Estimating Shadows with the Bright Channel Cue

Alexandros Panagopoulos[1], Chaohui Wang[2,3], Dimitris Samaras[1] and Nikos Paragios[2,3]

[1]Image Analysis Lab, Computer Science Dept., Stony Brook University, NY, USA
[2]Laboratoire MAS, École Centrale Paris, Châtenay-Malabry, France
[3]Equipe GALEN, INRIA Saclay - Île-de-France, Orsay, France

Abstract. In this paper, we introduce a simple but efficient cue for the extraction of shadows from a single color image, the bright channel cue. We discuss its limitations and offer two methods to refine the bright channel: by computing confidence values for the cast shadows, based on a shadow-dependent feature, such as hue; and by combining the bright channel with illumination invariant representations of the original image in a flexible way using an MRF model. We present qualitative and quantitative results for shadow detection, as well as results in illumination estimation from shadows. Our results show that our method achieves satisfying results despite the simplicity of the approach.

1 Introduction

Shadows are an important visual cue in natural images. In many applications they pose an additional challenge, complicating tasks such as object recognition. On the other hand, they provide information about the size and shape of the objects, their relative positions, as well as about the light sources in the scene. It is however difficult to take advantage of the information provided by shadows in natural images, since it is hard to differentiate between shadows, albedo variations and other effects.

The detection of cast shadows in the general case is not straightforward. Shadow detection, in the absence of illumination estimation or knowledge of 3D geometry is a well studied problem. [1] uses invariant color features to segment cast shadows in still or moving images. [2] suggests a method to detect and remove shadows based on the properties of shadow boundaries in the image. In [3, 4], a set of illumination invariant features is proposed to detect and remove shadows from a single image. This method is suited to images with relatively sharp shadows and makes some assumptions about the lights and the camera. Camera calibration is necessary; if this is not possible, an entropy minimization method is proposed to recover the most probable illumination invariant image. In [5], a method for high-quality shadow detection and removal is discussed. The method, however, needs some very limited user input. Recently, [6] proposed a method to detect shadows in the case of monochromatic images, based on a series of features that capture statistical properties of the shadows.

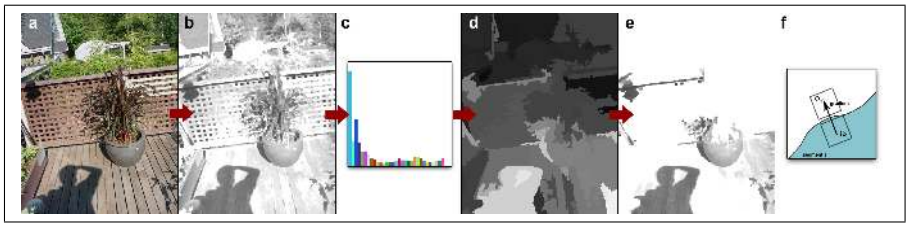

Fig. 1. Bright channel: a. original image (from [3]); b. bright channel; c. hue histogram; d. confidence map; e. refined bright channel; f. confidence computation: for a border pixel i of segment s, we compare the two patches oriented along the image gradient

In this paper, we discuss the estimation of cast shadows in a scene from a single color image.

We first propose a simple but effective image cue for the extraction of shadows, the *bright channel*, inspired from the dark channel prior [7]. Such a cue exploits the assumption that the value of each color channel of a pixel is limited by the incoming radiance, but there are pixels in an arbitrary image patch with values close to the upper limit for at least one color channel.

Then we describe a method to compute confidence values for the cast shadows in an image, in order to alleviate some inherent limitations of the bright channel prior. We process the bright channel in multiple scales and combine the results. We also present an alternative approach for refining the bright channel values, utilizing a Markov Random Field (MRF) model. The MRF model combines the initial bright channel values with a number of illumination-invariant representations to generate a labeling of shadow pixels in the image.

