Using Specularities to Recover Multiple Light Sources in the Presence of Texture

Pascal Lagger
Computer Vision Laboratory
EPFL, Switzerland
pascal.lagger@epfl.ch

Pascal Fua
Computer Vision Laboratory
EPFL, Switzerland
pascal.fua@epfl.ch

Abstract

Recovering multiple point light sources from a sparse set of photographs in which objects of unknown texture can move is challenging. This is because both diffuse and specular reflections appear to slide across surfaces. What is seldom demonstrated, however, is that it can be taken advantage of to address the light-source recovery problem. In this paper, we therefore show that, if 3D models of the moving objects are available or can be computed from the images, we can solve the problem without any a priori constraints on the number of sources, on their color, or on the surface albedos.

Our approach involves finding local maxima in individual images, checking them for consistency across images, retaining the apparently specular ones, and having them vote in a Hough-like scheme for potential light source directions. The precise directions of the sources and their relative power are then obtained by optimizing a standard lighting model. As a byproduct we also obtain an estimate of various material parameters such as the unlighted texture and specular properties.

We show that the resulting algorithm can operate in presence of arbitrary textures and an unknown number of light sources of possibly different unknown colors. We also estimate its accuracy using ground-truth data.

1. Introduction

Let us consider the following scenario. We are given a set of images acquired by a fixed camera of a scene in which rigid objects may move between exposures. Assuming that no additional information is provided, we can use commercially available software packages to reconstruct the geometry of the moving objects and to estimate their 3D poses. By contrast, recovering the number and direction of the possibly multiple light sources and the object textures remains a difficult problem. This is because, for any given surface patch projected in several images, none of the photometric angles are preserved between views. This implies that both diffuse and specular reflections move across the surface.

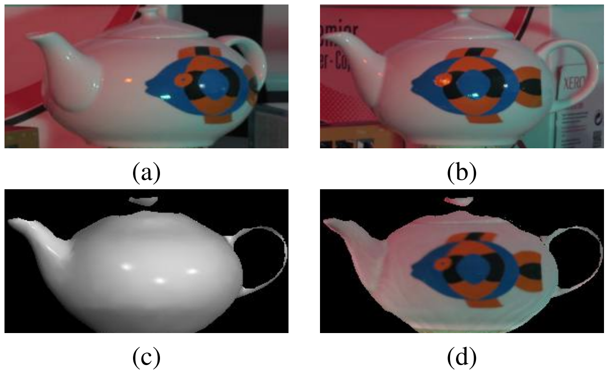


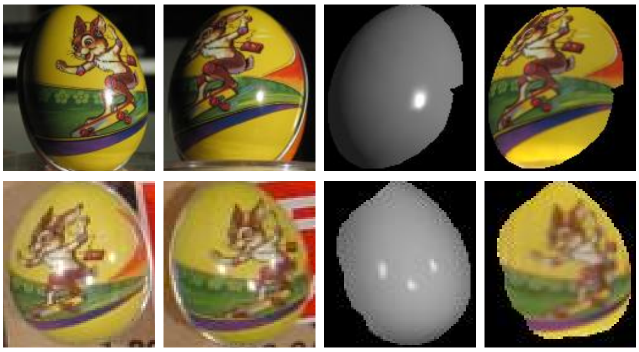
Figure 1. Recovering number of sources and their directions in a scene lighted by three different light sources of unknown and different colors. To see this, we suggest that the reader view these images in color. (a,b) Two of nine input photographs. (c) Shaded view of the 3D model of specular artifacts.

The above scenario is highly relevant to many real-world applications such as delighting and religthing scene objects or adding new objects that blend correctly in the scene. For example, in the remastered versions of the *Star Wars* movie, new synthetic ships and robots were added to scenes that had been filmed a long time ago using real props [1]. Since both the original props and lighting parameters are probably long lost, a technique able to recover those parameters automatically from the old images of the moving props could have saved a great deal of time and money.

In this paper, we present an approach to recovering multiple point light sources out of a sparse set of registered photographs, provided that 3D models of some moving scene objects are available or can be computed. Our approach involves finding local gray-level maxima in several images and deciding whether or not they represent specularities. The detected specularities are then used to count how many point light sources are in the scene and estimate their location. The precise directions of the sources, their relative power and their color are then obtained by optimizing a standard lighting model. As a byproduct, we also obtain an estimate of various material parameters such as the unlighted texture and specular properties.

