Illuminant Estimation by Voting

Christian Riess[1], Eva Eibenberger[1,2], Dr. Elli Angelopoulou[1,2]
[1]Chair of Pattern Recognition
Department of Computer Science
Friedrich-Alexander University Erlangen-Nuremberg,
Martensstr. 3, 91058 Erlangen, Germany
[2]Erlangen Graduate School in Advanced Optical Technologies (SAOT)
{eva.eibenberger, christian.riess, elli.angelopoulou}@informatik.uni-erlangen.de

Abstract

Obtaining an estimate of the illuminant color is an important component in many image analysis applications. Due to the complexity of the problem many restrictive assumptions are commonly applied, making the existing illuminant estimation methodologies not widely applicable on natural images. We propose a methodology which analyzes a large number of regions in an image. An illuminant estimate is obtained independently from each region and a global illumination color is computed by consensus. Each region itself is mainly composed by pixels which simultaneously exhibit both diffuse and specular reflection. This allows for a larger inclusion of pixels than purely specularity-based methods, while avoiding, at the same time, some of the restrictive assumptions of purely diffuse-based approaches. As such, our technique is particularly well-suited for analyzing real-world images. Experiments with laboratory data show that our methodology outperforms 75% of other illuminant estimation methods. On natural images, the algorithm is very stable and provides qualitatively correct estimates.

1. Introduction

Estimating the color of the illuminant has an impact on a variety of computer vision problems and imaging applications, including computational color constancy e.g. [9, 13, 23, 26, 31], image database retrieval [20, 35, 38], segmentation [2, 5] and color normalization accross images [6, 37]. Funt *et al*. concluded on early color constancy algorithms, that there is a high demand in the community for better and more stable methods [19]. Hence, a number of diverse techniques for extracting the illuminant chromaticity has been developed, ranging from statistical e.g. [4, 13, 40] to physics-based ones e.g. [9, 10, 21, 24, 26, 29, 39]. Because the problem of determining the illuminant color of an arbitrary image is underconstrained, most of these methods have to either learn illuminant classes which, however, do not generalize well, or make restrictive assumptions.

Previous work on illuminant estimation on arbitrary scenes is mostly based on machine learning techniques, e.g. [1, 7, 13, 12]. Depending on their training samples, machine learning approaches like the one of Cardei *et al*. [7] are not easily generalizable [3]. The methods of Brainard and Freeman [4], Finlayson *et al*. [13] and Geusebroek *et al*. [21] assume purely diffuse reflectance. Others look explicitly for specular highlights e.g. [24, 26, 29, 39]. Another family of techniques, which indirectly estimates the illuminant color, works best for Mondrian worlds e.g. [11, 18, 25]. There is also some previous work on estimating the type of illuminant present in a natural scene [40]. However, the majority of these methods either make restrictive assumptions or are by-design not applicable to images typically found on the web. For example, the methodology proposed by Tominaga and Wandell [40] assumes that the filter responses of the camera are known, which is usually not the case for the majority of images. Though making such assumptions is justifiable, it often limits the applicability of these methods on natural images. Recently, the group of gray-world related algorithms experienced a revival [14, 41, 22]. Modern versions of the gray-world algorithms assume certain image statistics to be "gray" in the sense of the gray-world assumption. Since every variant has its own advantages and disadvantages, Gijsenij and Gevers [22] presented a hybrid algorithm that selects with a machine learning approach the best suited gray-world-based algorithm.

There are a number of successful image analysis methods, which are based on physical and/or geometrical models (e.g. edge detection, stereo, tracking). Of course, most of them are tackling simpler problems, either by definition (e.g. circle detection), or by imposing constraints that may not be applicable to a large percentage of the images typ-

ically found on the web. Nonetheless, such physics-based methods can be used in augmenting semantic image analysis, if one is willing to replace the precise quantitative measurements, like metric distances, spectra, angle or velocity values, with more abstract descriptions. In this paper we show how a state-of-the-art physics-based technique for estimating the illuminant-color based on specular highlights [39] can be adapted for obtaining an estimate of the type of illumant in arbitrary images found on the web. This involves the treatment of difficulties in the segmentation of specularities and ways to assess the compliance of the obtained intensities with our model.

We propose the transfer of physics-based methodologies to semantics-oriented image analysis for three main reasons. First of all, when looking at images, people are able to compensate for variations in appearance caused by changes in the illuminant color (color constancy). There is also evidence that illumination clues, like specular highlights, are explicitly used in the perceptual analysis of scenes [32, 15]. Secondly, as in semantic analysis, previous work on illuminant estimation on arbitrary scenes is mostly based on machine learning techniques, e.g. [7, 1, 13, 30]. As such, their success is closely tied to the training samples. Lastly, one of the most popular families of illumination estimation techniques are either specularity based [24, 26, 29, 39] or work on purely diffuse pixels. To the best of our knowledge, there exists no physics-basd methodology that explicitly exploits a mixture of diffuse and specular pixels.

We propose a methodology for estimating the illuminant color, which is explicitly designed to take advantage of the variety of illumination cues that are present in real images. We exploit the distribution of color pixels in chromaticity space like [24, 26, 28, 39]. Unlike these methods, we focus on pixels that simultaneously exhibit both diffuse and specular reflection. This significantly increases the number of useful pixels for illumination analysis in a natural image, compared to pure specularity-based methodologies. Furthermore, it avoids many of the assumptions made by illuminant-estimation methods which are based on diffuse reflection. Under uniform illumination, local independent analysis of different regions in the image generates distinct sources of information on the illuminant. By combining these independent illuminant estimates our method counterbalances two of the bigger drawbacks in natural images: noise and violations of the reflectance assumptions. Our algorithm does not need training and typically only few parameters need to be adjusted. Experiments on publicly available ground truth data of laboratory images show that our method performs in the top 25% of other tested illuminant estimation methods. When applied on images found on the web, our technique produces stable and reasonable estimates of the illuminant color.