We evaluate our method on the dataset described in [6] and measure the accuracy of pixel classification. We also provide results for qualitative evaluation on other images, and demonstrate an example use of our results to perform illumination estimation with a very simple voting procedure.

This paper is organized as follows: Sec. 2 introduces the bright channel cue; Sec. 3 presents a way to compute confidences for cast shadows and refine the bright channel; Sec. 4 describes an MRF model to combine the bright channel with illumination-invariant cues for shadow estimation, followed by experimental results in Sec. 5. Sec. 6 concludes the paper.

2 Bright channel cue concept

To define the bright channel cue, we consider the following observations:

- The value of each of the color channels of the image has an upper limit which depends on the incoming radiance. This means that, if little light arrives at the 3D point corresponding to a given pixel, then all color channels will have low values.
- In most images, if we examine an arbitrary image patch, the albedo for at least some of the pixels in the patch will probably have a high value in at least one of the color channels.

From the above observations we expect that, given an image patch, the maximum value of the r, g, b color channels should be roughly proportional to the incoming radiance. Therefore, we define the *bright channel*, I_{bright} for image $\mathbf{I}$ in a way similar to the definition of the dark channel [7]:

$$I_{bright}(i) = max_{c\in\{r,g,b\}} \left(max_{j\in\Omega(i)}(I^c(j))\right) \tag{1}$$

where $I^c(j)$ is the value of color channel c for pixel j and $\Omega(i)$ is a rectangular patch centered at pixel i. We form the bright channel image of $\mathbf{I}$ by computing $I_{bright}(i)$ for every pixel i.

2.1 Interpretation

Let us assume that a scene is illuminated by a finite discrete set $\mathcal{L}$ of distant light sources. Each light source j ($j \in \mathcal{L}$) is described by its direction $\mathbf{d}_j$ and intensity α_j. We assume that the surfaces in the scene exhibit Lambertian reflectance. Let $\mathcal{G}$ be the 3D geometry of the scene and $\mathbf{p}$ be a 3D point imaged at pixel i. We can express the intensity $I(i)$ of pixel i as the sum of the contributions of the light sources that are not occluded at point $\mathbf{p}$:

$$I(i) = \rho(\mathbf{p})\eta(\mathbf{p}), \tag{2}$$

$$\eta(\mathbf{p}) = \sum_{j\in\mathcal{L}} \alpha_j \left[1 - c_{\mathbf{p}}(\mathbf{d}_j)\right] \max\{-\mathbf{d}_j \cdot \mathbf{n}(\mathbf{p}), 0\}, \tag{3}$$

where $\rho(\mathbf{p})$ is the reflectance (albedo) at $\mathbf{p}$, $\mathbf{n}(\mathbf{p})$ is the normal vector at $\mathbf{p}$ and $c_{\mathbf{p}}(\mathbf{d}_j)$ is the occlusion factor for direction $\mathbf{d}_j$ at $\mathbf{p}$:

$$c_{\mathbf{p}}(\mathbf{d}_j) = \begin{cases} 1, \text{ if ray from a light to } \mathbf{p} \text{ along } \mathbf{d}_j \text{ intersects } \mathcal{G} \\ 0, \text{ otherwise} \end{cases} \tag{4}$$

Here we are interested in the illumination component $\eta(\mathbf{p})$. One should note, though, that it cannot be calculated directly since the reflectance $\rho(\mathbf{p})$ above is unknown. The definition of the bright channel, $I_{bright}(i)$ produces a natural lower bound for $\eta(\mathbf{p})$:

$$I(i) \leq I_{bright}(i) \leq \eta(\mathbf{p}). \tag{5}$$

Eq. 5, combined with our observations above, means that the bright channel $I_{bright}(i)$ can provide an adequate approximation to the illumination component $\eta(\mathbf{p})$.

An example of the bright channel of an image is shown in Fig. 1.

