

Illumination Estimation from Shadow Borders

Alexandros Panagopoulos, Tomás F. Yago Vicente, Dimitris Samaras
Stony Brook University
Stony Brook, NY, USA

{apanagop, tyagovicente, samaras}@cs.stonybrook.edu

Abstract

In this paper we discuss illumination estimation from a single image in general scenes and associate it with the existence of shadow edges, avoiding several pitfalls that burden previous illumination estimation approaches, which rely on associating a parametrization of illumination with the per pixel intensity of shadows or shading. We show a way to couple shadow and illumination estimation, relying only on the subset of shadow edges that is relevant to the provided geometry. In our approach, illumination estimation is posed as the minimization of an energy function that penalizes the matching between the expected shadow outline and observed image edges. Minimizing this energy function is strongly tied to selecting the appropriate set of potential shadow edges in the image. Our approach leads to an illumination estimation algorithm that performs on par with or better than the state of the art, even when scene geometry knowledge is limited, while having much lower computational complexity than state-of-the-art methods. We demonstrate the effectiveness of this approach both with quantitative results on synthetic data and qualitative evaluation on real images.

1. Introduction

The process of image formation depends on three components: the 3D geometry of the scene, the reflectance properties of the surfaces in it, and illumination. The interaction among these three components means that estimation of one or two of them requires knowledge or strong assumptions about the rest ([10, 13, 15, 17, 18]). Most previous work in illumination estimation depends on strong assumptions on 3D geometry and reflectance, such as accurate knowledge of scene geometry, while only recently some approaches have appeared that allow the estimation of a set of illuminants having only coarse knowledge of geometry [12].

In this paper we discuss illumination estimation in general scenes and associate it with the existence of shadow edges. Most general illumination estimation methods from

shadows (or shading) associate a parametrization of illumination with the per pixel intensity of shadows or shading. As a result, estimating illumination from shadows in a general scene generally needs a way to separate shadows from scene albedo and other effects as initial input. Two significant types of errors can be introduced this way: errors in the initial shadow estimate propagate throughout the illumination estimation process, altering the final results, even in the case where the shadow estimate is refined during illumination estimation [12]; on the other hand, the knowledge of scene structure may not be adequate to explain a lot of correctly detected shadows in complex scenes, leading to erroneous illumination solutions that try to explain every observed shadow with inadequate geometry data. In this paper we propose a way to couple shadow and illumination estimation, trying to detect only the shadow edges that are relevant to the provided geometry, as part of the illumination estimation process. This leads to illumination estimation algorithms that perform on par with or better than the state of the art, even when scene geometry knowledge is limited, while not requiring an initial and potentially difficult to obtain shadow estimate, and having lower computational complexity than state-of-the-art methods. In our approach, illumination estimation is posed as the minimization of an energy function, and coupled with the detection of salient shadow edges.

As mentioned earlier, illumination estimation from cast shadows in a single image usually necessitates obtaining an estimate of the cast shadows in the image, which can be challenging in complex natural images. Shadow detection, in the absence of illumination estimation or knowledge of 3D geometry is a well studied problem. Salvador et al [14] use invariant color features to segment cast shadows in still or moving images. Finlayson et al [2, 3] propose a set of illumination invariant features to detect and remove shadows from a single image, making several assumptions about the lights and the camera. Recently, Zhu et al [20] combined a number of cues in a complex method to recognize shadows in monochromatic images, while in [9], Lalonde et al propose a learning approach to detect shadows in consumer-

grade photographs, focusing on shadows on the ground. The above methods detect the majority of shadow pixels, but they are not always accurate since they are based only on image statistics.

Extracting illumination from shading, specular reflections or cast shadows has been studied extensively in the past. Yang and Yuille ([18]) detect a small number of light source directions using critical points, and Wang et al [17] extend it to an image of an arbitrary object with known shape and estimate not only light source directions but also virtual lights. A method for estimating the illumination distribution of a real scene from shadows is proposed by Sato et al [15], assuming known geometry illuminated by infinitely distant light sources, casting shadows onto a planar lambertian surface. Hara et al [6] remove the distant illumination assumption, while simultaneously estimating illumination and reflectance. In [19], Zhou et al propose a unified framework to estimate both distant and point light sources.

