

Illuminant Color Estimation for Real-World Mixed-Illuminant Scenes

Christian Riess, Eva Eibenberger, Elli Angelopoulou
Pattern Recognition Lab,
University of Erlangen-Nuremberg
<http://www5.cs.fau.de>

Abstract

We present a physics-based approach for illuminant color estimation of arbitrary images, which is explicitly designed for handling images with multiple illuminants. The majority of techniques that extract the illuminant color assume that the illumination is constant across the scene. This, however, is not often the case. We propose an illuminant-color estimation method which is based on robust local illuminant estimates. There are no assumptions on the number or type of illuminants. An illuminant color estimate is obtained independently from distinct image mini-regions. From these mini-regions a robust local illumination color is computed by consensus. These local estimates are then used in deriving the chromaticity of the dominant illuminants. Experiments on an established benchmark database of real-world images show that our technique performs comparably to uniform-illuminant estimation methods. Furthermore, extensive tests on real-world images show that we can reliably process mixed illuminant scenes.

1. Introduction

Whenever an imaging algorithm relies on surface color information, the color of the illumination must be neutralized. Such a process typically requires knowledge about the color of the illuminant. Hence, a number of diverse techniques for extracting the illuminant chromaticity have been developed, ranging from statistical e.g. [4, 13, 32] to physics-based ones e.g. [16, 22, 31]. Because the problem of determining the illuminant color of an arbitrary image is underconstrained, these methods generally assume that the scene is uniformly illuminated. For an overview on existing algorithms, see e.g. [18].

Many images, however, exhibit a mixture of illuminants with distinct chromaticities. Consider, for example indoor scenes which are lit by both indoor light sources and light coming through the windows, or pictures taken using a camera-flash. In outdoor scenes, parts of the image may

be in shadow while others are lit by sunlight or street lamps etc. These multiple illuminant effects are further exaggerated by inter-reflections. Illuminant estimation methods that assume uniform illumination can not accurately recover the illuminant chromaticity in these cases.

There are illuminant estimation methods explicitly designed to handle varying illumination. In 1997, Barnard et al. [1] were the first ones to develop a methodology that automatically detects non-uniform illumination. They then proceeded with removing the illumination variation, at which point they could apply any gamut-based color constancy method such as [10]. Though this method was pioneering at that time, its assumptions of smooth illumination and of a set of common illuminants, restrict its applicability on real world images. Ebner [8] followed a different approach of applying a diffusion-based approach on the pixel intensities. However, he too assumes a smoothly varying illumination, which together with his underlying theory of regional gray-world can result in inaccuracies, especially in colorful scenes [20]. More recently, Kawakami et al. [21] proposed a physics-based method specifically designed to handle illumination variations between shadowed and non-shadowed regions in outdoor scenes. Due to its explicit assumption of hard shadows and sky-light/sunlight combination (or even more general Planckian illuminants), this method does not generalize well on arbitrary images. Thus, by construction, none of the existing multi-illuminant estimation methods can handle arbitrary images and as such, none of them has been extensively tested on a large variety of real images.

We propose an illuminant color estimation method which can explicitly handle multiple illuminants and is designed for real-world images, such as the ones typically found on the web. Like Ebner [8] we use a local approach, but employ the more physically-accurate dichromatic model which, together with multiple sampling, results in statistically-robust illuminant-color estimates. In order to extract the local illuminant-color information we extended the method of Tan et al. [31] so as to a) avoid specular pre-segmentation and b) expressly test for conformity to

the underlying theoretical assumptions. Since our method is physics-based, its effectiveness is independent of the quality of training data, a known weakness of the very successful learning-based techniques [2]. We identify regions of different illuminants by creating local estimates and merging them together to major areas of highly similar illumination. Finally, we end up with a small number of dominant illuminants in the scene.

Contributions. Our proposed technique for illuminant-color estimation under multiple illuminants makes the following contributions:

- It introduces a localized variant of the Tan et al. [31] method which avoids specular pre-segmentation.
- It explicitly takes advantage of image redundancy (i.e. many small regions are illuminated by the same light source) through the computation of multiple independent estimates.
- It groups the estimates for identifying multiple illuminants. In this method, the image regions where these illuminants have their strongest influence, as well as the color of these illuminants, are estimated under minimal additional assumptions.

Our methodology has been extensively tested on benchmark databases of both laboratory and real-world images. To our knowledge, the proposed technique is the first physics-based technique tested on a large database of real images (more than 11,000 images). In each of the databases we got a median angular error of 4.4° . In the real-world database we outperformed all comparable methods. Please note, that all reported results are for single color constancy methods. It has been shown [17, 27, 29], that the fusion of multiple color constancy methodologies significantly improves performance. Our goal is to introduce a new illuminant estimation method which could then be incorporated into a multi-method scheme. Furthermore, we show how information about scenes with multiple illuminants can be extracted. To the best of our knowledge, no ground-truth dataset exists for mixed illuminants. Thus, we present qualitative results to demonstrate the feasibility of this approach. Our estimates determine only the dominant illuminant for a particular image region. Thus, in the current state, a mixture of two illuminants is not explicitly modeled.