2.2 Post-processing

Assuming that at least one pixel in a patch $\Omega(i)$ is fully illuminated, one would observe high values in at least one color channel. However, due to low reflectance or exposure, only in few cases this maximum value is actually the full intensity (1.0). As a result, the values of I_{bright} appear slightly darker than our expectation

for η. Thus it is natural to assume that, for any image $\mathbf{I}$, at least β % of the pixels are fully illuminated, and their correct values in the bright channel should be 1.0. This assumption can be easily encoded through sorting the values $I_{bright}(i)$ of pixels in descending order, and choosing the value lying at β %, I^{β}_{bright}, as the white point. Then, we can adjust the bright channel values as:

$$\dot{I}_{bright}(i) = min \left\{ \frac{I_{bright}(i)}{I^{\beta}_{bright}}, 1.0 \right\} \tag{6}$$

The second concern of the bright channel is that the dark regions in the bright channel image appear shrunk by $\kappa/2$ pixels, where $\kappa \times \kappa$ is the size of the rectangular patches $\Omega(i)$. This can be explained if the *max* operation in Eq. 1 is seen as a dilation operation. We correct this by expanding the dark regions in the bright channel image by $\kappa/2$ pixels, using an *erosion* morphological operator [8]. An example of the adjusted bright channel is shown in Fig. 1.b.

3 Robust bright channel estimation

The value of the bright channel cue heavily depends on the scale of the corresponding patch and does not always provide a good approximation of $\eta(\mathbf{p})$ at scene point $\mathbf{p}$. For example, a surface with a material of dark color, which is larger in the image than the patch size used to compute the bright channel cue, will appear dark in the bright channel, even if it is fully illuminated. On the other hand, shadows that are smaller than half the patch size will not appear in the bright channel. We present a method to remedy these problems by computing the bright channel cue in multiple scales, and by computing a confidence value for each dark area in the bright channel image.

3.1 Computing confidence values

Since surfaces with dark colors can appear as dark areas in the bright channel, even if they are fully illuminated, we seek a way to compute a confidence that each dark area is indeed dark because of illumination effects. In this paper we are particularly interested in cast shadows.

We first obtain a segmentation Υ of the bright channel image, and we seek to compute a confidence value for each segment. This computation is based on the following intuition: Let Ω_1 and Ω_2 be two $m \times n$ patches in the original image, lying on the two sides of a border caused by illumination conditions (Fig. 1.f). If we compute the values of some feature f_I, which characterizes cast shadows, for both patches and compare them, we expect to find that the difference $\Delta f = f_I(\Omega_1) - f_I(\Omega_2)$ is consistent for all such pairs of patches taken across shadow borders in the scene. On the other hand, the difference Δf will be inconsistent across borders that can be attributed to texture or other factors.

The use of a simple feature like hue is enough to effectively compute a set of confidence values for each segment of the segmentation Υ of the bright channel.

Let $\Delta f_I^{hue}(\Omega_1, \Omega_2)$ be the difference in hue between neighboring patches Ω_1 and Ω_2, where Ω_1 lies inside a cast shadow while Ω_2 lies outside. We expect $\Delta f_I^{hue}(\Omega_1, \Omega_2)$ to be consistent for all pairs of patches Ω_1 and Ω_2 on the border of that shadow.

If patches Ω_1 and Ω_2 are chosen to lie on the two sides of the border of a shadow, then all $\Delta f_I^{hue}(\Omega_1, \Omega_2)$ along this border will lie close to a value μ_k that depends on the hue of the light sources that are involved in the formation of this shadow border. If we model the deviations from this value μ_k due to changes in albedo, image noise, etc., with a normal distribution $\mathcal{N}(0, \sigma_k)$, the hue differences $\Delta f_I^{hue}(\Omega_1, \Omega_2)$ will follow a normal distribution:

$$\Delta f_I^{hue}(\Omega_1, \Omega_2) \sim \mathcal{N}(\mu_k, \sigma_k) \tag{7}$$