The paper is organized as follows. In section 2, we review the image formation process and show how our algorithm relates to other illuminant estimation methodologies. In section 3, the inverse intensity chromaticity space introduced by Tan *et al*. [39] is explained. The main idea of our methodology is presented in section 4, followed by a detailed description in section 5. Experimental results are presented in section 6.

2. Image Formation

Most surfaces exhibit a mixture of diffuse and specular reflectance. According to the dichromatic reflection model [36], the amount of light L reflected from a point, $\mathbf{x}$, of a dielectric, non-uniform material is a linear combination of body (or diffuse) reflection L_d and surface (or specular) reflection L_s:

$$L(\lambda, \mathbf{x}) = w_d(\mathbf{x}) L_d(\lambda, \mathbf{x}) + w_s(\mathbf{x}) L_s(\lambda, \mathbf{x}) \ , \quad (1)$$

where λ denotes the wavelength, and $w_d(\mathbf{x})$ and $w_s(\mathbf{x})$ model the object geometry for diffuse and specular reflection in $\mathbf{x}$.

When an image is taken by a camera with a finite number of color filters (typically red, green, blue), the captured intensity, or spectral response, for filter c in point $\mathbf{x}$, $I_c(\mathbf{x})$, can be modeled as

$$I_c(\mathbf{x}) = \int_\Omega S(\mathbf{x}, \lambda) E(\mathbf{x}, \lambda) q_c(\lambda) d\lambda \ . \quad (2)$$

In this equation, $S(\mathbf{x}, \lambda)$ denotes the spectral reflectance function over $\mathbf{x}$ and the wavelength λ, the spectral power distribution of the illumination is $E(\mathbf{x}, \lambda)$, and the sensor sensitivity is $q_c(\lambda)$. The integration is done over the visible spectrum of wavelengths, Ω. Note that this equation implicitly assumes a linear camera response. For the remainder of this paper we define $i \in \{R, G, B\}$ and use this index for summing over intensities for the red, green and blue component.

Additionally, we define the sensor response to the illuminant intensity L_c as

$$L_c = \int_\Omega E(\mathbf{x}, \lambda) q_c(\lambda) d\lambda \ . \quad (3)$$

The application of the dichromatic reflection model on equation (2) leads to a split of $S(\mathbf{x}, \lambda)$ into a diffuse reflectance function $S_d(\mathbf{x}, \lambda)$ and a specular reflectance function $S_s(\mathbf{x}, \lambda)$,

$$I_c(\mathbf{x}) = w_d(\mathbf{x}) \int_\Omega S_d(\mathbf{x}, \lambda) E(\mathbf{x}, \lambda) q_c(\lambda) d\lambda + w_s(\mathbf{x}) \int_\Omega S_s(\mathbf{x}, \lambda) E(\mathbf{x}, \lambda) q_c(\lambda) d\lambda \ . \quad (4)$$

For dielectric materials, the spectral reflectance function of the specularity is assumed to be similar to the spectral energy distribution function of the incident light, which is known as the *neutral interface assumption*. Then, the spectral reflectance function $S_s(\lambda, \mathbf{x})$ can be replaced by a constant $k_s(\mathbf{x})$. Furthermore, it is typically assumed that the color of the illumination is uniform over the image, which makes $E(\mathbf{x}, \lambda)$ independent from $\mathbf{x}$ and we put just $E(\lambda)$. Then, $I_c(\mathbf{x})$ becomes

$$I_c(\mathbf{x}) = w_d(\mathbf{x}) \int_\Omega S_d(\mathbf{x}, \lambda) E(\lambda) q_c(\lambda) d\lambda + \tilde{w}_s(\mathbf{x}) \int_\Omega E(\lambda) q_c(\lambda) d\lambda \ , \tag{5}$$

with $\tilde{w}_s(\mathbf{x}) = w_s(\mathbf{x}) k_s(\mathbf{x})$. For the purpose of representation, this equation can be written as

$$I_c(\mathbf{x}) = w_d(\mathbf{x}) B_c(\mathbf{x}) + w_s(\mathbf{x}) G_c, \tag{6}$$

where $w_d(\mathbf{x})$ and $w_s(\mathbf{x})$ are the geometric parameters of diffuse and specular reflection respectively. $B_c(\mathbf{x})$ is the sensor response to the diffuse reflectance. G_c is the sensor response to the illuminant spectral distribution and is assumed to be constant over the image.

Many algorithms for illuminant estimation use a normalized color representation, which is called the *chrominance* or *chromaticity* σ_c of an intensity I_c.

$$\sigma_c(\mathbf{x}) = \frac{I_c(\mathbf{x})}{\sum_i I_i(\mathbf{x})}, i \in \{R, G, B\} \ . \tag{7}$$

Remember that we defined $i \in \{R, G, B\}$ as index for the red, green and blue component.

In a similar manner, one can define the diffuse chromaticity $\Lambda_c(\mathbf{x})$ and the specular chromaticity $\Gamma_c(\mathbf{x})$ as

$$\Lambda_c(\mathbf{x}) = \frac{B_c(\mathbf{x})}{\sum_i B_i(\mathbf{x})} \ , \tag{8}$$

$$\Gamma_c = \frac{G_c}{\sum_i G_i} \ . \tag{9}$$

Many of the existing methods of color constancy and illuminant estimation typically either assume diffuse reflectance, e.g. [4, 13, 21], or base their analysis on regions of the image with purely specular reflectance, e.g. [26, 29, 39]. Our algorithm works on points which exhibit a mixture of specular and diffuse reflection. Similarly to many illuminant-estimation methods, which use specular highlights [24, 26, 29, 39], we, too, base our analysis on the distribution of pixels in chromaticity space. However, by considering in our analysis all pixels with a specular component our methodology has several advantages over the purely specular methods.

