PHYSICS-BASED ILLUMINANT COLOR ESTIMATION AS AN IMAGE SEMANTICS CLUE

Christian Riess and Elli Angelopoulou

Chair of Pattern Recognition, Department of Computer Science,
Friedrich-Alexander University Erlangen-Nuremberg,
Martensstr. 3, 91058 Erlangen, Germany

ABSTRACT

Most algorithms for extracting illuminant chromaticity from arbitrary images, such as the images found on the web, are based on machine learning techniques. We will show how a physics-based methodology can be adapted to provide relative illumination information on real images. More specifically, we use the inverse-intensity chromaticity representation and show how the analysis of the histograms of illumination-chromaticity candidates provides information about the type of illumination(s) present in a scene. Experiments indicate that the estimate is quite robust towards noise, and that simple measurements on the histogram peak can be used to counter-check the reliability of the estimate.

Index Terms— specularities, inverse-intensity chromaticity

1. INTRODUCTION

The majority of semantic image analysis methods are based on machine learning [13], partially because it is very difficult to define mathematically or algorithmically what constitutes a semantic entity. There are, however, 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 typically 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 will show how a state-of-the-art methodology for illuminant color estimation can be extended, so that it can infer general illuminants information in arbitrary images.

The illuminant estimation problem was selected as an example application for transfer of physics-based methodologies to semantics-oriented 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 [15, 6]. Secondly, as in semantic analysis, previous work on illuminant estimation on arbitrary scenes is mostly based on machine learning techniques, e.g. [2, 1, 5]. Though recent work by Gijsenij et al. [8] augments machine learning with physics-based analysis, the majority of such methods is still closely tied to the training samples and some of them, e.g. [2], are not easily generalizable. algorithms which, under strong constraints, can estimate the color of the illuminant, e.g. [3, 4, 14, 9, 10]. One of the most popular family of these techniques is specularity based [9, 10, 12, 17].

We will show how a state-of-the-art technique for estimating the illuminant-color based on specular highlights [17] can be adapted for obtaining an estimate of the type of illuminant in arbitrary images found on the web. This involves the treatment of difficulties in the segmentation of specularities, ways to assess the compliance of the obtained intensities with our model and the handling of multiple illuminants.

2. ESTIMATION OF THE ILLUMINANT COLOR

Most of the methodologies, which use specular highlights to infer the illuminant color, e.g. [9, 10, 12], are based on the dichromatic reflection model [16]. This states that the amount of light reflected from a point, $\vec{p}$, of a dielectric, non-uniform material is a linear combination of diffuse reflection and specular reflection. Furthermore, they all assume that the color of specularities in such materials is the same as the illuminant color. When an image is taken with a trichromatic camera, the sensor response at each channel can be modeled as:

$$I_c(\vec{p}) = w_d(\vec{p})B_c(\vec{p}) + w_e(\vec{p})G_c \qquad (1)$$

where $c = r, g$, or b corresponds to the color channel, $w_d(\vec{p})$ and $w_e(\vec{p})$ are the geometric parameters for diffuse and specular reflection respectively. $B_c(\vec{p})$ is the sensor response to the diffuse spectral reflectance. G_c is the sensor response to the illumination spectral distribution and is assumed to be constant over the image. Thus, the color values recorded by a camera are a mixture of diffuse, $w_d(\vec{p})B_c$, and specular, $w_e(\vec{p})G_c$, components. The algorithms which extract illumi-

nation information from specular highlights typically analyze the scatter plots of these reflectance mixtures in color space.

2.1. Inverse-Intensity Space

Recently, Tan et al. [17] introduced the inverse-intensity chromaticity space as a new color space for analyzing these specularity clusters. Their technique is a very general method since it does not impose constraints on the type of illumination, does not require multi-colored or single-colored objects and does not assume known camera spectral responses.