The distribution of all $\Delta f_I^{hue}(\Omega_1, \Omega_2)$ across all segment borders in segmentation Υ is modeled by a mixture of normal distributions. The parameters of this mixture model are, for each component k, the mean μ_k, the variance σ_k and the mixing factor π_k. We use an Expectation-Maximization algorithm to compute these parameters, while the number of distributions in the mixture is selected by minimizing a quasi-Akaike Information Criterion (QAIC). The confidence for segment $s \in \Upsilon$ is then defined as:

$$p(s) = \frac{1}{|\mathcal{B}_s|} \max_k \sum_{i \in \mathcal{B}_s} P_k \left(\Delta f_I^{hue}(\Omega_1(i), \Omega_2(i)) \right), \tag{8}$$

where $\mathcal{B}_s$ is the set of all border pixels of segment s, k identifies the mixture components, and, for patches $\Omega_1(i)$ and $\Omega_2(i)$ on the two sides of border pixel i, $P_k \left(\Delta f_I^{hue}(\Omega_1(i), \Omega_2(i)) \right)$ is the probability density corresponding to Gaussian component k (weighed by the mixture factor π_k).

We take advantage of one more cue to improve the estimation of $p(s)$: we expect that, for every neighboring pair Ω_1, Ω_2, with Ω_1 lying inside the shadow and Ω_2 outside, the value of each of the three color channels will be decreasing to the direction of Ω_1:

$$\frac{1}{|\Omega_1|} \sum_{i \in \Omega_1} I^c(i) - \frac{1}{|\Omega_2|} \sum_{i \in \Omega_2} I^c(i) < 0, \forall c \in \{r, g, b\} \tag{9}$$

If the percentage of patch pairs that violate this assumption for segment s is bigger than θ_{dec}, we set $p(s)$ to 0.

3.2 Multi-scale computation

We mentioned earlier the trade-off associated with the patch size κ used to compute the bright channel cue. One can overcome this limitation through computating the bright channel in multiple scales and combining the results. The term "scale" refers here to the patch size $\kappa \times \kappa$.

For each scale j of a total N_s scales, a confidence value is computed for each pixel. We combine the confidences from all scales in a final confidence map, by

setting the final confidence of each pixel i to

$$p_s(i) = \left(\prod_{j=0}^{N_s} p_s^{(j)}(i) \right)^{\frac{1}{N_s}}, \tag{10}$$

where $p_s^{(j)}(i)$ is the confidence of segment s at scale j, and s is the segment to which pixel i belongs at scale j. Notice that the segmentation is different at each scale, since it is performed on the bright channel values, which depends on κ. We set the bright channel value of any pixel i with confidence $p_s(i) < \xi$ to 1.0. For the rest of the pixels, the final bright channel value is the value computed with the smallest patch size κ_j.

Fig. 1 shows that the use of confidence values significantly improves the results of the bright channel. While the unfiltered bright channel included every dark surface in the image, the result after computing the confidence values includes mainly values related to shadows. These measurements will be used for a global formulation that involves optimal cast shadows detection and illumination estimation in the next section.

4 An MRF Model for Shadow Detection

In this section we present an alternative method to refine the bright channel values, by combining them with well-known illumination-invariant representations of the input image. Graphical models can efficiently fuse different cues within a unified probabilistic framework. Here we describe an MRF model which fuses a number of different shadow cues to achieve higher quality shadow estimation. In this model, the per-pixel shadow values are associated on one hand with the recovered bright channel values, and on the other with a number of illumination invariant representations of the original image.

4.1 Illumination Invariants

Separating shadows from texture is a difficult problem. In our case, we want to reason about gradients in the original image and attribute them to either changes in shadow or to texture variations. For this purpose, we use three illumination-invariant image representations. Ideally, an illumination-invariant representation of the original image will not contain any information related to shadows. Having such a representation, we can compare gradients in the original image with gradients in the illumination-invariant representation to attribute the gradient to either shadows/shading or texture. Having identified shadow borders this way, we can produce a set of labels identifying shadows in the original image.

