Volume xx (200y), Number z, pp. 1–12

Multiple light source estimation in a single image

Jorge Lopez-Moreno[1,4], Elena Garces[1], Sunil Hadap[2], Erik Reinhard[3], Diego Gutierrez[1]
[1]Universidad de Zaragoza [2]Adobe Systems Inc [3]MPI for Informatics [4]REVES / INRIA Sophia-Antipolis

Abstract

Many high-level image processing tasks require an estimate of the positions, directions and relative intensities of the light sources that illuminated the depicted scene. In image-based rendering, augmented reality and computer vision, such tasks include matching image contents based on illumination, inserting rendered synthetic objects into a natural image, intrinsic images, shape from shading and image relighting. Yet, accurate and robust illumination estimation, particularly from a single image, is a highly ill-posed problem. In this paper, we present a new method to estimate the illumination in a single image as a combination of achromatic lights with their 3D directions and relative intensities. In contrast to previous methods, we base our azimuth angle estimation on curve fitting and recursive refinement of the number of light sources. Likewise, we present a novel surface normal approximation using an osculating arc for the estimation of zenith angles. By means of a new dataset of ground-truth data and images, we demonstrate that our approach produces more robust and accurate results, and show its versatility through novel applications such as image compositing and analysis.

Categories and Subject Descriptors (according to ACM CCS): I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Photometry

1. Introduction

In this paper we address the difficult problem of robustly estimating the illumination in a natural scene off a single photograph, without the constraints of a prescribed workflow such as the introduction of a physical light probe. We limit our input to a user selected arbitrary object in a single image, which is used as a *virtual* light probe instead. Our goal is to estimate the illumination in terms of a few distinct light sources, along with their directions, and possibly their positions and relative intensities.

However, the task at hand is ill-posed, even in the simplified case when only a single light source illuminates the object we have more unknowns than knowns for each pixel. This is due to the fact that we do not know the geometry of our chosen light probe nor its material properties. In addition, we would like to be able to handle the general case of illumination from multiple light sources.

Such under-constrained problems necessarily lead to approximate solutions. Recent psychophysical experiments, however, have quantified the accuracy with which humans can generally spot flaws in rendered illumination [LMSSG10]. We will show that our method yields valid illumination estimates that remain within those thresholds. Furthermore, we have built a ground truth dataset in order to validate and compare our method with previous and future light source estimation techniques. These methods allow for a wide range of applications related to image compositing, such as image editing and classification, image analysis and augmented reality.

2. Previous Work

The visual effects, animation and games industries have very successfully used light probes to accurately capture the incident lighting in a scene. As part of their standard workflow, the scene is photographed after inserting the light probe at one or multiple key locations. The light probe is a simple calibration object of known size, shape and reflectance properties. For instance, a Lambertian sphere inserted into the scene can be analyzed for estimating directions of multiple light sources [HA93, ZY01]. Further, multiple specular spheres can be used to effectively triangulate the accurate positions of the lights [PSG01, LF06]. A combination of Lambertian and specular spheres can also be employed [ZK02]. The human cornea has also proven to be useable as a light probe [NN04]. Finally, high dynamic range images of specular light probes are successfully used to acquire very detailed illumination environments, which can be used

to render synthetic objects — a technique known as image-based lighting [RGWP*10].

For many workflows, however, placing a light probe into a scene is not practical. If a scene was filmed or photographed without also capturing the illumination by means of a light probe, then compositing and relighting tasks become much more difficult. The solutions for estimating light sources then typically involve making significant and restrictive assumptions about the nature of the scene.

Assuming, for instance, that a given object in the scene is illuminated by a single light source, the analysis is significantly simplified and we can use the object itself to detect the incident lighting. In this case, a local analysis of the surface and image derivatives could be used to estimate the direction of the light source [Pen82, BH85]. Alternatively, occluding contours of a single object [Hor86, NE01, JF07] or even texturing of the object [KP03, VZ04] provide clues as to where the light is coming from. Different visual cues can naturally be combined, for instance to obtain the position of the sun in outdoor scenes [LEN11].

If the geometry of the object in the scene is known, or can be specified with a certain degree of accuracy, the light source positions or directions could be estimated [GHH01, PSP09]. Conversely, to estimate the geometry of an object in a photograph, an ill-posed problem known as shape-from-shading [ZTCS99], we need to know the incident illumination. Further, in either case, the unknown reflectance properties of the object have a significant impact on the results. One way to overcome the under-constrained nature of the problem is to use a range camera to record depth, allowing light sources to be inferred from the combination of the photograph and the range data [MG97]. Known geometry can be used to the same effect [WS02, SSI99, XW08].

Our proposed method is free of the restrictions described above, while additionally allowing us to detect multiple light sources. In particular, there is no need for a calibration object or a subject with a known geometry. To our knowledge, only the work by Lopez-Moreno and colleagues allows for such an unconstrained scenario [LMHRG10]. However, we present several key improvements. First, their method wrongly clusters multiple lights if they appear close to one another. In contrast, we propose a curve-fit optimization technique that identifies these lights as independent sources, yielding more accurate results. Second, our zenith estimation is based on an osculating arc, which is significantly more accurate than their simple ellipsoid approximation. Finally, we introduce a novel albedo removal technique at contour pixels which increases the accuracy of the results. Our method is fairly robust with respect to reflectance properties and albedo variations, and is able to correctly detect multiple lights in complex configurations.

3. Overview

The input of our algorithm is a single image (with unknown camera response) in which the user has marked an object which will act as our light probe. The output of the algorithm will then be one or more light sources, each identified either by a direction or a 3D position. A directional light source is specified in polar coordinates, relative to the image plane: the azimuth ϕ is the rotation in the image plane, whereas zenith ψ indicates the elevation above the image plane. As we cannot compute the actual reflectances or the camera response from such input we will estimate the relative (instead of absolute) strength of each light source.

A first pre-processing step transforms the selected object's luminance values into a valid light probe by removing texture, highlights and noise (Section 4). Then, the subsequent light detection algorithm follows a two-step process. First, the silhouette of the object provides sufficient information to infer the number of light sources, and to detect their position in screen space (Section 5). Second, the object's interior is used to infer the zenith (Section 6).

To perform light detection using a wide range of objects commonly found in images, our method follows the same assumptions as the work by [LMHRG10]: the object is globally convex, the estimated lighting environment consists of an unknown number of white light sources (point or directional) with unknown intensities and finally, the 3D normals at the silhouette of the light probe are assumed to lie in the screen plane [Hor86, NE01, JF07]. These assumptions are required to make light source detection of a single image tractable. Moreover, we argue that this set of assumptions is not unduly restrictive: it allows a large enough set of scenes and objects for our algorithm to be practical.

4. Pre-processing

In order to make the input image more suitable for our purposes, some pre-processing is needed to remove both texture and highlights. There are many different ways to remove texture details from images, such as using bilateral filtering [DD02] or a bilateral Laplacian pyramid [FAR07]. We choose an approach based on bilateral filtering of the luminance signal, due to its simplicity. We follow the approach of Bae et al. [BPD06] to automatically estimate the most suitable kernels for each image. Alternatively, recent advances in intrinsic image decomposition make this a viable option as well. However, very few methods exist that work well from a single image and no user input; we discuss this issue further in Section 7, and provide a preliminary comparison with the recent work by Garces et al [GMLMG12] to highlight the potential of this approach. There are also many different methods to remove highlights [ABC11]. We use the method by Mallick et al. [MZBK06] to remove highlights, producing an object which is roughly diffuse in appearance.

