Photometric Registration using Specular Reflections and Application to Augmented Reality

Salma Jiddi
Technicolor/Inria
salma.jiddi@technicolor.com

Philippe Robert
Technicolor
philippe.robert@technicolor.com

Eric Marchand
Université de Rennes 1/Inria
eric.marchand@irisa.fr

ABSTRACT

Photometric registration consists in blending real and virtual scenes in a visually coherent way. To achieve this goal, both reflectance and illumination properties must be estimated. These estimates are then used, within a rendering pipeline, to virtually simulate the real lighting's interaction with the scene. In this paper, we are interested in indoor scenes where light bounces off of objects with different reflective properties (diffuse and/or specular). In these scenarios, existing solutions often assume distant lighting or limit the analysis to a single specular object. We address scenes with various objects captured by a moving RGB-D camera and estimate the 3D position of light sources. Furthermore, using spatio-temporal data, our algorithm recovers dense diffuse and specular reflectance maps. Finally, using our estimates, we demonstrate photo-realistic augmentations of real scenes (virtual shadows, specular occlusions) as well as virtual specular reflections on real world surfaces.

Author Keywords

Photometry, Reflectance, Lambertian, Specular, Illumination, Augmented Reality, Relighting.

ACM Classification Keywords

I.4.8. Image Processing and Computer Vision: Scene Analysis-Photometry; I.4.1. Image Processing and Computer Vision: Digitization and Image Capture-Reflectance

INTRODUCTION

The image brightness of a 3D object is a function of three components: object geometry, surface reflectance and illumination distribution. Provided these three components are accurately estimated, one is able to photo-realistically render a virtual scene that is visually coherent with a camera's capture of the real scene. This output is of interest for many applications such as Augmented Reality (AR) and scene relighting.

Geometry estimation is a well-known problem in the computer vision community. In the last decade, the field of active RGB-D sensors has experienced significant improvements. Hence, sensors such as Microsoft Kinect, Google Tango and Intel RealSense are nowadays commonly used in photometry registration problems. In our method, we use the R200 sensor that provides a satisfactorily accurate model of the scene.

Problem statement. As far as reflectance and illumination estimation is concerned, existing solutions often address scenes with relatively simple reflectance (e.g. Lambertian and/or non-textured surfaces) and use cast shadows [1], light probes [2][3] or fisheye cameras [4] to estimate illumination.

On the contrary, when only specular reflections are considered, they are either recovered as highlights using a single frame [5][6] or the analysis is limited to a small region (e.g. a single object) [7][8]. Also, lighting is often assumed to be distant. Nonetheless, in indoor environments where the size of the scene is not small with regard to the distance to light, this approach is likely to deliver inaccurate renderings.

Our contributions. In this paper, we focus on specular reflections to estimate diffuse and specular reflectance properties of indoor scenes with arbitrary geometry and texture. Moreover, we aim at recovering the 3D position of existing light sources without using any light probes or external assistance. We only consider as input the RGB video and 3D geometry acquired with an active sensor. Our estimation is achieved via spatio-temporal processing of the input image sequence and 3D scene analysis. Our key contributions are: (a) diffuse and specular reflectance estimation of 3D objects with arbitrary surface properties; (b) 3D position recovery of light sources responsible for specular reflections.

RELATED WORK

Illumination and reflectance estimation requires fitting a Bidirectional Reflectance Distribution Function (BRDF) and lighting distribution to the color intensity of an image. In this task, also called inverse rendering, several reflection models [9][10] and rendering techniques [2][11] have been considered. In the following, we give an overview of related work that takes into account specular reflections.

Specular surfaces and distant lighting. Nishino et al. [7] separate the reflectance components based on pixel intensity variations throughout a registered image sequence. In [8], Jachnik et al. capture a surface light field using a hand-held camera. This light field is split into its diffuse component and specular component. The latter is used to recover the environment map. Both [7] and [8] consider a single specular object. Using a fisheye camera to capture lighting, Knecht et al. [12] detect specular reflections as highlight in each frame independently. The scene is segmented using color informa-

tion and an albedo value is attributed to each segment. Hence, considered surfaces are mainly textureless.

Specular surfaces and non-distant lighting. In [5], Plopski et al. estimate the position of a dominant light source by recovering the intersection point of reflection rays raised from the center of detected highlights. In [6], Neverova et al. separate an image into its specular and diffuse components then use them as inputs of an optimization process to recover the 3D position of light sources. Both methods use [13] to estimate reflectance. The latter can deliver inaccurate diffuse estimates when the size of the specularity is significant.