Illumination-invariant image cues are not sufficient in the general case, however, and more complicated reasoning is necessary for more accurate shadow detection. An example of this can be seen in Fig.2, which shows the illumination invariant features we use for an example image. Edges due to illumination, although dimmer, are still noticeable, while some texture edges are not visible.

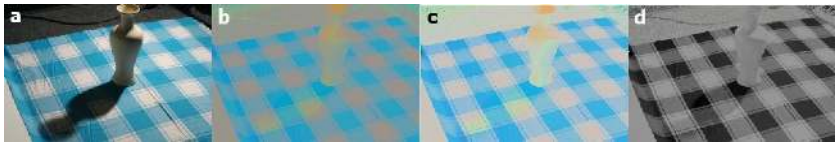

Fig. 2. Illumination invariant images: a) original image, b) normalized rgb, c) $c_1c_2c_3$, d) the 1d illumination invariant image obtained using the approach in [4]. Notice that in all three illumination invariant images, the shadow is much less visible than in the original.

4.2 Illumination-invariant cues

Photometric color invariants are functions which describe each image point, while disregarding shading and shadows. These functions are demonstrated to be invariant to a change in the imaging conditions, such as viewing direction, object's surface orientation and illumination conditions. Some examples of photometric invariant color features are normalized RGB, hue, saturation, $c_1c_2c_3$ and $l_1l_2l_3$ [9]. A more complicated illumination invariant representation specifically targeted to shadows is described in [4]. Other interesting invariants that could be exploited are described in [10], [11], [12]. In this work, three illumination-invariant representations are integrated into our model: normalized rgb, $c_1c_2c_3$ and the representation proposed in [4] (displayed in Fig. 2). It is however very easy to add or substitute more illumination invariant representations.

The $c_1c_2c_3$ invariant color features are defined as:

$$c_k(x,y) = \arctan \frac{\rho_k(x,y)}{max\{\rho_{(k+1)mod3}(x,y), \rho_{(k+2)mod3}(x,y)\}} \tag{11}$$

where $\rho_k(x,y)$ is the k-th RGB color component for pixel (x,y).

We only use the $1d$ illumination invariant representation proposed in [4]. For this representation, a vector of illuminant variation e is estimated. The illumination invariant features are defined as the projection of the log-chromaticity vector x' of the pixel color with respect to color channel p to a vector $e^{\perp}$ orthogonal to e:

$$I' = \mathbf{x}'^T e^{\perp} \tag{12}$$

$$x'_j = \frac{\rho_k}{\rho_p}, k \in 1,2,3, k \neq p, j = 1,2 \tag{13}$$

and ρ_k represents the k-th RGB component.

These illumination invariant features assume narrow-band camera sensors, Planckian illuminants and a known sensor response, which requires calibration. We circumvent the known sensor response requirement by using the entropy-minimization procedure proposed in [3] to calculate the illuminant variation direction e. Futhermore, it has been shown that the features extracted this way are sufficiently illumination-invariant, even if the other two assumptions above are not met ([4]).

4.3 The MRF Model

In this section we describe an MRF model that models the relationship of a brightness cue such as the bright channel with the illumination invariant cues, in order to obtain a shadow label for each pixel. Intuitively, through this MRF model we seek to obtain labelings that correspond to shadow edges where there is a transition in the bright channel value, but no significant transition/edge appears at the same site of an illumination-invariant representation of the image.