It is well known that lighting maxima glide across the surface of a moving object according to the normals, whereas texture maxima tend to be detected at the same surface location in many images. However, this has not been explicitly exploited by any previously published technique we are aware of. The originality of our method is therefore to take advantage of this physical fact to distinguish between lighting and texture maxima and to implement a light-source detection scheme that relies upon consistency across images. Although, the current algorithm is not designed to deal with very extended light sources, it *is* designed to handle sources that are not at infinity and are not truly point-light-sources.

As shown in Figs. 1 and 2, the resulting method can operate in presence of arbitrary textures and multiple light sources of possibly different unknown colors. This unknown chromaticity completely rules out the use of methods based on the dichromatic model [3, 2, 8, 9]. Similarly there are almost no shadows and no obviously visible critical boundaries, which would surely handicap methods [11, 4] that require them. In short, given a set of images containing moving objects, our method makes it possible to recover the lighting parameters under far less restrictive assumptions than state-of-the-art methods.



(a) (b) (c) (d)

Figure 2. Lighting parameters recovery in presence of texture. The first row shows the recovery for one light source from *high* resolution images. The second row shows the recovery for three light sources from *low* resolution images. (a,b) Two out of nine input images. (c) Using the recovered parameters and the perspective of (b) yield a shaded view in which the specularity is very similar to the true one. (d) The recovered albedo map is free of specular artifacts.

The limitations are that there must be specularities and that we require multiple images and a 3D model, which does not however have to be extremely accurate. For many applications, this is not particularly onerous given the prevalence of video cameras and the growing number of structure-from-motion techniques, both manual and automated, that can be used to build the required models. Here, the fact that our approach can handle texture is essential since many of these geometric reconstruction techniques rely on it.

In the remainder of the paper, we first review some of the most representative approaches to light source recovery that have been proposed with a view to understanding what the kind of assumptions they make. We then present the light source estimation algorithm and describe the refinement process. The results section shows the ability of the method to handle multiple light source recovery and discuss its accuracy.

2. Related Work

Specularities have bedeviled Computer Vision algorithms for many years. Many researchers have attempted either to handle them by introducing calibration objects or controlling the image acquisition process, or to eliminate them by treating them as statistical outliers. Others have attempted to actually use them as information sources but often make very strong assumptions, as discussed below.

Among methods that require a single image, chromaticity-based ones [3, 2, 8, 9] are among the most popular. However, they typically require the material chromaticity to be different from the lighting chromaticity and all sources to be of the same color, except in very specific cases where it is possible to segment the contributions of individual sources. An alternative to using color is to rely on shadows and critical boundaries [11, 4] but they are not necessarily easy to find in the presence of texture.

When a static scene can be filmed using a moving camera, specularities have been successfully used to find the light sources [5, 13]. In that case the lighting maxima do *not* glide across the surface and the problem is comparatively easier than when objects can move. The detection is based on the fact that the intensity variations between images can only be caused by specularities. More sophisticated models that explicitly represent specularities have been proposed. For example, the signal-processing framework introduced in [7] describes the reflection operator as a convolution and formulates the recovery of a BRDF and lighting distribution as a deconvolution. In addition to giving a mathematical tool to analyze the well-posedness and conditioning of inverse-rendering problems, the method can recover complete BRDFs, thus handling both specular materials and the lighting environment. It has, however, only been demonstrated in cases where the albedo is either constant or known *a priori*, which is a severe limitation.

3. Method

It is generally accepted that, once the specularities have been detected, counting and positioning the light sources is comparatively easy. However, this detection is often hard and the simple thresholding techniques that are often reported in the literature cannot be expected to work very well

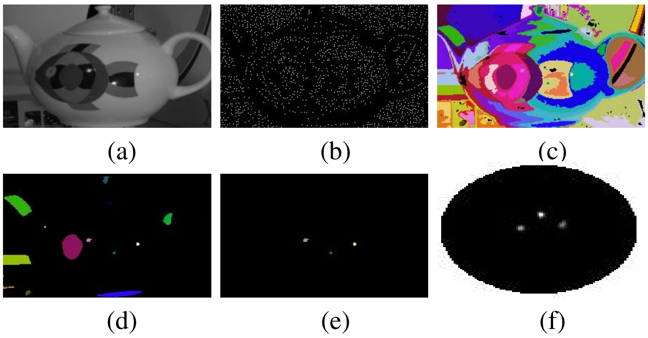


Figure 3. Specularity detection workflow. (a) The input image is converted to grayscale. (b) The maxima are detected and neighboring ones are merged. (c) The maxima are grown by including neighboring pixels of lower intensity to form regions. (d) Only the region that are brighter than their neighbors are retained. (e) Enforcing consistency across images lets us discard non-specular regions. (f) The remaining specular regions vote in an accumulator yielding maxima that corresponds to the number and direction of the light sources.

in general. For example, in a textured scene lit by multiple light-sources, a diffuse area of large albedo can be brighter than a specular one of smaller albedo. In such a case, no simple threshold can be used to extract specular gray-level maxima while eliminating others.