Algorithms that work on pure specularities require a specularity segmentation, which is a non-trivial problem (see e.g. [8, 27, 33]). When working on natural images taken in a non-research context, there are a lot of scenes that do not contain specular highlights at all. It may also happen that specular regions in natural images are often clipped due to camera settings of untrained users. Therefore, one of the weaknesses of specularity based methodologies on natural images is the number of available pixels. We believe that by extending the number of pixels available for illuminant analysis is very important, particularly for extending illuminant estimation from laboratory images to natural images. Purely diffuse techniques also have limitations. Some of them e.g. [13, 21] assume Lambertian reflectance which does accurately express the diffuse reflectance present in natural scenes. Other methods e.g. [18, 25, 31] make assumptions about the type of illuminants or surfaces or overall scene composition. Natural images are often too general for these constraints.

3. Inverse-Intensity Space

A convenient color space for analyzing the chromaticity relationship between the purely specular, the purely diffuse and the combined specular and diffuse reflections is the inverse-intensity chromaticity space introduced by Tan *et al.* [39]. We will briefly restate the justification for the representation of specularities in inverse-intensity space and give an intuition about its properties.

Equation (6) can be rewritten as

$$I_c(\mathbf{x}) = m_d(\mathbf{x}) \Lambda_c(\mathbf{x}) + m_s(\mathbf{x}) \Gamma_c \ , \tag{10}$$

where

$$m_d(\mathbf{x}) = w_d(\mathbf{x}) \sum_i B_i(\mathbf{x}) \ , \tag{11}$$

$$m_s(\mathbf{x}) = w_s(\mathbf{x}) \sum_i G_i \ . \tag{12}$$

Dividing equation (10) by $\sum_i I_i(\mathbf{x})$, (also represented according to equation (10)), the image chromaticity σ_c becomes

$$\sigma_c = \frac{m_d(\mathbf{x}) \Lambda_c(\mathbf{x}) + m_s(\mathbf{x}) \Gamma_c}{m_d(\mathbf{x}) \sum_i \Lambda_i(\mathbf{x}) + m_s(\mathbf{x}) \sum_i \Gamma_i} \ .$$

Solving this equation for $m_s(\mathbf{x})$ and inserting it in equation (10),

$$I_c = m_d(\mathbf{x}) (\Lambda_c - \Gamma_c) \left(\frac{s_c}{s_c - \Gamma_c} \right)$$

which leads to the definition of $p_c(\mathbf{x})$

$$p_c(\mathbf{x}) = m_d(\mathbf{x}) (\Lambda_c(\mathbf{x}) - \Gamma_c) \ , \tag{13}$$

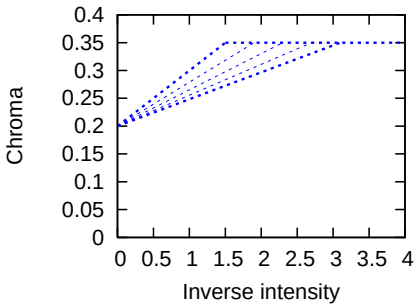


Figure 1. Idealized model for a monochrome object in inverse-intensity space. The specular pixels pixels aim at the illuminant color on the vertical axis, while the diffuse pixels form a horizontal line.

such that a linear relationship between the image chromaticity σ_c and $1/\sum_i I_i(\mathbf{x})$, can be established

$$\sigma_c(\mathbf{x}) = p_c(\mathbf{x})\frac{1}{\sum_i I_i(\mathbf{x})} + \Gamma_c \ . \tag{14}$$

In this representation, $p_c(\mathbf{x})$ can be seen as the slope of a line with intercept Γ_c. The domain of the line is determined by $1/\sum_i I_c(\mathbf{x})$, the range is given by $0 \leq \sigma_c \leq 1$. This space is called *inverse-intensity space* [39]. The chromaticity value where this line intersects the vertical axis gives the illuminant chroma estimate.

The inverse-intensity diagram is a direct graphical representation of the inverse-intensity space. The horizontal axis defines the inverse-intensity $1/\sum_i I_c(\mathbf{x})$, and the vertical axis σ_c, the illuminant chromaticity.

In general, it is not possible to obtain $p_c(\mathbf{x})$ directly in order to estimate the illuminant color. Typically, within one specular region of the image, different values for $p_c(\mathbf{x})$ can occur: according to equation equation (13), different values of $p_c(\mathbf{x})$ are determined by different geometrical parameters m_d. On the other hand, specular pixels with the same underlying albedo and the same geometric factors $w_d(\mathbf{x})$ and $w_s(\mathbf{x})$ share the same slope in the inverse-intensity histogram. In Fig. 1 one can visualize the ideal shape of a monochromatic cluster of pixels in inverse-intensity space. The diffuse pixels lie on a single horizontal line, and specular pixels align along their specific $p_c(\mathbf{x})$ between the illuminant color on the vertical axis and the diffuse line. For a complete discussion, see [39]. It is assumed that sufficiently many specular pixels with sufficiently similar slopes occur in one image, such that one can estimate the Γ_c intercept and hence, the illuminant color. Tan *et al*. used a Hough transform of the specular pixels with parameters $p_c(\mathbf{x})$ and Γ_c in order to estimate the illuminant color.

Fig. 2 illustrates the difference in specular and diffuse reflection. The pixels in the red rectangle contain part of the wooden bar in the background. Since there is only diffuse reflection, the distribution of those pixels in the inverse-intensity diagram (plotted in red) is roughly horizontal. The pixels in the green rectangle, although not fully specular, contain a specular portion. Therefore, assumed that the dichromatic model is fulfilled, the diagram representation (plotted in green) is a line that moves towards the illuminant chromaticity.