This is similar to the work by Lopez-Moreno et al. [LMHRG10], although here we take advantage of the fact that we know the surface normals at the contours in order to improve texture removal at the silhouette: they lie in the image plane and are orthogonal to the tangent. In this manner, large texture variations (a limitation of bilateral filtering) can

be detected by comparing the difference in luminance with the variation of curvature of the pixels at the silhouette. Intuitively, in a diffuse surface (note that we have removed highlights) a large variation of luminance should be accompanied by a large variation in curvature: should that not happen, then we can assume that the change is due to reflectance variations. We can observe a visual example of this phenomenon in Figure 1. We can observe two large discontinuities (corresponding to the pair of dark tiles) which should not appear: if the reflectance is uniform, neighbouring azimuth values should have a similar luminance.

We identify these uncorrelated variations by segmentation: a new segment is started if there is more than 30% variation with respect to the local average (in a five pixel wide local window along the contour) of the ratio between the angular difference of normals at the silhouette and the difference of luminance values in consecutive pixels. Secondly, we assume that these variations are produced by an albedo factor, which is obtained by subtracting (in log space) the luminance values at the connecting pixels of each pair of neighbouring segments. Finally we add these factors to the pixels at the silhouette, balancing the luminance values (See Figure 1). This method removes most of the texture (and cast shadows) at the silhouette of the light probe, which is important since we assume uniform reflectance in our computations from this point on.

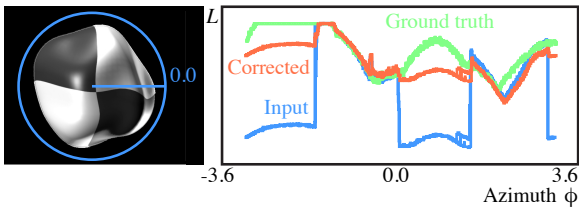

Figure 1: *Example of albedo removal at the silhouette pixels of the input image (a). In (b) we show in blue the luminance at the silhouette before pre-processing, plotted by azimuth angle of the normal at each point. In red, we show how our method reduces albedo discontinuities. Ground truth lambertian data is shown in green. To demonstrate the results of our method, we have not used the bilateral filter before processing the silhouette pixels.*

5. Estimating Lights Sources and Azimuth Angles

For multiple light sources, it would be possible to use known geometry of the surface and to rely on locating *critical points*, which are the points at the boundary of surface areas that are affected by a different combination of lights [ZY01, WS02, BB04]. However, these techniques require the surface geometry to be known. In our case, we only have reliable surface normals at the contour, so that these algorithms are less suitable. Furthermore, these algorithms will detect directional lights only, and may become less effective in the presence of noisy input.

The silhouettes of objects have surface normals that are approximately perpendicular to the viewing direction. It is therefore reasonable to assume that the surface normals of the contour of objects lie in the image plane. This assumption is termed *occluding contours*, and has previously been successfully used to detect light sources [Hor86, NE01, LMHRG10]. The method by Lopez-Moreno et al. [LMHRG10], for instance, is based on a voting scheme and K-means clustering. By its greedy nature, it tends to detect two light sources as a single one if they are separated by less than 90° in azimuth angle. Figure 2 shows a failure case of the method by Lopez-Moreno et al. [LMHRG10]: note how the two top light sources have been collapsed into one. For fair comparison (both at this example and following results), we apply the method by Lopez-Moreno et al. after our preprocessing step, given that both are theoretically exchangeable once texture and highlights are removed.

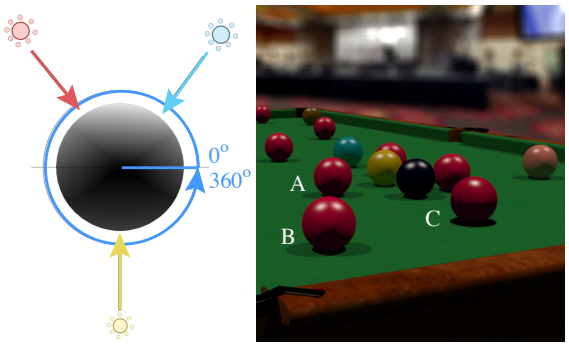

Figure 2: *Left: Description of the illumination in a synthetic scene. Right: Input image showing (A) the ball used as virtual light probe, (B) a ball illuminated by the light sources detected by our method and (C) a wrong result obtained by the method of Lopez-Moreno et al. [LMHRG10].*

To detect separate light sources in such cases, we present a novel method which is able to locally refine each light source, checking if combinations of multiple light sources explain the observed image better than a single larger light source. We will show how such a refinement scheme can reduce the estimation error.

The first step of our method is to light the object with one directional light source and iteratively add more light sources until the process converges to a set of lights that explain approximately the illumination of the light probe (see Section 5.1).

Since the initial number of lights may be under-estimated, for each light source found, the result is refined in two ways: First, we check if the recently added directional light can be replaced with a pair of light sources with directions either side of this light. This process is explained in detail in Section 5.2. Second, we test whether the directional light source would be better explained with a point light source at a finite distance from the object, as discussed in Section 5.3.

The points on the contour are given by $\mathbf{p}_i$. The pixel values can be converted to luminance, indicated by $L_{\mathbf{p}_i}$. Their surface normals are given by $\mathbf{n}_i = [\cos(\phi_i), \sin(\phi_i)]$. Thus, each surface normal could also be represented by azimuthal angle ϕ_i. If multiple pixels share the same surface normal ϕ_i, we will represent this set of pixels with their median luminance value, and therefore run our calculations on fewer pixels. This helps streamline the optimization process. Finally, the total number of pixels on the contour is assumed to be $N_{\mathbf{p}}$, while the number of contour pixels on which calculations are carried out, is given by N_ϕ.

During the estimation process, an estimated light source $\mathbf{l}_k$ is characterized by its direction (defined by azimuth ϕ_k and zenith ψ_k angles) and the amount of light that reaches the object's contour L_k^{in}. After rendering the 3D model of the contour, the current set of N estimated light sources gives rise to a set of N_ϕ pixel luminances L'_{ϕ_i}.

5.1. Finding light source candidates

In this step, the algorithm finds the candidate light sources that best explain the luminance variation along the contour of the object taken as light probe.

After pre-processing, the chosen light probe can be assumed as diffuse and globally convex, and thus the amount of reflected light will depend on the angle between the surface normal and the direction and intensity of the light sources. This energy can be expressed for a set of N light sources $\mathbf{l}_k$ with unknown luminances L_k^{in} and azimuth angles $\mathbf{l}_k^\phi$ as follows:

$$L'_{\phi_i} = K_d^i \sum_{k=1}^{N} L_k^{\text{in}} \cos\left(\mathbf{n}_i^\phi - \mathbf{l}_k^\phi\right) \tag{1}$$

where K_d^i is the albedo of the pixels and $\mathbf{n}_i^\phi$ is the azimuth angle component of the surface normal $\mathbf{n}_i$. Both values are known due to our previous assumptions: $K_d^i = 1$ since the albedo is assumed to be diffuse and the surface normals are given by the contour of the object (see Section 4).