The paper is organized as follows. Section 2 provides a justification for selecting the Tan et al. [31] method as the starting point of our methodology and briefly describes the inverse intensity chromaticity space. In Section 3 we present our method for local illumination-color estimation. Section 4 describes how the local estimates are used in extracting the number of illuminants. Our experimental results are shown in Section 5, followed by a brief discussion in Section 6.

2. Background on Illuminant-Color Estimation

There is a large body of work on illuminant estimation under the assumption of uniform illumination, which could be used as a building block for a multiple-illuminant estimation method. Most of the uniform illuminant methods have to either learn illuminant classes, which however do not generalize well, or make restrictive assumptions. For example, some techniques, which indirectly estimate the illuminant color, work best for Mondrian worlds e.g. [11, 15, 23] or make assumptions about the overall scene composition e.g. [5, 23]. Others, like the methodology proposed by Tominaga and Wandell [32], assume prior knowledge, like known camera filter sensitivities, which can often not be obtained. The methods of Brainard and Freeman [4], Finlayson et al. [13] and Geusebroek et al. [16] assume the scene is made of diffuse surfaces, while others, e.g. [22, 24, 26, 31] exploit specular highlights. Though such assumptions are justifiable, they often limit the applicability of these methods on real-world images. Thus, previous work on uniform illuminant estimation on arbitrary scenes is mostly based on machine learning techniques, e.g. [6, 12, 13, 17, 27].

One could develop a multiple-illuminant estimation method as an extension of one of the learning-based methods. However, its effectiveness would depend on the training samples [2], which for the large variety of possible mixed illuminants would be challenging to obtain. Instead, we chose to use Tan et al.’s [31] physics-based method as the springboard for our multiple-illuminant estimation algorithm. We selected [31] out of the various physics-based methods for two main reasons: a) it exploits surfaces that exhibit a mixture of diffuse and specular reflectance, which are typically more common than purely diffuse surfaces and b) more importantly, when evaluated on the Ciurea-Funt benchmark database [7] its median angular error between 5.12° and 5.9° outperformed all other physics-based methods and was comparable with the best-performing learning-based methods (for more details see Section 5).

The method of Tan et al. [31] is based on the dichromatic reflection model [30], which states that the amount of light reflected from a point, \mathbf{x} , of a dielectric, non-uniform material is a linear combination of diffuse reflection and specular reflection. Furthermore, it is assumed that the color of the specular reflection approximates the color of the incident light. Thus, the camera response $I_c(\mathbf{x})$ for each color filter c is:

$$I_c(\mathbf{x}) = m_d(\mathbf{x})\Lambda_c(\mathbf{x}) + m_s(\mathbf{x})\Gamma_c \quad (1)$$

where $m_d(\mathbf{x})$ and $m_s(\mathbf{x})$ subsume the geometric parameters of diffuse and specular reflection respectively. $\Lambda_c(\mathbf{x}) = B_c(\mathbf{x})/\sum_i B_i(\mathbf{x})$ and $\Gamma_c = G_c/\sum_i G_i$, $i \in \{R, G, B\}$, are

the diffuse and specular chromaticities with $B_c(\mathbf{x})$ and G_c being the sensor response to the diffuse and specular reflectance accordingly. For the remainder of the paper, we define $i \in \{R, G, B\}$ and use this index for summing over the color channels.

The image chromaticity, σ_c , is similarly defined: $\sigma_c(\mathbf{x}) = I_c(\mathbf{x}) / \sum_i I_i(\mathbf{x})$. Tan et al. [31] investigated the relationship between diffuse, specular and image chromaticities and showed that there is a linear relationship between the image chromaticity σ_c and the inverse-intensity $1 / \sum_i I_i(\mathbf{x})$:

$$\sigma_c(\mathbf{x}) = p_c(\mathbf{x}) \frac{1}{\sum_i I_i(\mathbf{x})} + \Gamma_c . \quad (2)$$

where $p_c(\mathbf{x})$ is defined as $p_c(\mathbf{x}) = m_d(\mathbf{x})(\Lambda_c(\mathbf{x}) - \Gamma_c)$. In this representation, $p_c(\mathbf{x})$ can be seen as the slope of a line with intercept Γ_c , i.e. the specular chromaticity, which is also the illuminant chromaticity. The domain of the line is determined by $1 / \sum_i I_i(\mathbf{x})$ and the range is given by $0 \leq \sigma_c \leq 1$.