The prior art on illumination estimation from shadows cast on textured surfaces is limited. Sato et al [15] require an extra image to deal with texture. Li et al [10] propose a method that integrates multiple cues from shading, shadow, and specular reflections. Kim et al [7] use regularization by correlation to estimate illumination from shadows when texture is present, but require extra user-specified information and assume lambertian reflectance and known geometry. Panagopoulos et al [11] proposed a method able to deal with inaccurate geometry and texture, but the shadow detection results when texture is present are limited. In the more limited case of daytime outdoor scenes, Lalonde et al [8] proposed an approach that combines cues from the sky, cast shadows on the ground and surface brightness to estimate illumination where the sun is the single light source. The previous approach that is more relevant to the work proposed in this paper was recently presented in [12], where a small set of distant light sources is estimated in the presence of inaccurate knowledge of geometry, by formulating the interaction of geometry and illumination for the creation of cast shadows as a Markov Random Field (MRF) model. That approach results in a difficult inference problem on a higher-order MRF, while at the same time requiring an initial shadow estimate and the initialization of the illumination solution through a separate algorithm.

In this paper we will explicitly associate illumination with the existence of shadow edges, instead of per-pixel shadow intensities. We do so by defining an energy that corresponds to the quality of the matching between the observed shadow edges in the image and the shadow edges expected by the illumination solution. Shadow edge detection is based on comparing gradients in the original image and two illumination invariant representations of it; in the limit, our approach can work without performing any shadow edge detection at all, assuming that all image edges

are potential shadow edges (as we demonstrate in figure 6). The potential shadow edges are encoded in a shadow edge confidence map, and a simple approach is described to minimize the solution energy given this map, obtaining the illumination parameters that correspond to a good matching of the expected shadow edges with observed image edges.

The contributions of this paper are the following:

- We explicitly associate illumination with shadow edges instead of per-pixel shadow intensities. This allows our approach to ignore errors in shadow detection, and concentrate only on potential shadow silhouettes that are meaningful given the scene geometry.
- This fact further allows our approach to estimate illumination using 3D geometry that only partially models a complex scene; for example, approximate knowledge of a single shadow-casting object and the rough shape of the surface its shadow is cast on can be adequate to estimate illumination in a larger, complex scene.
- Our approach is robust to inaccurate knowledge of 3D geometry, allowing us to model objects in real images using very coarse geometry, such as 3D bounding boxes. Our quantitative results demonstrate the robustness of our method with regard to geometry inaccuracies.

We present both quantitative and qualitative results. Quantitative results show the accuracy of our approach when estimating illumination in a synthetic dataset, demonstrating that our approach performs better than the state of the art in the case of errors in the shadow estimation and inaccuracies in the modeling of 3D geometry. Qualitative results show how our method performs in a set of real images collected from Flickr, and are compared to the results obtained by [12]. At the same time, the computational cost of our approach is significantly lower than comparable approaches, corresponding to a much simpler implementation.

This paper is organized as follows: in section 2 we formulate illumination estimation as the minimization of an energy function that measures the quality of the match between the expected shadow gradient and an edge map extracted from the image. Section 3 describes how we obtain this edge map, while in section 4 we describe energy minimization and we extend our solution to the case of multiple light sources. Results are presented in section 5, while section 6 concludes the paper.

2. Formulation

We will first examine the case where the scene is illuminated by a single distant light source, with direction \mathbf{d}_0 and intensity α_0 . Let $\mathcal{E} = \{e_i\}$ be a set of edges detected from the original image, and $Q(e_i) \in [0, 1]$ be a confidence

value that edge e_i is generated by a cast shadow (larger values indicate higher confidence). A geometric model \mathcal{G} is also known. Geometry \mathcal{G} may model only a small part of the scene and may be approximate - e.g. in many of our experiments we approximate objects by 3D bounding boxes.

Our goal is to find the light parameters $\theta_{\mathcal{L}} = (\mathbf{d}_0, \alpha_0)$ that produce a shadow with shadow borders $\hat{\mathcal{E}}(\theta_{\mathcal{L}}|\mathcal{G})$ that:

- best coincide with image edges that have high confidence values $Q(e_i)$ to belong to shadows, and
- have a similar orientation with the corresponding observed image edges.

We express this requirement by defining an energy for each set of light parameters:

$$E_{match}(\theta_{\mathcal{L}}) = \frac{1}{|\hat{\mathcal{E}}(\theta_{\mathcal{L}}|\mathcal{G})|} \left(\sum_{i \in \hat{\mathcal{E}}(\theta_{\mathcal{L}}|\mathcal{G})} (1 - Q(i)) + \sum_{i \in \hat{\mathcal{E}}(\theta_{\mathcal{L}}|\mathcal{G})} \widehat{\nabla I_i, \mathbf{e}_i}^2 \right), \quad (1)$$

where $\widehat{\nabla I_i, \mathbf{e}_i}$ is the angle between the observed image gradient ∇I_i at pixel i and the direction of the synthetic shadow edge at i , \mathbf{e}_i .