To estimate the light sources we render with Equation (1) an image of the silhouette using each candidate light source and compare it with the observed values. We minimize the following function:

$$O = \underset{\phi_{N+1}, L_{N+1}^{\text{in}}}{\operatorname{argmin}} \sum_{i=1}^{N_\phi} \omega_i^{N+1} \left(L_{\phi_i} - L'_{\phi_i}\right)^2 \tag{2}$$

$$= \underset{\phi_{N+1}, L_{N+1}^{\text{in}}}{\operatorname{argmin}} \sum_{i=1}^{N_\phi} \omega_i^{N+1} \left(L_{\phi_i} - K_d^i \sum_{k=1}^{N+1} L_k^{\text{in}} \cos\left(\mathbf{n}^{\phi_i} - \mathbf{l}_k^\phi\right)\right)^2 \tag{3}$$

where $\omega_i^1 = 1$ for the first light and for the remainder is given by:

$$\omega_i^{N+1} = \frac{1}{2N} \sum_{k=1}^{N} 1 - \cos(\mathbf{n}^{\phi_i} - \mathbf{l}_k^\phi) \tag{4}$$

The weight $\omega_i \in [0,1]$ favors adding new light sources at directions that are maximally different from the directions of existing light sources. This increases the speed of convergence. Note that O is minimized only for azimuth of the new light source ϕ_{N+1}.

Previous approaches rely on least squares optimization, simultaneously exploring for the number of light sources and their properties [ZY01]. This, however, may lead to an incorrect number of light sources and wrongly estimated directions due to poor signal-to-noise ratio (SNR) [BB04]. Since we have fewer samples on the light probe contour than required for least squares optimization, and even poorer SNR due to remaining albedo variations, we opt for a top-down approach, aiming to detect the number of lights ordered by their importance as expressed by the contribution to the luminance of the contour.

To find the best candidate for the first light, that is, the light source with the highest energy contribution to the luminance at the contour L_k^{in}, we minimize the function O (Equation (3)) subject to the following maximum luminance requirement:

$$\max L'_{\phi_i} = \max L_{\phi_i} \qquad \forall\{p_i \mid \mathbf{n}_i^\phi \in [\mathbf{l}_k^\phi - \pi/2, \mathbf{l}_k^\phi + \pi/2]\} \tag{5}$$

The intuition behind this requirement is two-fold. First, the estimated virtual light source should not produce energy values higher than the highest observed values in the image. Second, as our algorithm proceeds in top-down fashion, the first virtual light source should be estimated to have an energy to exactly correspond to the light reflected at the contour. Later refinements can replace this light source with multiple light sources each having less energy. Given this constraint, we can use a simple direct search method (Hooke-Jeeves) optimizing for the azimuth of the light $\mathbf{l}_k^\phi$.

This optimization is a basic building block in our algorithm, allowing us to find optimal azimuthal light source directions given the light probe's contour and a fixed number of light sources. To decide whether more light sources would lead to a better estimate of the observed image, we developed a light source splitting algorithm which is described next.

5.2. Splitting a light source

Several methods have been proposed to detect multiple light sources given a set of known surface normals and their corresponding luminance values [ZY01, WS02, BB04]. Although we lack 3D information at the surface, in our case we have approximated normals at the contour pixels, together with their observed luminances (see Figure 3).

The main challenge is to identify combinations of light sources which affect each set of samples. For instance, in Figure 3 the set of normals (for the sake of clarity now represented by their azimuth angle ϕ) colored in yellow are illuminated by $\mathbf{l}_1$, the ones in blue by $\mathbf{l}_2$ and the green area is illuminated by both $\mathbf{l}_1$ and $\mathbf{l}_2$. In luminance-direction space,

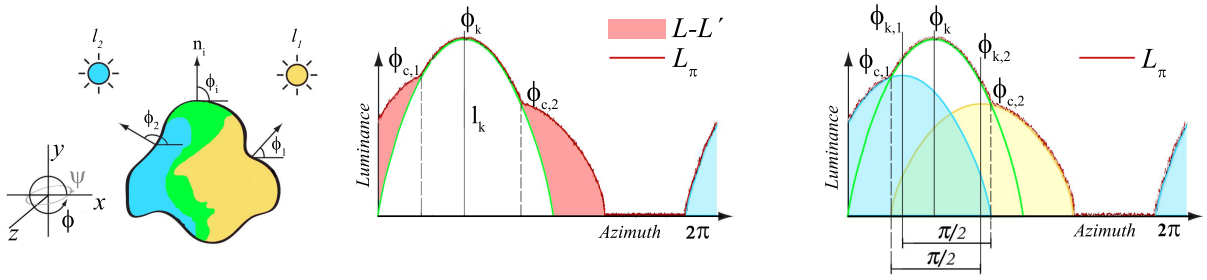

Figure 3: *Left: Object lit by two light sources, also showing the angular coordinate system used in this paper. Middle: Luminance at the contour, plotted by azimuth angle of the normal at each point. First light estimation and error yielded (in red). Note how the critical points are located at the crossing of L and L′. Right: Detected light sources after our splitting approach.*

the boundaries of these groups are characterized by geometrical inflections called critical points (for one dimensional analysis such as the contour in our case) or critical lines (in two dimensions, when a 3D surface is available). In the plots in Figure 3 they are indicated as $\phi_{c,1}$ and $\phi_{c,2}$.

In practice, the detection of these points remains a challenge due to the poor signal-to-noise ratio (SNR) of real image luminances. For instance Wang and Samaras [WS02] combine two geometrical error estimators, a ratio threshold and least squares optimization to find them. Given our even poorer SNR, the limited number of samples and our approximation of the normals at the contour, instead of solving simultaneously for an unknown number of lights with unreliable critical points, we will try to find a geometrical solution and limit the optimization to a close range near this first guess.

Such a solution is possible due to the fact that in a Lambertian model, any observed luminance peak (at ϕ_k) can be fitted by a cosine function within a range of π radians which is produced by either a single light source (defined by that cosine function) or the cosine produced by the intersection of two light sources with cosine functions centered in the range $\phi_k \pm 90$ at $\phi_{k,1}$ and $\phi_{k,2}$ with lower luminance values $L^{\text{in}}_{k,1}$ and $L^{\text{in}}_{k,2}$ (see Figure 3).

By starting with a single light source fitted to the observed peak, the location of the critical points is given by the cross-over points of the observed luminance L and the approximated luminance L' (see Figure 3, middle). Thus, the first critical angle is $\phi_{c,1} = \phi_k - \phi_{\min}$, where $0 < \phi_{\min} < 90$ is the smallest angle for which we have:

$$L^1_{\text{diff}} = L(\phi_k - \phi_{\min}) - L'(\phi_k - \phi_{\min}) > 0 \qquad (6)$$

Moreover, to account for noise we require that in a small neighborhood of directions around this critical point, this difference is larger, i.e.:

$$\int_{-5}^{5} \left(L(\phi_k - \phi_{\min} + \phi) - L'(\phi_k - \phi_{\min} + \phi)\right)\, d\phi > L^1_{\text{diff}} \qquad (7)$$

The second critical angle $\phi_{c,2} = \phi_k + \phi_{\max}$ is found similarly:

$$L^2_{\text{diff}} = L(\phi_k + \phi_{\max}) - L'(\phi_k + \phi_{\max}) > 0 \qquad (8)$$

$$\int_{-5}^{5} \left(L(\phi_k + \phi_{\min} + \phi) - L'(\phi_k + \phi_{\min} + \phi)\right)\, d\phi > L^2_{\text{diff}} \qquad (9)$$

where $0 < \phi_{\max} < 90$ degrees.