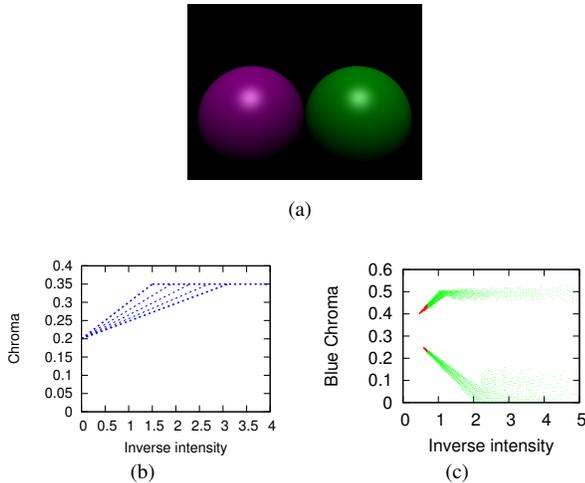


Figure 1. Sample pixel distributions in IIC space (blue chromaticity). Top: synthetic image. Bottom left: ideal distribution. Bottom right: The highly specular pixels are shown in red.

This linear relationship can be easily visualized in the *inverse-intensity chromaticity* (IIC) space [31]. Fig. 1 shows sample distributions of pixels in IIC space. The horizontal axis depicts the inverse-intensity $1 / \sum_i I_i(\mathbf{x})$, and the vertical axis σ_c , the illuminant chromaticity. Fig. 1(b) is an ideal distribution of pixels of a monochrome object in IIC space. The diffuse pixels lie on a single horizontal line, while pixels that exhibit specular reflection align, according to their specific $p_c(\mathbf{x})$ -values, in lines between the illuminant color on the vertical axis and the diffuse line. Fig. 1(c) is the pixel distribution for the synthetic image with two distinct albedos shown in Fig. 1(a). Note that the chromaticity value

where the specular clusters converge and intersect the vertical axis is the illuminant chroma estimate.

Though the formulation is mathematically elegant, it is, in general, not possible to directly compute $p_c(\mathbf{x})$ and consequently Γ_c . One could, however, exploit the distribution of pixels in IIC space in order to detect the σ_c intercept. Tan et al. [31] developed such a methodology. They relied on a specularity-segmentation pre-processing step to identify pixels that lie in the specular locus (in Fig. 1(c), highly specular pixels are plotted red for the purpose of illustration). Each such highly specular pixel contributed to the estimation of a single illuminant color via a Hough transform with parameters $p_c(\mathbf{x})$ and Γ_c . For a complete discussion on the properties of IIC-space, see [31].

3. Local Illuminant-Color Estimates

The investigation of chromaticity distributions in IIC-space can also be performed per image region. Such a localized investigation has several advantages, besides allowing for the extraction of local illumination estimates:

1. It is best suited for **complex colorful scenes**, whose chromaticity distributions do not form clearly separable clusters (see Fig. 2).
2. One can explicitly test whether the selected region forms a non-horizontal (i.e. non-diffuse) cluster in IIC-space. As a consequence: a) one **avoids the specular segmentation** step, which can be unreliable and thus lead to imprecise illuminant color estimation [14, 28] and b) typically, a larger number of pixels is used in the illuminant-color estimation, making the result **less sensitive to outliers**.
3. Local independent analysis of different regions generates distinct sources of information on the illuminant. By combining these independent illuminant estimates one can obtain a **statistically robust** illuminant-color estimate.

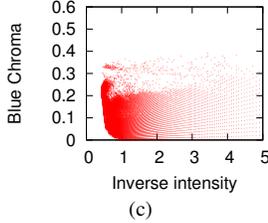
3.1. Region Screening

Consider a small image region r_j which can be used for estimating the illuminant color. According to eq.(2), purely diffuse pixels should be excluded from the computation of the σ_c intercept. Thus, if r_j is a purely diffuse region (i.e. it forms a horizontal cluster in IIC space), it should not be used in the Tan et al. [31] method. Furthermore, if the non-diffuse pixels have the same underlying albedo, they will form tighter clusters and thus tighter convergence to σ_c intercept. Thus, instead of pre-segmenting the image specularities, one can select an image region r_j and verify whether it satisfies the underlying assumptions of [31]. We propose the following conformity criteria:

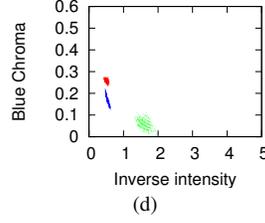


(a)

(b)



(c)



(d)

Figure 2. (a) Example real-world image. (b) Selected regions. (c) The distribution of all the pixels of image (a) in IIC space (blue chromaticity). (d) The distribution of pixels of the regions from (b) in IIC space (blue chromaticity).

