

Illumination Estimation using Cast Shadows for Realistic Augmented Reality Applications

Salma Jiddi, Philippe Robert, Eric Marchand

▶ **To cite this version:**

Salma Jiddi, Philippe Robert, Eric Marchand. Illumination Estimation using Cast Shadows for Realistic Augmented Reality Applications. IEEE Int. Symposium on Mixed and Augmented Reality (ISMAR-Adjunct), Oct 2017, Nantes, France. 2017. hal-01668701

HAL Id: hal-01668701
https://inria.hal.science/hal-01668701

Submitted on 22 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

[POSTER] Illumination Estimation using Cast Shadows for Realistic Augmented Reality Applications

Salma Jiddi* Technicolor IRISA

Philippe Robert† Technicolor

Eric Marchand‡ Université de Rennes 1 IRISA

Abstract

Augmented Reality (AR) scenarios aim to provide realistic blending between real world and virtual objects. A key factor for realistic AR is thus a correct illumination simulation. This consists in estimating the characteristics of real light sources and use them to model virtual lighting. In this paper, we briefly introduce a novel method for recovering both 3D position and intensity of multiple light sources using detected cast shadows. Our algorithm has been successfully tested on a set of real scenes where virtual objects have visually coherent shadows.

Index Terms: I.4.1 [Image Processing and Computer Vision]: Scene Analysis—Photometry;

1 Introduction

The image brightness of a 3D scene is a function of three components: scene geometry, surface reflectance and illumination distribution. Provided that these three components are accurately estimated, one is able to photo-realistically render a virtual scene. This output is of interest for many applications such as Augmented Reality (AR). In this paper, we focus on estimating the characteristics of the illumination distribution. Specifically, we aim at recovering the 3D position and intensity of light sources in 3D real scenes with arbitrary geometry and texture. Existing solutions often use additional devices such as light probes [2] or fisheye cameras [4] to recover lighting conditions. On the contrary, when only image and geometry information are considered, lighting is often assumed to be distant [3] [5] or reduced to a single point light [1]. In this paper, we propose a novel framework to jointly detect cast shadows and estimate real world lighting from a single image, given coarse 3D geometry. The proposed approach makes two main assumptions. First, scene surfaces are supposed to be quasi Lambertian. The second assumption consists in having a principal planar surface on which cast shadows are detected. Our main contributions are: **(a)** detection of cast shadows on textured surfaces using both geometry and image information; **(b)** estimation of the 3D position of multiple light sources represented by point lights; **(c)** estimation of the intensity of recovered illumination using shadow pixels observations.

2 Related Work

Early work on photometric reconstruction using shadows was presented by Sato et al. [5]. They used shadows cast by an object of known shape to recover distant illumination distribution. In [5], cast shadows are detected by manually removing occluding objects and capturing a shadow-free image. Panagopoulos et al. [3] used a graphical model to detect shadows in scenes with textured surfaces and recovered the direction of distant lighting. In [1], Arief et al. analyzed the shadow produced by a cube to estimate the position of a single strong light source. The surface on which the shadow is cast needs to be of a single color. Our methodology differs from previous illumination estimation work in three main points: First, our approach does not require any user intervention or reference objects to recover illumination. On the contrary, it targets middle scale scenes including various 3D objects with arbitrary geometry. Secondly, our algorithm handles shadows cast on textured surfaces. Finally, we do not make the assumption of distant lighting and instead, estimate both 3D position and intensity of real world lighting.

*e-mail: salma.jiddi@technicolor.com
†e-mail: philippe.robert@technicolor.com
‡e-mail: eric.marchand@irisa.fr

3 Our Approach to Illumination Estimation

In the following, we choose Phong reflection model [8] to describe the way a point p in the scene reflects light:

$$\mathbf{I}^p = \mathbf{k}_d^p(\mathrm{L}_a + \sum_{i=1}^{M}(\mathbf{N}^p \cdot \boldsymbol{\omega}_i^p)\mathrm{L}_i\mathrm{O}_i^p) \quad (1)$$

where $\mathbf{I}^p$ is the color intensity of point p, $\mathbf{k}_d^p$ is its albedo and $\mathbf{N}^p$ is its normal vector. L_a and L_i are respectively the intensity of ambient lighting and light source i, $\boldsymbol{\omega}_i^p$ is the incoming light direction vector of light i, and M is the number of light sources present in the scene. O_i^p is a binary visibility term that is equal to 1 if light i is visible from the 3D point corresponding to pixel p and equal to 0 if occluded.