Since variations in image brightness can affect image analysis, it is quite common to normalize color representation. We used the same chromaticity definition, σ_c, as [17]:

$$\sigma_c(\vec{p}) = I_c(\vec{p}) / \sum_i I_i(\vec{p}) \qquad (2)$$

In a similar manner, one can define diffuse chromaticity as $\Lambda_c(\vec{p}) = B_c(\vec{p})/\sum_i B_i(\vec{p})$ and specular chromaticity as $\Gamma_c = G_c/\sum_i G_i$. Following [17], eq. (2) can be rewritten in terms of diffuse and specular chromaticity as follows:

$$I_c(\vec{p}) = m_d(\vec{p})\Lambda_c(\vec{p}) + m_s(\vec{p})\Gamma_c \qquad (3)$$

where $m_d(\vec{p}) = w_d(\vec{p})\sum_i B_i(\vec{p})$ and $m_s(\vec{p}) = w_e(\vec{p})\sum_i G_i$.

Tan et al. [17] have shown, that from these equations one can derive the following non-linear relationship between chromaticity σ_c and image intensity I_c:

$$\sigma_c(\vec{p}) = a(\vec{p})\frac{1}{\sum_i I_i(\vec{p})} + \Gamma_c \qquad (4)$$

where $a(\vec{p}) = m_d(\vec{p})(\Lambda_c(\vec{p}) - \Gamma_c)$. This last equation captures the core idea of their methodology: In order to extract the illumination chromaticity Γ_c, one need only derive the value of $a(\vec{p})$. The image chromaticity, $\sigma_c(\vec{p})$, and the sum of all the colors per pixel, $\sum_i I_i(\vec{p})$, can be directly computed from the image.

Analytic estimation of $a(\vec{p})$ is in most cases not feasible. However, in chromaticity space, specular pixels are typically grouped together into clusters that have approximately the same value of $a(\vec{p})$ and different $\sum_i I_i$. As [17] has shown, all the pixels with the same m_d value form a line in inverse-intensity chromaticity space, i.e. the 2D space where the horizontal axis represents $1/\sum_i I_i(\vec{p})$ and the vertical axis represents σ_c. Different m_d values form distinct straight lines in inverse-intensity space, which intersect at a single point in the vertical axis, which is, in turn, the illumination chromaticity Γ_c (see Figure 1a). By using Hough Transform, Tan et al. can compute the intersection point Γ_c more robustly. Additionally, in order to further increase the reliability of their algorithm, instead of examining only the single point in Hough space with the largest number of occurences, they create a histogram of the occurences of the different Γ_c candidates (see Figure 1b). We will show in section 3 how one can analyze

this histogram of candidate illumination-chromaticities in order to obtain illumination information for arbitrary images.
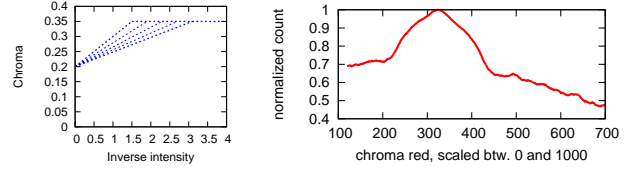

Fig. 1. Left: Expected color distrubution in inverse intensity space. Right: Red channel intersection counting distribution for the image in Figure 3.

2.2. Specularity Detection

In order to easily identify the chromaticity lines in color space, one should first obtain a rough estimate of the specular region. We use the method of Lehmann and Palm [11] for identifying such regions. The method is simple to implement and requires minimal user-interaction. A pixel is considered specular if it is among the brightest pixels in the given image but its color is still not saturated:

$$\mathcal{I} = \frac{I_r + I_g + I_b}{3} > t_b \mathcal{I}_{max} \qquad (5)$$

$$\mathcal{S} = 1 - \frac{min(I_r, I_g, I_b)}{\mathcal{I}} < t_s \mathcal{S}_{max} \qquad (6)$$

where $\mathcal{I}$ and $\mathcal{S}$ are measures of a pixel's brightness and color saturation respectively. $\mathcal{I}_{max}$ and $\mathcal{S}_{max}$ are the highest brightness and saturation values in the image and t_b and t_s are their corresponding thresholds. Lehmann and Palm [11] set $t_b = t_s = 0.5$. Tan et al. [17] used a range of threshold values $0.4 \leq t_b \leq 0.6$ and $0.4 \leq t_s \leq 0.6$. We used threshold values that were more restrictive, with ranges $0.55 \leq t_b \leq 0.8$ and $0.2 \leq t_s \leq 0.45$. The treshold values influene the extend of the specular regions, but the number of detected specular regions and their center remain largely unaffected [11].

