

# Lighting Simulation of Augmented Outdoor Scene Based on a Legacy Photograph

Guanyu Xing<sup>1,2</sup> Xuehong Zhou<sup>1</sup> Qunsheng Peng<sup>1</sup> Yanli Liu<sup>3</sup> Xueying Qin<sup>4</sup>

<sup>1</sup>State Key Laboratory of CAD & CG, Zhejiang University, Hangzhou, China

<sup>2</sup>School of Computer Science & Engineering, University of Electronic Science and Technology of China, Chengdu, China

<sup>3</sup>College of Computer Science, Sichuan University, Chengdu, China

<sup>4</sup>School of Computer Science and Technique, Shandong University, Ji'nan, China



**Figure 1:** Our system simulates the illumination of an outdoor scene based on a legacy photograph. We recover an environment map based on the image, and model its illumination with six lighting parameters. Shadow casting between real world and virtual object can then be synthesized, ensuring the shadow consistency of the augmented scene image. This figure shows two different scenes, for each scene, the left one is the input image and the right one demonstrates our synthetic result.

## Abstract

We propose a novel approach to simulate the illumination of augmented outdoor scene based on a legacy photograph. Unlike previous works which only take surface radiosity or lighting related prior information as the basis of illumination estimation, our method integrates both of these two items. By adopting spherical harmonics, we deduce a linear model with only six illumination parameters. The illumination of an outdoor scene is finally calculated by solving a linear least square problem with the color constraint of the sunlight and the skylight. A high quality environment map is then set up, leading to realistic rendering results. We also explore the problem of shadow casting between real and virtual objects without knowing the geometry of objects which cast shadows. An efficient method is proposed to project complex shadows (such as tree's shadows) on the ground of the real scene to the surface of the virtual object with texture mapping. Finally, we present an unified scheme for image composition of a real outdoor scene with virtual objects ensuring their illumination consistency and shadow consistency. Experiments demonstrate the effectiveness and flexibility of our method.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Augmented reality—Illumination consistency

## 1. Introduction

Lighting simulation of real scenes is necessary for ensuring the illumination consistency when inserting a synthetic object into their images or videos. For outdoor illumination simulation based on a legacy photograph, there are two prob-

lems to be solved: (1). recovering the illumination distribution of the scene; (2). casting shadow between real and virtual objects. Nevertheless, the color at each pixel of the photo represents the result of comprehensive interactions between objects and the incident light, making the lighting simulation based on a legacy photograph difficult.

There have been several methods for estimating the illumination distribution of natural scenes [JL06] based on their images or videos. Some works try to estimate scene's illumination according to lighting related prior information. [Deb98] photos a mirror ball to record scene's illumination, [SSI99] captures HDR omni-directional images to achieve the goal. Lalonde's method [LEN09] uses Perez's sky model [PMI93] as prior information to estimate the outdoor illumination, it generates an environment map containing high-frequency components of real lighting environment. However, this method requires a time-lapse image sequence and the intensity of light source is estimated based on an empirical model. A simple illumination model of outdoor scenes is adopted by [LQX\*09], which assumes that the sun is a directional light source, and the sky is a uniform area light source, the illumination can therefore be estimated with the illumination related statistical parameters or basis images learned from sample images; [LG12] adopts the same assumptions, they take spatial and temporal coherence of illumination as prior to recover illumination from videos captured with moving cameras. Nevertheless, the simple light source assumption of outdoor scenes prevent these approaches to achieve highly realistic rendering result.

Alternatively, some works recover scene's illumination according to surface radiosity, they deduce an optimization problem regarding the appearance of objects in the scene adhering to the rendering equation. Nevertheless, most of these approaches need to know the 3D geometry of the scene. Sato et al. estimate illumination of lambertian scene from shadows [SSI03]. [ZY01] selects several critical points to calculate the direction and intensity of the light source. [YWAC06, HFC\*09] use multiple images to recover the reflectance model and illumination environment of real scenes. In general, the methods based on scene geometry provide perfect theoretical basis for illumination estimation of real scenes, but the request of known scene geometry limits their application. Moreover, the illumination recovered by these methods contains only low-frequency lighting components of real environment, which may fail in rendering virtual objects with strong specular reflection. Obviously, all methods discussed above cannot apply to illumination simulation based on a legacy photograph. Recently, Karsch et al. [KHFH11] provided an interaction based approach to estimate the illumination of a legacy photograph, they assume the scene in a box, the size of the box and the position of the light source are specified interactively, then their method automatically rebuilds the scene's material and intensity of light source adhering to a lambertian based rendering equation. Nevertheless, the recovery of material and lighting condition necessitates solving a large-scale nonlinear optimization problem, which is computation-intensive; the intensity of the extremely far away light source need be assigned by the user; besides, the method assumes that all the light

sources(except the sunlight) are distributed on a plane. No wonder, Karsch's method is more suitable for indoor scenes.

There are two problems need to be solved if we want to simulate shadow casting between real scene and virtual objects: (1). how to cast real shadows on the virtual objects; (2). how to avoid overlap between virtual and real shadows. [KTM\*10, NGM\*11] proposed several rendering methods to cast shadows of virtual objects onto real world rapidly, however, they overlook the interactions of real shadows on the virtual objects. [GM00] uses shadow map to get the right synthetic result. To achieve a similar effect, [JAL05, MJA06] use shadow volume. Nevertheless, a 3D model of the scene should be available as the basic information, which is very difficult to be recovered from a legacy photograph. Karsch's method [KHFH11] uses light shafts to simulate the extremely far away light source, they set the intensity of light shaft corresponding to shadow area as zero, hence capable of producing shadow effect on the surface of virtual objects. However, the shape of shadows in nature scenes are always grotesque, which makes the creation of light shafts difficult; in addition, the light shafts will introduce extra cost of rendering.

In this paper, we propose a method to realistically insert synthetic objects into a legacy photograph of outdoor scene which accounts for not only the accurate environment illumination but also the interactive shadow casting between virtual objects and real scene. Our method requires no illumination capturing devices, nor multiple photographs. Unlike previous methods which only take surface radiosity or lighting related prior information as the basis of illumination estimation, our method employs skylight distribution and lighting condition extracted from the legacy image as the illumination related prior information and incorporates them into rendering equation. By adopting spherical harmonics, we set up a linear model with only six illumination parameters, this simple model provides us an easy and efficient approach to synthesize high-quality environment maps of real scenes. We first estimate a rough geometry of the scene with a few easy interactions; then the material of sample points are solved linearly (diffused) or iteratively (specular); the illumination parameters are finally calculated by solving a linear least square problem with the color of the sunlight and the skylight as its constraint. We also propose an efficient method to re-project the shadow on the ground of the real scene to the surface of the virtual object with texture mapping. To avoid the overlap between virtual and real shadows, we present an unified scheme to generate the composite image of augmented outdoor scene. Unlike previous works, our method need not to know the geometry of objects which cast shadows nor to apply complex rendering technique.