- uniform albedo,
- elongated, non-horizontal clusters in IIC space.

The following region selection process increases the probability that a selected region r_j satisfies the conformity criteria.

1. **Superpixel segmentation** Segment the image in superpixels of approximately uniform chromaticity values. A superpixel is a locally connected region of pixels that share low-level properties, like in our case similar chromaticity values. We use the graph-based segmentation by Felzenszwalb and Huttenlocher [9], but any segmentation method that decomposes an image into regions with approximately the same albedo could also be employed.
2. **Select mini-regions within each superpixel** Sample with replacement small regions within superpixels with probability proportional to the size of the superpixel. Any iid sampling which results in small regions of approximately uniform albedo could be employed.
3. **Exclude horizontal symmetric clusters** Examine the shape of the distribution of pixels in the candidate region. One way of doing this is via PCA. Let P_{IIC} be the set of pixels under investigation in IIC space, λ_1 its largest eigenvalue, λ_2 its second largest eigenvalue.

Then the eccentricity $\text{ecc}(P_{IIC})$ is

$$\text{ecc}(P_{IIC}) = \sqrt{1 - \frac{\sqrt{\lambda_2}}{\sqrt{\lambda_1}}} . \quad (3)$$

We consider only sets P_{IIC} that have at least an order of magnitude difference between the minor and major axes of the covariance ellipse, i.e. $\text{ecc}(P_{IIC}) > 0.94$. In order to avoid purely diffuse pixels we compute also the slope of the eigenvector \mathbf{v}_1 of λ_1 . A set P_{IIC} must also satisfy a minimum slope (0.003, in our experiments). See Section 5 for further discussion on the region size.

A mini-region r_j selected through this process generates an illuminant estimate $\hat{\mathbf{g}}_j$ by computing the point of intersection of \mathbf{v}_1 with the σ_c axis. Please note that like [31] we exclude pixels with duplicate values when generating the distribution of pixels in IIC space. We also exclude any pixels that are very close to the limits of the dynamic range of the camera (i.e. saturated and very dim pixels). Note that the mathematical formulation of the algorithm assumes linear camera response. Since real-world images contain typically a gamma factor, it might be necessary to correct for this, using *e.g.* the method by Lin *et al.* [25].

3.2. Multiple Samples

One of the goals of our local illumination methodology is to provide statistically robust estimates. We, thus, take advantage of the redundancy of information that is typically available in an image: nearby mini-regions are often illuminated by approximately the same illuminant. Hence, one of the key ideas of the proposed methodology is the derivation of a robust local illuminant estimate $\mathbf{\Gamma} = (\Gamma_R, \Gamma_G, \Gamma_B)$ per superpixel through the use of multiple mini-region estimates $\hat{\mathbf{g}}_j$. This superpixel estimate is obtained from k independent and identically distributed (iid) samples.

The overall goal is to minimize the estimation error E ,

$$E = \cos^{-1} \left(\frac{\mathbf{\Gamma} \cdot \hat{\mathbf{\Gamma}}}{\|\mathbf{\Gamma}\| \|\hat{\mathbf{\Gamma}}\|} \right), \quad (4)$$

between the true illuminant $\mathbf{\Gamma}$ and the final estimate $\hat{\mathbf{\Gamma}}$. Our approach is to sample over each superpixel. This leads to a set $\mathcal{E} = \{\hat{\mathbf{g}}_j | j = 1..k\}$ of iid estimates. This set consists of “good” samples $\mathcal{G} = \{\hat{\mathbf{g}}_j | \mathbf{n}_j < \epsilon\}$, $\hat{\mathbf{g}}_j = \mathbf{g}_j + \mathbf{n}_j$ (where \mathbf{n}_j is noise), and of “bad” samples $\mathcal{N} = \{\hat{\mathbf{g}}_j | \mathbf{n}_j \geq \epsilon\}$,

$$\mathcal{E} = \mathcal{G} \cup \mathcal{N} . \quad (5)$$

Then, the elements of \mathcal{G} form a unimodal distribution around the true illuminant $\mathbf{\Gamma}$, such that

$$\lim_{|\mathcal{G}| \rightarrow \infty} \text{argmax}_{\mathbf{\Gamma}} \text{Hist}(\mathcal{G}) = \mathbf{\Gamma} , \quad (6)$$

where $\text{Hist}(\mathcal{G})$ denotes the histogram of the illuminant estimates in \mathcal{G} . The elements of \mathcal{N} can be arbitrarily distributed. Our goal is to reduce the influence of \mathcal{N} while preserving \mathcal{G} , so that finally

$$\lim_{|\mathcal{E}| \rightarrow \infty} \operatorname{argmax} \text{Hist}(\mathcal{E}) = \Gamma. \quad (7)$$

By verifying that each sample region r_j satisfies the conformity criteria (through the process described in 3.1) we increase the probability that the estimate \mathbf{g}_j obtained from such a local region r_j will be a good estimate (i.e. $\mathbf{g}_j \in \mathcal{G}$).