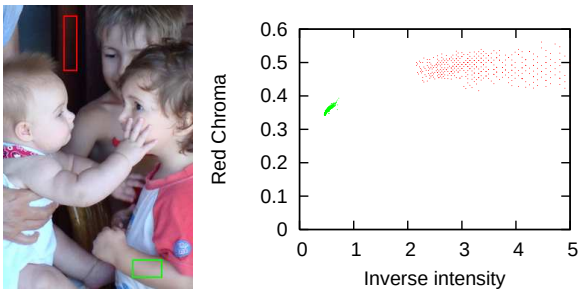
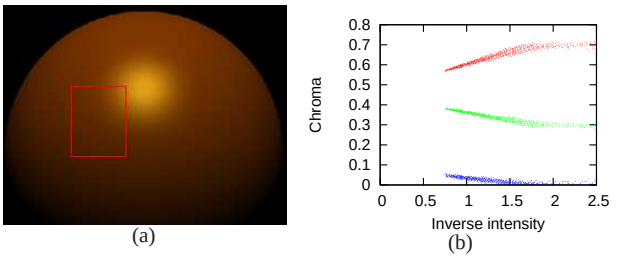


Figure 2. Real-world image with diffuse and specular reflection (left). The pixels in the red and green rectangle are plotted in the inverse-intensity diagram (right), for illustrating purposes only the red channel is shown.



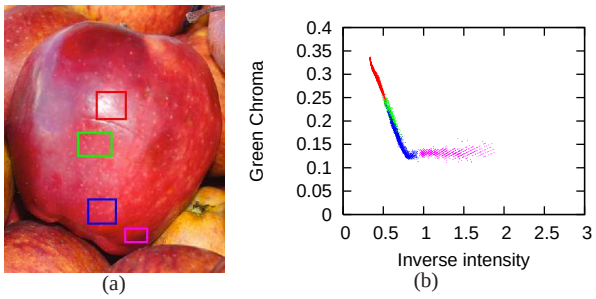
(a) (b)

Figure 3. (a) Synthetic image with normalized material chromaticities $(0.7, 0.3, 0)$ and illumination chromaticities $(0.45, 0.45, 0.09)$. The red rectangle shows the selected pixel. Note that the image as well as the selected area contains no specularities. (b) The inverse-intensity diagram of this image. The point colors correspond to the original color channel.

4. Chromaticity Distributions of Natural Images

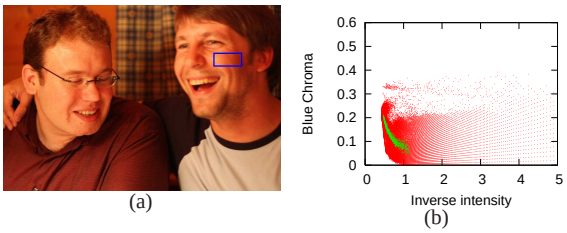
For images with few distinct albedo values, locating the purely specular pixels and identifying the different line segments in inverse-intensity space is relatively straightforward. As the scenes become more complex, the line segments and purely specular become increasingly more difficult to identify. We claim that by analyzing one small image region at a time and including all pixels that exhibit at least partial specular reflectance, we can still obtain an estimate of the Γ_c intercept.

Let us first demonstrate that it is not necessary to collect pure specularities in order to obtain an estimate of the illuminant color. Fig. 3 shows a synthetic monochromatic



(a) (b)

Figure 4. Real-World image with specularities (left). Different regions contain different amounts of specular and diffuse color. The magenta colored region contains almost only diffuse reflection, and therefore appears as a horizontal line in the diagram. Picture courtesy of Don Piefcone [16].



(a) (b)

Figure 5. Full inverse-intensity diagram of all pixels (red), and the shape of a single locally connected cluster (green), taken from the blue box in the image on the left.

example image without true specularities. The object chromaticities are $(0.7, 0.3, 0)$, the illumination chromaticities are $(0.45, 0.45, 0.09)$. The area within the red rectangle was used to create the inverse-intensity diagram in Fig. 3(b). The colors of the points in the inverse-intensity plot correspond to the originating color channel. Note that the green and the red curves tend towards their illuminant chromaticity 0.45, while the blue curve tends towards the blue illuminant chromaticity 0.09, although the chosen pixels contain no pure specularities.

To some extend, the same holds for natural images. Fig. 4(a) shows an apple. On its surface, several groups of pixels were handpicked and plotted in the inverse-intensity histogram in Fig. 4(b). In this (admittedly very clean) example, it can be nicely seen how the specular part decays from the red over the green to the blue patch. Finally, in the magenta colored patch, the typical horizontal shape for diffuse reflection shows up.

The central idea of our algorithm is to exploit the shape of pixels in the diagram that behave similar to the red, green and blue pixel groups in Fig. 4(b). However, for an automated process, it is not immediately clear how the pixels can be divided into groups with specular reflectance and groups with purely diffuse reflectance, as shown in Fig. 5. We plotted all pixels from the image in inverse-intensity space (red pixels). However, the green locally connected pixels show the characteristic specular shape. The shape is not as clear as it might be for truly specular pixels. In fact, if the underlying material does not follow the dichromatic model it does not need to have any useful interpretation within this framework. But if these pixels roughly adhere to the model and belong to a monochromatic object, their shape points roughly towards the illuminant chromaticity of the scene.

One central concept of our algorithm is to pick such local patches from the image, obtain from each suitable patch an illuminant estimate and make the final estimation by voting between the estimates. The full algorithm is presented in section 5. We emphasize on properties of the slopes of the pixels in these patches in inverse-intensity space.

Assuming that these pixels belong to our dichromatic model, the slope of these pixels is described by $p_c(\mathbf{x})$ from equation (14). We look at some properties of $p_c(\mathbf{x})$ as a foundation for our algorithm.

We rewrite $p_c(\mathbf{x})$ as

$$p_c(\mathbf{x}) = I_c(\mathbf{x}) - \Gamma_c \sum_i I_i(\mathbf{x}) \ . \tag{15}$$

Recall that $p_c(\mathbf{x}) = m_d(\mathbf{x})(\Lambda_c(\mathbf{x}) - \Gamma_c)$. For a single distinct pixel m_d is constant. Thus we can state that the slope within this pixel is determined by the difference

$$\delta_{\sigma_c} = \Lambda_c(\mathbf{x}) - \Gamma_c \ . \tag{16}$$

If $\Lambda_c(\mathbf{x})$ is similar to Γ_c, therefore δ_{σ_c} is small, $p_c(\mathbf{x})$ is also small. On the other hand, if δ_{σ_c} is large, we know that $p_c(\mathbf{x})$ is large, too.