The main contributions of this work include: (1). we adopt both lighting related prior information and surface radiosity as the basis of illumination estimation, which provides an easy and efficient approach to synthesize high-quality

environment map of real scene from a legacy photograph; (2). we propose an image based method to cast shadows of real scene onto the synthetic objects without knowing the geometry of objects which cast shadows, while at the same time, avoiding the overlap between virtual and real shadows. These contributions lead to a robust and efficient approach for outdoor lighting simulation based on a single image with good application potentials in many fields.

## 2. Outdoor Lighting Simulation

This section will introduce our illumination simulation algorithm. We first present the illumination model of outdoor scenes. Based on this model, a new method is proposed to solve the scene's material and the illumination parameters. This approach requires a few interactions to partially rebuild the scene, and involves the techniques of environment map synthesis and linear constrained least squares.

### 2.1. Model of Outdoor Illumination

Objects in outdoor scenes receive light from the sun and the environment, such as the sky and surrounding objects, the radiosity at point  $p$  in the scene can be expressed as

$$I_p(\lambda) = I_p^{sun}(\lambda) + I_p^{env}(\lambda) \quad (1)$$

where  $\lambda$  denotes the R, G, B channels.

The sun can be simulated as a directional light source, therefore

$$I_p^{sun}(\lambda) = s_p^{sun} L^{sun}(\lambda) [\rho_p(\lambda) (\mathbf{n}_p \cdot \mathbf{l}) + k_p(\lambda) (\mathbf{n}_p \cdot \mathbf{h}_p)^{\alpha_p}] \quad (2)$$

where  $L^{sun}$  is the incident intensity of the sunlight,  $\mathbf{l}$  denotes its direction,  $\mathbf{n}_p$  is the surface normal at  $p$ ,  $\mathbf{h}_p$  is the angular bisector between  $\mathbf{l}$  and the view direction at  $p$ ,  $\rho_p$  is the diffuse coefficients,  $k_p$  and  $\alpha_p$  are the specular properties of  $p$ , and  $s_p^{sun} \in [0, 1]$  is the sun occlusion coefficient.

The environment light can be regarded as an area light source distributed over a sphere, we ignore the specular components of the environment light. Our experiments show that this approximation can reasonably keep the accuracy of the simulation results, while at the same time save a lot of computations. We record the environment light at  $p$  by an environment map and use  $L_p^{env}$  to denote its distribution. Then, we have

$$I_p^{env}(\lambda) = \rho_p(\lambda) \int_{\Omega(\mathbf{n}_p)} L_p^{env}(\omega, \lambda) (\mathbf{n}_p \cdot \omega) d\omega \quad (3)$$

where  $\Omega(\mathbf{n}_p)$  is the upper hemisphere over the surface at which  $p$  locates,  $\omega$  is an unit direction vector. We adopt spherical harmonics  $Y_{lm}$ , with  $l \geq 0$  and  $-l \leq m \leq l$ , to approximate the integral. It has been mentioned that  $I_p^{env}$  can be represented by a linear combination of  $Y_{lm}$  [RH01]

$$I_p^{env}(\lambda) = \rho_p(\lambda) \sum_{l,m} \hat{A}_l L_{lm,p}^{env}(\lambda) Y_{lm}(\mathbf{n}_p) \quad (4)$$

where  $\hat{A}_l$  is related with the combination coefficients of  $A = (\mathbf{n}_p \cdot \omega)$ , which is a constant,  $L_{lm,p}^{env}(\lambda)$  can be calculated as follow [RH01]:

$$L_{lm,p}^{env}(\lambda) = \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} L_p^{env}(\lambda, \theta, \phi) Y_{lm}(\theta, \phi) \sin \theta d\theta d\phi \quad (5)$$

The integral can be viewed as inner-products of the functions  $L_p^{env}$  and  $Y_{lm}$ . On the other hand, the unit sphere surface  $\Omega_p$  can be divided into  $\Omega_p^{sky}$  and  $\Omega_p^{obj}$  corresponding to the sky and the surrounding objects respectively. Therefore

$$\begin{aligned} L_{lm,p}^{env}(\lambda) &= L_{lm,p}^{sky}(\lambda) + L_{lm,p}^{obj}(\lambda) \\ &= \sum_{(\theta, \phi) \in \Omega_p^{sky}} L_p^{env}(\lambda, \theta, \phi) Y_{lm}(\theta, \phi) \sin \theta \Delta \theta \Delta \phi \\ &\quad + \sum_{(\theta, \phi) \in \Omega_p^{obj}} L_p^{env}(\lambda, \theta, \phi) Y_{lm}(\theta, \phi) \sin \theta \Delta \theta \Delta \phi \end{aligned} \quad (6)$$

The sky is a non-uniform area light source, its distribution is hard to be calculated from images. Luckily, Perez et al. [PMI93] explored the sky illumination distribution and proposed a model of sky. This model expresses the absolute luminance  $L_q$  of a sky element  $q$  as a function of the actual zenith luminance  $L_z^{sky}$ , the sun's zenith angle  $\theta_s$ ,  $q$ 's zenith angle  $\theta_q$  and  $q$ 's angle  $\gamma_q$  with respect to the sun.

$$L_q = L_z^{sky} \frac{f(\theta_q, \gamma_q)}{f(0, \theta_s)} \quad (7)$$

where  $f(\theta_q, \gamma_q)$  is the function to calculate the relative luminance of  $q$ . For an incident direction  $(\theta, \phi)$ , denoting

$$F(\theta, \phi) = \frac{f(\theta, \gamma(\theta, \phi))}{f(0, \theta_s)}$$

$\gamma(\theta, \phi)$  is a function to approximate the angle with respect to the sun, therefore

$$L_{lm,p}^{sky}(\lambda) = L_z^{sky}(\lambda) \sum_{(\theta, \phi) \in \Omega_p^{sky}} F(\theta, \phi) Y_{lm}(\theta, \phi) \sin \theta \Delta \theta \Delta \phi \quad (8)$$

