the adequacy and conciseness of the mathematical formulation, and a reference to evaluate human performance.

First, we segment the input image using Quick Shift [Vedaldi and Soatto 2008], and pick the scribble in each segmented region to be as long as possible by fitting an ellipse to the segmented region, and finding its maximal axis. Each segmented region is then assigned a chromaticity value equal to the mean chromaticity of the pixels composing it. By comparing to ground truth, and simple thresholding, segments can be determined to be "neutral" or "correct" in the single-light white balance image. Further, pairs of segments can be deemed to be "same color". This segmentation and assignment gives us our set of potential scribbles.

For each potential scribble, we evaluate the error, the CIE Lab L2-norm, of the image, compared to ground truth, if we applied that scribble. We then greedily apply the scribble that decreases the error most. Then we iterate, currently using a brute-force evaluation of all possible scribbles at each iteration.

The green curve in Figure 5 plots the error of the greedy evaluation. Even though it is greedy, and not optimal, it is lower than any current approach. In particular, these greedily-chosen scribbles, based on full ground truth knowledge, show that our user did not make optimal choices, but got close with additional scribbles.

4.2 Discussion and limitations

Since our approach is user-driven, the quality of the result depends on the amount of interaction that the user performs. We found that simple scenes require only a few scribbles, e.g., 9 in Figure 2, and more complex scenes may need up to 50 scribbles (Figure 1). Although it is difficult to quantify, our scribbles are relatively easy to use. Qualities such as "this is a white wall" and "this region has a uniform color" are easy to determine for users.

The usefulness of the "correct single white balance" strokes depends on the choice of the white balance settings. However, such strokes could be extended to allow the user to explore multiple options for single white balance and use constraints from these multiple versions.

We designed our algorithm using linear RGB values, which can be easily obtained from the RAW image files produced by DSLR cameras. Although JPEG files produced by cameras are processed to make them more appealing, e.g., to increase their contrast and saturation, we found that our approach is robust enough to handle them. For instance, Figure 12 shows a sample result on a JPEG photo for which we assumed a standard gamma of 2.2.

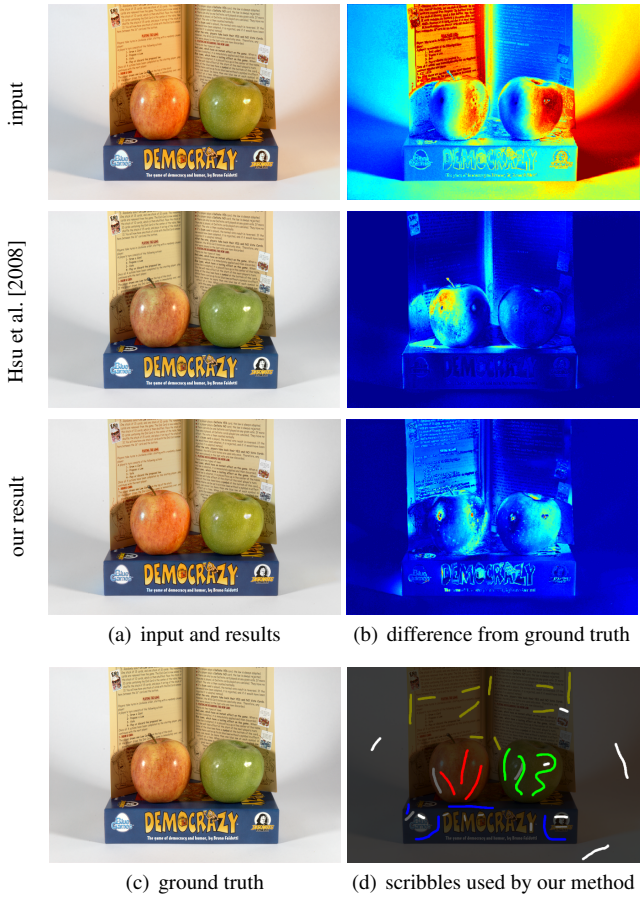
Our current implementation uses a vanilla solver. The first few strokes result in interactive feedback, but later passes take about a minute per color channel. A multi-grid solver would dramatically reduce this cost.

Our user-driven approach to the many-lights white balance problem relies on the subjective judgement of the users. This provides important constraints for our interpolation scheme, which tends to produce results that are plausible or pleasing, but not necessarily physically accurate. For example, in Figure 8, the ground-truth left and right pages have slightly different reflectances, which the user did not know, which leads to numerical error.

5 Conclusions

We present a practical method for high-quality white balancing in scenes with complex lighting based on user-provided scribbles. It relies on what is most intuitive to humans, reflectance properties. Our contributions are a new formulation of the white-balance problem based on intensity preservation, a study that shows that important canonical local configurations are in the null space of the Matting Laplacian, an interpolation energy that performs better for white balance than generic approaches and strategy to extend constraints and reduce the required user interaction.