4. Multiple Illuminants

Once local illuminant estimates are obtained per superpixel, the local information can be combined as follows for the final computation of the number and color of the dominant illuminants in the scene.

1. Group local estimates into regions with consistent/similar illuminant color.
2. Obtain a new estimate per illuminant region.

An example of this process is shown in Fig. 3. The details of each of the aforementioned steps are provided in the following subsections.

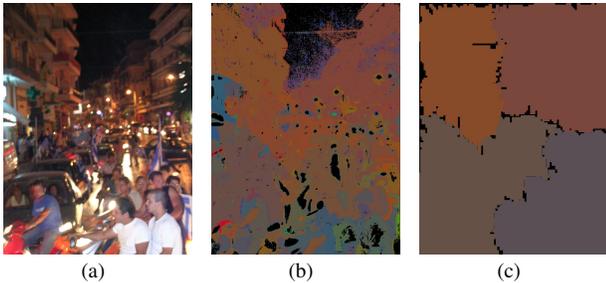


Figure 3. (a) Original image (b) Local illuminant estimation (c) Segmented regions, colored according to the illuminant estimate.

We extended our algorithm as described in Sec. 3 to handle multiple illuminants by examining the estimates per superpixel more closely. Note two assumptions. First, multiple illuminants are often clearly visible in the superpixel map, see Fig. 3 (b) for an example. Second, outlier estimates occur typically isolated, both spatially and in the distribution of estimated colors. In order to extract the regions of the dominant regions, we do the following steps.

1. Create an illuminant map by recoloring every superpixel by its local illuminant estimate.
2. Downscale the map, such that the larger dimension of this image is only 140 pixels.

3. Group regions of similar estimates with the Quick Shift algorithm [33].

The downscaling suppresses a large amount of relatively small noisy regions. Its purpose is to speed up the Quick Shift algorithm. Quick Shift is a method for seeking modes in densities, which is why we preferred it over [9] for grouping similar estimates. In our case, we obtained the best results by applying it on the joint spatial and chromaticity domain, using red and blue chromaticities. Quick Shift creates trees of data points and distances between these nodes, such that similar regions can be segmented by separating subtrees from this graph. By discarding smaller segments, we typically obtain three to six major regions in the downscaled image.

For refining the estimation, we use the estimated illuminant regions for iid sampling instead of the whole image. The resulting per-region illuminant estimates can further be merged. In this work, we merged regions that were smaller than a predefined threshold of 10% of the image region.

5. Experiments

We quantitatively evaluated our methodology on two widely used datasets. The first dataset contains images taken indoors under tightly controlled imaging conditions. The second database contains real-world images of both indoor and outdoor scenes which are more representative of the type of arbitrary images often found on the web. For the detection of multiple illuminants, we present qualitative results on images downloaded from websites like flickr. We also examined the images used by Hsu et al. [20]. The code for our method can be downloaded from the web¹.

5.1. Error measure for benchmark data

The error metric used in the evaluation of the two benchmark datasets is the angular error e ,

$$e = \cos^{-1} \left(\frac{\Gamma_l \cdot \Gamma_e}{\|\Gamma_l\| \|\Gamma_e\|} \right), \quad (8)$$

between the ground truth illuminant color Γ_l and the estimated color Γ_e . To summarize the angular errors for the different images of a dataset, the median of the estimates is computed, as recommended in [19].

5.2. Parameter selection

For the segmentation of the chromaticity images by Felzenszwalb and Huttenlocher [9], the parameters were fixed by visual inspection to $\sigma = 0.3$, $k = 200$, and minimum segment size $m = 15$. The sampling rectangle size was set to 7×31 pixels. Our tests, however, indicated that the lab database was more challenging for our methodology,

¹<http://www5.cs.fau.de/>



Figure 4. Examples of benchmark laboratory images.

Scene	Median e
Gamut mapping	3.1°
Gray-World	8.8°
White-Patch	5.0°
Color-by-Correlation	8.6°
Original IIC Method	-
Physics-based diff+spec	4.4°

Table 1. Algorithm performance on benchmark laboratory images.

since it did not closely satisfy our design criteria. Hence, for the lab images we tried different rectangles and concluded that a larger size of 30×55 pixels gave the best performance.

