Can we Invert a Local Reflectance Model From a Single Specular Highlight with Known Scene Geometry and Camera Pose?

S. Hadj-Said[1,2], M. Tamaazousti[1] and A. Bartoli[2]

[1]CEA, LIST, Gif-sur-Yvette, France [2] Institut Pascal, UMR 6602, Clermont-Ferrand, France

Abstract

Recovering the scene's illumination from images is a crucial step in Augmented and Diminished Reality. We present an experimental investigation to assess the well-posedness of this problem (i) using a known scene geometry and camera pose and (ii) assuming that a local reflectance model such as Blinn-Phong's holds. Based on results on synthetic and real data, we establish two major observations. First, the problem of retrieving the full local reflectance model's parameters and light source position from the original image is ill-posed. Second, the specular parameters including the light source position can be stably estimated from the image's specular component using a single specular highlight region.

1. Introduction

The recent advances in SLAM approaches have facilitated the introduction of Augmented Reality (AR) and Diminished Reality (DR) in several fields including the industry. In many applications, estimating the scene's illumination is a crucial step. In spite of the various works on retrieving illumination from images [BG01,ZCC16], the minimal required visual cues to solve this problem are still unknown. One of the widely used approaches to solve this problem is the local reflectance inversion method. While local reflectance models are usually used in Computer Graphics to simulate light reflection and produce photo-realistic images [HVDF*14], its inversion consists in using the illumination intensity as input data and retrieving some of or all the scene's parameters. Working with SLAM, we consider a known geometry (a surface mesh and its normal map), known camera parameters (pose and calibration) and the image of a unique specular region. As output, we estimate the reflectance parameters, the roughness and the light position. We conduct several tests on synthetic and real images using two different scenarios. First, we estimate the full model's parameters from the original image modeled as the sum of three components: ambient, diffuse and specular. Second, we estimate only the specular component's parameters from the specular-diffuse decomposition of the original image. While the first scenario does not require any preliminary processing of the original image, the second scenario has the advantage of estimating a smaller number of parameters. The comparison between the results of these two experiments provides an answer for the well-posedness of the local reflectance inversion problem.

2. Method

Standard reflectance model. The ambient component is represented as $\mathbf{k}_a(\mathbf{P}) \odot \mathbf{i}_a$ where $\mathbf{i}_a \in [0,1]^3$ is the ambient light intensity in RGB, $\mathbf{k}_a(\mathbf{P})$ is the ambient reflectance coefficient of a 3D point $\mathbf{P}$ and $\odot$ is the Hadamard product. This term approximates the effect of indirect lighting. The diffuse and the specular components are described, respectively, by the terms $\mathbf{N}(\mathbf{P}) \cdot \mathbf{L}_n(\mathbf{P})\mathbf{k}_d(\mathbf{P}) \odot \mathbf{i}_n$ and $\mathbf{J}_s(\mathbf{P}, \mathbf{k}_s \odot \mathbf{i}_n, m, \mathbf{L})$ where $\mathbf{i}_n$ is the intensity of the light source s_n located in the position $\mathbf{S}_n$ and $\mathbf{k}_d(\mathbf{P})$, $\mathbf{k}_s(\mathbf{P})$ represent, respectively, the diffuse and specular reflectance coefficients, $\mathbf{N}(\mathbf{P})$ is the surface normal in $\mathbf{P}$ and m is the roughness of the surface. We compare three models used in the literature in the context of local reflectance inversion: Blinn-Phong which we refer to as BP, a simplified version of Torrance-Sparrow which we refer to as TS and Ward for isotropic surfaces which we refer to as WI. These models differ solely by their specular component $\mathbf{J}_s(\mathbf{P}, \mathbf{K}_s, m, \mathbf{L})$.