PHOTOMETRIC REGISTRATION

We choose Phong reflection model [9] to describe the way a point p on a surface reflects light as a combination of diffuse and specular reflections:

$$\mathrm{I}^p = \mathrm{I}^p_d + \mathrm{I}^p_s \tag{1}$$

where I^p, I^p_d and I^p_s are respectively the color, diffuse and specular intensities. Diffuse reflectance is the property that defines an ideal "matte" surface, also called Lambertian surface. Its apparent brightness to an observer is the same regardless of his angle of view. On the other hand, specular reflection is the mirror-like reflection of light from a surface, in which light from a single incoming direction is reflected into a single outgoing direction. Using [9], diffuse and specular intensities can be described as:

$$\mathrm{I}^p = \sum_{m=1}^{l} [\mathbf{k}^p_d \mathrm{i}_m (\mathbf{N}^p \cdot \mathbf{L}^p_m) + \mathbf{k}^p_s \mathrm{i}_m (\mathbf{R}^p_m \cdot \mathbf{V}^p)^{\alpha_p}] \tag{2}$$

where $\mathbf{k}^p_d$ is the diffuse reflectance parameter of point p, $\mathbf{k}^p_s$ is its specular reflectance parameter, $\mathbf{N}^p$ is its normal vector, $\mathbf{V}^p$ is its viewpoint vector, α_p is its shininess parameter, i_m is the intensity of the light source m, $\mathbf{L}^p_m$ is its light direction vector, $\mathbf{R}^p_m$ is its reflection vector, and l is the number of light sources present in the scene. Our goal is to estimate these parameters for each 3D point in the scene.

Reflectance Estimation

We take advantage of dense 3D data provided by the sensor to spatio-temporally analyze surface color variations throughout camera trajectory. First, we sub-sample the input image sequence with respect to camera poses in order to avoid redundancies. Then we split the resulting sequence into subsets (this allows us to avoid registration errors resulting from distant camera poses). Finally, we register each subset with regard to a keyframe and provide a spatio-temporal color profile of each 3D point (Fig.1) represented by a pixel in the image.

In this work, the lighting and scene are supposed to be static. Thus, the intensity variations present in the profile can only originate from the specular component as described in [9]. In [8], the median of the profile's values is used to compute I^p_d. Although this value is robust in presence of shadows or registration errors, it often gives an over-estimation of the diffuse component. We have chosen, as in [7], to estimate the diffuse intensity as the minimum observed value in the profile as it should be closer to the correct value (Fig.2). Moreover, we constrain the recovered diffuse intensity to have a minimum number of occurrences throughout the subset.

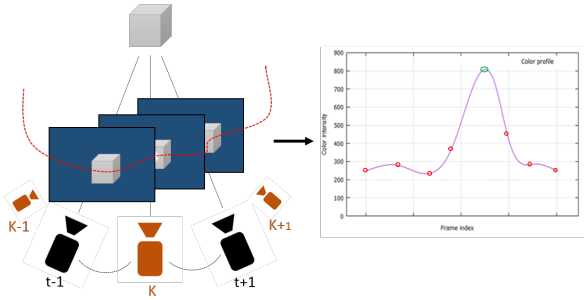


Figure 1: **Left.** 3D/2D registration : t-1 and t+1 are frames registered with regard to keyframe K. **Right.** Recovered color intensity profile of a point p using a frame subset (red dots) registered with respect to a keyframe K (green dot).



Figure 2: **Left.** Input image of a specular book in the scene. **Right.** Estimated diffuse reflectance map of the book.

The specular intensity is retrieved, at each frame t, as the difference between I^p(t) and I^p_d. The unknown specular parameters are then $(\mathbf{k}^p_s \mathrm{i}_m)$, $\mathbf{R}^p_m$ and α_p (Eq. 2). If pixel intensity at the profile peak is not saturated, the product $(\mathbf{k}^p_s \mathrm{i}_m)$ is retrieved as the lobe's peak intensity value. In fact, specular reflection occurs when the vectors $\mathbf{R}^p_m$ and $\mathbf{V}^p$ are aligned (their scalar product is equal to 1). Hence, we set $\mathbf{R}^p_m$ to be roughly equal to the peak's corresponding view vector $\mathbf{V}^p$. Finally, estimated $(\mathbf{k}^p_s \mathrm{i}_m)$ values are reported in the chosen keyframe images.