Notice that in this formulation, we have already removed several important requirements of traditional illumination estimation methods:

1. We do not need to know or estimate the intensity of ambient illumination
2. We are not defining the energy over all possible shadow edges in the scene, but only for that set of edges that is generated by the geometry \mathcal{G} and the set of light parameters $\theta_{\mathcal{L}}$.
3. We do not need to estimate the intensity of light sources while estimating light source directions, because the set of edges $\hat{\mathcal{E}}(\theta_{\mathcal{L}}|\mathcal{G})$ depends only on the light source direction. The light source intensity can be included in the energy minimization (see Eq.3) or, as we preferred here, it can be estimated after the light source directions have been estimated.

Therefore, the matching cost $E_{match}(\theta_{\mathcal{L}})$ only depends on the light directions and the confidences assigned to observed shadow edges.

If we wish to estimate light source intensity α_0 concurrently with light source direction, we can minimize the sum of $E_{match}(\theta_{\mathcal{L}})$ and a term $E_{\alpha}(\theta_{\mathcal{L}})$:

$$E_{\alpha}(\theta_{\mathcal{L}}) = \sum_{i \in \hat{\mathcal{E}}(\theta_{\mathcal{L}}|\mathcal{G})} \left(\alpha_0 - \frac{\bar{I}_{out} - \bar{I}_{int}}{\max\{\mathbf{n}_i \cdot \mathbf{d}_0, 0\}} \right)^2, \quad (3)$$



Figure 1. a. The original image; b. The edge confidence map Q (right) extracted from the original image. Brighter pixels indicate higher confidence; c. The gradient directions after smoothing (used to penalize the expected shadow gradient). The x and y components of the gradient are encoded in the red and green channels.

where \bar{I}_{in} and \bar{I}_{out} are the mean pixel intensities of two image patches placed on the two sides of pixel i , in the inside and outside of the expected shadow respectively, \mathbf{n}_i is the normal vector at pixel i , as given by the provided 3D geometry (if any for that pixel), and \mathbf{d}_0 is the light source direction. Eq.3 assumes Lambertian reflectance - our estimates, however, do not deteriorate significantly when this assumption is violated, and no such assumption is necessary to estimate only the light source directions.

We therefore manage to associate the light source directions with only the subset of observed edges in the image that matches the shadow borders $\hat{\mathcal{E}}(\theta_{\mathcal{L}}|\mathcal{G})$ produced by the current illumination estimate. One obvious issue that arises, though, is that in some cases there are trivial solutions that do not produce any cast shadows and such solutions will be preferred because they lead to minima of $E_{match}(\theta_{\mathcal{L}})$. To avoid this, we encourage solutions that have a larger number of well-explained shadow edge pixels, by defining the final energy to be minimized as:

$$E(\theta_{\mathcal{L}}) = \frac{1}{1 + |\hat{\mathcal{E}}_g|} (E_{match}(\theta_{\mathcal{L}}) + w_{\alpha} E_{\alpha}(\theta_{\mathcal{L}})), \quad (4)$$

where w_{α} is a weight and the term $w_{\alpha} E_{\alpha}(\theta_{\mathcal{L}})$ can be omitted if there is no need to estimate light directions and intensities *concurrently*. The set $\hat{\mathcal{E}}_g$ is the set of all the expected shadow edge pixels in $\hat{\mathcal{E}}(\theta_{\mathcal{L}}|\mathcal{G})$ that coincide with observed edges of high confidence:

$$\hat{\mathcal{E}}_g = \{e_i \in \hat{\mathcal{E}}(\theta_{\mathcal{L}}|\mathcal{G}) \mid Q(e_i) > \theta_Q\}, \quad (5)$$

where θ_Q is a confidence threshold. We set $\theta_Q = 0.5$ in our experiments.

3. Extracting the edge map

The main term of the energy we want to minimize is $E_{match}(\theta_{\mathcal{L}})$, which is mainly a sum of confidence values along the expected shadow borders, the form of the confidence map is important for finding the light parameters that minimize $E(\theta_{\mathcal{L}})$.