3.1 Detection of Cast Shadows

In this section, our goal is to detect shadow pixels p and estimate their respective occlusion attenuation values $\delta(p)$. The occlusion attenuation represents the brightness ratio between shadow pixels p and their corresponding non-shadow pixels $\hat{p}$ which have the same reflectance properties but are subject to different illumination conditions. The proposed shadow detection approach is threefold. First, we segment the 3D geometric model of the scene into a set of 3D clusters using normals deviation and Euclidean distance between 3D points. Secondly, as shadows are caused by the occlusion of light due to occluding geometry, we define a proportional region of interest (ROI) with respect to each detected 3D object (Figure 1). These regions represent potential cast shadows regions. Finally, to recover shadow pixels inside ROIs, we compare inside-region pixels p and outside-region pixels $\hat{p}$ using shadow-variant and shadow-invariant features as follows: we first compute the Modified Specular Free

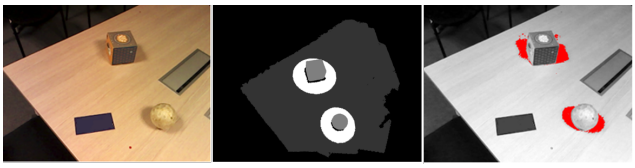


Figure 1: **Col.1**: Scene capture. **Col.2**: shows the 3D region of interest (white pixels) where potential cast shadows are located. Dark grey pixels represent the detected plane and light grey pixels are occluders. **Col.3**: Initial detection of cast shadows (red pixels).

(MSF) image [6] as it allows us to handle weak specular reflections. Then, we consider the chromaticity C of the MSF image as the

shadow-invariant feature since it eliminates the effect of varying intensities resulting from lighting occlusion. As far as the shadow-variant feature is concerned, we use the value channel V of the HSV color space as it represents scene brightness. Finally, a voting scheme for all pixels p inside ROI is applied as follows:

$$s(p) = \begin{cases} +1, & \text{if } |C^p - C^{\hat{p}}| \leq \varepsilon_C \text{ and } (V^{\hat{p}} - V^p) \geq \varepsilon_V \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where ε_C and ε_V are respectively thresholds with respect to MSF chromaticity and value differences. Pixels with high voting values $s(p)$ are detected as shadows. The outputs of the algorithm are a mask of detected cast shadows (Figure 1) and an occlusion attenuation value $\delta(p)$ computed as follows:

$$\delta(p) = \frac{V^p}{\bar{V}} \qquad \text{with:} \qquad \bar{V} = \frac{\sum_{\hat{p}} V^{\hat{p}}}{s(p)} \quad (3)$$

where $\bar{V}$ is the mean brightness value of selected pixels $\hat{p}$. The occlusion attenuation $\delta(p)$ represents the percentage of light that is occluded at shadow pixel p.

3.2 Estimation of the 3D Position of Lighting

Illumination distribution is first approximated by a set of point light sources equally distributed in the scene. In this section, we aim at extracting a subset of point lights whose shadow maps correlate with the mask of detected shadows. Our approach is two-fold: to begin with, our method assumes that surface materials are opaque; thus, only the 3D space above the detected plane is considered. The latter is equally sampled and represents a three-dimensional grid composed of a set of voxels $\mathscr{V}_j$. Point light candidates are positioned at the center of each voxel. The estimation of an accurate 3D lighting position depends on the resolution of the grid. Hence, to avoid considering all light sources resulting from dense sampling and to speed up data processing, we evaluate visibility for a set of shadow pixels: rays originating from the 3D location of shadow pixels are cast in a discrete number of directions. Finally, light sources laying inside voxels that are intersected at least once are selected. This first step of light candidates selection does not take into account the effect of light sources 3D position. Hence, the identification of final point lights subset is based on matching their respective shadow maps with detected shadow pixels. This matching is carried out via the computation of correlation between two binary variables: the binary mask of the detected shadows and the binary mask of the rendered shadows. The correlation corresponds to the *phi* coefficient value Φ also referred to as the mean square contingency coefficient.

3.3 Estimation of the Intensity of Lighting

In this section, our goal is to recover light sources intensity. The inputs are detected shadow pixels (Sec 3.1) and the recovered selection of point lights (Sec 3.2). To simplify Phong reflection model [8], we assume that $((\mathbf{N}^p \cdot \omega^p) = \cos\theta)$ is equal across shadow/non-shadow pairs, and use L_i to represent $L_i \cos\theta$. Subsequently:

$$\mathrm{I}^p = \mathbf{k}_d^p (\mathrm{L}_a + \sum_{i=1}^{M} \mathrm{L}_i \mathrm{O}_i^p) \quad (4)$$

Let us consider the ratio of color intensities of both a shadow pixel p and its corresponding non-shadow pixel $\hat{p}$:

$$\frac{\mathbf{I}^p}{\mathbf{I}^{\hat{p}}} = \frac{\mathbf{k}_d (\mathrm{L}_a + \sum_{i=1}^{M} (\mathrm{L}_i \mathrm{O}_i^p))}{\mathbf{k}_d (\mathrm{L}_a + \sum_{i=1}^{M} \mathrm{L}_i)} = \delta(p) \quad (5)$$