5.3. Benchmark laboratory images under uniform illumination

The database of Barnard et al. [3] contains images of relatively simple scenes (few objects, uniform background) illuminated by a single light source and captured with a high quality 3-CCD camera (Sony DXC-930). There are 31 different scenes taken under 11 different illuminants. Out of the 31 scenes, only 9 include items which are not purely diffuse. Thus, we tested our methodology on these 9 scenes under all available illuminants, for a total of 99 test images. A subset of these images is shown in Fig. 4.

Table 1 summarizes the performance of the presented methodology in comparison to state-of-the-art algorithms. Our physics-based technique is only outperformed by the gamut mapping, which, however, is dependent on a training stage. With the original method by Tan et al. [31], it turned out that the specular segmentation parameters are not easy to handle. Manual adjustment of the parameters for every image individually gave very good results. However, we were unable to find a single fixed parameter set that gave satisfying results on the whole database.

5.4. Benchmark real-world images under uniform illumination

The database of Ciurea and Funt [7] includes images that are more representative of the pictures taken by arbitrary users. The dataset contains around 11,000 images from 15 quite diverse scenes. Sample images are shown in Fig. 5.

As can be seen in Table 2, the proposed physics-based method achieves a significant improvement over different state-of-the-art methods. The referenced angular errors



Figure 5. Examples of benchmark real-world images.

Scene	Median e
Regular gamut with offset-model	5.7°
Gray-World	7.0°
White-Patch	6.7°
Color-by-Correlation	6.5°
1 st -order Gray-Edge	5.2° (*)
2 nd -order Gray-Edge	5.4° (*)
Original IIC-based method	5.1° (*)
Physics-based diff+spec	4.4°

Table 2. Algorithm performance on benchmark real-world images.

Scene	Γ_r	Γ_g	Γ_b
Town	0.37 ± 0.02	0.34 ± 0.01	0.29 ± 0.02
Woman	0.33 ± 0.01	0.34 ± 0.01	0.33 ± 0.00
Pool	0.42 ± 0.04	0.37 ± 0.02	0.21 ± 0.02
Sculpture	0.31 ± 0.07	0.36 ± 0.05	0.33 ± 0.08

Table 3. Stability of the algorithm results on single-illuminant real-world images. For the images of Fig.6, the mean chromaticities and standard deviations of ten estimation runs are listed.

marked with an asterisk (*) are taken from [27] and are evaluated only on a subset of 711 images. The remaining measurements are extracted from [17] and are, like our evaluation, computed on the entire set of 11,000 images.

Out of the 15 provided scenes, the best result was obtained for “FalseCreek1” ($e = 1.57^\circ$), while “CIC2002.3” resulted in the worst performance ($e = 11.46^\circ$).

5.5. Arbitrary real-world images under mixed illumination

Since our algorithm was designed for illuminant estimation of images typically found on the web, we performed a qualitative evaluation on a set of almost 300 images we downloaded from various websites. The database contains images both of indoor and outdoor scenes (see Fig. 6), and different subjects, such as nature, people, animals and architecture. Within this set, we collected also about 30 mixed-illuminant scenes.

To illustrate the performance of the physics-based estimation method on different lighting and scene contents, we present estimates we obtained on a subset of representative images. The example images in Fig. 6 are all captured under different illumination conditions. Table 3 lists the corresponding estimates. For each scene, the mean estimate of

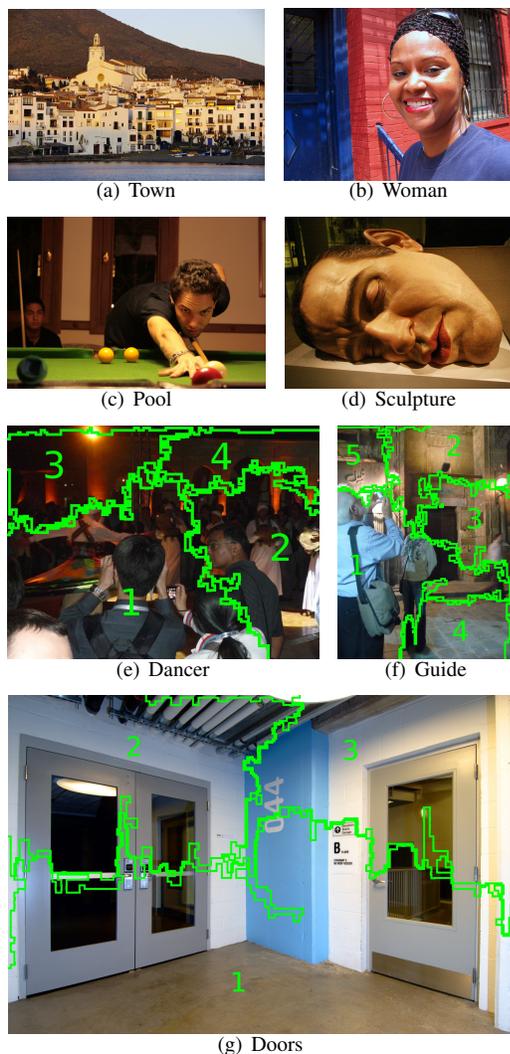


Figure 6. Subset of the selected real-world images. Single-illuminant images are shown in the first two rows. In the two bottom rows, the multi-illuminant images are annotated with the segment numbers.