The proposed MRF has the topology of a 2D lattice and consists of one node for each image pixel $i \in \mathcal{P}$. The 4-neighborhood system [13] composes the edge set $\mathcal{E}$ between pixels. The energy of our MRF model has the following form:

$$E(\mathbf{x}) = \sum_{i \in \mathcal{P}} \phi_i(x_i) + \sum_{(i,j) \in \mathcal{E}} \psi_{i,j}(x_i, x_j), \tag{14}$$

where $\phi_i(x_i)$ is the singleton potential for pixel nodes and $\psi_{i,j}(x_i, x_j)$ is the pairwise potential defined on a pair of neighbor pixels. The singleton potential has the following form:

$$\phi_i(x_i) = \left(x_i - \dot{I}_{bright}(i)\right)^2, \tag{15}$$

where $\dot{I}_{bright}(i)$ is the value of the bright channel for pixel i. The pairwise potential has the form:

$$\psi_{i,j}(x_i, x_j) = (x_i - x_j)^2 \left(min_k\{I^{(k)}_{invar}(i) - I^{(k)}_{invar}(j)\}\right)^2, \tag{16}$$

where $I^{(k)}_{invar}(i)$ is the value of the k-th illumination invariant representation of the image at pixel i. Note that our MRF model is modular with respect to the illumination invariants used. Other cues can easily be integrated.

The latent variable x_i for pixel node $i \in \mathcal{P}$ represents the quantized shadow intensity at pixel i. We can perform cast shadows detection through a minimization over the MRF's energy defined in Eq. 14:

$$\mathbf{x}^{opt} = \arg\min_{\mathbf{x}} E(\mathbf{x}) \tag{17}$$

To minimize the energy of this MRF model we can use existing MRF inference methods such as TRW-S [14], the QPBO algorithm [15, 16] with the fusion move [17], etc. The latter was used for the experimental results presented in the next section.

5 Experimental Validation

In this section we present qualitative and quantitative results with the bright channel, and we show further results in an example application in illumination estimation from shadows.

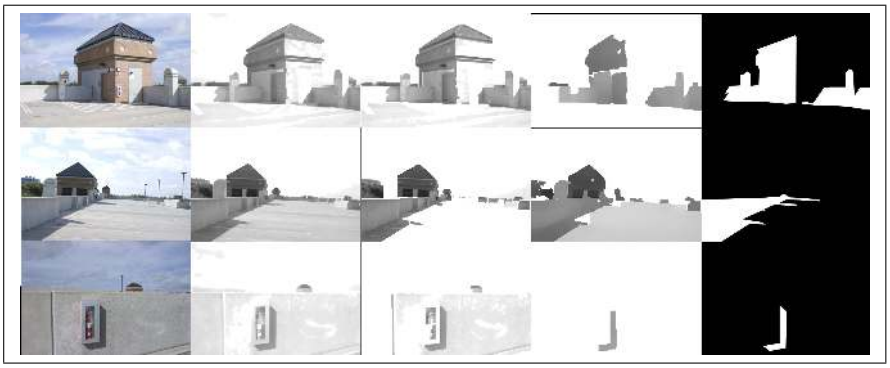

Fig. 3. Results with images from the dataset by [6]. From left to right: the original image; the (unrefined) bright channel; the bright channel refined using confidence estimation; the bright channel refined using the MRF model; the ground truth. These examples show advantages and weaknesses of the two refinement methods.

5.1 Quantitative Evaluation

We evaluated our approach on the dataset provided by [6], which contains 356 images and the corresponding ground truth for the shadow labels. In order to convert the bright channel values to a 0-1 shadow labeling, we used simple thresholding. The pixel classification rates are presented in table 1. Example results can be found in Fig. 3. Fig. 5 shows a case where our algorithm fails, due to very large uniformly dark surfaces.

method	classification rate (%)	false positives (%)	false negatives (%)
bright channel	83.52	13.16	3.31
bright channel + confidence	84.61	11.21	4.17
bright channel + MRF	85.88	8.83	5.28
brightness + MRF	52.53	46.31	1.15

Table 1. Pixel classification results for the unrefined bright channel (using a single patch size $\kappa = 6$ pixels); the bright channel refined using confidence values and 4 scales; our MRF model with the bright channel (using a single patch size $\kappa = 6$ pixels); and our MRF model with pixel brightness in the LAB color space instead of the bright channel for the singleton potentials.