Light Sources 3D Position Estimation

In this section, we aim at estimating the 3D position of light sources represented by point lights. For each keyframe, we use surface normals $\mathbf{N}^p$ and recovered reflection vectors $\mathbf{R}^p_m$ to compute the light direction vectors ($\mathbf{L}^p_m = 2.(\mathbf{R}^p_m \cdot \mathbf{N}^p) \cdot \mathbf{N}^p - \mathbf{R}^p_m$). Light rays originating from specular reflections that fall within the same 3D region in the scene are clustered using Euclidean distance. Hence, small clusters are processed as outliers since they generally result from registration errors or inaccurate normals. Then, a mean light direction vector is computed for each significant cluster. Finally, the problem of finding the position of light sources is similar to computing the intersection points of a set of 3D lines.

Optimization

Specular reflectance parameters have been previously recovered only for 3D points roughly viewed along the observed reflection direction. Our goal is to estimate dense reflectance

maps. To achieve this, we initially assume uniform specular reflectance for each 3D object in the scene. A first step consists then in clustering the 3D mesh of the scene using Euclidean distance between vertices and normals smoothness constraint [14]. Then, provided that each cluster contains at least one 3D point with its specular intensity estimate $\mathbf{k}_s\mathrm{i}_m$, we spread its value to the entire cluster points. Finally, considering all keyframes, we compute the specular reflectance of each object as the maximum encountered $\mathbf{k}_s\mathrm{i}_m$ value.

We now address the possibility that an object/cluster may not exhibit a unique specular reflectance. First, we render specular reflections using previously recovered $\mathbf{k}_s\mathrm{i}_m$ values and light sources position. Rendered specular maps are correlated with observed specular maps. If correlation fails, we proceeed to a color-based segmentation using the k-means algorithm and set final $\mathbf{k}_s\mathrm{i}_m$ values to each color segment as follows: observed points in the direction of light sources that do not exhibit specular effects are considered to be Lambertian ($\mathbf{k}_s = 0$), points that are left keep the recovered $\mathbf{k}_s$ value. Finally, we refine $\mathbf{k}_s\mathrm{i}_m$ and estimate α using the Levenberg Marquardt algorithm with the following cost function:

$$F_j = \sum_i [\mathrm{I}_i - (\mathrm{I}_{d,i} + \sum_m \mathbf{k}_s\mathrm{i}_m(\mathbf{R}_m^i \cdot \mathbf{V}^i(\mathrm{t}))^\alpha)]^2 \quad (3)$$

where i and m respectively iterate over pixels that belong to cluster j and over recovered light sources. I_i correspond to observed pixel intensities. In (3), the diffuse component $\mathrm{I}_{d,i}$ is fixed and only $\mathbf{k}_s\mathrm{i}_m$ and α can be varied by the solver.

EXPERIMENTAL RESULTS

We use the Intel RealSense R200 sensor that provides camera positions and a 3D model of the scene. Also, we photometrically calibrate our sensor as in [15] by taking three images of a color checker with patches of known reflectance at three different shutter speeds. Finally, a calibrated camera with fixed aperture, shutter speed and color gain browses the scene.

Reflectance Evaluation

Figure 4 shows our photometric registration results for three different scenes (S1, S2 and S3). Our algorithm recovers dense diffuse (Fig.4, col-2) and specular (Fig.4, col-4) maps. To evaluate the accuracy of these maps, we achieve two main applications: relighting and Augmented Reality (AR). In Fig.3, we use recovered reflectance and illumination to virtually relight the specular book. One can notice that the position and intensity of the specularity are well estimated as they are visually coherent with the input image. Furthermore, column 5 in Figure 4 shows different augmented scenes where virtual shadows and specular occlusions are correctly rendered. In Fig.4, row-1/col-5, one can notice a correct shape of virtual shadows as well as consistent specular occlusions by the virtual green and red spheres. In Figure 4. row-3/col-5, the number of light sources is correctly recovered as well.