Let  $R_{lm,p}^{sky}$  denote the result of the summation in Eq. 8, from Eq. 6 we have

$$L_{lm,p}^{env} = L_z^{sky}(\lambda) R_{lm,p}^{sky} + L_{lm,p}^{obj}(\lambda) \quad (9)$$

Taking it back into Eq. 4, then

$$\begin{aligned} I_p^{env}(\lambda) &= \rho_p(\lambda) [L_z^{sky}(\lambda) \sum_{l,m} \hat{A}_l R_{lm,p}^{sky} Y_{lm}(\mathbf{n}_p) \\ &\quad + \sum_{l,m} \hat{A}_l L_{lm,p}^{obj}(\lambda) Y_{lm}(\mathbf{n}_p)] \end{aligned} \quad (10)$$

Denoting  $P_p^{sky} = \sum_{l,m} \hat{A}_l R_{lm,p}^{sky} Y_{lm}(\mathbf{n}_p)$  and  $E_p^{obj}(\lambda) = \sum_{l,m} \hat{A}_l L_{lm,p}^{obj}(\lambda) Y_{lm}(\mathbf{n}_p)$ , our final illumination model of outdoor scenes is

$$\begin{aligned} I_p(\lambda) &= s_p^{sun} [\rho_p(\lambda) (\mathbf{n}_p \cdot \mathbf{l}) + k_p(\lambda) (\mathbf{n}_p \cdot \mathbf{h}_p)^{\alpha_p}] L^{sun}(\lambda) \\ &\quad + \rho_p(\lambda) P_p^{sky} L_z^{sky}(\lambda) + \rho_p(\lambda) E_p^{obj}(\lambda) \end{aligned} \quad (11)$$

Tab. 1 lists part of the variables in Eq. 11.

**Table 1:** Part of the variables in Eq. 11

Variables	Definition
$P_p^{sky}$	$\sum_{l,m} \hat{A}_l R_{lm,p}^{sky} Y_{lm}(\mathbf{n}_p)$
$E_p^{obj}$	$\sum_{l,m} \hat{A}_l L_{lm,p}^{obj} Y_{lm}(\mathbf{n}_p)$
$R_{lm,p}^{sky}$	The summation in Eq. 8, $I_{lm,p}^{sky} = L_z^{sky} R_{lm,p}^{sky}$
$L^{sun}, L_z^{sky}$	Illumination parameters(Eq. 2, Eq. 7)
$\hat{A}_l, L_{lm,p}^{env}$	Combination coefficients of $I_p^{env}$ (Eq. 4)
$L_{lm,p}^{sky}, L_{lm,p}^{obj}$	$L_{lm,p}^{env} = L_{lm,p}^{sky} + L_{lm,p}^{obj}$ (Eq. 6)

## 2.2. Initialization Data Acquisition

This section will describe the process we conduct to acquire the initialization data which will be used to calculate the parameters of the outdoor illumination model.

**Geometry Estimation.** To partially recover the coarse geometry of the scene, the camera needs to be calibrated. We estimate the vanishing points of the scene automatically [HHF09] or interactively, the camera parameters can then be obtained. To simplify the geometry reconstruction, we set the equation of the ground plane as  $z = 0$ . Then the user draws several lines on the ground, each corresponding to the intersection of the ground and a vertical plane. These planes represent a coarse geometry of the scene. With these geometric information and camera pose, the item of  $(\mathbf{n}_p \cdot \mathbf{h}_p)$  in Eq. 11 can be calculated.

**Environment Map Synthesization.** To estimate  $E_p^{obj}(\lambda)$  in the model, we should synthesize an environment map without sky area. A sky mask is approximated [HEH07] to remove the sky area. Then the desired environment map at  $p$  can be generated by Lalonde's method [LE10]. This environment map will be used to calculate  $L_{lm,p}^{obj}(\lambda)$  according to Eq. 6, after that,  $E_p^{obj}(\lambda)$  can be acquired (Eq. 10).

**Sun Direction Estimation.** The sun is a directional light source. As indicated by Perez's sky model, the illumination distribution of the sky is determined by the sun's position. Landonde et al. [LEN11] proposed an automatic approach to estimate the sun's position based on a single outdoor image, which is also adopted by our method. However, this method requires sky, shadows on the ground and multiple vertical planes in the scene. For an image that does not contain these three parts, the user should point the sun direction manually, which is also very simple. The user just needs to draw two lines to indicate a vertical object and its shadow. After this step, both  $P_p^{sky}$  and  $(\mathbf{n}_p \cdot \mathbf{l})$  are available.

**Sun Occlusion Coefficient Approximation.** Let  $C^{sun}$  be an image of the scene illuminated only by the sunlight without shadow. Adopting the symbols used in Eq. 1 and Eq. 2, we have

$$\begin{aligned} I_p(\lambda) &= s_p^{sun} C_p^{sun}(\lambda) + I_p^{env}(\lambda) \\ &= s_p^{sun} (C_p^{sun}(\lambda) + I_p^{env}(\lambda)) + (1 - s_p^{sun}) I_p^{env}(\lambda) \end{aligned} \quad (12)$$

Obviously,  $I$  is composed by  $(C^{sun} + I^{env})$ (shadow removed image) as foreground and  $I^{env}$ (shadow image) as background. Therefore, the sun occlusion coefficient can be solved by a matting algorithm. We use Levin's method [LLW08] to get a shadow matte image, this method requires the user to indicate the foreground(unshadowed areas) and background(shadow areas) by a few scribbles. Alternatively, we can also adopt an automatic shadow matting algorithm to estimate the occlusion coefficient, however, the automatic method appears less stable than Levin's method.

## 2.3. Illumination Estimation

Now the unknowns in Eq. 11 are  $L^{sun}(\lambda)$ ,  $L_z^{sky}(\lambda)$ ,  $\rho_p(\lambda)$  and  $\alpha_p$ . This section will describe how to estimate them. We first propose an algorithm to estimate the reflectance of the scene, then the illumination parameters will be solved with the recovered material.

### 2.3.1. Recovering the Material Reflectance

We start the algorithm by assuming a perfect diffuse case, then consider the case of specular surfaces.