Scene	Segment 1	Segment 2
Dancer	(0.327, 0.336, 0.337)	(0.330, 0.334, 0.336)
Guide	(0.312, 0.343, 0.345)	(0.347, 0.336, 0.316)
Doors	(0.309, 0.339, 0.352)	(0.294, 0.337, 0.369)
Scene	Segment 3	Segment 4
Dancer	(0.415, 0.306, 0.279)	(0.354, 0.319, 0.327)
Guide	(0.343, 0.327, 0.330)	(0.331, 0.334, 0.335)
Doors	(0.379, 0.334, 0.287)	-

Table 4. Per segment illuminant chromaticity estimates for the multi-illuminant images.

ten randomized runs is given in combination with the standard deviation.

One can expect that the red component of the illuminant color in the outdoor scenes decreases between “Town”, Fig. 6(a) and “Woman” Fig. 6(b). At the same time, the blue component is increasing. This tendency is captured quite well in the estimation results (Table 3). For the indoor images (Fig. 6(c) and Fig. 6(d)), the estimation for “Pool” contains a strong red portion as expected, while the greenish estimation for “Sculpture” is most likely a failure case.

The two next rows of Fig. 6 contain example scenes with multiple illuminants. The spatial location of the segments is denoted by their overlaid respective numbers in the images. In Fig. 6(e), flash light illuminates the heads of the spectators, while the remaining scene is mainly reddish illuminates. In Fig. 6(f), the tourist in the foreground is illuminated from behind by a blueish light source. The rest of the scene contains mainly light from the lamps. Finally, Fig. 6(g) is taken from the dataset by Hsu et al. [20]. Table 4 shows the illuminant estimates per segment. The tendency of the illuminant colors is well captured by the localized estimates.

6. Discussion

We consider these results as a starting point for investigating the recovery of multiple, spatially distributed illuminants. Estimating on smaller image regions leads naturally to locally weaker results. We aim to compensate this loss with a combination of local estimation, grouping of similar estimates, and reestimation on a coarser scale. Qualitative results on a number of images looked promising. One drawback is the lack of ground truth for scenes under non-uniform illumination. In future work, we aim to capture and use such data. A limitation of the presented method is that only one dominant illuminant can be estimated per region. A more elegant formulation, also subject to future work, would probably estimate a set of scene illuminants, and a per-region contribution of each of these.

7. Conclusions and future work

In this paper we have introduced a physics-based method for the automated estimation of the illuminant color which was specifically designed for handling real-world images with multiple illuminants. Its building block is the computation of statistically robust local illuminant estimates which are then used in deriving the number and color of the dominant illuminants. When tested on a large benchmark database for uniform illuminant tests, it performed comparably to state-of-the-art uniform illuminant methods. Qualitative experiments on real-world images with mixed illuminants demonstrated the effectiveness of our method. In future work, we plan to collect ground truth data under non-uniform illumination for quantitative evaluation. Additionally, we seek a formulation to estimate the per-segment

contribution of a light source, as stated in Section 6.