Acknowledgements. Ivaylo Boyadzhiev and Kavita Bala acknowledge funding from NSF (CAREER 1041534 and IIS 1011919) and Adobe. Frédo Durand acknowledges funding from Quanta, NSF grant 0964004, and gifts from Adobe and Cognex. We thank A. Bousseau and M. Ebner for helping us comparing to their systems.



(a) input and results (b) difference from ground truth

(c) ground truth (d) scribbles used by our method

Figure 8: *On a ground truth data set, provided by Hsu et al. [2008], our scribble based method produces results, comparable to their automatic system. However, notice that we do not assume anything about the lighting in the scene, while their approach is limited to two known light sources.*

References

AN, X., AND PELLACINI, F. 2008. Appprop:All-pairs appearance-space edit propagation. *ACM Trans. on Graphics 27*, 3.

BLEIER, M., RIESS, C., BEIGPOUR, S., EIBENBERGER, E., ANGELOPOULOU, E., TRÖGER, T., AND KAUP, A. 2011. Color constancy and non-uniform illumination: Can existing algorithms work? In *IEEE Color and Photometry in Comp. Vision Workshop*.

BOUSSEAU, A., PARIS, S., AND DURAND, F. 2009. User-assisted intrinsic images. *ACM Trans. on Graphics 28*, 5.

CARROLL, R., RAMAMOORTHI, R., AND AGRAWALA, M. 2011. Illumination decomposition for material recoloring with consistent interreflections. *ACM Trans. on Graphics 30*, 3.

CHEN, J., PARIS, S., AND DURAND, F. 2007. Real-time edge-aware image processing with the bilateral grid. *ACM Trans. on Graphics 26*, 3.

CHEN, Q., LI, D., AND TANG, C. 2012. KNN matting. In *IEEE Conf. on Computer Vision and Pattern Recognition*.

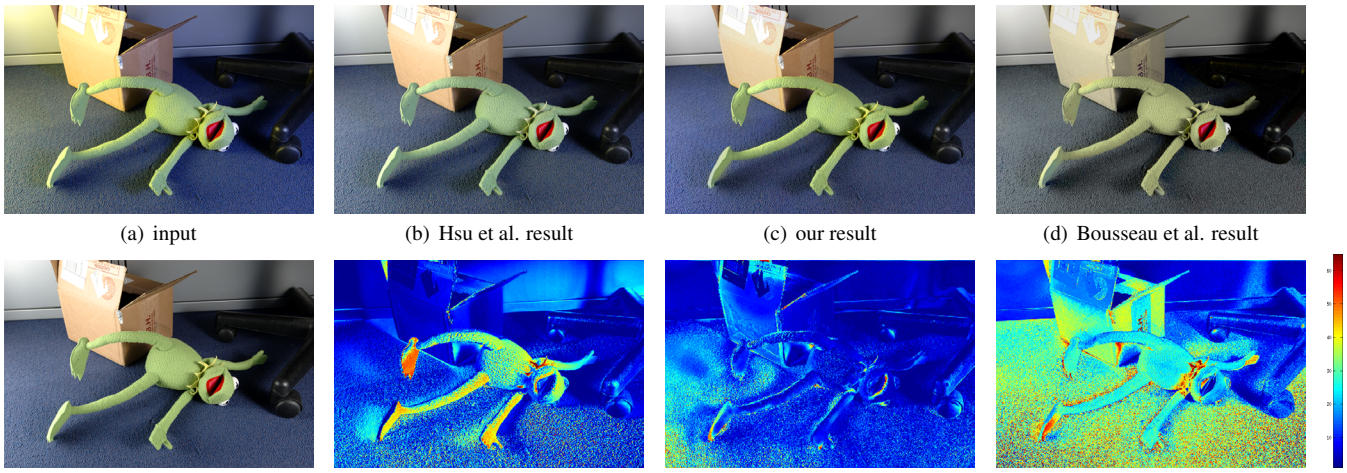
CHONG, H., GORTLER, S., AND ZICKLER, T. 2007. The von Kries hypothesis and a basis for color constancy. In *IEEE International Conf. on Computer Vision*.

EBNER, M. 2004. Color constancy using local color shifts. In *European Conf. on Computer Vision*.

EBNER, M. 2009. Color constancy based on local space average color. *Machine Vision and Applications Journal 20*, 5.

FINLAYSON, G. D., HORDLEY, S. D., AND TASTL, I. 2006. Gamut constrained illuminant estimation. *International Journal of Computer Vision 67*, 1.

GEHLER, P. V., ROTHER, C., BLAKE, A., MINKA, T., AND SHARP, T. 2008. Bayesian color constancy revisited. In *IEEE Conf. on Computer Vision and Pattern Recognition*.