#### The Case of Diffuse Material

The illumination model can be written as follow if we ignore the specular component

$$I_p(\lambda) = \rho_p(\lambda) [P_p^{sun} L^{sun}(\lambda) + P_p^{sky} L_z^{sky}(\lambda) + E_p^{obj}(\lambda)] \quad (13)$$

where  $P_p^{sun} = s_p^{sun} (\mathbf{n}_p \cdot \mathbf{l})$ . Divided by  $\rho_p(\lambda)$  for both sides of Eq. 13

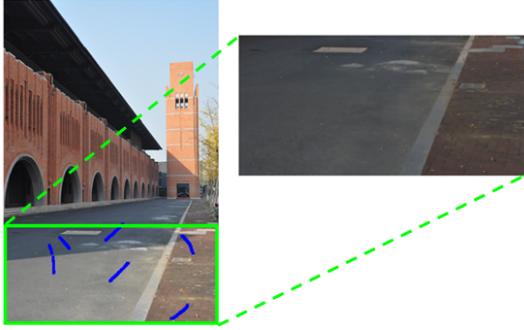
$$P_p^{sun} L^{sun}(\lambda) + P_p^{sky} L_z^{sky}(\lambda) - I_p(\lambda) \frac{1}{\rho_p(\lambda)} = -E_p^{obj}(\lambda) \quad (14)$$

Note that this is a linear equation with unknowns of  $L^{sun}(\lambda)$ ,  $L_z^{sky}(\lambda)$  and  $1/\rho_p(\lambda)$ , if we choose  $n$  sample points in the scene, a system of linear equations can be set up. However, there are  $3n$  equations with  $3n+6$  unknowns, which are under constraint. Fortunately, different points in the scene may share the same material, this will reduce the number of unknowns dramatically. Similar as Rousseau's method [BPD09], we let the user choose points with the same material by a brush. For simplicity, we also assume that all points covered by a stroke share the same environment map.

Additional constraints are also employed to improve the accuracy of the approximation. We first explore the color constraints due to the spectral distribution of the daylight. For sunlight we have  $L^{sun}(R) > L^{sun}(G) > L^{sun}(B)$ , and  $L_z^{sky}(R) < L_z^{sky}(G) < L_z^{sky}(B)$  for the skylight;  $\rho_p(\lambda)$  is the diffuse coefficient of  $p$ , so it must be  $1/\rho_p(\lambda) \geq 1$ . In practice, setting  $1/\rho_p(\lambda) \geq 5$  will get a better visual effect of the estimated sky.

This linear system of equations can be solved by the

method of liner constrained least squares. The diffuse coefficients of more points in the scene can be solved from Eq. 13 with the estimated  $L^{sun}(\lambda)$  and  $L_z^{sky}(\lambda)$ . Nevertheless, it is unnecessary to use the whole image to calculate the illumination, so we only recover the diffuse coefficients of the points within the bounding box of the strokes. Fig. 2 shows our interactions and the recovered material.



**Figure 2:** The blue strokes in the images are drawn by user, indicating that points covered by the same stroke consist of the same material. The recovered diffused albedo of points on the ground are shown in the right image.

### The Case of Specular Material

Our method can also account for material of less strong specular reflectance. Similar to the method by Boivin et al. [BG01], we first ignore the specular reflection component of these surfaces to get an approximate material parameters  $\rho^0$  and illumination parameters  $L^0$  as the initial values, the final material is then solved by an procedural refinement approach. There are two procedures: (1). updating the material parameters; (2). updating the illumination parameters. These two procedures are implemented alternatively until convergence.

**Updating The Material Parameters.** The material parameters of specular surfaces include  $\rho_p(\lambda)$ ,  $k_p(\lambda)$  and  $\alpha_p$ . We make an approximation here, setting  $\alpha_p$  as a constant for these surfaces in the scene and assuming that  $k_p(\lambda)$  varies spatially slow. We also adopt the assumption of intrinsic image that the diffuse component consists of piecewise constant patches with potentially sharp boundaries. Using the initial material and illumination parameters, an objective function can be developed to solve  $\rho_p(\lambda)$ ,  $k_p(\lambda)$  and  $\alpha_p$ .

$$\begin{aligned} \arg \min_{\rho, k, \alpha} & \sum_{p \in Q} \sum_{\lambda \in C} |I_p(\lambda) - I_p^*(\lambda)|^2 + \mu_1 w_p |\nabla \rho_p(\lambda)|^2 \\ & + \mu_2 (\rho_p(\lambda) - \rho_p^0(\lambda))^2 + \mu_3 |\nabla k_p(\lambda)| \end{aligned} \quad (15)$$

where  $C = \{R, G, B\}$  and  $Q$  is the selected point set.  $\mu_1, \mu_2, \mu_3$  are weights, and they are set as  $\mu_1 = 5.0, \mu_2 = 3.0, \mu_3 = 5.0$  in our experiments. The first term ensures the synthesized image  $I^*$  conforming with the original image; the third term

makes sure that the estimated  $\rho_p(\lambda)$  stays near its initial value; the second and last terms enforce the smoothness of  $\rho$  and  $k$ ,  $w_p$  is a weighting function to penalize the pixels with large gradient magnitude, it is defined as:  $w_p = 1 - 1/(1 + \exp(-s(|\nabla I|^2 - c)))$ , setting  $s = 20, c = 0.15$  in our implementation.

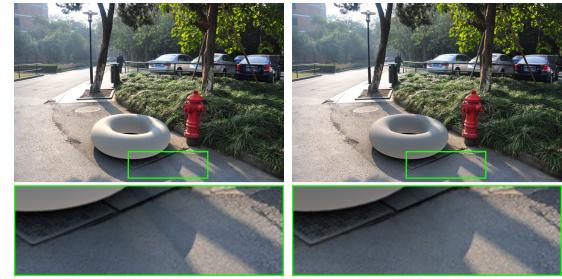
**Updating The Illumination Parameters.** After the material parameters are available, we rewrite Eq. 11 as

$$M_p^{sun}(\lambda) L^{sun}(\lambda) + M_p^{sky}(\lambda) L_z^{sky}(\lambda) = B_p(\lambda) \quad (16)$$

where

$$\begin{aligned} B_p(\lambda) &= I_p(\lambda) - \rho_p(\lambda) E_p^{obj}(\lambda) \\ M_p^{sky}(\lambda) &= \rho_p(\lambda) P_p^{sky} \\ M_p^{sun}(\lambda) &= s_p^{sun} [\rho_p(\lambda) (\mathbf{n}_p \cdot \mathbf{l}) + k_p(\lambda) (\mathbf{n}_p \cdot \mathbf{h}_p)^{\alpha_p}] \end{aligned}$$

We sample points uniformly from  $Q$  and embed them into a linear system of equations. Adopting the color constraints again, the illumination parameters can be obtained by solving this linear constrained least squares problem. Fig. 3 shows the synthetic results adopting the recovered illumination with and without considering the specular attribute of the material.