The specularity-based methodology by Tan *et al*. works very good under laboratory conditions. It is also well applicable in constrained real-world applications. However, it suffers two drawbacks that we would like to address: First, the image itself needs to have pure specularities. This is especially difficult for images that were taken on cloudy days or when mainly matte material is captured on the picture. Second, as a general problem for specularity based methods, it requires a robust specularity segmentation that works on natural images.

We present an algorithm that works also for images with moderate highlights and without explicit specularity segmentation at all. The starting point of our considerations is equation 6. It states that a pixel intensity at position $\mathbf{x}$, $I_c(\mathbf{x})$, is a linear combination of specular and diffuse influences. While this is best exploited in the work by Tan *et al*. by considering only points with a very high specular portion (ideally what we refer to as specularities), equation 6 still holds for pixels with a much lower influence of the specular part. The idea of our method is to detect and exploit this lower influence in an automated approach.

Therefore, this algorithm can work on much more pixels than pure specularity based approaches, but does not suffer from the problematic segmentation of specularities, as reported for several algorithms working on matte reflection in [3].

5. Illuminant Estimation by Voting

The outline of our algorithm is as follows: We select many patches as shown by example in section 4. We test every patch on a number of criteria that are described below. If it passes all tests, we take the slope of the pixels in inverse-intensity space, approximated by the principal component of these pixels, and take the intersection of this slope with the vertical axis as a vote for the illuminant chromaticity at the intersection. Finally, we smooth the votes carefully and take the chromaticity with the maximum smoothed votes as the illuminant estimate.

If selected pixels come from dielectric surfaces with the same albedo and varying specular reflectance, we assume according to our model that these pixels form the triangular shape, as sketched in Fig. 1. Therefore, one central concept in our algorithm is to pick local groups of pixels, in the following referred to as patches. We filter these patches and reject every patch that fails in one of the filtering steps.

In the following, we explain the single steps of the algorithm and conclude this section with pseudo-code for it.

We generate rectangular patches of fixed size with significant horizontal or vertical elongation. The patch position and orientation are randomly chosen. We have to test these patches for their suitability to our algorithm. In the following, we first describe the overall strategy for the suitability conditions, and give then in detail the justification for the tests and the elongation of the patches.

In order to obtain optimal results for the illuminant color estimate, we need to evaluate image regions with pixels that roughly adhere to the specularity detection model. We apply several easily computable features that aim to ensure two basic properties for our model. Note that the structure is very modular, and can therefore easily be enhanced to more sophisticated tests, that extend the algorithm with prior knowledge. The two properties are:

Property 1 Operate on pixels with some *specular reflectance*. Geometrically, the *slope* of a group of pixels must fulfill two basic conditions: First, in absolute it may not be too large in order to intersect the vertical axis between 0 and 1. Second, it may not be close to 0, since then it might be the case that we picked a purely diffuse region. However, it is possible that by application of the latter rule we reject also patches that could lead to good estimates: Recall that the slope $p_c(\mathbf{x})$ of a pixel is defined as $p_c(\mathbf{x}) = m_d(\mathbf{x})(\Lambda_c(\mathbf{x}) - \Gamma_c)$. Therefore, if we locally pick pixels with similar m_d, and $\Lambda_c(\mathbf{x})$ and Γ_c are accidentally almost equal for a group of pixels, their common slope is almost 0, although they are specular.

Property 2 Operate on pixels with the same (or similar) *albedo*. Geometrically, the *shape* of a group of pixels must also fulfill two conditions, based on the following observation: If a patch is chosen that goes over object boundaries with different underlying albedo, the points form at least two clusters in inverse-intensity space or are even completely dispersed. On the converse, pixels that fit our model well are typically compactly distributed and significantly elongated along the principal axis, which can be found out by virtue of second order moments analysis. This property also excludes small circular clusters of pixels that come from monochromatic, flat surfaces where no significant intensity differences exist.

We perform the following simple checks and operations on the patches to fulfill the two properties:

1. Run an Laplacian of Gaussian edge detector on the patch and reject it if too many edge pixels are on the patch to encompass Property 2.
2. Exclude saturated and too dark pixels from the patch to reduce effects due to image capturing nonlinearities.
3. Compute the standard deviation on the chromaticity values. It must be larger than a minimum value to exclude diffuse reflectance patches (Property 1), and lower than a minimum value to prevent the breaking of object boundaries (Property 2).
4. Compute the standard deviation on the intensity values. It must be larger than a minimum value in order to measure the elongation along the x-axis (Property 1).
5. Reduce duplicated entries in the inverse-intensity diagram.
6. Compute the eccentricity of the remaining pixels by virtue of principal component analysis. Let λ_1 the largest eigenvalue, λ_2 the second largest, then the eccentricity is $\sqrt{1 - \frac{\lambda_2}{\lambda_1}}$ (Property 1). Exclude the patch if its excentricity is too small.
7. Compute the slope from the eigenvectors of the PCA and exclude too large or too small values (Property 1).

We support the patch selection process with a simple hypothesis by limiting the randomly picked patches to a large horizontal or vertical elongation, since we are looking for smooth intensity progressions in the image that lead to nice specular curves. Therefore, we anticipate convex curved objects in the scene and aim to cover them perpendicular

to the object curvature with the patch. This could also be replaced by a more sophisticated combination of an analysis of the patch content and a selection of a subset of patch pixels.

The chromaticity estimate of every patch is taken as a single vote. After a number of patches was accepted, the histogram of votes is smoothed with a Gaussian filter and the chromaticity value with the maximum votes is picked.

Although there are many parameters involved, we adjusted during our experiments almost none of them, the overall methodology appears to be surprisingly robust, as long as a sufficiently large number of patches passes the test in the end.