Let $Q = \{Q(i)\}$ be the confidence map. For each image pixel i , the confidence $Q(i)$ expresses the probability that

i belongs to a shadow edge, if there is an edge at i , or the probability that a shadow edge lies in the vicinity of i . Map Q must contain confidence values that smoothly increase as we approach observed edges in the image, in order to allow effective minimization of $E(\theta_{\mathcal{L}})$.

Therefore, to compute Q , we first detect edges in the image and compute the probability that they correspond to a shadow. Then we perform a series of smoothing operations to propagate the appropriate confidence values to pixels in Q that do not lie on image edges. The form of the final confidence map can be seen in Figure 1.

We first apply the Sobel edge detector to the original image I , obtaining a set of gradients, ∇I . We also calculate a set of illumination invariant representations of the original image I . We refer to the k -th illumination invariant representation of I as $I^{(k)}$. An illumination-invariant representation of the original image I will, ideally, not contain any effects of illumination, such as cast shadows and shading [4, 16, 1, 3]. Having such a representation, we can compare the gradients in the original image with gradients in the illumination-invariant representation to attribute the gradient to either shadows/shading or texture. The illumination invariants we chose to use for our experiments are the normalized RGB and $c_1c_2c_3$ representations [5]. We apply the Sobel edge detector to each illumination-invariant image representation $I^{(k)}$ to obtain the corresponding gradients $\nabla I^{(k)}$.

To compute a confidence that each pixel i belongs to a shadow border, we compare the gradients from the original image and each illumination invariant. We define the confidence value for pixel i as:

$$Q(i) = \max_k \left\{ \max \left\{ \|\nabla I(i)\| - w_I^k \|\nabla I^{(k)}(i)\|, 0 \right\} \right\}. \quad (6)$$

Because in practice some gradients related to illumination appear in the illumination invariant representations, we take the maximum of the differences between gradients in the original image and illumination invariants. The weights w_I^k were learned from a training set of images, which was a subset of the dataset of images with hand-annotated cast shadows provided by [20].

After obtaining this initial set of confidences, we apply a smoothing operation for a fixed number of iterations to propagate the confidences to pixels that do not belong to detected image edges. In this smoothing operation, the new confidence value $\hat{Q}(i)$ of pixel i with previous confidence $Q(i)$ is set to be:

$$\hat{Q}(i) = \begin{cases} (1 - \lambda)Q(i) + \lambda\bar{Q}(i), & \text{if } \|\nabla I(i)\| < \theta_e \\ \max\{Q(i), \bar{Q}(i)\}, & \text{otherwise} \end{cases}, \quad (7)$$

where $\bar{Q}(i)$ is the average of confidence values in a 3×3 neighborhood centered at pixel i . The value of λ was set to

0.5 and θ_e is a small threshold, so that edges with gradient magnitudes less than θ_e are not significant.

Similarly, we create a smoothed version of the edge gradients by setting the new gradient direction of each pixel to be the average of itself and its neighbors, weighted by their relative confidence values. The resulting confidence map and gradient directions can be seen in Fig.1.

4. Energy minimization

To find the optimal light parameters, we need to minimize the energy $E(\theta_{\mathcal{L}})$ in Eq.4. This energy contains multiple local minima, while we also cannot get a good approximation to its gradient. The evaluation of the energy for different parameters, however, is relatively fast. We therefore use a move-making approach, where we start from a random initial set of parameters, and perform a number of iterations, examining at each iteration a random step from the current parameter values:

Algorithm 1 Minimization of $E(\theta_{\mathcal{L}})$

Light parameters: $\theta_{\mathcal{L}} \leftarrow$ random parameters

Energy minimum: $E_{min} = E(\theta_{\mathcal{L}})$

loop

 generate proposed parameters $\hat{\theta}_{\mathcal{L}}$ given $\theta_{\mathcal{L}}$

if $E(\hat{\theta}_{\mathcal{L}}) < E(\theta_{\mathcal{L}})$ **then**

$E_{min} \leftarrow E(\hat{\theta}_{\mathcal{L}})$

$\theta_{\mathcal{L}} \leftarrow \hat{\theta}_{\mathcal{L}}$

end if

end loop

For the first K iterations, the generation of proposed parameters $\hat{\theta}_{\mathcal{L}}$ is done randomly, to randomly sample the whole parameter space. After the first K iterations, $\hat{\theta}_{\mathcal{L}}$ is generated by choosing the proposed light direction by sampling a von Mises-Fisher distribution centered at the previous estimate of light direction (if we want to estimate intensities at the same time, we also choose a proposed intensity as a sample from a normal distribution around the previous intensity estimate).