(a) input (b) Hsu et al. result (c) our result (d) Bousseau et al. result

(f) ground truth (g) Hsu et al. error (h) our error (i) Bousseau et al. error

Figure 9: *Ground truth evaluation with image data from [Hsu et al. 2008], shows that our approach produces the result (c), closer to the ground truth image (e), compared to competing approaches, (b) and (d). We plot the numerical error, using the L2-norm of the Lab difference between each method and the ground truth data.*

(a) input (b) Ebner et al. [2009] (c) Hsu et al. [2008] (d) our result

Figure 10: *Comparison for images from Hsu et al. [2008] (a). Ebner's method [2009] tends to produce desaturated results (b). These results are reproduced from Ebner et al. [2009]. Hsu et al. [2008] (c) and our method (d) produce equivalent results with no color cast. The two approaches differ in that Hsu et al. is automatic but assume two lights of known colors, whereas our method is user-assisted and does not make assumptions about the illumination. The results by Hsu et al. are reproduced from their article.*

(a) input (b) single-light white balance (c) Hsu et al. result (d) our scribbles (e) our result

Figure 11: *The scene is lit by outdoor and tungsten lights, making the cat looks yellow (a). The single-light white balance improves the white part of the fur, but makes the gray suitcase appear blue (b). The technique of Hsu et al. [2008] does not produce a good result because of the complex appearance of the fur (c). Using our scribbles (d), our result correctly captures the white fur and the gray suitcase (e).*

(a) input (b) neutral and correct color scribbles (c) after adding neutral and correct color scribbles (d) adding same color scribbles (e) final output

Figure 12: *Birthday. Sample result on a JPEG photo (a). We assumed a standard 2.2 gamma correction to convert the RGB values into linear space. Our method is sufficiently robust to produce a satisfying result (e).*

GIJSENIJ, A., LU, R., AND GEVERS, T. 2011. Color constancy for multiple light sources. *IEEE Trans. on Image Processing*.

HEDGECOE, J. 2009. *New Manual of Photography*. Dorling Kindersley.

HSU, E., MERTENS, T., PARIS, S., AVIDAN, S., AND DURAND, F. 2008. Light mixture estimation for spatially varying white balance. *ACM Trans. on Graphics 27*, 3.

JACOBSON, R. 2000. *The manual of photography: photographic and digital imaging*. Media Manual Series. Focal Press.

KOPF, J., COHEN, M. F., LISCHINSKI, D., AND UYTTENDAELE, M. 2007. Joint bilateral upsampling. *ACM Transactions on Graphics 26*, 3.

LEE, P., AND WU, Y. 2011. Nonlocal matting. In *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, Washington, DC, USA, CVPR '11, 2193–2200.

LEVIN, A., LISCHINSKI, D., AND WEISS, Y. 2006. A closed form solution to natural image matting. In *IEEE Conf. on Computer Vision and Pattern Recognition*.

LISCHINSKI, D., FARBMAN, Z., UYTTENDAELE, M., AND SZELISKI, R. 2006. Interactive local adjustment of tonal values. *ACM Trans. on Graphics 25*, 3.

RIESS, C., EIBENBERGER, E., AND ANGELOPOULOU, E. 2011. Illuminant color estimation for real-world mixed-illuminant scenes. In *IEEE Color and Photometry in Computer Vision Workshop*.

SHEN, L., TAN, P., AND LIN, S. 2008. Intrinsic image decomposition with non-local texture cues. In *IEEE Conf. on Computer Vision and Patten Recognition*.

SHEN, J., YANG, X., JIA, Y., AND LI, X. 2011. Intrinsic images using optimization. In *IEEE Conf. on Computer Vision and Pattern Recognition*.

VEDALDI, A., AND SOATTO, S. 2008. Quick shift and kernel methods for mode seeking. In *European Conf. on Comp. Vision*.

Appendix: Duochromatic scenes under 2D lighting. We show that the W factors of scenes made of two reflectance values R_1 and R_2 lit by a combination of two lights L_1 and L_2 can be expressed as an affine combination of the input chromaticities. We also assume that the observed intensities are not affected by the lights, only the chromaticities vary. That is, $\sum L_1 R_1 = \sum L_2 R_1$ and $\sum L_1 R_2 = \sum L_2 R_2$. We name these two quantities i_1 and i_2. We show that the W_r factor is an affine combination of the $(C^I_\mathrm{r}, C^I_\mathrm{g}, C^I_\mathrm{b})$ triplet; similar results can be derived for W_g and W_b. From Equation 1, we have:

$$I_\mathrm{r} = \begin{cases} (\lambda_1 L_{1\mathrm{r}} + \lambda_2 L_{2\mathrm{r}}) R_{1\mathrm{r}} & \text{if in } R_1 \text{ region} \\ (\lambda_1 L_{1\mathrm{r}} + \lambda_2 L_{2\mathrm{r}}) R_{2\mathrm{r}} & \text{if in } R_2 \text{ region} \end{cases} \tag{15}$$