We overcome this problem by explicitly using the fact that specularities glide over moving surfaces in a predictable fashion. This lets us distinguish between non-specular maxima and specular ones and take advantage of the later to count the light-sources and to estimate the lighting parameters as follows.

3.1. Detecting Specularities and Positioning the Light Sources.

This involves the following three-step algorithm:

1. In each greyscale converted image, we select all the pixels of greater or equal intensity than their eight-connected neighbors. Neighboring local maxima of equal intensity are merged. To estimate the local support of these maxima, Fig. 3(b), we sequentially consider them, starting with the brightest ones. We grow areas around them by including all the neighboring pixels whose difference in intensity with the corresponding maximum remains below a predefined value. This local thresholding adapts to the gray level of the maxima and absorb local texture maxima. This results in a set of regions such as those depicted by Fig. 3(c). Among these regions, we only retain those that correspond to peaks of higher intensity than those of their neighbors, such as those of Fig. 3(d). The remaining maxima are the ones that will be used as input to the voting scheme described below.

2. To classify the maximal intensity areas as either texture or lighting maxima, we compute surface normal and line of sight for each corresponding pixel. If a maximum region, instead of moving with respect to the mesh, stays at the same mesh location even though the object rotates, the algorithm labels it a texture maximum.

3. Specularities are located at places where an object surface behaves like a mirror so that the viewer sees a reflection of the light source. Therefore, for each potentially specular pixel that have passed the consistency test, we compute the reflected line of sight mirrored off the 3D model. We express the line's orientation in terms of its (ϕ, θ) spherical coordinates, which we use to increment a 2D orientation accumulator such as the one of Fig. 3(f). The resulting number of peaks in the accumulator gives us the number of sources and an estimate of the corresponding light source's orientation, Fig. 3(f).

3.2. Refining the Light Source Estimates.

Given these lighting parameters estimates, we refine the directions of the sources and their relative powers by fitting the Ward [12] lighting model to gray levels at all locations in all images where the projected 3D Model gives us normal vectors. In addition, this yields values for the albedos and the specular coefficients. The lighting distribution and the material parameters are separatively optimized in an iterative loop. We chose the Ward model because it is one of the simplest model that, unlike the Phong model [6], is based on physical measures. An alternative would have been to use an even more accurate model, such as the Torrance-Sparrow one [10]. However, such models have far more parameters, making them more difficult to instantiate with the imagery at our disposal.

Our specularity detection method is designed to handle small area light sources, which are neither perfect point light sources nor at infinity. Note that we do not make assumptions about the presence or absence of texture and do not require any foreknowledge of the albedos. Because we can use grayscale images, our method is insensitive to the fact that the light sources may be of different and unknown colors, which sets it apart from most color-based techniques that place restrictive assumptions on the color of the lighting. However, we can also use color images to estimate the chromaticity of the light sources.

4. Results

To demonstrate the performance of our algorithm, we use three very different objects, the shiny teapot of Fig. 1 we have used to illustrate its workflow, the highly texture Easter egg of Fig. 2, and the human face of Fig. 4. In all these examples, even though the objects have very different

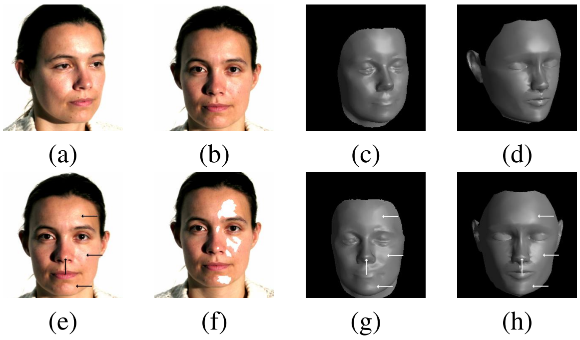


Figure 4. Insensitivity to the quality of the model. (a,b) Two out of nine input images. (c) The high-resolution scan. (d) A much lower resolution model computed using a structure from motion technique. (e) The arrows point towards the main specularities. (f) Specular areas detected using the 3D scan overlaid in white. (g) Shaded view, using the parameters recovered with the high resolution 3D scan. Note that the synthetic specularities appear at the right places. (h) Shaded view, using the parameters recovered with the low resolution model, to be compared to (g). Note that the specularities still appear at the right places.

physical properties—the teapot is shiny, the egg is very textured, and the face produces specularities that are less well defined—we used the same parameter settings. The only things that change are the 3D model and calibration matrices we feed to our system.