**Figure 3:** The torus is the virtual object inserted into the image and rendered with the recovered illumination. It can be found that the overlapping area of tour's shadow and tree's shadow on the ground in (a) appears less realistic than that in (b).

### 2.3.2. Illumination Parameters Estimation

We use the same method for updating illumination parameters described in Sec. 2.3.1 (Eq. 16) to solve the illumination with the estimated surface reflectance. While  $L^{sun}$  indicates the intensity of the directional sunlight,  $L_z^{sky}(\lambda)$  is used to generate a complete HDR environment map to simulate the environment light. These two light sources are used to simulate the outdoor illumination. Fig. 4 shows the environment maps of different scenes synthesized by our method.



**Figure 4:** The final environment maps generated by our method, we only demonstrate LDR maps corresponding to a particular exposure time. The appearance of the sky is determined by  $L_z^{\text{sky}}(\lambda)$  and Perez's sky model, while the surrounding objects are generated by Lalonde's method [LE10].

### 3. Shadow Modeling

The sun is the primary light source of outdoor scenes, when the sunlight incident to the scene is partially or totally occluded, shadows will occur. A synthetic object will cast shadow onto the real environment, also it may receive shadows cast from real objects. This section tries to simulate this phenomenon. We first propose a new approach to process the real and virtual shadows, then illustrate a method to insert the virtual objects as well as their shadows into the video.

#### 3.1. Virtual Object in Real Shadow

Shadows on the virtual object can be synthesized if the geometry of the occluding objects is known. Unfortunately, this information is unavailable on the legacy image. We present a novel method to cast the real shadows onto the surface of a virtual object. The coarse geometry and the shadow matte image(a scalar image that is 0 at shadow pixels, and  $(0,1]$  otherwise) calculated in Sec. 2.2 will be employed. We denote the shadow matte image as  $M_S^{\text{img}}$  for convenience.

We only consider the real shadows that are on the ground, the method can be extended to complex situation if the geometry of the shadowed surface is known. Supposing we have inserted a virtual object into the scene. We cast a ray  $r$  from a vertex  $v$  on the virtual object along the sunlight's direction, and calculate its intersection with the ground plane, denoting the intersection point as  $i$ . If  $i$  is in shadow,  $v$  must be in shadow provided that there is no other objects between  $v$  and  $i$ . It implies that whether  $v$  is in shadow or not depends

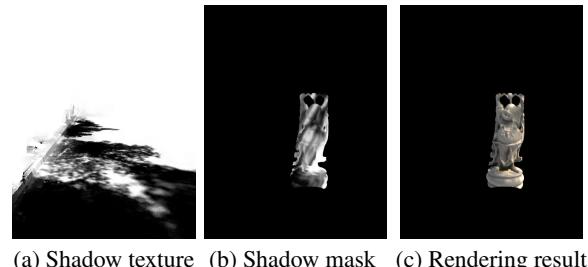
solely on the status of  $i$ , which can be detected by checking the corresponding pixel  $p_i$  in  $M_S^{\text{img}}$ .

The next question is how to synthesize the shadow on the occluded virtual object. Note that the relationship among  $v$  and  $i$  can be defined as a texture map. With this map, we can project the shadow matte image onto the surface of the virtual object automatically. Since the shadow area is illuminated by environment light only, this area can be represented by a shadow mask image denoted as  $M_S^{\text{obj}}$ . Then we render the virtual object twice to achieve a complete rendering result of the virtual object, one accounts for the illumination of the sunlight, and the other accounts for that of the environment light, let the two components be denoted as  $I_{\text{obj}}^{\text{sun}}$  and  $I_{\text{obj}}^{\text{env}}$  respectively. The final image  $I_{\text{obj}}$  can then be expressed by:

$$I_{\text{obj}} = M_S^{\text{obj}} \odot I_{\text{obj}}^{\text{sun}} + I_{\text{obj}}^{\text{env}} \quad (17)$$

where  $\odot$  is the Hadamard product.

Note that, the above method will fail if the shadow on the ground is casted by some object  $O$  located between  $v$  and  $i$ , which is a limitation of our method. To avoid this, the user needs to indicate these shadow areas manually, and remove them from  $M_S^{\text{img}}$ . The final shadow image is called shadow texture. Fig. 5(a) and (b) shows the shadow texture and the shadow mask estimated by our method, the rendering result of the virtual object is illustrated in Fig. 5(c).



**Figure 5:** We try to integrate the stanford buddha into a scene. Our shadow mask is generated without knowing the geometry of the occluded object. The rendering result demonstrates that shadows on the ground are cast onto the virtual object naturally.

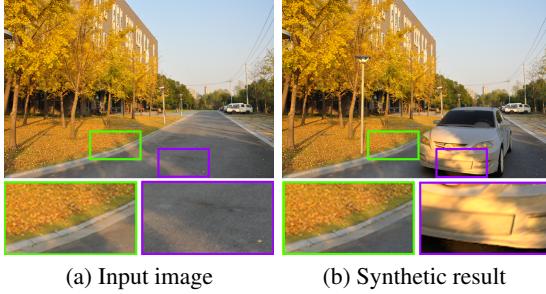
#### 3.2. Virtual Shadow Processing

Alternatively, virtual object may also cast shadow on the environment, but it can take effect only on the pixels whose  $M_S^{\text{img}}$  value is not 0. Since we have got the shadow matte image  $M_S^{\text{img}}$ , the error of overlap between virtual and real shadows can be avoided. Our method composes four images: the first two are images of the original scene illuminated by the sunlight and the environment light, which are

denoted by  $I_{scene}^{sun}$  and  $I_{scene}^{env}$  respectively. The second two, denoted by  $I_{v-scene}^{sun}$  and  $I_{v-scene}^{env}$ , are images of the augmented scene containing virtual object, which are illuminated by the sunlight and the environment light respectively. With the shadow matte image  $M_S^{img}$ , we can generate an augmented shadow mask of the original image, its value at pixel  $p$  is calculated by

$$I_S(p) = \frac{I_{v-scene}^{sun}(p)M_S^{img}(p) + I_{v-scene}^{env}(p)}{I_{scene}^{sun}(p)M_S^{img}(p) + I_{scene}^{env}(p)} \quad (18)$$

For a pixel  $p_s$ , if  $M_S^{img}(p_s) = 0$ , no matter the virtual objects will cast shadows on  $p_s$  or not, the value of  $I_S(p_s)$  is  $I_{v-scene}^{env}(p_s)/I_{scene}^{env}(p_s)$ . The  $I_{v-scene}^{sun}$  will take effect for the pixel  $p$  if  $M_S^{img}(p) \neq 0$ . (Fig. 6).