Illumination Evaluation

Using a fish-eye lens, we capture the environment map and qualitatively compare it to the recovered lighting distribution (Fig.5) for three scenes (S1, S2 and S3 with respectively one, two and three light sources). Our algorithm estimates correct lighting distribution as it recovers the correct number of light sources responsible for specular reflections. Furthermore, we use salient control points and evaluate the accuracy of recovered positions. Table 1 shows a comparison between measured (Measured D) and recovered distances to light sources (Estimated D). Our algorithm is tested on various indoor scenes under various lighting (e.g single and multiple spot lights and/or led lights) and recovers light sources positions with an average error of 21cm for a mean distance of 1.95m to the light source.

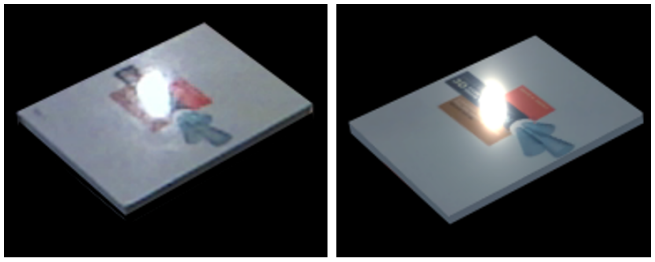


Figure 3: Comparison between input image (left) and virtually rendered image using estimated reflectance and lighting.

Scenes	Measured D (cm)			Estimated D (cm)		
S1	72.2	-	-	86.1	-	-
S2	94.7	105.6	-	88.9	127.1	-
S3	314.7	293.6	306	297.1	261.6	283.2

Table 1. Comparison between measured and estimated distances to light sources for scenes under different lightings.

CONCLUSION AND FUTURE WORK

We presented an algorithm that estimates dense diffuse and specular reflectance maps. Furthermore, our proposed approach recovers the 3D position of light sources for indoor scenes with arbitrary geometry and texture. For future work, we are interested in estimating the color and shape of the illumination in order to produce both hard and soft shadows.

REFERENCES

1. I. Sato, Y. Sato, and K. Ikeuchi, "Illumination from shadows," in *PAMI*, 2002.
2. P. Debevec, "Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography," in *ICCV*, 1999.
3. D. Nowrouzezahrai, S. Geiger, K. Mitchell, and M. Gross, "Light factorization for mixed-frequency shadows in augmented reality," in *ISMAR*, 2011.
4. K. Rohmer, W. Buschel, D. Raimund, and T. Grosch, "Interactive near-field illumination for photorealistic augmented reality on mobile devices," in *VGTC*, 2014.
5. A. Plopski, T. Mashita, K. Kiyokawa, and H. Takemura, "Reflectance and light source estimation for indoor ar applications," in *VR*, 2014.