If the light intensity is not estimated as part of the energy minimization, we estimate it afterwards, using the estimate $\hat{\mathbf{d}}_0$ of light direction we obtained. The intensity estimate $\hat{\alpha}_0$ is the median of local intensity estimates along the expected shadow edges:

$$\hat{\alpha}_0 = \text{median}_{i \in \hat{\mathcal{E}}(\theta_{\mathcal{L}}|\mathcal{G})} \left\{ \frac{\bar{I}_{out} - \bar{I}_{int}}{\max\{\mathbf{n}_i \cdot \hat{\mathbf{d}}_0, 0\}} \right\}. \quad (8)$$

This very simple approach to minimize the energy $E(\theta_{\mathcal{L}})$ proved effective because it samples the whole parameter space, avoiding many of the local minima, and then concentrates its effort to the area around the best solution so far. However, it would be very desirable in the future

to examine other approaches that can give some guarantees about optimality, while also reducing the number of times the energy $E(\theta_{\mathcal{L}})$ has to be evaluated during minimization.

4.1. Dealing with multiple light sources

In our discussion so far we have examined only the case of a single light source. When multiple light sources are present, there will be multiple shadow outlines that can be explained by the provided geometry \mathcal{G} . We can deal with this case by discovering light sources one-by-one: We estimate the direction and intensity of each light source j , and then remove the corresponding edges from the edge confidence map \mathcal{Q} (removing the corresponding edge pixels and then re-applying the smoothing operation). We then repeat, estimating the next light source from the new, reduced edge confidence map. The process stops when the last discovered light source has very low average confidence values along its projected shadow border, or has near-zero intensity. This procedure can allow not only the estimation of the parameters of multiple light sources, but also to determine the number of light sources illuminating the scene.

5. Results

We evaluated our approach quantitatively using a synthetic dataset of 3D models rendered with a known distant point light source, as well as qualitatively with images collected from Flickr [12]. A total of 1000 iterations was performed for each image.

Results on the synthetic dataset are shown in Table 1. Examples of the synthetic images and models used are shown in Fig.2. The direction and intensity of the light source was chosen randomly. We examined four different cases:

Exact geometry: We used the same 3D model to render the image and to estimate illumination.

Approximate geometry: We used a 3D model that coarsely approximated the original geometry by a bounding box and a ground plane to estimate illumination.

Exact geometry and noisy shadow input: We used the same 3D model to render the image and to estimate illumination, as above, and a noisy initial shadow estimate. The latter was obtained by adding random dark patches to the rendered shadow (Fig.2). We used this form of noise because, on one hand our methods are relatively insensitive to spatially-uniform random noise, and on the other, in real data the errors generally affect large image regions which get mislabeled, which is emulated by this patch-based noise.

Approximate geometry and noisy shadow input: We estimated illumination parameters using a coarse 3D model and a noisy initial shadow estimate, as described above.

Table 1 shows the average error in light direction estimation, in degrees. It is easy to notice that our approach is almost unaffected by errors in the initial shadow estimate,

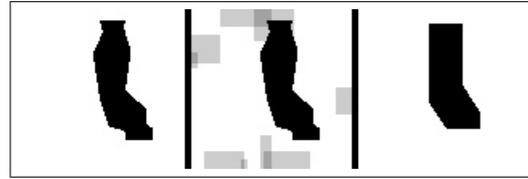


Figure 2. Examples of synthetic images used for quantitative evaluation. Image size was 100x100 pixels. From left to right: a) the full-resolution 3D model, rendered with 1 point light source; b) the full-resolution 3D model, after the addition of noise, as described in the paper (noise is random and may coincide with the shadow); c) the approximate 3D model corresponding to the full-resolution model on the left, rendered under the same illumination (for demonstration; this model is used for the illumination estimation only, in the experiments with approximate geometry).

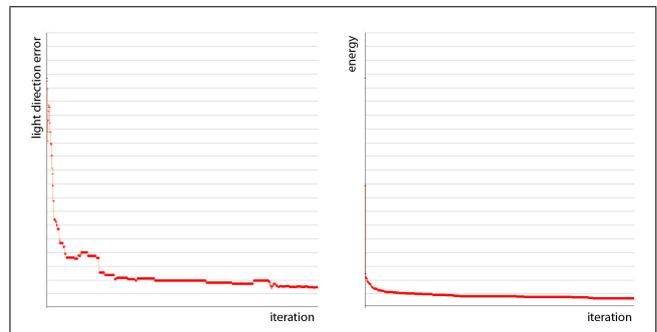


Figure 3. Convergence of our algorithm. Left: the error in the estimated light direction, averaged over a set of synthetic examples, per iteration; right: the average energy per iteration.

which have been simulated by the noisy shadow input. On the other hand, [12] shows a reduction in accuracy, especially when inaccurate shadow estimates are combined with inaccurate geometry. Notice that the synthetic dataset we used is different than that used in [12].