**Figure 6:** Our method avoids the problem of overlap between real and virtual shadows. (a) the input image, (b) the synthetic result, the detail of the shadows are shown at the bottom of the image, the shadow of the synthetic car takes effect on the ground which is not covered by real shadow.

### 3.3. Inserting Synthetic Object

After our approach obtains the parameters of outdoor illumination, detects shadow areas on the ground and reconstructs partially the geometry of the scene, the user can freely insert synthetic objects into the scene. Let the legacy photograph of the outdoor scene be  $I$ , the final composite image  $I_{final}$  can be generated by

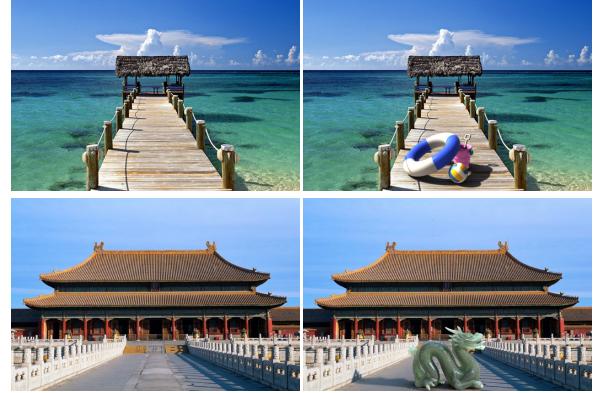
$$I_{final} = M \odot I_{obj} + (1 - M) \odot I_S \odot I \quad (19)$$

where  $M$  is the mask of the virtual object(a scalar image that takes 0 at pixels outside the projected area of the virtual object, and (0,1] otherwise).

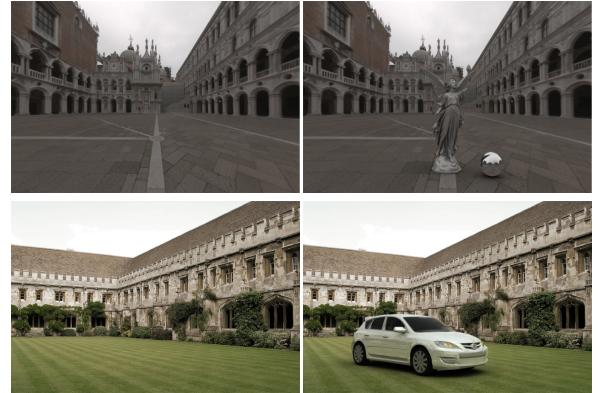
## 4. Experiments

We implemented the proposed approach on a PC with Core2 Duo E7400 2.80GHz CPU and 3GB of RAM. We selected several images of outdoor scenes, and recovered their lighting environments with our approach.

We rendered some virtual objects with the estimated illumination and integrated them into outdoor images captured either in sunny days (Fig. 7) or in overcast days (Fig. 8). The recovered illumination contains the environment information, enabling the generation of high quality rendering results. Two animation sequences of virtual objects were produced to examine the shadow casting on moving objects, Fig. 9 shows some video frames in the synthetic video sequences.



**Figure 7:** Virtual objects with complex material are inserted into images captured in sunny days.



**Figure 8:** Virtual objects are integrated into outdoor images captured in overcast days.

We compared the shading of the virtual objects under the illumination recovered by our method with those by assuming a uniform hemisphere light source, using light probe and rendered by [KHFH11]'s method (Fig. 10). It is observed that the uniform light source cannot accurately simulate the illumination distribution of the real scene, the bottom of the synthetic objects is very dark, as the uniform light hemisphere does not account for the light reflected by ground. Light probe tries to record real illumination of the scene. We

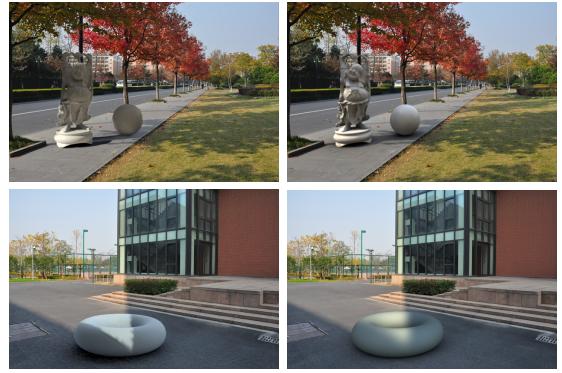


**Figure 9:** We embed animations into two background images. The first row shows a rolling basketball, and the second row demonstrates a jumping football.

employed a software named HDR SHOP to synthesize the environment map of the scene corresponding to light probe. Since the sun and its surrounding area are too bright to be captured by a camera, consequently, the sun was recorded as an area light source in the HDR environment map, leading to unrealistic shadows and bright areas on the surface of virtual objects. On the other hand, [KHFH11]’s method sets the intensity of the sunlight manually, therefore, it is very hard to keep the color of virtual shadow areas consistent with the surrounding environment. Our method recovers both the directional light source and an environment map, hence achieving a better rendering result. In addition, we use environment map as the light source, which accelerates the rendering of virtual objects. Even for complex model, the rendering procedure costs only a few minutes. [KHFH11]’s method needs to render the whole scene accounting for global illumination, and it usually takes several hours.

We also compare the shadow interactions between virtual objects and real scene simulated by our method and by [KHFH11]’s method (Fig. 11). We insert the stanford buddha and a ball into the image shown in the first row. Since [KHFH11]’s method cannot accurately recover the intensity of the sunlight, the appearance of buddha’s shadow appears different from the real case. The virtual object in the second row is a torus, apparently, our method generates better effect in soft shadow processing.