3. FROM THE LAB TO THE WEB

The methodology introduced by Tan et al. had only been shown to give accurate illuminant-color estimates under laboratory conditions with controlled illumination, high quality linear camera and convex objects. Images from the internet introduce new challenges. The image quality is typically lower than under laboratory conditions. Specularity handling is more difficult. An image often contains very few clean specularities, but many pixels may be falsely segmented as specular. We call this noise, although it is not image noise in the stricter sense. Finally, specularities often do not fulfill the theoretical constraints presented in section 2. All these factors affect the histograms described in section 2.1.

3.1. Extensions to the Histogram Creation

Among the first things to notice is that the specularity segmentation of real world images often labels large bright regions, like the sky, as specular. So noise reduction is a desireable preprocessing step. Assuming that clean specularities typically cover small image regions, we extended the Lehmann and Palm method as follows.

We did segmentations with increasingly strict thresholds, ending up with $t_b = 0.80$ and $t_s = 0.20$, and computed connected regions on these pixels. According to our experiments, large connected noise elements like parts of the sky are more compact and decay slower than specularities. By thresholding on the decay and the minimum size of the region, many noise pixels can be removed. The connected components can be reused to remove small isolated regions and to detect different illuminant sources, see section 3.4.

3.2. Towards Automated Self-Assessment

Due to the large variety of deviations from laboratory conditions, we need to have tools to judge how well the histogram satisfies our constraints.

All further processing is done on the (linearly) smoothed histogram with normalized values between 0 and 1. The histogram peak can be identified by computing its first derivative. We define the peak region as the area between the derivative's extreme values, expanded until the curve drops below (raises above, resp.) a threshold of 0.1.
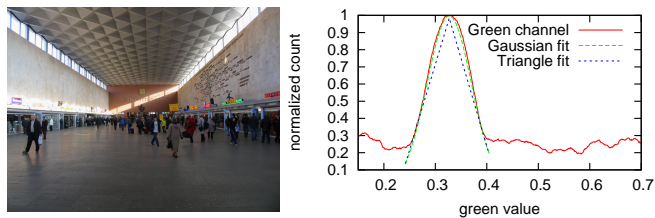

Fig. 2. Left: Image without true specularities. Right: The Gaussian fits better than the triangle.

To assess the quality of the peak regions we fitted triangles and Gaussian curves. Ideally, the topmost part of the peak should clearly follow a triangular shape. Noise or missing specularities often lead to plateaus that can be better characterized by a broader curve. By comparing the area of the peak region with the triangle and the Gaussian, we get a measure of the estimation's quality. Figure 2 left shows an indoor image without specularities (picture courtesy of Alexsey Troshin [18]). Although we obtain a clear histogram (right), the Gaussian fits the shape better than the triangle.

3.3. Influence of the segmentation performance and noise

We compared automatically and (semi-)manually segmented images. Within the same image, the results were surprisingly

stable, although the width of the peak region changed. In Figure 3 (a) and (b), two images are shown that were captured at almost the same time and place. The first segmentation used $t_b = 0.55$ and $t_s = 0.45$, the 2nd $t_b = 0.65$ and $t_s = 0.35$, and the 3rd used the same thresholds as the 2nd, but with manual assistance. The location of the peak is barely influenced by the rough automated segmentation (c).
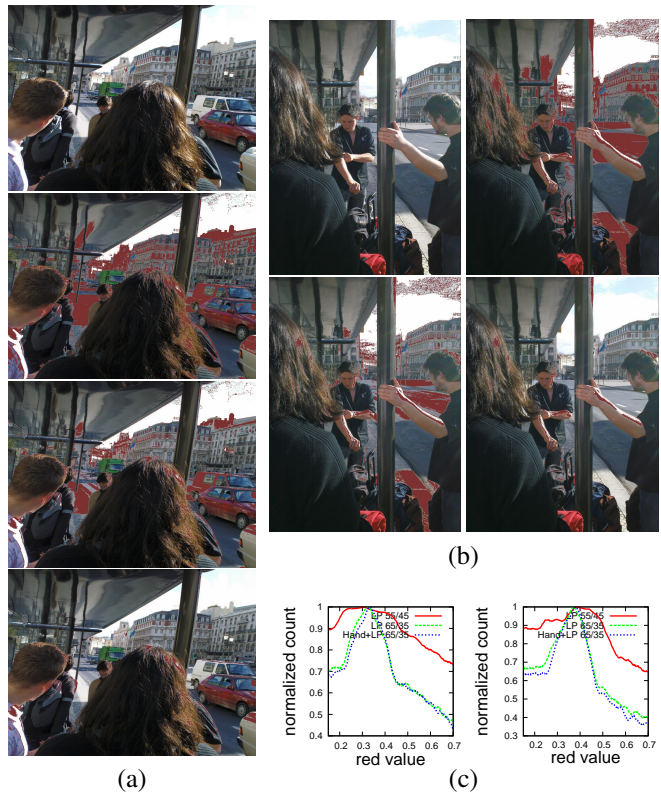