Simplified model. Assuming that *(i) a single light source* s_1 *contributes to the observed specularity* and that *(ii) the object's surface has a constant albedo and roughness*, the reflectance model can be simplified as:

$$\mathbf{I}_e(\mathbf{P}, \mathbf{S}_1, \mathbf{K}_a, \mathbf{K}_d, \mathbf{K}_s, m) = \mathbf{K}_a + \mathbf{N}(\mathbf{P}) \cdot \mathbf{L}_1(\mathbf{P})\mathbf{K}_d + \mathbf{J}_s(\mathbf{P}, \mathbf{K}_s, m, \mathbf{L}_1), \quad (1)$$

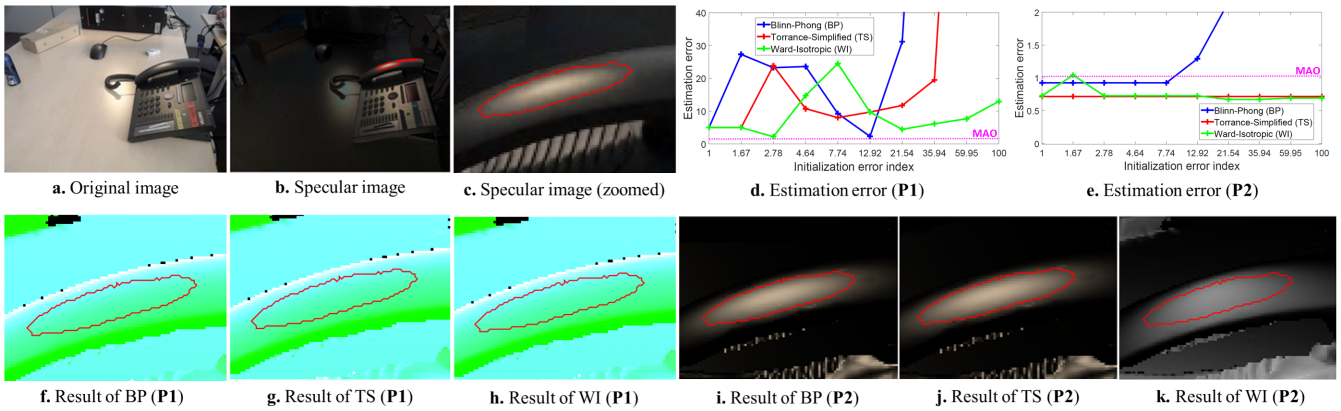
where $\mathbf{K}_a = \mathbf{k}_a(\mathbf{P}) \odot \mathbf{i}_a + \sum_{n=2}^{N} \mathbf{N}(\mathbf{P}) \cdot \mathbf{L}_n(\mathbf{P})\mathbf{k}_d(\mathbf{P}) \odot \mathbf{i}_n$, $\mathbf{K}_d = \mathbf{k}_d \odot \mathbf{i}_1$ and $\mathbf{K}_s = \mathbf{k}_s \odot \mathbf{i}_1$.

Inversion approach. We consider an image $\mathbf{I}_r$ of a specular surface. An image region Ω is extracted as the reference data. This region corresponds to the largest specular highlight in the image. We perform a non-linear optimization distinctly for each RGB channel. This means that the values of the parameters $\mathbf{K}_s$, $\mathbf{K}_d$ and $\mathbf{K}_a$, are estimated separately per channel. Using Levenberg-Marquardt, we minimize the following cost:

$$C^2_{\text{photo}} = \frac{1}{|\Omega|} \sum_{\mathbf{P} \in \Omega} \left(I^c_r(\mathbf{P}) - I^c_e(\mathbf{P}, \mathbf{S}^*_1, K^{c*}_a, K^{c*}_d, K^{c*}_s, m^*) \right)^2. \quad (2)$$

The notation ^{c} indicates the color channel c and * indicates the estimated parameters. In total, three separate minimizations are carried

DOI: 10.2312/egp.20191046



a. Original image **b.** Specular image **c.** Specular image (zoomed) **d.** Estimation error (**P1**) **e.** Estimation error (**P2**)

f. Result of BP (**P1**) **g.** Result of TS (**P1**) **h.** Result of WI (**P1**) **i.** Result of BP (**P2**) **j.** Result of TS (**P2**) **k.** Result of WI (**P2**)

Figure 1: *Results of the scenarios **P1** and **P2** on a real example. The target specular region is delineated by the red curve. The estimation errors shown in figures (d) and (e) illustrate the stability of the approach on scenario **P2** compared to **P1**. The colors in f, g and h are due to the flawed estimation of the reflectance coefficients* $\mathbf{k}_a$ *and* $\mathbf{k}_d$ *in scenario **P1**.*

out for each test. We obtain three different values for $\mathbf{S}_1$ and m and use the median as final estimate. We initialize the estimation by the value of the groundtruth parameters.