6. Experiments

In this section the different experimental results of our illuminant estimation method are presented. The applicability of the estimation approach to natural images found on the web is shown.

6.1. Image Data Sets

The evaluation of natural images was performed on a set of 39 images found on the web. The database contains images of indoor and outdoor scenes. The image set includes a variety of content, such as nature, people, animals and architecture. The outdoor images were acquired at different daytimes and under various weather conditions. To analyze the stability of the results, the database also contains three image series where the single images were acquired in the same environment and with only little time offset.

6.2. Algorithm

As described in the previous section our approach randomly selects patches of the input image and assesses the quality of them. The parameters which are needed for this decision were thoroughly selected once and then left unchanged for all experiments.

For the natural images the mask size for the patch selection was 11×61 pixels. Compared to the large image sizes (from about 700×1000 pixels to 2000×3000 pixels), this mask is very small. However, by selecting smaller patches, the performance of the algorithm becomes less dependent on the image content. For each image the voting for an illuminant color is based on a set of 200 patches that pass the quality criteria described in section 5. If 1000 patches were tested without reaching enough accepted patches, the algorithm stops nevertheless.

6.3. CIE standard illuminants

The *Commission Internationale de l'Eclairage* (CIE) has defined different standard illuminats in order to make images acquired under different lighting conditions comparable. The standard illuminats are specified by spectral power distributions. For being able to assess the results of our illuminant estimation technique, we compare the estimates to the standard illuminants. Therefore, it is necessary to compute the chromaticities of the standard illuminants. This transformation is described in the following.

	x_2	y_2	X_2	Y_2	Z_2
A	0.4476	0.4075	109.828	100.000	35.547
D50	0.3457	0.3585	96.422	100.000	82.521
D55	0.3324	0.3476	95.642	100.000	92.085
D65	0.3127	0.3291	95.017	100.000	108.813
D75	0.2990	0.3150	94.939	100.000	122.558

Table 1. Values of different CIE standard illuminants in the CIE XYZ and CIE xyY color spaces computed for the CIE 1931 2° standard observer [42, 34].

If the spectral power distribution $P(\lambda)$ of an emissive light source is known, the X, Y and Z tristimulus values of the *CIE XYZ color space* can be computed with

$$X = \int_0^\infty P(\lambda)\bar{x}(\lambda)d\lambda \ , \tag{17}$$

$$Y = \int_0^\infty P(\lambda)\bar{y}(\lambda)d\lambda \ , \tag{18}$$

$$Z = \int_0^\infty P(\lambda)\bar{z}(\lambda)d\lambda \ , \tag{19}$$

where $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{x}(\lambda)$ are the color matching functions, which are defined by the *Commission Internationale de l'Eclairage* as the CIE 1931 (2° standard observer) and CIE 1964 (10° standard observer) standard colorimetric observer. The set of tristimulus values of an illuminant represent the white point of the illumination.

By computing the CIE-chromaticities x, y and z with

$$x = \frac{X}{X+Y+Z} \ , \tag{20}$$

$$y = \frac{Y}{X+Y+Z} \ , \tag{21}$$

$$z = \frac{Z}{X+Y+Z} = 1 - x - y \ , \tag{22}$$

the *CIE xyY color space* can be derived. For the different CIE standard illuminants the tristimulus and corresponding chromaticity values for the CIE 1931 2° standard observer are listed in Table 1.

In order to be able to compare the results of the illumination estimation of natural images with the CIE standard illuminants, it is necessary to convert the tristimulus values of the reference white point of the standard illuminants to RGB values. For our calculations we have chosen the standard RGB color space (sRGB). The reference white point of

	x	y	Y
red	0.64	0.33	0.212656
green	0.3	0.6	0.715158
blue	0.15	0.06	0.072186

Table 2. Chromaticity coordinates x_c, y_c, Y_c with $c \in \{r, g, b\}$of the sRGB color system [34].

the sRGB color space is the CIE D65. The transformation to sRGB can be determined with

$$\begin{bmatrix} r & g & b \end{bmatrix} = \begin{bmatrix} X & Y & Z \end{bmatrix} \mathbf{M}^{-1} \tag{23}$$

and, finally,

$$I_c = \begin{cases} 12.92 \times c & \text{if } c \leq 0.0031308 \\ 1.055 \times c^{1/2.4} & \text{if } c > 0.0031308 \end{cases} \quad \text{for } c \in \{r, g, b\}\,. \tag{24}$$

The matrix $\mathbf{M}$ is defined as

$$\mathbf{M} = \begin{bmatrix} C_r X_r & C_r Y_r & C_r Z_r \\ C_g X_g & C_g Y_g & C_g Z_g \\ C_b X_b & C_b Y_b & C_b Z_b \end{bmatrix}, \tag{25}$$

where

$$\begin{bmatrix} C_r & C_g & C_b \end{bmatrix} = \begin{bmatrix} X_W & Y_W & Z_W \end{bmatrix} \begin{bmatrix} X_r & Y_r & Z_r \\ X_g & Y_g & Z_g \\ X_b & Y_b & Z_b \end{bmatrix}^{-1} . \tag{26}$$

The tristimulus values $X_c = x_c/y_c$, $Y_c = 1$ and $Z_c = (1-x_r-y_r)/y_r$, where $c \in \{r, g, b\}$, can be computed with the chromaticity coordinates (x_r, y_r), (x_b, y_b) and (x_b, y_b) of the sRGB system (see Table 2, [17]). The values X_W, Y_W, Z_W denote the tristimulus values of the reference white of the destination RGB color space. For the sRGB system, these are the tristimulus values of the D65 standard illuminant (see Table 1).

Afterwards, the chromaticities of the RGB values are computed using Equation 7. For the different CIE standard illuminants the values of the sRGB color space and the corresponding chromaticities are listed in Table 3.

6.4. Experiments on natural images

Fig. 6, Fig. 7 and Fig. 8 show a subset of the images which are part of the evaluation set. The corresponding estimations of the illuminant chromaticity are listed in Table 4.