If no critical points are found, then the single existing light source explains the data better than if it were split into two separate light sources. If critical points are found, however, then we will estimate two new light sources at angles $\phi_{k,1}$ and $\phi_{k,2}$. Given that the directional influence of a light source is 180 degrees, and that a critical point denotes the boundary where a light source begins to contribute, initial estimates of the angles of the two light sources are given by:

$$\phi_{k,1} = \phi_{c,2} - 90 \qquad (10)$$

$$\phi_{k,2} = \phi_{c,1} + 90 \qquad (11)$$

The luminances associated with these two light sources are then estimated to be:

$$L^{\text{in}}_{k,j} = L_{\phi_{k,j}} = \frac{\cos(\phi_{k,j})}{L(\phi_{k,j})}, \quad \forall j \in [1,2] \qquad (12)$$

Each of the new light sources in this geometrical solution is further refined using Hooke-Jeeves curve-fitting, thus locally optimizing their luminance and azimuth components in a maximum of 10% and 10 degrees respectively from the initial solutions $(L_{\theta_{k,1}}, \phi_{k,1})$ and $(L_{\theta_{k,2}}, \phi_{k,2})$. This typically produces a small correction on the initial estimates, compensating for local noise at a low computational cost. Note that this local optimization does not impose any limitation on the brightness balance or angles between the pair of light sources and the original light.

Finally, we replace the original light source at ϕ_k with these two new light sources if the error ε at direction ϕ_k is less for the pair of new lights than for the original light. This error is computed with:

$$\varepsilon = \left(L_{\phi_k} - \sum_{j=1}^{N+2} L^{\text{in}}_j \cos\left(\phi_k - \phi_j\right)\right)^2 \qquad (13)$$

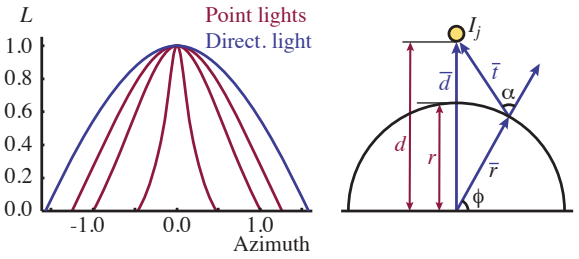

Figure 4: ***Left:*** *Shading created on the contour of the light probe by a directional light source, plotted in angle-luminance space, as well as the shading due to corresponding point light sources at different distances (in number of times the radius of the light probe's bounding circle).* ***Right:*** *Diagram showing the trigonometric ratios between the angles.*

We note that this approach is designed to avoid overestimating the number of lights, which in turn helps reduce noise in the zenith estimation (Section 6). Further, this technique is amenable to an optimization that chooses either a direction or a point light source, as discussed in the following section.

5.3. Detecting point light sources

Each time a directional light $\mathbf{l}_k$ is chosen as a candidate, we check that a near point light source is not a better option. If placed infinitely far away, a point light source behaves as a directional light and its corresponding curve in luminance-azimuth space $L(\phi_i)$ corresponds to a scaled cosine. However, as the point light source gets closer to the object, this curve changes to a Gaussian-like function (Figure 4 left). Let us assume that the light probe is enclosed within a bounding circle of radius r (a value that can be computed from the image) and that the point light source is located a distance d above this circle, a configuration that is shown in Figure 4 (right). Also shown is that for a given angle ϕ we can define intermediate vectors $\bar{r}$, $\bar{d}$ and $\bar{t}$, which have the following relations:

$$\begin{aligned} \bar{d} &= \bar{r} + \bar{t} \\ \bar{r} \cdot \bar{t} &= |\bar{t}||\bar{r}|\cos(\alpha) \\ \bar{r} &= [r\cos(\phi_i), r\sin(\phi_i)] \\ \bar{d} &= [d, 0] \end{aligned} \tag{14}$$

Assuming a Lambertian shading model, we can use these relations to model luminance L as function of angle ϕ as follows:

$$L(\phi_i) = \frac{d \cdot \cos(\phi_i) - r}{\sqrt{r^2 + d^2 - 2d \cdot r \cdot \cos(\phi_i)}} \tag{15}$$

where $\phi_i \in [\phi_k - \pi/2, \phi_k + \pi/2]$. Thus, $L(\phi)$ is a function only of the distance d of the point light, the radius r of the bounding circle and the angle ϕ.

Note that we started by assuming a directional light source of intensity L_k^{in} and angle ϕ_k which is equivalent to a point light source of intensity L_k^{in} and angle ϕ_k placed at an infinite distance d. Using those values as starting points, we optimize the distance d parameter in Equation 15 in order to minimize the error function in Equation 5, by means of the Hooke-Jeeves optimization method. The initial value of d is set to 10^3 times the value of the radius r. Given that distance, the difference of luminance at any pixel on the contour of the light probe between a directional and a point light source is below the error introduced by the compression of the luminance in 8-bit images.

If the directional light assumption is correct the method stops immediately, otherwise after a few iterations the parameter d is estimated. In general the accuracy of the method depends on the size in pixels of our assumed light probe as well as the distance of the point light source and it is limited by the radius in pixels of the light probe: the smaller this radius, the closer the probe has to be to the light source for it to be classified as a point light.

6. Estimating Zenith Angles

For each of the $\mathbf{l}_k$ light sources previously estimated we now compute its corresponding elevation angle (zenith) ψ_k to introduce the third dimension to its direction.

To estimate the zenith angle ψ per light we cannot rely on the contour pixels alone as they are assumed to lie in the screen plane. To overcome this limitation Lopez-Moreno et al. [LMHRG10] locally approximate the geometry of the light probe by an ellipse and analyze the luminance of the shading $S(p_i)$ at the pixels $\{p_i\}$ enclosed by the contour in the original image. Their idea is to find the maxima and minima in the gradient of the luminance; points which provide direct information of the light source's zenith elevation (their surface normals are respectively parallel and perpendicular to the direction of the light). Specifically, for each light source $\mathbf{l}_k$ they march from the silhouette to the interior of the object, following the direction given by ϕ_k. In the presence of multiple light sources this directional derivative of the luminance is the main indicator of the shading due to a particular light aligned to its direction.

Due to the way in which surface normals are approximated in their work [LMHRG10], the zenith estimation is prone to accumulate error for two reasons: first, any arbitrary surface will present considerable local deviations from an ideal ellipse, and second, these deviations or bumps in the surface will present local self-cast shadows, affecting the computation of the zenith angle (see Figure 5). Here, we introduce a curve fitting approach, the *osculating arc* (see Section 6.1) as an alternative for a better normal approximation. In order to reduce the surface estimation error produced by local protuberances or eccentric volumes, we limit our approximation to the area between the contour and the first detected maxima (or minima). Local self-shadows become global at this scale. By relaxing the global convexity assumption this method allows for a wider range of input shapes, even asymmetric ones.