References

- [1] K. Barnard, G. Finlayson, and B. Funt. Color Constancy for Scenes with Varying Illumination. *Computer Vision and Image Understanding*, 65(2):311–321, Feb. 1997.
- [2] K. Barnard, L. Martin, A. Coath, and B. Funt. A Comparison of Computational Color Constancy Algorithms – Part II: Experiments With Image Data. *IEEE Transactions on Image Processing*, 11(9):985–996, Sept. 2002.
- [3] K. Barnard, L. Martin, B. Funt, and A. Coath. A Data Set for Color Research. *Color Research and Application*, 27(3):147–151, 2002.
- [4] D. H. Brainard and W. T. Freeman. Bayesian Color Constancy. *Journal of the Optical Society of America A*, 14(7):1393–1411, 1997.
- [5] G. Buchsbaum. A Spatial Processor Model for Color Perception. *Journal of the Franklin Institute*, 310(1):1–26, July 1980.
- [6] V. C. Cardei, B. Funt, and K. Barnard. Estimating the Scene Illumination Chromaticity Using a Neural network. *Journal of the Optical Society of America A*, 19(12):2374–2386, Dec. 2002.
- [7] F. Ciurea and B. Funt. A Large Image Database for Color Constancy Research. In *Color Imaging Conference*, pages 160–164, 2003.
- [8] M. Ebner. Color Constancy Using Local Color Shifts. In *European Conference in Computer Vision*, pages 276–287. Springer-Verlag, 2004.
- [9] P. F. Felzenszwalb and D. P. Huttenlocher. Efficient Graph-based Image Segmentation. *International Journal of Computer Vision*, 59(2):167–181, 2004.
- [10] G. D. Finlayson. Color Constancy in Diagonal Chromaticity Space. In *IEEE International Conference on Computer Vision*, pages 218–223, 1995.
- [11] G. D. Finlayson, B. V. Funt, and K. Barnard. Color Constancy Under Varying Illumination. In *IEEE International Conference on Computer Vision*, pages 785–790, 1995.
- [12] G. D. Finlayson, S. D. Hordley, and P. M. Hubel. Color by Correlation: A Simple, Unifying Framework for Color Constancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1209–1221, Nov. 2001.
- [13] G. D. Finlayson, S. D. Hordley, and I. Tastl. Gamut Constrained Illuminant Estimation. *International Journal of Computer Vision*, 67(1):93–109, 2006.
- [14] G. D. Finlayson and G. Schaefer. Convex and Non-convex Illuminant Constraints for Dichromatic Colour Constancy. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 598–604, 2001.
- [15] D. Forsyth. A Novel Algorithm for Color Constancy. *International Journal of Computer Vision*, 5(1):5–36, 1990.
- [16] J.-M. Geusebroek, R. Boomgaard, A. Smeulders, and T. Gevers. Color Constancy from Physical Principles. *Pattern Recognition Letters*, 24(11):1653–1662, July 2003.
- [17] A. Gijsenij, T. Gevers, and J. van de Weijer. Generalized Gamut Mapping using Image Derivative Structures for Color Constancy. *International Journal of Computer Vision*, 86(2–3):127–139, Jan. 2010.
- [18] A. Gijsenij, T. Gevers, and J. van de Weijer. Computational Color Constancy: Survey and Experiments. *IEEE Transactions on Image Processing*, 20(9):2475–2489, 9 2011.
- [19] S. D. Hordley and G. D. Finlayson. Re-evaluating Color Constancy Algorithm Performance. *Journal of the Optical Society of America A*, 23(5):1008–1020, May 2006.
- [20] E. Hsu, T. Mertens, S. Paris, S. Avidan, and F. Durand. Light Mixture Estimation for Spatially Varying White Balance. *ACM Transactions on Graphics*, 27(3):70:1–70:7, Aug. 2008.
- [21] R. Kawakami, K. Ikeuchi, and R. T. Tan. Consistent Surface Color for Texturing Large Objects in Outdoor Scenes. In *IEEE International Conference on Computer Vision*, pages 1200–1207, 2005.
- [22] G. J. Klinker, S. A. Shafer, and T. Kanade. The Measurement of Highlights in Color Images. *International Journal of Computer Vision*, 2(1):7–26, 1992.
- [23] E. H. Land. Lightness and the Retinex Theory. *Scientific American*, 237(6):108–129, Dec. 1977.
- [24] H.-C. Lee. Method for Computing the Scene-Illuminant Chromaticity from Specular Highlights. *Journal of the Optical Society of America A*, 3(10):1694–1699, 1986.
- [25] S. Lin, J. Gu, S. Yamazaki, and H.-Y. Shum. Radiometric Calibration from a Single Image. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 938–945, 2004.
- [26] S. Lin and H.-Y. Shum. Separation of Diffuse and Specular Reflection in Color Images. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 341–346, 2001.
- [27] R. Lu, A. Gijsenij, T. Gevers, V. Nedovic, D. Xu, and J.-M. Geusebroek. Color Constancy using 3D Scene Geometry. In *IEEE International Conference on Computer Vision*, 2009.
- [28] C. Riess and E. Angelopoulou. Physics-Based Illuminant Color Estimation as an Image Semantics Clue. In *International Conference on Image Processing*, 2009.
- [29] G. Schaefer, S. D. Hordley, and G. D. Finlayson. A Combined Physical and Statistical Approach to Colour Constancy. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 148–153, 2005.
- [30] S. A. Shafer. Using Color to Separate Reflection Components. *Journal Color Research and Application*, 10(4):210–218, 1985.
- [31] R. Tan, K. Nishino, and K. Ikeuchi. Color Constancy through Inverse-Intensity Chromaticity Space. *Journal of the Optical Society of America A*, 21(3):321–334, 2004.
- [32] S. Tominaga and B. A. Wandell. Natural Scene-Illuminant Estimation using the Sensor Correlation. In *Proceedings of the IEEE*, pages 42–56, 2002.
- [33] A. Vedaldi and S. Soatto. Quick Shift and Kernel Methods for Mode Seeking. In *European Conference on Computer Vision*, pages 705–718, 2008.