Fig. 3. (a), (b) Two images captured under almost identical conditions and the segmented pixels in red. (c) plots of the red-channel histograms of all segmentations together.

Quantitatively, the estimated illuminant color differs between the two images. For image (a) we obtain as RGB $(0.378, 0.303, 0.318)$, in (b) $(0.339, 0.341, 0.319)$, which shows how problematic it can be to to use the absolute values to analyze the scene content. However, one can still relate these values and the histogram analysis to other images to obtain an indirect scene content estimate.

3.4. Multiple Illuminants

Multiple illuminants become problematic when all segmented regions are analyzed together. In Figure 4, the statue is illuminated by an orange and a blue light source. While there is a clear difference when each image subregion is analyzed independently, the combined histogram contains a stronger third peak. Noise can further obscure the valley between the peaks. This motivates grouping of similar illuminant estimates.
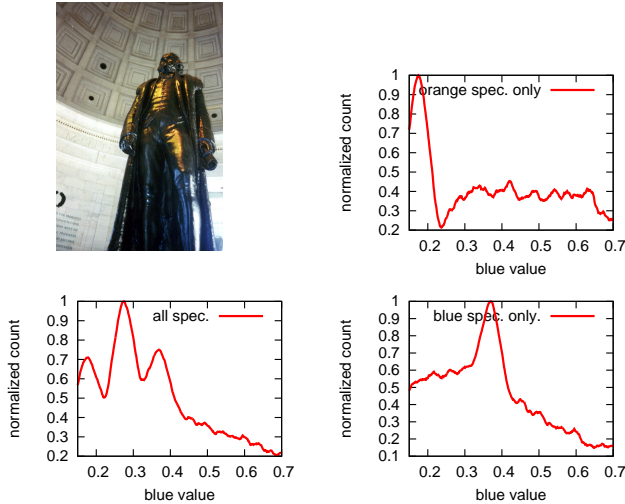
Fig. 4. Image illuminated by distinct light sources and the blue channel histograms for the blueish region, orange region and both together. Picture courtesy of Wallace Gobetz.

4. DISCUSSION AND CONCLUSIONS

This is preliminary work. Further testing on the relationship between the accuracy of specularity detection and our histogram analysis has to be done. We tested various images obtained from the web, mostly from Flickr [7]. We worked with 10 images whose specularities were both hand- and automatically segmented and 20 further images whose specular regions were only automatically detected. The results with manual assistance showed crisper histogram peaks, but the overall estimate was surprisingly robust to segmentation-inferred noise. Therefore, the differences in the absolute estimates are most likely introduced from other sources. Please note, that other specularity detection methods besides Lehmann and Palm [11] can also be used, as long as they do not impose additional constraints on the image.

Tan's et al. [17] method, and by extension our analysis too, assumes cameras with linear response ($\gamma = 1$). On images found on the web one can not know whether the cameras were truly linear, and constraints like non-dielectric surfaces also influence the result. Nevertheless, we believe that by analyzing the histogram shapes, conclusions about the content of an image can be drawn, as seen on the example with multiple illuminants. We are currently developing a system for grouping images based on histogram information.

5. 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.

[2] 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, 2002.

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

[4] 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(11):3011–3019, 1994.

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

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

[7] Flickr.com. http://www.flickr.com.

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

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

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

[17] 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.

[18] Walks.ru. http://www.walks.ru.