Figure 3 shows the convergence of our approach in the case of synthetic data.

Figure 4 shows results with our approach on a set of images of cars from Flickr. The geometry in these images consists of a ground plane and a 3D bounding box representing the car. We compare our results with the results obtained in [12]. The estimated illumination is shown by rendering a synthetic orange sun dial, illuminated by the illumination estimate obtained with our method, into the original image. Although accurate comparisons are difficult since there is no ground truth for illumination in these images, a visual inspection shows that our results are equally or more convincing than those obtained by [12].

Figure 6 demonstrates the flexibility of our method with regard to shadow detection. In this case, we compare the illumination estimate obtained when using all image edges (obtained with a Sobel detector) compared to using only

method	exact geometry	exact geom.+noise	approximate geom.	approximate geom.+noise
[12] (MRF)	0.39	5.06	2.55	17.87
our method, all samples	1.82	1.67	4.00	4.76
our method, 20% of edges	2.06	2.03	5.86	6.28

Table 1. Average error in light direction estimation for a set of synthetic images. The images were rendered using 1 known point light source, and the displayed error is the angle between the real and estimated light directions, in degrees. We compare with the state-of-the-art approach recently proposed in [12]. Examples of the images and geometry used are shown in Fig.2. The second row shows results with our approach when all expected shadow border pixels are used to evaluate the energy, and the third row shows results with our approach when only 20% of expected shadow border pixels is used, achieving a 5-fold speedup with small deterioration of the results. In every case, our approach is influenced much less than [12] by noisy shadow input, while being robust to inaccurate geometry.



Figure 4. Results with images of cars from Flickr. The results of illumination estimation are presented by rendering a synthetic orange sundial to the original image, using the estimated illumination. Top: the results with our method; bottom: results with the much more computationally intensive method proposed in [12]. Our results are equivalent or better than the results from [12], although our method uses only simple shadow edge detection and a much more efficient optimization to estimate illumination parameters.

potential shadow edges (by utilizing illumination invariants as described earlier). Our approach can select those image edges that correspond to plausible cast shadows, and obtain a good illumination estimate, even when no initial shadow edge detection is performed.

Examples of the 3D geometry used for illumination estimation in the case of Flickr images is shown in figure 7.

Excluding the cost of raytracing shadows, the computational cost of our method is linear to the number of edge samples; the number of edge samples used can be reduced without significant impact to the final results, in order to improve performance. In our unoptimized implementation, on a system with an Intel i5 CPU, each iteration took 4-10msec depending on image size, or 1-3msec when using one of every 5 shadow edge samples (in the case of synthetic results in Table 1). Paired with a GPU raytracer, the total time for our algorithm can be limited to 5-60 seconds per image, which is a significant improvement compared to previous, computationally intensive methods such as [12]. As our results show, this improvement usually comes at no cost to the accuracy of the estimated results.

One big difference compared to [12] is that, in the proposed approach, potential shadow edges that are far from the shadow edges generated by the known geometry are

not penalized at all. The advantages of this approach are that our method will ignore real shadows that cannot be explained by the geometry, if the geometry models only small part of the actual scene, while at the same time wrongly detected shadow edges need not be accounted for, and have no effect on the energy of the final solution. This advantage of our approach is demonstrated in figure 5. In this example only one of two nearby objects is modeled, causing the approach of [12] (and probably most approaches that are based on an error computed over all shadow pixels) to try to explain all shadows using the provided geometry, resulting in erroneous illumination estimation. Our approach, on the other hand, correctly estimates illumination by associating it only with a subset of the observed shadow edges. One drawback is that this kind of approach could potentially ignore real shadows when the geometry differs substantially, but in our experiments modeling objects with 3D bounding boxes, even with inaccurate modeling, was enough to associate the illumination solution with the correct set of shadow edges.

We only used light sources that produce sharp shadows in our experiments. A limitation of the proposed approach is that it cannot handle well soft shadows produced by area light sources.