where λ_1 and λ_2 are the spatially varying intensities of the lights L_1 and L_2. We seek affine coefficients $(a_\mathrm{r}, a_\mathrm{g}, a_\mathrm{b})$ and b independent of λ_1 and λ_2 such that $W_\mathrm{r} = \sum_{c\in\{\mathrm{r,g,b}\}} a_c C^I_c + b$. We first derive useful relationships in the R_1 region, similar formulas can be obtained with R_2. We divide Equation 15 by $\sum I$, and name $\alpha = \lambda_1/(\lambda_1+\lambda_2)$ and $k_1 = \sum_c R_{1c} \,/\, i_1$ to get:

$$C^I_\mathrm{r} = \frac{(\lambda_1 L_{1\mathrm{r}} + \lambda_2 L_{2\mathrm{r}}) R_{1\mathrm{r}}}{\sum_c (\lambda_1 L_{1c} + \lambda_2 L_{2c}) R_{1c}} = \frac{(\lambda_1 L_{1\mathrm{r}} + \lambda_2 L_{2\mathrm{r}}) R_{1\mathrm{r}}}{(\lambda_1+\lambda_2) i_1} \tag{16a}$$

$$= \left(\alpha L_{1\mathrm{r}} + (1-\alpha) L_{2\mathrm{r}}\right) \frac{\sum_c R_{1c}}{i_1} \frac{R_{1\mathrm{r}}}{\sum_c R_{1c}} \tag{16b}$$

$$= k_1 \left(\alpha L_{1\mathrm{r}} + (1-\alpha) L_{2\mathrm{r}}\right) C^{R_1}_\mathrm{r} \tag{16c}$$

Using Equation 4, we have $W_\mathrm{r} = k_1 \left(\alpha L_{1\mathrm{r}} + (1-\alpha) L_{2\mathrm{r}}\right)$. We first show the result for $\alpha = 0$ and $\alpha = 1$ and use superposition to extend it to other α values. That is, we seek $(a_\mathrm{r}, a_\mathrm{g}, a_\mathrm{b})$ and b such that $W_\mathrm{r} = \sum_{c\in\{\mathrm{r,g,b}\}} a_c C^I_c + b$ for $\alpha = 0$ and $\alpha = 1$ in both R_1 regions and R_2 regions. Using a matrix formulation, this gives:

$$\begin{pmatrix} k_1 L_{1\mathrm{r}} C^{R_1}_\mathrm{r} & k_1 L_{1\mathrm{g}} C^{R_1}_\mathrm{g} & k_1 L_{1\mathrm{b}} C^{R_1}_\mathrm{b} & 1 \\ k_1 L_{2\mathrm{r}} C^{R_1}_\mathrm{r} & k_1 L_{2\mathrm{g}} C^{R_1}_\mathrm{g} & k_1 L_{2\mathrm{b}} C^{R_1}_\mathrm{b} & 1 \\ k_2 L_{1\mathrm{r}} C^{R_2}_\mathrm{r} & k_2 L_{1\mathrm{g}} C^{R_2}_\mathrm{g} & k_2 L_{1\mathrm{b}} C^{R_2}_\mathrm{b} & 1 \\ k_2 L_{2\mathrm{r}} C^{R_2}_\mathrm{r} & k_2 L_{2\mathrm{g}} C^{R_2}_\mathrm{g} & k_2 L_{2\mathrm{b}} C^{R_2}_\mathrm{b} & 1 \end{pmatrix} \begin{pmatrix} a_\mathrm{r} \\ a_\mathrm{g} \\ a_\mathrm{b} \\ b \end{pmatrix} = \begin{pmatrix} k_1 L_{1\mathrm{r}} \\ k_1 L_{2\mathrm{r}} \\ k_2 L_{1\mathrm{r}} \\ k_2 L_{2\mathrm{r}} \end{pmatrix} \tag{17}$$

Since the matrix is square, the system is either well-posed or under-constrained, which guarantees that there is at least one solution. Further, because the values of C^I and W for an arbitrary α are a linear interpolation of the values at $\alpha = 0$ and $\alpha = 1$, this ensures that a solution of Equation 17 is valid for any α value.

Discussion. The assumption $\sum L_1 R_1 = \sum L_2 R_1$ and $\sum L_1 R_2 = \sum L_2 R_2$ may not always be satisfied. Nevertheless, since we freely scale up and down L_1 and L_2, as long as we apply the inverse scale factors to λ_1 and λ_2, we can use these degrees of freedom to minimize the differences between $\sum L_1 R_1$ and $\sum L_2 R_1$, and $\sum L_1 R_2$ and $\sum L_2 R_2$.