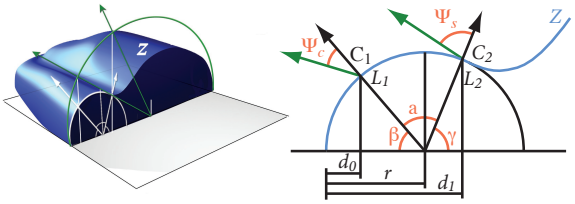

Figure 5: *Approximation of normals at a given scanline of the unknown surface* **Z** *by fitting a circle which has the closest gradient of luminance to the original image between the contour and the second change of curvature (the end of a convex section). In green we show a naive ellipse fitting.*

6.1. Normal Approximation

To compute the zenith angle we need to approximate the surface normal at the points of interest (maxima p^{hi} and minima p^{lo} in the shading $\nabla L_{\mathbf{p}_i}$). To find a valid solution for such an under-constrained problem, we assume convexity near the silhouette, and fit osculating arcs along the scanline of the figure in the image plane (see Figure 5). Each arc (with unknown radius for now) intersects the silhouette at the start of the scanline and fits the curvature of the surface at two points: the first point c_1 represents the first maximum p^{hi} in the luminance gradient for the given scanline (which we can assume has its normal in the direction of the light [WH99]) and it has a luminance value of L_1. The second is the point c_2 where the luminance gradient changes sign and has luminance L_2. The projection of those two points onto the image plane define the distances d_0 and d_1 respectively.

In order to determine the radius r of the semicircle we assume that L, the luminance ratio L_2/L_1 between points c_1 and c_2, is due only to a variation of shading (Lambertian) and thus is directly related to the cosine of the angle between the corresponding normals. Given this assumption and the known values L_1, L_2, d_0 and d_1 from the points c_1 and c_2 we obtain with reference to Figure 5 the following constraints:

$$\begin{aligned} \alpha+\beta+\gamma &= \pi, 0<\alpha<\frac{\pi}{2}, 0<\beta<\frac{\pi}{2} \\ \cos(\alpha) &= L \\ r-d_0 &= r\cos(\beta) \\ d_1-r &= r\cos(\gamma) \end{aligned} \tag{16}$$

where α, β and γ are angles used to establish a trigonometric relation between the radius r and the known values L, d_0 and d_1. From this relation it can be shown that the radius r is then given by:

$$r = \frac{d_0+d_1 \pm \sqrt{2\cdot d_0\cdot d_1\cdot(L+1)}-d_0-d_1}{1-L} \tag{17}$$

This equation yields two solutions as the shading ratio between both points can be achieved with two osculating arcs given the set of constraints. We constrain the solution to the smaller radius r with a center value between d_0 and d_1. In the case of back-lighting we use the point at the silhouette and the minima p^{lo} as fitting points c_1 and c_2 respectively.

The advantages of our approach w.r.t. previous work [LMHRG10], are shown in Figure 5. The normal computed at the first maximum is biased ψ_c by an overestimation of the overall roundness of the light probe. Similarly, with the ellipse fitting, the normal at the first minimum produced by a local self-cast shadow near the silhouette, would be computed with a considerable error ψ_s.

6.2. Grouping lights and luminances

A first estimate of the intensity of each light source L_k^{in} was previously obtained along with the azimuth value (See Equation (12)), however we need to update its value given the detected zenith angle ψ_k according to:

$$L_k^{in} = L_k^{in}\cdot(1+\cos(\psi_k)) \tag{18}$$

Given that zenith information is available at this step, we reconsider our previous estimation of lights derived from the analysis of azimuth values at the contour. Some light configurations (e.g.: frontal light sources aligned with the camera axis) do not provide enough shading information at the silhouette, resulting in an initial overestimation of the number of light sources. To avoid these cases, and in order to derive the simplest possible solution that explains the shading in the image, we perform pairwise comparisons between all the detected candidates in 3D; for each pair of light directions on a plane, we collapse them into one direction if the inner angle is less than 15°. Given the average signal to noise ratio (SNR) that we observed in our experiments, there is no guarantee that this pair of sources are, indeed, two different lights, however with better image inputs (HDR, better texture removal) this threshold could be reduced or even removed. The zenith angles are averaged and their intensities are re-computed by using Equation (18) and the new zenith value. Note that previous splitting was done in azimuth space, and only when zenith values have being computed, we can refine the final number of light sources.

After all light sources are detected, we add a final term to take into account ambient illumination. Its light contribution is assumed to be constant for all pixels and we simply approximate its intensity by analyzing pixels in the shadow regions (note that we have already detected shadow edges when looking for minima in the zenith estimation, from which shadow regions can trivially be estimated). We average the set of samples along these boundaries. We cannot rely on the regions contained by them as they cannot be assumed to be fully covered in shadows (e.g. an extruding bump in the middle of a shadowed area can be brightly lit while its surroundings are not). This ambient intensity estimate is also relative to the previously detected lights.

7. Results

We have tested our algorithm on photographs, captured under known illuminations (measured with a mirror sphere), using multiple objects with significantly different BRDFs as light probes (see Figure 6). Our first test consisted of seven

Figure 6: *Some input images used in our tests (see website for full dataset). The top row shows different illuminations for one object. Middle and bottom rows show the different objects used as input. Bottom-right corner: Ground truth mirror spheres for our single (five positions) and two-light sources tests (right and left images respectively).*

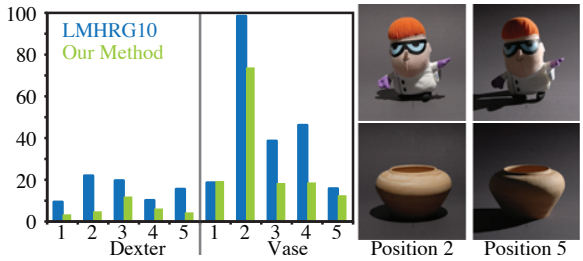

Figure 8: *Comparison of errors between two different objects:* Dexter*, with ideal geometry and* Vase*, which violates our assumption of global convexity. The plot represents the average error for one directional light comparing our work with previous work [LMHRG10]. Right, examples of input images corresponding to the configuration of lights* Position2 *and* Position5*, which are, respectively, the worst and the best case for the error of the* Vase*. It is clear on the plots that the overall error increases when the object does not hold our assumption.*

objects illuminated by a single light source from five different positions. The plot on Figure 7, left, shows that our method obtains an error below 20 degrees, being considerably smaller than the previous method by Lopez-Moreno et al. [LMHRG10]. The small standard deviations of our method confirm its robustness under different objects as inputs (precise errors for each object can be found in the additional materials).