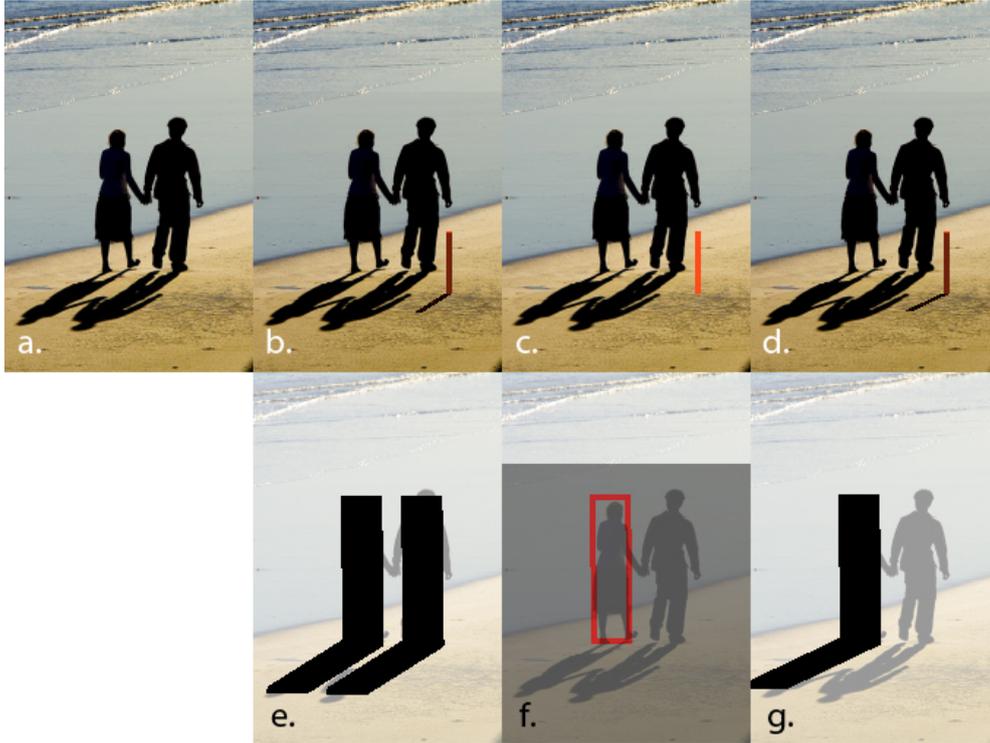


Figure 5. Advantages of our approach: in this figure we compare the behavior of our method with [12] when the geometry is only partially modeled. a) original image from Flickr [11], depicting 2 persons; b) the illumination estimation result when both people are modeled, with our method (illustrated by rendering an orange sundial into the original image with the estimated illumination).; c) illumination estimate with [12] when *only one* of the two people is modeled. The algorithm tries to explain both shadows with one object, resulting in a light source placed under the scene; d) our approach using **the same** 3D model as in (c), when *only one* of the two people is modeled - the illumination estimate is convincing and almost the same as in (b) where full geometry was given. In (e), (f) and (g) we show the 3D model used to estimate illumination in (b),(c) and (d) respectively, rendered with the estimated illumination. Notice that (f) and (g) show **the same** 3D model, but because the estimated light is under the scene in (f), we marked the model with a red outline to make it visible.

method	running time (sec)
[12] (MRF)	244
our method, all samples	4.4
our method, 20% of samples	1.2

Table 2. Running times for our algorithm compared to the state-of-the-art approach of [12], for a 500x300 pixel image. The times *exclude* the time spent raytracing (which is the same for both approaches and can be reduced to less than 1 sec using hardware acceleration).

6. Conclusions

In this paper, we presented an approach to estimate illumination from a subset of shadow borders. The advantages of this approach, as we demonstrated, are that illumination estimation relies much less in the quality of shadow detection, while at the same time allowing the partial and coarse modeling of the 3D geometry of the scene. Our results show that our approach can estimate illumination even

when no shadow detection is performed (Fig.6), since the sets of image edges that match potential shadow outlines are limited. The accuracy of our results is comparable or better than state of the art approaches, as the comparison with the method recently proposed in [12] demonstrates, while achieving relatively low computational complexity, which can be further controlled by limiting the number of edge samples used in our computations as desired. In the future, we would like to examine optimization approaches that can give us guarantees on the optimality of the solution and to examine other ways to build the shadow edge confidence map and the corresponding map of edge orientations, for example through a diffusion process. Another exciting direction for future work would be to associate the matching of the synthetic shadow silhouette with deformable geometric models, allowing some refinement of geometry concurrently with illumination estimation.