To show the performance of the method under different illumination conditions and scene content, Fig. 6 contains an indoor scene, a portrait of a woman and two outdoor scenes. For the indoor scene, the estimated illuminant chromaticity fits very good the chromaticities of a tungsten light bulb (CIE standard illuminant A, see Table 3), where the red component dominates significantly compared to blue. The portrait of the woman (see Fig. 6(b)) was acquired at noon. The corresponding results confirm that the illuminant has the same intensities in all channels (CIE standard illuminant D65, see Table 3). An interesting observation is that the chromaticities of the three channels are well balanced, although the surfaces in the scene are significantly dominated by blue and red. This aspect indicates that the illumination estimation is not disturbed by pure diffuse patches. The outdoor scene with the dog (Fig. 6(c) is illuminated by warm light coming from the horizon. The reduced blue and increased red component of the illuminant is well included in the estimation. In contrast to the previous images, Fig. 6 contains a scene in the evening. The results of the estimation again are reasonable (CIE standard illuminant D50, see Table 3).

An important aspect of the evaluation on natural images is the stability of the voting process. The standard deviations of the results listed in Table 4 are comparably small. As the voting scheme is based on the randomized selection of patches, this stability in the results indicates that the number of selected patches and the parameters are suitable. However, to show the general applicability of the approach on a variety of images the parameters used for the evalu-

	I_r	I_g	I_b	σ_r	σ_g	σ_b
A	1.307	0.919	0.520	0.476	0.335	0.189
D50	1.074	0.989	0.866	0.367	0.338	0.296
D55	1.044	0.994	0.918	0.353	0.336	0.310
D65	1.000	1.000	1.000	0.333	0.333	0.333
D75	0.968	1.003	1.061	0.319	0.331	0.350

Table 3. Values of different CIE standard illuminants in the sRGB and chromaticity color spaces computed for the CIE 1931 2° standard observer


(a) Indoor


(b) Woman


(c) Dog


(d) Castle

Figure 6. Subset of the evaluated natural images

Scene	Γ_r	Γ_g	Γ_b
Indoor	0.552±0.041	0.308±0.055	0.140±0.016
Woman	0.331±0.002	0.331±0.002	0.338±0.004
Dog	0.361±0.019	0.345±0.005	0.294±0.016
Castle	0.283±0.014	0.325±0.004	0.392±0.016
Shadow	0.308±0.004	0.341±0.002	0.351±0.002
Chaple	0.172±0.058	0.269±0.020	0.560±0.077
Highland	0.330±0.008	0.330±0.006	0.340±0.012

Table 4. Stability of the algorithm performance on natural images. For the images of Fig.6 and Fig.7 the mean estimation results of three randomized voting processes are listed together with the standard deviation.

(a) Chapel (b) Highlands

Figure 7. Images on which our method showed limitations.

ation were not tuned. However, the highly modular structure of our algorithm allows the smooth integration of prior knowledge, if available, for further improved accuracy.

Comparing the images in Fig. 7 with their estimation results (Table 4), the limitations of the proposed method can be observed. One requirement of the voting approach is an uniform illumination of the scene. If this assumption is not fulfilled, as in Figure 7(b), the illuminant estimation fails. For this example image, the estimated illuminant is composed of the same proportion of red, green and blue. However, it is obvious that the value of the red chromaticity should dominate slightly. Besides the non-uniformly illuminated scene, two further aspects can be observed in this image. First, in outdoor scenes, the sky can sometimes be seen as a second light source in addition to the sun. Our proposed method, however, can only handle scenes with a single light source. Secondly, as the grass on the ground and the coat of the buffalo are highly textured, it is difficult to find a sufficient quantity of patches which pass the selection step. This drawback can be reduced by an appropriate preprocessing of the images and a reduced patch size.

Another limitation of the method is the applicability to non-dielectric surfaces. As the voting scheme is based on the assumption of dielectric materials the illuminant estimation results for Fig. 7(a) drop off. Furthermore, parts of the towers and the front of the chapel are shadowed.

Without further preprocessing the results of the voting scheme on images containing shadows contain a bias towards an increased blue chromaticity. This relationship

Figure 8. Scene which is covered by a shadow.

could be observed in several images and is also illustrated in the results of Fig. 8, where a group of people is sitting in a slight shadow of a marquee. The corresponding illuminant chromaticities have an increased blue chromaticity although the scene was captured at a sunny day at noon. Although the voting scheme by definition can not handle shadows, the results of the estimation show a very high stability.

7. Conclusions

We presented an automated illuminant color estimation algorithm that can handle specularities and highlighted non-specular regions.

For a reliable estimation of the illuminant color we ideally need patches of convex opaque surfaces that are directly illuminated by the light source. Although we employ a randomized algorithm, our results were within the same scene surprisingly stable. Slight deviations from our assumptions still lead to good results, while others, like multiple light sources or completely flat scenes, can currently not be handled by our methodology.

Multiple illuminants are subject to further work: In many cases, multiple illuminants can possibly be detected by a cluster analysis of the votes.

8. Acknowledgements

The authors gratefully acknowledge funding of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German National Science Foundation (DFG) in the framework of the excellence initiative.