To quantify the accuracy of our estimates, we placed a white sphere on the same table as the teapot. It behaves almost as a mirror, making it very easy to precisely locate the specularities it produces. We used them to accurately estimate the locations of the light sources using a 3D spherical model, which is very easy to register to the actual sphere, thus making these estimates highly believable. We therefore treat them as a baseline against which we compare results obtained using other objects. In Table 1 (a), we give the recovery error for the orientations of each one of the three light sources. This yields errors ranging from 0.72 to 2.50 degrees. Note that light source number 2 is the one recovered most accurately, presumably because it produces more specularities that falls on parts of the images where the angle between the normal and light source direction is small.

Light source	Error after detection	Error after fitting
1	1.82	1.43
2	1.18	0.72
3	2.64	2.50

Table 1. Accuracy of the estimated light source directions in the teapot case. For each light source, we compare the estimates to those obtained using a white sphere, first after the step of Section 3.1 and then after the step of Section 3.2. The numbers we give are differences between directions expressed in degrees.

5. Conclusion

We have shown that, given a set of images and a 3D model of a moving object, which may be neither very precise nor complete, we can accurately recover the number and direction of multiple light sources of potentially different colors, even when the object is very textured.

This is achieved by explicitly using the fact that specularities glide over moving surfaces in a predictable fashion, which is an original approach even though the basic underlying physical facts are well known. This lets us distinguish specular areas from those surrounding other kinds of image intensity maxima. As a result, we can accurately recover lighting parameters and produce unlighted texture maps by removing specular artifacts, even in situations where other state-of-the-art techniques are not applicable.

In future work, to further increase robustness, we will explore ways to track the maximum intensity areas in video sequences to exploit spatio-temporal constraints. We will also incorporate shadows that we currently ignore into our estimation framework.

References

[1] http://www.thedigitalbits.com/reviews3/starwarschanges.html.

[2] A. K. K. Schluens. Global And Local Highlight Analysis In Color Images. 2000.

[3] G. J. Klinker, S. Shafer, and T. Kanade. The measurement of highlights in color images. *IJCV*, 2(1):7–32, June 1988.

[4] Y. Li, S. Lin, H. Lu, and H.-Y. Shum. Multiple-cue illumination estimation in textured scenes. In *ICCV*, page 1366. IEEE Computer Society, 2003.

[5] K. Nishino, Z. Zhang, and K. Ikeuchi. Determining reflectance parameters and illumination distribution from a sparse set of images for view-dependent image synthesis. In *ICCV*, pages 599–606, 2001.

[6] B. T. Phong. Illumination for computer generated pictures. *Commun. ACM*, 18(6):311–317, 1975.

[7] R. Ramamoorthi and P. Hanrahan. A signal processing framework for inverse rendering. In *SIGGRAPH*, pages 117–128, 2001.

[8] K. Schluens and M. Teschner. Analysis of 2d color spaces for highlight elimination in 3d shape reconstruction. In *Proc. ACCV'95, Singapore, vol. II*, pages 801 – 805, 1995.

[9] R. Tan and K. Ikeuchi. Iterative framework for separating reflection components of textured surface using a single image. In *CREST03*, pages 67–81, 2003.

[10] K. Torrance and E. Sparrow. Theory for Off-Specular Reflection from Roughened Surfaces. *Journal of the Optical Society of America*, 67(9), 1967.

[11] Y. Wang and D. Samaras. Estimation of multiple directional light sources for synthesis of mixed reality images. In *PCCGA*, pages 38–47, 2002.

[12] G. Ward. Measuring and modeling anisotropic reflection. In *SIGGRAPH*, pages 265–272, 1992.

[13] W. Zhou and C. Kambhamettu. Estimation of illuminant direction and intensity of multiple light sources. In *ECCV02*, page IV: 206 ff., 2002.