Exceptionally, *Position2* presents a higher degree of error than the rest of the objects. The reason is the following: the object *Vase* of our dataset violates our assumption of global convexity. Hence, when the object is lit by an upper light (*Position2*), the unexpected luminance values produced by the concavity of the hole produce a strong bias for the analysis in that direction. Moreover, the lower half of the object is rather cylindrical, capturing the light sources only in one angle. Figure 8 shows this fact comparing an ideal globally convex object like *Dexter* with the non-convex one, *Vase*. As we can see in the plot, the overall error is much larger for *Vase* in all the positions.

Additionally, a secondary light was detected by both methods in several cases due to light bouncing from the ground. It was identified (by its direction and significantly lower intensity values) and discarded in our plots. Note that although the detection of secondary sources is not wrong per se, these are not calibrated for error measure (e.g.:direction, intensity). The bias introduced at the estimation of the main light sources in this test is almost negligible, as both methods start by fitting high energy light sources, while leaving the low energy values for the end.

In our second experiment, we analyzed how the presence of a second light source affects the accuracy of the light detection. Four objects from the previous test were selected and illuminated by two fixed lights at five different configurations. In each configuration, we changed the relative intensity of the two lights, for example, *Config2* has a ratio of 0.5:1 which means that one light is half as bright as the other (specific ratios are shown in the upper table of Figure 7, middle). We measured both error of light position and error of luminance intensity for each configuration. The first is shown in Figure 7, middle, and although the results show somewhat lower accuracy, previous studies suggest that this overall error level is below human perceptual threshold for illumination inconsistencies [LMSSG10]. We note that as the configuration tested with these images could be considered a best-case scenario for the method by Lopez-Moreno et al. (low degree of overlap), the current results represent an improvement. Finally, although both methods compute the intensity values of the light sources in a similar fashion, the greater accuracy of our direction estimation reduces the average error of the luminance intensity, being 27.72% for the method by [LMSSG10] and 14.75% for our algorithm (detailed error plot in the additional materials).

In the third experiment, we have analyzed and compared with previous work [LMHRG10] six spatial combinations of three light sources for three different objects (see Figure 7). We note that our method shows a better performance on average, as our method does not tend to cluster neighbouring light sources (See Figure 2). Also note that, in general, our method predicts the majority of the lights while the previous method tends to collapse two lights with less than 90 degrees of azimuthal difference. See additional material for an exhaustive analysis of the error.

In Section 5.3, we introduced an optimization process for point light sources. Although in practice, due to distance, indirect light, quantization of luminance levels and limited dynamic range of the cameras, most point light sources are difficult to distinguish from a directional light source, with our method we can deal with extreme cases such as those in Figure 9. In this test we show that our algorithm is able to

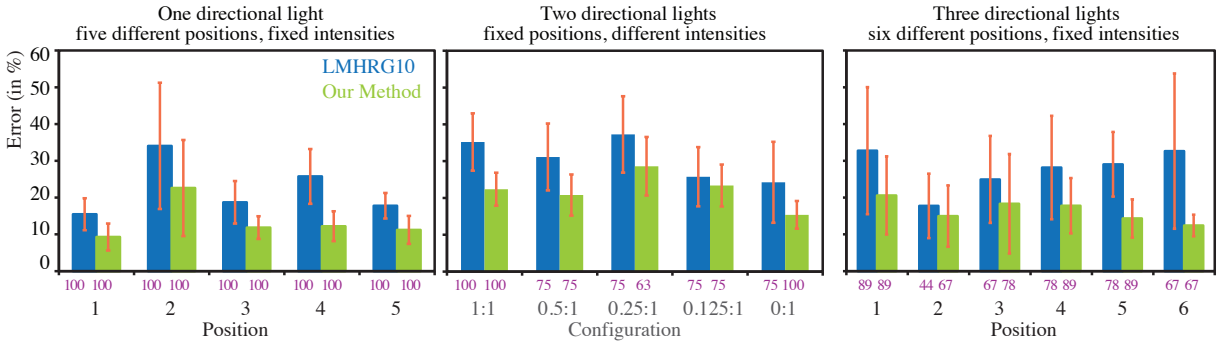

Figure 7: *Average error and standard deviation (red lines) obtained by our method and the previous algorithm by [LMHRG10]. The number located at the bottom of the bars represent the percentage of lights found by the method w.r.t the ground truth number of lights. Left, one directional light from five different positions and fixed intensity. Middle, two directional lights at fixed positions but different intensities. Intensity values are defined in the top table of the plot by their ratios e.g in* Config2 *one light is half bright the other which means a ratio of 0.5:1. Right, three directional lights at six different positions and fixed intensity. Exact positions of the lights and images can be found at the webpage.*

estimate the relative distance of the light source to the object at five different positions.

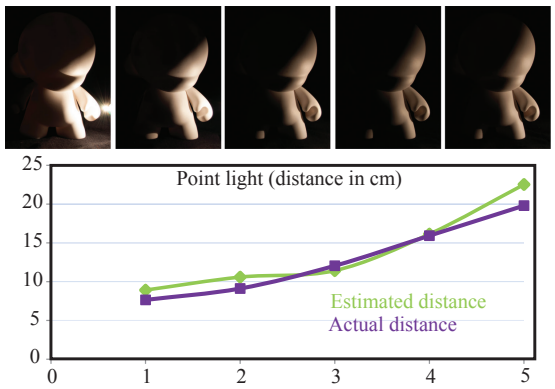

Figure 9: *Top row: Input images. Photographs acquired with a point light source at a distance of 7.64, 9.118, 12.05, 15.92 and 19.8 cm from the head of the mannequin. Bottom row: estimated distances for the five images: 8.90, 10.58, 11.42, 16.13 and 22.52 cm.*

Figures 10 and 11 show an application of our method to image compositing: in both cases we use an object from the target image as light probe, and relight the composited object with the estimated illumination. For relighting purposes we estimate a plausible depth map using the approach by Khan et al. [KRFB06], who approximate it by analyzing shading variations. Although there are no limitations on the complexity of the reflectance models employed, we use a simple combination of Lambert and Phong's models to represent surface reflectance. We note that Figure 10 shows a particularly difficult example, given the Scottish quilt which was used to detect lighting directions has a spatially varying albedo. Nevertheless, the final composited result is visually

Figure 10: *Top left: Original background image with a fluffy toy super-imposed. As the lighting on the toy is not corrected, the result looks out-of-place. Top right: A light probe rendered with colored light sources derived from the detected directions, a normal map recovered from the toy image, and the resulting crab relit with the recovered illumination. Bottom: Final image with the toy coherently integrated in the image.*

plausible, and a significant improvement over naïve insertion of the toy into the scene.

Our method has potential as an image analysis tool, helping to classify areas of images depending of their estimated light sources (or even spot lighting inconsistencies). In Figure 12, we show a side-by-side comparison between an input image (left) and the result of running our algorithm in three different objects of that image and relighting those objects according to our estimation (right). The light configurations

Figure 11: *A result of compositing images through relighting by using the information from our light detector. The soldier and one of the elephants were relighted (with the light sources detected in the Venetian mask and the wooden mannequin) and inserted in the image. See additional materials for the detailed process.*

Figure 12: *Top Left: Input image. Top right: Result of relighting the three toys with their estimated light environments. The image was obtained by re-capturing those objects with new light configurations and compositing them on top of the input image. Bottom: Pixel areas selected as light probes (highlighted in white).*

of each light probe are slightly different due to local interreflections and occlusions, but we can observe their consistency with the global light environment in the image of the right. Note that the relighted objects present slightly darker shadows. This is partially due to the impossibility of adding the estimated ambient luminance with this relighting technique (0.24, 0.28 and 0.22 w.r.t. the main lights in spaceman, cow and monkey toys respectively).