References

[1] V. Agarwal, A. V. Gribok, A. Koschan, and M. A. Abidi. Estimating Illumination Chromaticity via Kernel Regression. In *International Conference on Image Processing*, pages 981–984, 2006. 1, 2

[2] R. Bajcsy, S. W. Lee, and A. Leonardis. Detection of Diffuse and Specular Interface Reflections and Interreflections by Color Image Segmentation. *International Journal of Computer Vision*, 13(3):241–272, 1996. 1

[3] 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. 1, 6

[4] D. H. Brainard and W. T. Freeman. Bayesian Color Constancy. *Journal of the Optical Society of America A*, 14(7):1393–1411, 1997. 1, 3

[5] M. H. Brill. Image Segmentation by Object Color: a Unifying Framework and Connection to Color Constancy. *Journal of the Optical Society of America A*, 7(10):2041–2047, Oct. 1986. 1

[6] V. C. Cardei and B. Funt. Color Correcting Uncalibrated Digital Images. *Journal of Imaging Science and Technology*, 44(4):288–294, 2000. 1

[7] 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. 1, 2

[8] M. S. Drew. Robust Specularity from a Single Multi-Illuminant Color image. *CVGIP: Image Understanding*, 59:320–327, 1994. 3

[9] M. D'Zmura and P. Lennie. Mechanisms of Color Constancy. *Journal of the Optical Society of America A*, 3(10):1622–1672, Oct. 1986. 1

[10] G. D. Finlayson, M. S. Drew, and B. V. Funt. Color Constancy: Generalized Diagonal Transforms Suffice. *Journal of the Optical Society of America A*, 11:3011–3019, 1994. 1

[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. 1

[12] G. D. Finlayson, S. D. Hordley, and P. Hubel. Color by Correlation: A Simple, Unifying Framework for Color Constancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1209–1221, 2001. 1

[13] G. D. Finlayson, S. D. Hordley, and I. Tastl. Gamut Constrained Illuminant Estimation. *International Journal of Computer Vision*, 67(1):93–109, 2006. 1, 2, 3

[14] G. D. Finlayson and E. Trezzi. Shades of Gray and Colour Constancy. In *Twelfth Color Imaging Conference*, pages 37–41, 2004. 1

[15] R. W. Fleming, A. Torralba, and E. H. Adelson. Specular Reflections and the Perception of Shape. *Journal of Vision*, 4(9):798–820, 2004. 2

[16] Flickr.com. http://www.flickr.com. 5

[17] J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes. *Computer Graphics – Principles and Practice*. Addison-Wesley, 2002. 8

[18] D. Forsyth. A Novel Algorithm for Color Constancy. *International Journal of Computer Vision*, 5(1):5–36, 1990. 1, 3

[19] B. Funt, K. Barnard, and L. Martin. Is Machine Color Constancy Good Enough? In *European Conference in Computer Vision*, pages 445–459. Springer-Verlag, 1998. 1

[20] B. V. Funt and G. D. Finlayson. Color Constant Color Indexing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17:522–529, 1995. 1

[21] J. Geusebroek, R. Boomgaard, A. Smeulders, and T. Gevers. Color Constancy from Physical Principles. *Pattern Recognition Letters*, 24(11):1653–1662, July 2003. 1, 3

[22] A. Gijsenij and T. Gevers. Color Constancy using Natural Image Statistics. In *Computer Vision and Pattern Recognition*, pages 1–8, June 2007. 1

[23] J. Ho, B. V. Funt, and M. S. Drew. Separating a Color Signal into Illumination and Surface Reflectance Components: Theory and Applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(10):966–977, 1990. 1

[24] 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. 1, 2, 3

[25] E. H. Land. Lightness and the Retinex Theory. *Scientific American*, 237(6):108–129, Dec. 1977. 1, 3

[26] H. Lee. Method for Computing the Scene-Illuminant Chromaticity from Specular Highlights. *Journal of the Optical Society of America A*, 3(10):1694–1699, 1986. 1, 2, 3

[27] S. W. Lee and R. Bajcsy. Detection of Specularity Using Color and Multiple Views. *Image and Vision Computing*, 10:643–653, 1992. 3

[28] T. M. Lehmann and C. Palm. Color Line Search for Illuminant Estimation in Real World Scene. *Journal of the Optical Society of America A*, 18(11):2679–2691, 2001. 2

[29] S. Lin and H.-Y. Shum. Separation of Diffuse and Specular Reflection in Color Images. In *Computer Vision and Pattern Recognition*, volume 1, pages 341–346, 2001. 1, 2, 3

[30] Y. Liu, D. Zhang, G. Lu, and W. Ma. A Survey of Content-Based Image Retrieval with High-Level Semantics. *Pattern Recognition*, 40(1):262–282, 2007. 2

[31] L. T. Maloney and B. A. Wandell. A Computational Model of Color Constancy. *Journal of the Optical Society of America A*, 3(1):29–33, 1986. 1, 3

[32] J. F. Norman, J. T. Todd, and G. A. Orban. Perception of Three-Dimensional Shape from Specular Highlights, Deformations of Shading, and Other Types of Visual Information. *Psychological Science*, 16(8):565–570, 2004. 2

[33] J. B. Park and A. C. Kak. A Truncated Least Squares Approach to the Detection of Specular Highlights in Color Images. In *IEEE International Conference on Robotics and Automation*, pages 1397–1402, 2003. 3

[34] D. Pascale. *A Review of RGB Color Spaces ... from xyY to R'G'B'*. The BabelColor Company, 2003. 7, 8

[35] R. Schettini, G. Ciocca, and I. Gagliardi. Content-Based Color Image Retrieval with Relevance Feedback. In *International Conference on Image Processing*, volume 3, pages 75–79, 1999. 1

[36] S. A. Shafer. Using Color to Separate Reflection Components. *Journal Color Research and Application*, 10(4):210–218, 1985. 2

[37] G. Sharma and H. J. Trussell. Digital Color Imaging. *IEEE Transactions on Image Processing*, 6(7):901–932, July 1997. 1

[38] M. J. Swain and D. H. Ballard. Color Indexing. *International Journal of Computer Vision*, 7(1):11–32, 1991. 1

[39] 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. 1, 2, 3, 4

[40] S. Tominaga and B. A. Wandell. Natural Scene-Illuminant Estimation using the Sensor Correlation. In *Proceedings of the IEEE*, pages 42–56, 2002. 1

[41] J. van de Weijer, T. Gevers, and A. Gijsenij. Edge-Based Color Constancy. *IEEE Transactions on Image Processing*, 16(9):2207–2214, Sept. 2007. 1

[42] G. Wyszecki and W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae*. John Wiley & Sons, 1982. 7