5.2 Simple Illumination Estimation

We can use the bright channel image to perform illumination estimation from shadows. As a proof of concept, we describe a very simple voting method in

Algorithm 1, which is in most cases able to recover an illumination estimate given simple 3D geometry of the scene.

Algorithm 1 Voting to initialize illumination estimate

Lights Set: $\mathcal{L} \leftarrow \varnothing$
Direction Set: $\mathcal{D} \leftarrow$ all the nodes of a unit geodesic sphere
Pixel Set: $\mathcal{P} \leftarrow$ all the pixels in the observed image
loop
  votes[$\mathbf{d}$] $\leftarrow 0, \forall \mathbf{d} \in \mathcal{D}$
  for all pixel $i \in \mathcal{P}$ **do**
    for all direction $\mathbf{d} \in \mathcal{D} \setminus \mathcal{L}$ **do**
      if $I_{bright}(i) < \theta_S$ **and** $\forall \mathbf{d}' \in \mathcal{L}, c_i(\mathbf{d}') = 0$ **then**
        if $c_i(\mathbf{d}) = 1$ **then** votes[$\mathbf{d}$] $\leftarrow$ votes[$\mathbf{d}$] $+ 1$
      else
        if $c_i(\mathbf{d}) = 0$ **then** votes[$\mathbf{d}$] $\leftarrow$ votes[$\mathbf{d}$] $+ 1$
  $\mathbf{d}^* \leftarrow \arg\max_{\mathbf{d}}(votes[\mathbf{d}])$
  $\mathcal{P}_{\mathbf{d}^*} \leftarrow \{i | c_i(\mathbf{d}^*) = 1 \text{ and } \forall \mathbf{d} \neq \mathbf{d}^*, c_i(\mathbf{d}) = 0\}$
  $\alpha_{\mathbf{d}^*} \leftarrow median \left\{ \frac{1 - I_{bright}(i)}{\max\{-\mathbf{n}(\mathbf{p}(i)) \cdot \mathbf{d}^*, 0\}} \right\}_{i \in \mathcal{P}_{\mathbf{d}^*}}$
  if $\alpha_{\mathbf{d}^*} < \epsilon_\alpha$ **then**
    stop the loop
  $\mathcal{L} \leftarrow \mathcal{L} \cup (\mathbf{d}^*, \alpha_{\mathbf{d}^*})$

The idea is that, shadow pixels that are not explained from the discovered light sources vote for the occluded light directions. The pixels that are not in shadow vote for the directions that are not occluded. After discovering a new light source direction, we estimate the associated intensity using the median of the bright channel values of pixels in the shadow of this new light source. The process of discovering new lights stops when the current discovered light does not have a significant contribution to the shadows in the scene. To ensure even sampling of the illumination environment, we choose the nodes of a geodesic sphere of unit radius as the set of potential light directions [18]. The results of the voting algorithm are used to initialize the MRF both in terms of topology and search space leading to more efficient use of discrete optimization. When available, the number of light sources can also be set manually.

We present results on illumination estimation on images of cars collected from Flickr (Fig. 4). The geometry used in this case was a 3D bounding box representing the car in each image, and a plane representing the ground. The camera parameters were matched by hand so that the 3D model's projection would roughly coincide with the car in the image.

6 Conclusions

In this paper, we presented a simple but effective image cue for the extraction of shadows from a single image, the bright channel cue. We discussed the lim-

Fig. 4. Results with images of cars collected from Flickr. Top row: the original image and a synthetic sun dial rendered with the estimated illumination; Bottom row: the refined bright channel. The geometry consists of the ground plane and a single bounding box for the car.

itations of this cue, and presented a way to deal with them, by examining the bright channel values at multiple scales and computing confidence values for each dark region using a shadow-dependent feature, such as hue. We further described an MRF model as an alternative way to refine the bright channel cue by combining it with a number of illumination-invariant representations. In the results, we computed the classification accuracy for shadow pixels on a publicly available dataset, we showed examples of the resulting shadow estimates, and we discussed one potential application of the bright channel cue in illumination estimation from shadows. In this application, the low false-negative rate and the relatively accurate shadow estimate we can get from this simple cue makes it possible to tackle a hard problem such illumination estimation with rough geometry information in natural images using simple algorithms such as the voting algorithm we described. In the future, we are interested in incorporating this cue in a more complex shadow detection framework.