Figure 6. Results comparing illumination estimation based on potential shadow edges (left) and all edges in the image (right). It is clear that our method can work even when we do not explicitly detect shadow edges, because few image edges match potential shadow silhouettes. The results of illumination estimation are presented in the bottom by rendering a synthetic orange sundial to the original image, using the estimated illumination. Maps Q are in the top row. This is an image where we failed to obtain any meaningful illumination estimate with the method of [12].

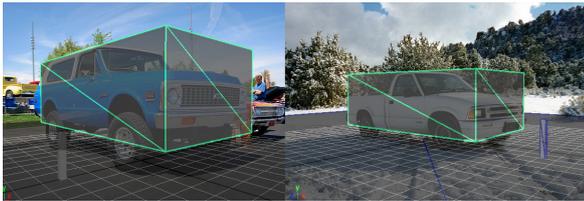


Figure 7. The geometry used to approximate the cars in images from Flickr (from [12]). The geometry consists of a bounding box (green) that encloses the body of the car, and a plane for the ground. Camera parameters were selected by hand to match each scene.

References

- [1] A. Diplaros, T. Gevers, and I. Patras. Combining color and shape information for illumination-viewpoint invariant object recognition. *IEEE Trans. on Image Processing*, 15:1–11, 2006.
- [2] G. Finlayson, M. Drew, and C. Lu. Intrinsic images by entropy minimization. In *ECCV*, 2004.
- [3] G. Finlayson, S. Hordley, C. Lu, and M. Drew. On the removal of shadows from images. *PAMI*, 28(1):59–68, 2006.
- [4] J. M. Geusebroek, R. van den Boomgaard, A. W. M. Smeulders, and H. Geerts. Color invariance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(12):1338–1350, 2001.
- [5] T. Gevers and A. W. M. Smeulders. Color based object recognition. *PR*, 32:453–464, 1999.
- [6] K. Hara, K. Nishino, and K. Ikeuchi. Light source position and reflectance estimation from a single view without the distant illumination assumption. *PAMI*, 27(4):493–505, 2005.
- [7] T. Kim and K. Hong. A practical approach for estimating illumination distribution from shadows using a single image. *IJIST*, 15(2):143–154, 2005.
- [8] J.-F. Lalonde, A. A. Efros, and S. G. Narasimhan. Estimating natural illumination from a single outdoor image. In *ICCV*, 2009.
- [9] J.-F. Lalonde, A. A. Efros, and S. G. Narasimhan. Detecting ground shadows in outdoor consumer photographs. In *European Conference on Computer Vision*, 2010.
- [10] Y. Li, S. Lin, H. Lu, and H.-Y. Shum. Multiple-cue illumination estimation in textured scenes. In *ICCV*, 2003.
- [11] A. Panagopoulos, D. Samaras, and N. Paragios. Robust shadow and illumination estimation using a mixture model. In *CVPR*, 2009.
- [12] A. Panagopoulos, C. Wang, D. Samaras, and N. Paragios. Illumination estimation and cast shadow detection through a higher-order graphical model. In *In CVPR*, pages 673–680, 2011.
- [13] R. Ramamoorthi, M. Koudelka, and P. Belhumeur. A fourier theory for cast shadows. *PAMI*, 27(2):288–295, 2005.
- [14] E. Salvador, A. Cavallaro, and T. Ebrahimi. Cast shadow segmentation using invariant color features. *CVIU*, 95(2):238–259, 2004.
- [15] I. Sato, Y. Sato, and K. Ikeuchi. Illumination from shadows. *PAMI*, 25(3):290–300, 2003.
- [16] J. van de Weijer, T. Gevers, and J.-M. Geusebroek. Edge and corner detection by photometric quasi-invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27, 2005.
- [17] Y. Wang and D. Samaras. Estimation of multiple directional light sources for synthesis of augmented reality images. *Graph. Models*, 65(4):185–205, 2003.
- [18] Y. Yang and A. Yuille. Sources from shading. In *CVPR*, 1991.
- [19] W. Zhou and C. Kambhampettu. A unified framework for scene illuminant estimation. *IVC*, 26(3):415–429, 2008.
- [20] J. Zhu, K. G. G. Samuel, S. Masood, and M. F. Tappen. Learning to recognize shadows in monochromatic natural images. In *CVPR*, 2010.