Our synthetic result is also compared with the ground truth. We employed several real plaster models as the reference objects, and took photos of different outdoor scenes containing those models as the ground truth. We then took photos of the same scenes but without these models as background image. These plaster models were inserted as virtual objects into the background images. The rendering results

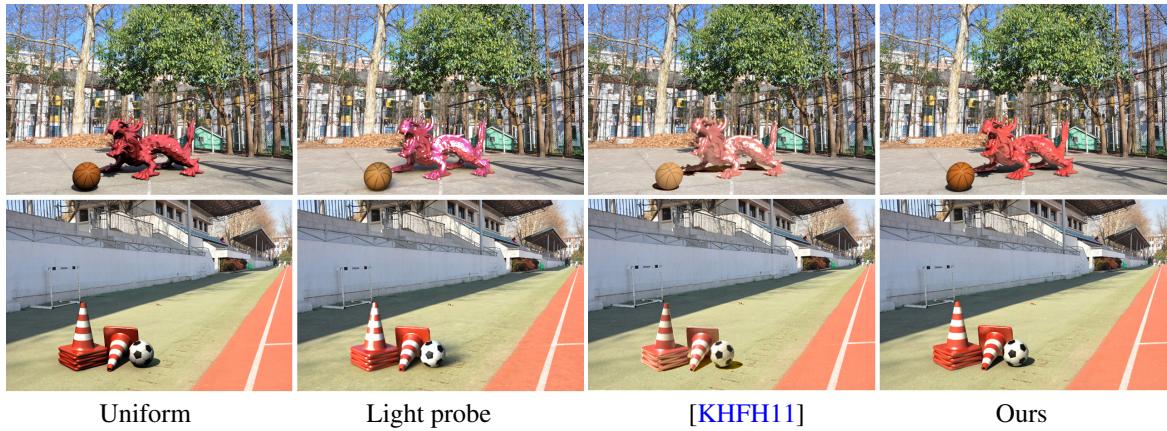


**Figure 11:** Comparison between our shadow modeling results and that by [KHFH11]’s method.

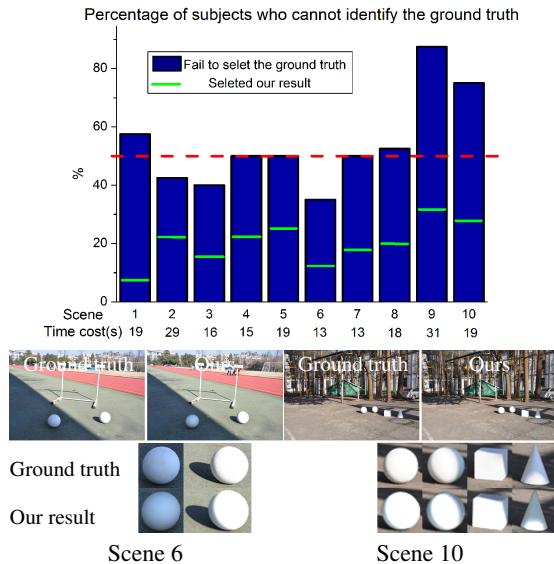
of the augmented scene are compared with the ground truth. We assume initially an uniform Lambert reflectance of the plaster models and adjust their reflectance coefficients interactively until they can match one chosen test image successfully. Fig. 12 shows two of our comparison results. More test scenes can be found in our demonstrate video.

No wonder our recovered illumination may differ from the real lighting condition, however, this should not prevent us to get a realistic synthetic result. An user study was conducted to measure how well user can differentiate the ground truth and the synthetic result. 40 graduate students were recruited as the subjects for this task, most of them had a graphics background. Our test adopted ten different scenes, all of these scenes were presented in a randomly permuted order. For each scene, both of the ground truth and the synthetic result are demonstrated, and the image placement (left or right side) was also randomized. The users were told that there are two situations for each pair of images: (1). one is real photo, the other is a synthetic result; (2). both are real photos. Therefore, there are three options for the users: the left image is a real photo, the right image is a real photo and both are real photos. The result of the user study is shown in Fig. 12. It is found that over half of the users can hardly identify the ground truth in most scenes, a part of the users even thought that the synthetic results are more realistic. For scene 6, 65% of the users can identify the real photo, the evidence for them is that the shadowed ball is too smooth compared to the ground truth. Most users acknowledge that it is very difficult to find out the abnormal of the synthetic image without comparing to the ground truth.

Note that our algorithm recovers the material at some regions of the image only, we sample pixels uniformly from the chosen pixel set (one for every 5 pixels in our test) for calculating the illumination parameters. The time costs of choosing different size of pixel set are shown in Tab. 2. The time cost regarding the specular material refers to one iter-



**Figure 10:** Comparison between our synthetic results and those by other methods'. From left to right, synthetic images produced by uniform light source, light probe, [KHFH11]'s and our method. We insert a dragon and a basketball into the first scene, while a roadblock and a football into the second scene.



**Figure 12:** The result of an user study. For each scene, the blue bar represents the percentage of users who can hardly identify the ground truth; while the green line represents the percentage of users who think our result is more realistic. The horizontal red line indicates the line of 50%. The average time cost to make a decision for each scene is also demonstrated. The images below the graph show two scenes we used in the user study.

**Table 2:** Time cost of illumination calculation

Number	Diffuse (s)	Specular (s)	Illumination (ms)
70 × 70	2.86	2.30	468
110 × 110	3.75	4.83	812
230 × 230	5.63	20.43	2158
340 × 340	7.83	41.03	3016

taining 50000 vertices, it costs only half a second, and 2.8 seconds for a model containing 263144 vertices.

Although the estimation of specular material in our approach is time consuming, the calculation for recovering the diffuse material and illumination parameters is efficient. Since, many objects in outdoor scenes are nearly diffuse, therefore, our approach runs reasonably fast for most outdoor images. Currently, we adopt uniform sampling, an adaptive sampling strategy will reduce the number of requested samples, allowing our method applicable for on-line video lighting recovery.

## 5. Conclusion and Future Work

We have presented a new method to simulate the illumination of an outdoor scene based on a legacy photograph. We employ lighting related prior information and incorporate them into the rendering equation to derive a model of outdoor illumination. Spherical harmonics is adopted to reduce the complex lighting information extracted from the image into a set of parameters, which makes the illumination recovery much easier. Our method estimates the reflectance of diffuse surfaces by solving a linear problem, while performing an iterative approach in the case of specular material. The illumination can then be recovered by solving

tion of the refinement procedure, the iterations are conducted five to ten times in most of our examples. For shadow casting, its efficiency depends mainly on the time of creating the texture map, which runs very quickly. For a model con-

a liner least square problem under some constraints. With these illumination parameters we can build a high quality environment map of the scene to support realistic rendering. We also explore the problem of shadow casting which may arise when virtual objects are inserted into real images. A novel texture mapping based approach is proposed to cast real shadows onto the synthetic objects. Finally, we adopt a unified scheme to generate the composite image with correct shadow casting between virtual and real objects. Experimental results demonstrate the validity and flexibility of our approach.