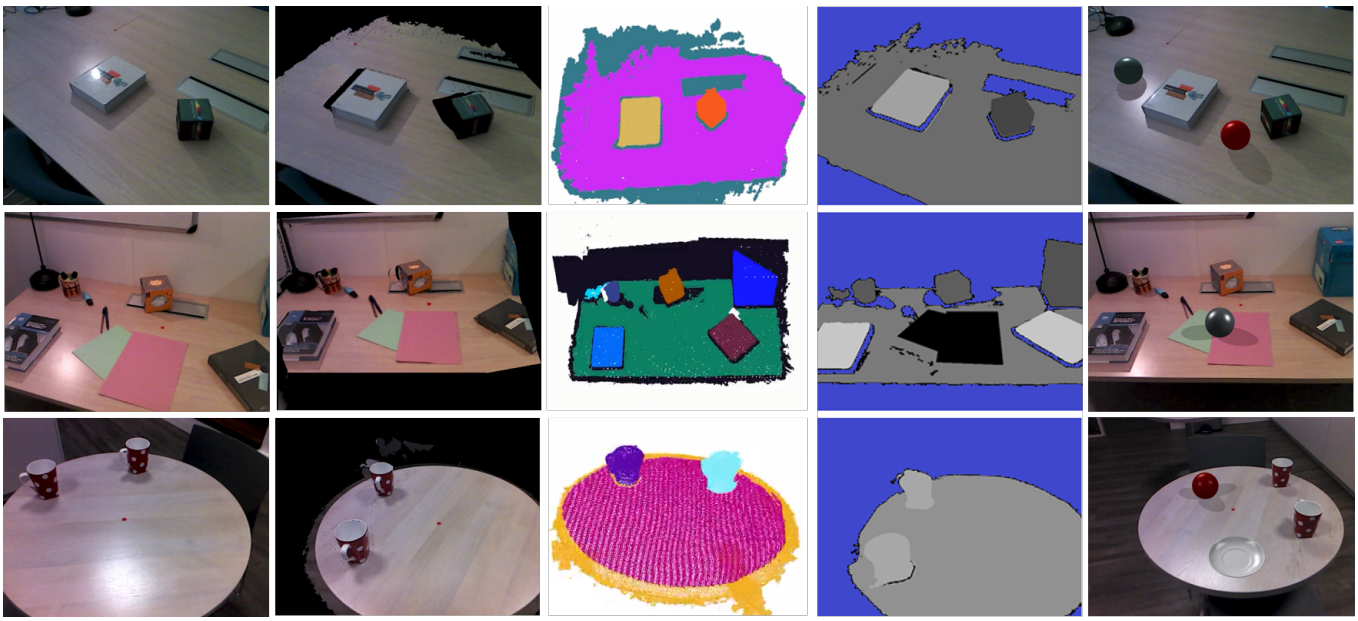


Figure 4: **Column 1.** Camera views of three indoor scenes (S1, S2 and S3). **Column 2.** Recovered diffuse maps for surfaces with various textures and reflective properties. **Column 3.** 3D mesh clustering. **Column 4.** Estimated specular reflectance parameter $\mathbf{k}_s^j$ for each cluster j. Blue pixels correspond to 3D points frequently occluded during scene browsing. Brighter $\mathbf{k}_s^j$ values correspond to more specular surfaces. **Column 5.** Augmented scenes, with different reflectance and illumination conditions. We demonstrate correctly rendered virtual objects as they occlude real specular effects and show realistic shadows.

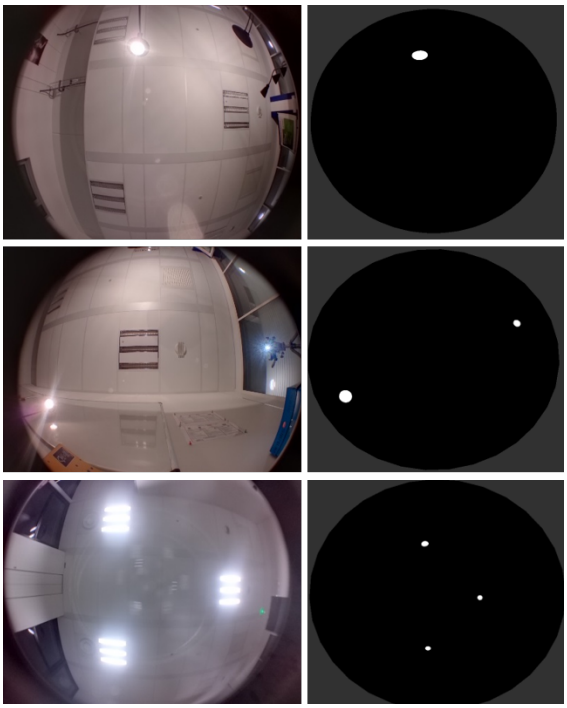


Figure 5: **Left**. Captured environment maps using a fish-eye lens. **Right**. Recovered lighting respectively for S1 (row-1), S2 (row-2) and S3 (row-3) with respectively one, two and three main light sources.

6. N. Neverova, D. Muselet, and A. Tremeau, "Lighting estimation in indoor environments from low-quality images," in *ECCV*, 2012.
7. 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*, 2001.
8. J. Jachnik, R. A. Newcombe, and A.J. Davison, "Real-time surface light-field capture for augmentation of planar specular surfaces," in *ISMAR*, 2012.
9. B.T. Phong, "Illumination for computer generated pictures," in *CACM*, 1975.
10. K.E. Torrance and E.M. Sparrow, "Theory for off-specular reflection from roughened surfaces," in *JOSA*, 1967.
11. A. Keller, "Instant radiosity," in *SIGGRAPH*, 1997.
12. M. Knecht, G. Tanzmeister, C. Traxler, and M. Wimmer, "Interactive brdf estimation for mixed-reality applications," in *WSCG*, 2012.
13. H.-L. Shen and Q.-Y. Cai, "Simple and efficient method for specularity removal in an image," in *Applied optics*, 2009.
14. Point Cloud Library, ," `http://pointclouds.org/`.
15. S. Robertson, Mark A.and Borman and R.L. Stevenson, "Dynamic range improvement through multiple exposures," in *ICIP*, 1999.