3. Experiments and Results

We conduct two types of experiments on synthetic and real data. For synthetic data, we use three 3D objects to generate our dataset of 81 images. For each image, we launch a total of 300 trials (10 initialization magnitudes $\times$ 30 random values). In the first scenario **P1**, we estimate the full model's parameters from the original image which are $\mathbf{S}_1^*$, K_a^{c*}, K_d^{c*}, K_s^{c*}, m^*. In the second scenario **P2**, we estimate the specular component's parameters $\mathbf{S}_1^*$, K_s^{c*} and m^* from the specular component of the image. We obtain the specular-diffuse decomposition by generating the images using only the specular term in equation (1). In **P1**, the inversion approach fails to provide good estimates for the parameters on synthetic data produced by the same reflectance model. However, in **P2**, we obtained satisfying results. The estimation error used to assess the success of each scenario is fixed as the normalized difference between the true and estimated values of the parameters:

$$E_g = \frac{1}{T_\mathbf{S}}\|\mathbf{S}_1 - \mathbf{S}_1{}^*\|_2 + \frac{1}{T_\mathbf{K}}\|\mathbf{K}_a - \mathbf{K}_a{}^*\|_2 + \frac{1}{T_\mathbf{K}}\|\mathbf{K}_d - \mathbf{K}_d{}^*\|_2 + \frac{1}{T_\mathbf{K}}\|\mathbf{K}_s - \mathbf{K}_s{}^*\|_2 + \frac{1}{T_m}\|m - m^*\|_2, \quad (3)$$

where $T_\mathbf{S}$, $T_\mathbf{K}$ and T_m are weights computed independently from the numerical error of each parameter in equation (3). Since the parameters have different orders of magnitude, this allows us to normalize the terms to a common scale. We define MAO, the Maximum Acceptance Offset as the maximum tolerated estimation error. Each term of equation (3) has to be lower than 1. So, MAO is 5 for **P1** (5 terms) and it is 3 for **P2** as $\mathbf{K}_a$ and $\mathbf{K}_d$ are not estimated. For real data, we use 4 images from a real scene. An example is showed in figure 1. The 3D reconstruction is obtained by the HandySCAN 3D scanner from Creaform. To separate the diffuse and specular, we use two polarizers, one in front of the camera and another in front of the light source. The same results as for synthetic data are obtained. Only the light position is used to compute the estimation error in this case. MAO is therefore 1 for both scenarios.

4. Discussions and Conclusions

We have addressed the problem of inverting a local reflectance model from a single specular highlight. We have not proposed a new computational solution for the problem but we investigated the *solvability* of the problem. The results have been conducted on two test scenarios using a single point light source, known surface geometry and homogeneous isotropic surfaces. These assumptions are usually met in real-case scenarios and they allow us to focus on our main objective of establishing the correlation between the input data and the solvability. They also form a base scenario for later studies on more complex light configurations and surfaces. Therefore, we have confirmed that through a specular-diffuse separation on real data associated with an optimal surface geometry estimation and using a single specular highlight, we can stably estimate the specular parameters of an object according to a simplified local reflectance model. We have also observed that the full reflectance model cannot be estimated directly from the original image. In the light of our findings, we recommend a specular-diffuse decomposition of the image associated with the single specularity approach as a flexible approach to reflectance recovery in AR and DR. This approach can be used without the need for any priors on the number of light sources since each specularity would be processed separately. In future work, We will carry an analysis of the robustness of our approach in the presence of significantly noisy data. We will also investigate methods for separating the diffuse and specular components using recent deep learning approaches.

References

[BG01] BOIVIN S., GAGALOWICZ A.: Image-based rendering of diffuse, specular and glossy surfaces from a single image. In *ACM SIGGRAPH* (2001), pp. 107–116. 1

[HVDF*14] HUGHES J. F., VAN DAM A., FOLEY J. D., MCGUIRE M., FEINER S. K., SKLAR D. F., AKELEY K.: *Computer graphics: principles and practice*. Pearson Education, 2014. 1

[ZCC16] ZHANG E., COHEN M. F., CURLESS B.: Emptying, refurnishing, and relighting indoor spaces. *ACM Transactions on Graphics (TOG) 35*, 6 (2016), 174. 1