Analysis of the influence of texture

In the above tests with photographs we obtained similar error values for very different illumination, albedo and geometry configurations. This leaves the question as to why both simple an complex shapes with both easy-to-decompose and difficult reflectance yield such similar error values for the estimated light sources? To analyze the influence of reflectance and surface complexity, while keeping the dimensionality of the experiment at a manageable yet meaningful scale, we have created a standard three-light source illumination configuration (quite ubiquitous in photography: key light, fill light and rim light). For the purpose of exercising control over the degree of surface complexity, we have generated a set of four synthetic models with a combination of fractal and Gaussian noise at different spatial scales and a set of four textures with varying spatial frequencies.

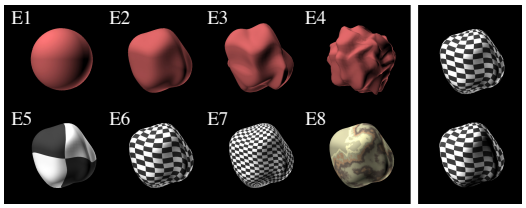

Figure 13: *Left image: Input objects used as light probes by our method to estimate the light source (see additional materials for the table with the numerical results). Light sources are located at* $(90°,0°)$*,*$(180°,45°)$ *and* $(315°,-45°)$*.* ***Top row****: From left to right, increasing complexity of the surface (E1-E4).* ***Bottom row****: Multiple reflectance variations generated with textures with different spatial frequencies (E5-E8). The rightmost column shows object E6 (top) and the result of rendering E6 with the light detected by our method (bottom).*

In the top row of Figure 13, we can observe how increasing the complexity of the object's geometry does not alter the accuracy of the method showing an azimuth and zenith average error of 8.55 and 8.84 degrees with a standard deviation of 3.76 and 4.79 degrees (in additional materials we show azimuth and zenith angle errors separately). This result is consistent with our previous results for photographic images. Our osculating arc fitting at the contour on the other hand, makes a weaker assumption about the geometry of the light probe, thus introducing less error on average than previous work [LMHRG10] (See Figure 7). The rightmost column shows how a failure case (three lights were detected as four by our method) still yields perceptually equivalent results.

We obtain similar results for multiple reflectance variations in the object, shown in the bottom row of Figure 13: an azimuth and zenith average error of 18.72 and 12.55 degrees respectively with a standard deviation of 10.43 and 8.53 degrees. These four textures were selected for their poor behaviour in bilateral filtering decomposition (due to high contrast), especially if the kernels are not perfectly chosen (our method uses an automatic kernel estimator based on image statistics, see Section 4). As expected, the method shows poor results (in the range of 20-30 degrees), yet there is no apparent correlation with the increase of spatial frequencies in the texture. The effectiveness of the method depends mainly on the right (automatic) choice of kernels in the bilateral filter decomposition. As expected, increasing the complexity of the surface has no impact on the zenith or azimuth

estimation and increasing the spatial frequency of the surface reflectance has no effect in the quality of the result.

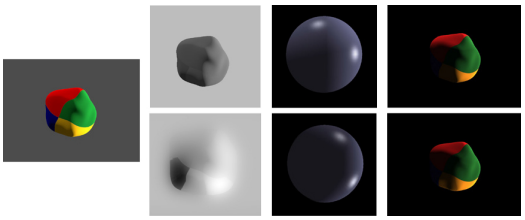

Figure 14: *From left to right: Input synthetic image generated with three light sources at (0,0),(90,0) and (180,45) degrees, shading extracted with intrinsic image decomposition [GMLMG12] (top) and bilateral filtering (bottom), virtual light probe with our detected light sources and relighting of the input image with the estimated light sources.*

In Figure 14 we show a comparison between our texture removal (partially based in bilateral filtering) and the intrinsic image decomposition by Garces et al. [GMLMG12]. In this case, although our light detection method fails with both approaches in capturing the fill ambient light (intensity of 0.11), it yields plausible results for the two directional lights, being slightly more accurate with the intrinsic image decomposition. Although it is still unclear which algorithm is the most robust for a wider set of scenes (e.g.:the method by Garces et al. is not adapted to low chromaticity images), novel approaches in intrinsic image decomposition [SY11, BPD09] could potentially reduce the influence of texture, improving the quality of the input and minimizing the error in our method.

Further, we have tested the effect of an area light source on our method (See additional materials). In such cases, our optimization method tries to find the minimum number of directional light sources that minimize the differences w.r.t. the shading of the light probe. In this case our algorithm approximates the solution with two light sources at varying distances from one another depending on the size of the area source. This is to be expected, as a greater light area will require a broader angular coverage by the corresponding directional sources.

8. Discussion and Future Work

Detecting how a scene was illuminated based on a single photograph is a valuable tool for many image processing tasks, including compositing and scene reconstruction. We have introduced a novel algorithm that estimates illumination strength, directions and positions for scenes containing multiple light sources. This is achieved with higher accuracy than the current state-of-the-art, while alleviating some of the restrictions as to which objects in a scene can be used as a light probe.

We have provided insight into the albedo removal process of our method, showing that even when our assumptions about constant albedo are broken (e.g.: a black and white checkerboard texture), most of the times the results are less accurate but still plausible.

Our method relies in the user selection of an appropriate silhouette. In general, unless background pixels are selected, the method is robust against small deviations from the actual silhouette. The user criteria should be to select an object as globally convex as possible. For instance if you select a cylindrical object, the method is missing all the information from one axis. Additionally, objects which have large normal discontinuities at the contour (e.g. a cube in isometric view) should be avoided, as they violate our "normals lie in the screen plane" assumption, thus reducing the accuracy of the detection (an example of this influence is available at the supplementary material). On the other hand, strong albedo variations such as a head with dark hair, are still suitable if there is enough dynamic range and luminance information in dark areas for the reflectance pre-processing.

Although it is an odd scenario, as most of the contour-based approaches, we rely on the contour values to set the initial number of light sources, therefore a limitation of our method is that multiple light sources with the same azimuth angle will be combined and detected as one. Future approaches could consider a splitting optimization method (see Section 5.2) to account for such light sources in the zenith estimation step.

Our method yields absolute errors usually below 20-30 degrees in light direction as suggested by Lopez-Moreno et al. [LMSSG10] in their study of human perception sensitivity to illumination inconsistencies. It improves the results of previous works in the area with similar computational costs. Specifically, the work by Lopez-Moreno et al. [LMHRG10] presents lower accuracy in general, while wrongly collapsing close lights into single sources due to the greedy nature of their algorithm. In contrast, our splitting process yields correct results in these cases. The results obtained by our osculating arc approach suggest that there is no need to approximate the whole surface of the light probe and future research should include novel partial surface approximation techniques.