We set $\mathrm{L}_a + \sum_{i=1}^{M} \mathrm{L}_i = 1$ and obtain a linear system ($\mathbf{AL} = \delta$) using shadow pixels observation, where:

$$\mathbf{A} = \begin{pmatrix} 1 & \mathbf{O}_1^{p_1} & \mathbf{O}_2^{p_1} & \cdots & \mathbf{O}_M^{p_1} \\ 1 & \mathbf{O}_1^{p_2} & \mathbf{O}_2^{p_2} & \cdots & \mathbf{O}_M^{p_2} \\ \vdots & \vdots & \ddots & \vdots & \\ 1 & \mathbf{O}_1^{p_N} & \mathbf{O}_2^{p_N} & \cdots & \mathbf{O}_M^{p_N} \end{pmatrix} ; \mathbf{L} = \begin{pmatrix} \mathrm{L}_a \\ \mathrm{L}_1 \\ \vdots \\ \mathrm{L}_M \end{pmatrix} ; \delta = \begin{pmatrix} \delta(p_1) \\ \delta(p_2) \\ \vdots \\ \delta(p_N) \end{pmatrix} \quad (6)$$

The linear system is solved using Reweighted Least Squares with bounds and equality constraints:

$$\widehat{L} = \min_L (\frac{1}{2} \|\mathbf{AL} - \delta\|^2) \quad \text{such as:} \begin{cases} 0 \leq \mathrm{L}_i \leq 1 \quad \text{and} \quad 0 \leq \mathrm{L}_a \leq 1 \\ \mathrm{L}_a + \sum_{i=1}^{M} \mathrm{L}_i = 1 \end{cases} \quad (7)$$

The weights are computed using Tukey's bisquare loss function. Hence, observations with small weights are marked as outliers and removed in the next iteration. Finally, initial shadow estimates are refined by excluding outliers from the shadow mask.

4 Experimental Results

Fig.2-col.2 demonstrates a correct mapping between detected and real shadows (red pixels correspond to refined shadow estimates and cyan pixels to rejected outliers using our robust estimator). Furthermore, figure 3 shows a correct lighting distribution recovery where both area lights (Fig.3-col.3) are approximated by a set of point lights (Fig.3-col.2). Our algorithm is tested on eighteen indoor scenes under various lighting and recovers light sources positions with an average error of 17cm for a mean distance of 2.55m to the light source. Finally, Fig.2-col.3 shows two AR scenarios where virtual shadows are visually coherent in terms of shape and intensity.

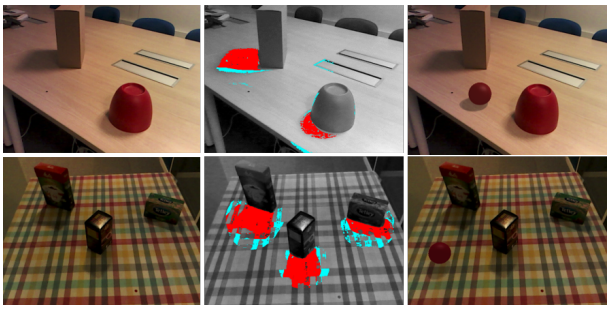


Figure 2: **Col.1**: Real scenes capture. **Col.2**: Refined cast shadows detection. **Col.3**: Augmented Reality with realistic virtual shadows.

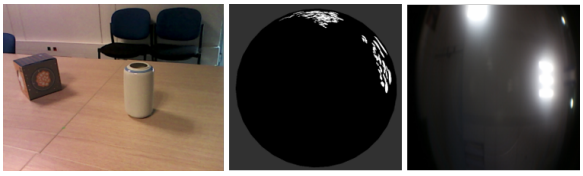


Figure 3: **Col.1**: Real scene capture. **Col.2**: Recovered illumination distribution for two area lights. **Col.3**: Fish-eye capture of lighting.

5 Conlusion and Future Work

The proposed algorithm jointly detects shadows and recovers illumination. Our estimates are used to render visually satisfying augmentations. Specifically, virtual and real shadows are consistent in terms of shape and intensity. Future work concerns extending our estimation to more complex lighting conditions (e.g, windows).

References

[1] I. Arief, S. McCallum, and J. Hardeberg. Realtime estimation of illumination direction for augmented reality on mobile devices. In *CIC'12*.

[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'99*.

[3] A. Panagopoulos, D. Samaras, and N. Paragios. Robust shadow and illumination estimation using a mixture model. In *CVPR'09*.

[4] K. Rohmer and T. Buschel, W. Grosch. Interactive near-field illumination for photorealistic augmented reality on mobile devices. In *VGTC'14*.

[5] I. Sato, Y. Sato, and K. Ikeuchi. Illumination from shadows. In *PAMI'02*.

[6] H.-L. Shen and Q.-Y. Cai. Simple and efficient method for specularity removal in an image. In *Applied optics'09*.