However, there are several limitations in our work. The camera calibration and the geometry recovery may fail if the available information is not enough to determine vanishing points. Our method cannot handle the scenes with complex texture, since it is very hard to choose points with similar material. More over, our method is not yet applicable to surfaces with strong specular reflection. Our shadow modeling method also needs some further enhancement to deal some particular cases.

### Acknowledgements

This research is supported by 973 Program of China (Grant No.2009CB320802), 863 Program of China (Grant No.2013AA013902) and National Natural Science Foundation of China (Grant No.61173070, 61103137).

### References

- [BG01] BOIVIN S., GAGALOWICZ A.: Image-based rendering of diffuse, specular and glossy surfaces from a single image. In *Proc. ACM Siggraph* (August 2001), Los Angeles, California, USA, pp. 107–116. 5
- [BPD09] BOUSSEAU A., PARIS S., DURAND F.: User-assisted intrinsic images. *ACM Transactions on Graphics* 28, 5 (December 2009), 130. 4
- [Deb98] DEBEVEC P.: Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *Proc. SIGGRAPH'98* (July 1998), Orlando, Florida, USA, pp. 189–198. 2
- [GM00] GIBSON S., MURTA A.: Interactive rendering with real-world illumination. In *Proceedings of the Eurographics Workshop on Rendering Techniques* (2000), pp. 365–376. 2
- [HEH07] HOIEM D., EFROS A. A., HEBERT M.: Recovering surface layout from a single image. *International Journal of Computer Vision* 75, 1 (October 2007), 151–172. 4
- [HFC\*09] HABER T., FUCHS CHRISTIAN, BEKAER P., SEIDEL H.-P. P., GOESELE M., LENSCHE H. P. A.: Relighting objects from image collections. In *Proceeding CVPR09* (june 2009), Miami Beach, Florida, USA, pp. 627–634. 2
- [HHF09] HEDAU V., HOIEM D., FORSYTH D.: Recovering the spatial layout of cluttered rooms. In *Proc. ICCV* (September 2009), Champaign, USA, pp. 1849–1856. 4
- [JAL05] JACOBS K., ANGUS C., LOSCOS C.: Automatic generation of consistent shadows for augmented reality. In *Proceedings of Graphics Interface 2005* (May 2005), Waterloo, Ontario, Canada, pp. 113–120. 2
- [JL06] JACOBS K., LOSCOS C.: Classification of illumination methods for mixed reality. *Computer Graphics Forum* 25, 1 (2006), 29–51. 2
- [KHFH11] KARSCH K., HEDAU V., FORSYTH D., HOIEM D.: Rendering synthetic objects into legacy photographs. In *Proc. ACM Siggraph Asia* (December 2011), Hong Kong, China, pp. 157:1–157:12. 2, 7, 8, 9
- [KTM\*10] KNECHT M., TRAXLER C., MATTAUTSCH O., PURGATHOFER W., WIMMER M.: Differential instant radiosity for mixed reality. In *Proceedings of ISMAR* (October 2010), Miami Beach, Florida, USA, pp. 99–107. 2
- [LE10] LALONDE J.-F., EFROS A. A.: *Synthesizing Environment Maps from a Single Image*. Tech. Rep. CMU-RI-TR-10-24, Robotics Institute, Carnegie Mellon University, July 2010. 4, 6
- [LEN09] LALONDE J.-F., EFROS A. A., NARASIMHAN S. G.: Webcam clip art: appearance and illuminant transfer from timelapse sequences. *ACM Transactions on Graphics* 28, 5 (December 2009), 131:1–131:10. 2
- [LEN11] LALONDE J.-F., EFROS A. A., NARASIMHAN S. G.: Estimating the natural illumination conditions from a single outdoor image. *International Journal of Computer Vision* 98, 2 (2011), 123–145. 4
- [LG12] LIU Y., GRANIER X.: Online tracking of outdoor lighting variations for augmented reality with moving cameras. *IEEE Trans. Vis. Comput. Graph* 18, 4 (April 2012), 573–580. 2
- [LLW08] LEVIN A., LISCHINSKI D., WEISS Y.: A closed-form solution to natural image matting. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30, 2 (2008), 228–242. 4
- [LQX\*09] LIU Y., QIN X., XU S., EIACHIRO N., PENG Q.: Light source estimation of outdoor scenes for mixed reality. *Vis. Comput.* 25, 5-7 (April 2009), 637–646. 2
- [MJA06] MADSEN C. B., JENSEN T., ANDERSEN M. S.: Real-time image-based lighting for outdoor augmented reality under dynamically changing illumination conditions. In *Proceedings: International Conference on Graphics Theory and Applications* (February 2006), Setubal, Portugal, pp. 364–371. 2
- [NGM\*11] NOWROUZEZHAI D., GEIGER S., MITCHELL K., SUMNER R., JAROSZ W., GROSS M.: Light factorization for mixed-frequency shadows in augmented reality. In *Proceedings of ISMAR* (Oct. 2011). 2
- [PMI93] PEREZ R., M R. S. J., INEICHEN P.: An all-weather model for sky luminance distribution. *Solar Energy* 50, 3 (March 1993), 235–245. 2, 3
- [RH01] RAMAMOORTHI R., HANRAHAN P.: On the relationship between radiance and irradiance: determining the illumination from images of a convex lambertian object. *Journal of the Optical Society of America A* 18, 10 (October 2001), 2448–2459. 3
- [SSI99] SATO I., SATO Y., IKEUCHI K.: Acquiring a radiance distribution to superimpose virtual objects onto a real scene. *IEEE TVCG* 5, 1 (March 1999), 542–547. 2
- [SSI03] SATO I., SATO Y., IKEUCHI K.: Illumination from shadows. *IEEE Trans. Pattern Analysis and Machine Intelligence* 25, 3 (March 2003), 290–300. 2
- [YWAC06] YU T., WANG H., AHUJA N., CHEN W.-C.: Sparse lumigraph relighting by illumination and reflectance estimation from multi-view images. In *Proceeding Siggraph2006* (july 2006), Boston, Massachusetts, USA, pp. 175–175. 2
- [ZY01] ZHANG Y., YANG Y.-H.: Multiple illuminant direction detection with application to image synthesis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23, 8 (August 2001), 915–920. 2