We have provided multiple validation tests with controlled lighting environments. Furthermore, we have created a dataset of images and their corresponding ground truth data in the hope that future light analysis methods might benefit from them. In our opinion, there is a need for such dataset, as there is a lack of photographs of common objects coupled with their corresponding (measured) illumination environments. The light probes in our database include a wide range of geometric and reflectance properties. This dataset is publicly available at `http://www.jorg3.com/2013/lightdetection/`. The images were taken under the following lighting configurations: with one light at five different positions, two lights with five different brightness balances, three lights with six different positions and two area

lights. We provide the corresponding ground truth in the form of mirror balls plus brightness levels.

References

[ABC11] ARTUSI A., BANTERLE F., CHETVERIKOV D.: A survey of specularity removal methods. *Computer Graphics Forum 30*, 8 (December 2011), 2208–Ű2230.

[BB04] BOUGANIS C.-S., BROOKES M.: Multiple light source detection. *IEEE Trans. Pattern Anal. Mach. Intell. 26*, 4 (2004), 509–514.

[BH85] BROOKS M., HORN B.: Shape and source from shading. In *Proc. Int. Joint Conf. Artificial Intell.* (1985), pp. 932–936.

[BPD06] BAE S., PARIS S., DURAND F.: Two-scale tone management for photographic look. *ACM Trans. Graph. 25*, 3 (2006), 637–645.

[BPD09] BOUSSEAU A., PARIS S., DURAND F.: User assisted intrinsic images. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia 2009) 28*, 5 (2009).

[DD02] DURAND F., DORSEY J.: Fast bilateral filtering for the display of high-dynamic-range images. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 2002), ACM, pp. 257–266.

[FAR07] FATTAL R., AGRAWALA M., RUSINKIEWICZ S.: Multiscale shape and detail enhancement from multi-light image collections. *ACM Transactions on Graphics (Proc. SIGGRAPH) 26*, 3 (Aug. 2007).

[GHH01] GIBSON S., HOWARD T., HUBBOLD R.: Flexible image-based photometric reconstruction using virtual light sources. *Computer Graphics Forum 19*, 3 (2001), 203–214.

[GMLMG12] GARCES E., MUNOZ A., LOPEZ-MORENO J., GUTIERREZ D.: Intrinsic images by clustering. *Computer Graphics Forum (Proc. EGSR 2012) 31*, 4 (2012).

[HA93] HOUGEN D., AHUJA N.: Estimation of the light source distribution and its use in integrated shape recovery from stereo shading. In *ICCV* (1993), pp. 29–34.

[Hor86] HORN B.: *Robot Vision.* McGraw-Hill, 1986.

[JF07] JOHNSON M. K., FARID H.: Exposing digital forgeries in complex lighting environments. *IEEE Transactions on Information Forensics and Security 2*, 3 (2007), 450–461.

[KP03] KOENDERINK J. J., PONT S. C.: Irradiation direction from texture. *Journal of the Optical Society of America 20*, 10 (2003), 1875Ű–1882.

[KRFB06] KHAN E. A., REINHARD E., FLEMING R., BÜLTHOFF H.: Image-based material editing. *ACM Transactions on Graphics (SIGGRAPH 2006) 25*, 3 (2006), 654–663.

[LEN11] LALONDE J.-F., EFROS A. A., NARASIMHAN S. G.: Estimating natural illumination from a single outdoor image. *International Journal on Computer Vision, to appear* (2011).

[LF06] LAGGER P., FUA P.: Using specularities to recover multiple light sources in the presence of texture. In *ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition* (2006), IEEE Computer Society, pp. 587–590.

[LMHRG10] LOPEZ-MORENO J., HADAP S., REINHARD E., GUTIERREZ D.: Compositing images through light source detection. *Computers & Graphics* (2010), 698 – 707.

[LMSSG10] LOPEZ-MORENO J., SUNDSTEDT V., SANGORRIN F., GUTIERREZ D.: Measuring the perception of light inconsistencies. In *APGV '10: Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization* (2010), ACM, pp. 25–32.

[MG97] MARSCHNER S. R., GREENBERG D. P.: Inverse lighting for photography. In *Fifth IST/SID Color Imaging Conference* (1997), pp. 262–265.

[MZBK06] MALLICK S. P., ZICKLER T., BELHUMEUR P. N., KRIEGMAN D. J.: Specularity removal in images and videos: A PDE approach. In *In Proc. of ECCV* (2006), pp. 550–563.

[NE01] NILLIUS P., EKLUNDH J.-O.: Automatic estimation of the projected light source direction. In *CVPR* (2001), pp. I:1076–1083.

[NN04] NISHINO K., NAYAR S. K.: Eyes for relighting. *ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH) 23*, 3 (2004), 704–711.

[Pen82] PENTLAND A.: Finding the illuminant direction. *Journal of the Optical Society of America A 72*, 4 (1982), 448–455.

[PSG01] POWELL M., SARKAR S., GOLDGOF D.: A simple strategy for calibrating the geometry of light sources. *IEEE Transactions on Pattern Analysis and Machine Intelligence 23*, 9 (2001), 1022–1027.

[PSP09] PANAGOPOULOS A., SAMARAS D., PARAGIOS N.: Robust shadow and illumination estimation using a mixture model. In *CVPR* (2009), pp. 651–658.

[RGWP*10] REINHARD E., GREG WARD P. D., PATTANAIK S., HEIDRICH W., MYSZKOWSKI K.: *High Dynamic Range Imaging.* Morgan Kaufmann Publishers, 2010, ch. 11, pp. 501–593.

[SSI99] SATO I., SATO Y., IKEUCHI K.: Illumination distribution from brightness in shadows: Adaptive estimation of illumination distribution with unknown reflectance properties in shadow regions. In *ICCV (2)* (1999), pp. 875–882.

[SY11] SHEN L., YEO C.: Intrinsic images decomposition using a local and global sparse representation of reflectance. In *Computer Vision and Pattern Recognition* (2011), IEEE.

[VZ04] VARMA M., ZISSERMAN A.: Estimating illumination direction from textured images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Washington, DC* (June 2004), vol. 1, pp. 179–186.

[WH99] WORTHINGTON P. L., HANCOCK E. R.: New constraints on data-closeness and needle map consistency for shape-from-shading. *IEEE Trans. Pattern Anal. Mach. Intell. 21*, 12 (Dec. 1999), 1250–1267.

[WS02] WANG Y., SAMARAS D.: Estimation of multiple illuminants from a single image of arbritary known geometry. In *ECCV02* (2002), vol. 3, pp. 272–288.

[XW08] XU S., WALLACE A. M.: Recovering surface reflectance and multiple light locations and intensities from image data. *Pattern Recogn. Lett. 29*, 11 (2008), 1639–1647.

[ZK02] ZHOU W., KAMBHAMETTU C.: Estimation of illuminant direction and intensity of multiple light sources. In *ECCV '02: Proceedings of the 7th European Conference on Computer Vision-Part IV* (2002), pp. 206–220.

[ZTCS99] ZHANG R., TSAI P., CRYER J., SHAH M.: Shape from shading: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence 28*, 8 (1999), 690–706.

[ZY01] ZHANG Y., YANG Y.-H.: Multiple illuminant direction detection with application to image synthesis. *IEEE Trans. Pattern Anal. Mach. Intell. 23*, 8 (2001), 915–920.