Acknowledgments: This work was partially supported by NIH grants 5R01EB7530-2, 1R01DA020949-01 and NSF grants CNS-0627645, IIS-0916286, CNS-0721701.

Fig. 5. A failure case: from left to right, the original image, the bright channel, and the refined bright channel. The uniformly dark road surface is identified as a shadow.

References

1. Salvador, E., Cavallaro, A., Ebrahimi, T.: Cast shadow segmentation using invariant color features. Computer Vision and Image Understanding **95** (2004) 238–259
2. Levin, A., Lischinski, D., Weiss, Y.: A closed-form solution to natural image matting. IEEE Transactions on Pattern Analysis and Machine Intelligence **30** (2008) 228–242
3. Finlayson, G., Drew, M., Lu, C.: Intrinsic images by entropy minimization. In: European Conference on Computer Vision (ECCV). (2004)
4. Finlayson, G., Hordley, S., Lu, C., Drew, M.: On the removal of shadows from images. IEEE Transactions on Pattern Analysis and Machine Intelligence **28** (2006) 59–68
5. Shor, Y., Lischinski, D.: The shadow meets the mask: Pyramid-based shadow removal. Computer Graphics Forum **27** (2008) 577–586
6. Zhu, J., Samuel, K.G.G., Masood, S., Tappen, M.F.: Learning to recognize shadows in monochromatic natural images. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2010). (2010)
7. He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). (2009)
8. Gonzalez, R.C., Woods, R.E.: Digital Image Processing (3rd Edition). Prentice-Hall, Inc. (2006)
9. Gevers, T., Smeulders, A.W.M.: Color based object recognition. Pattern Recognition **32** (1997) 453–464
10. Geusebroek, J.M., van den Boomgaard, R., Smeulders, A.W.M., Geerts, H.: Color invariance. IEEE Transactions on Pattern Analysis and Machine Intelligence **23** (2001) 1338–1350
11. van de Weijer, J., Gevers, T., Geusebroek, J.M.: Edge and corner detection by photometric quasi-invariants. IEEE Transactions on Pattern Analysis and Machine Intelligence **27** (2005)
12. Diplaros, A., Gevers, T., Patras, I.: Combining color and shape information for illumination-viewpoint invariant object recognition. IEEE Trans. on Image Processing **15** (2006) 1–11
13. Boykov, Y., Funka-lea, G.: Graph cuts and efficient n-d image segmentation. International Journal of Computer Vision **70** (2006) 109–131
14. Kolmogorov, V.: Convergent tree-reweighted message passing for energy minimization. IEEE Transactions on Pattern Analysis and Machine Intelligence **28** (2006) 1568–1583
15. Hammer, P.L., Hansen, P., Simeone, B.: Roof duality, complementation and persistency in quadratic 0-1 optimization. Mathematical Programming **28** (1984) 121–155
16. Kolmogorov, V., Rother, C.: Minimizing nonsubmodular functions with graph cuts-a review. IEEE Transactions on Pattern Analysis and Machine Intelligence **29** (2007) 1274–1279
17. Lempitsky, V., Rother, C., Roth, S., Blake, A.: Fusion moves for markov random field optimization. IEEE Transactions on Pattern Analysis and Machine Intelligence **32** (2010) 1392–1405
18. Sato, I., Sato, Y., Ikeuchi, K.: Illumination from shadows. IEEE Transactions on Pattern Analysis and Machine Intelligence **25** (2003) 290–300