Vis Comput (2009) 25: 637–646
DOI 10.1007/s00371-009-0342-4

ORIGINAL ARTICLE

Light source estimation of outdoor scenes for mixed reality

Yanli Liu · Xueying Qin · Songhua Xu · Eihachiro Nakamae · Qunsheng Peng

Published online: 3 March 2009

Abstract Illumination consistency is important for photo-realistic rendering of mixed reality. However, it is usually difficult to acquire illumination conditions of natural environments. In this paper, we propose a novel method for evaluating the light conditions of a static outdoor scene without knowing its geometry, material, or texture. In our method, we separate respectively the shading effects of the scene due to sunlight and skylight through learning a set of sample images which are captured with the same sun position. A fixed illumination map of the scene under sunlight or skylight is then derived reflecting the scene geometry, surface material properties and shadowing effects. These maps, one for sunlight and the other for skylight, are therefore referred to as basis images of the scene related to the specified sun position. We show that the illumination of the same scene under different weather conditions can be approximated as a linear combination of the two basis images. We further extend this model to estimate the lighting condition of scene images under deviated sun positions, enabling virtual objects to be seamlessly integrated into images of the scene at any time. Our approach can be applied for online video process and deal with both cloudy and sun shine situations. Experiment results successfully verify the effectiveness of our approach.

Keywords Light source recovery · Outdoor scenes · Augmented reality

Y. Liu · X. Qin (✉) · S. Xu · Q. Peng
State Key Lab of CAD&CG, Zhejiang University, Hangzhou, People's Republic of China
e-mail: xyqin@cad.zju.edu.cn

Q. Peng (✉)
e-mail: peng@cad.zju.edu.cn

X. Qin
School of Computer Science & Technology, Shandong University, Jinan, People's Republic of China

E. Nakamae
Sanei Co., Hiroshima, Japan

1 Introduction

Augmented Reality (AR) is a rapidly growing research field of the mixed reality, in which the seamless integration of synthetic objects into a live video is highly desired. To meet this goal, not only the geometry but also the illumination of the virtual objects should be consistent with those of the real environment displayed in the video.

Achieving illumination consistency in augmented reality is a challenging task since the lighting conditions of the real environment in the live video are usually unknown. In outdoor scenes, the only direct light source is the sunlight, which lit the atmosphere to produce the skylight. As the orbit of the sun is fixed and its position is deterministic at any time during a day, the light conditions mainly due to sunlight can be easily calculated. However, during cloudy days the skylight plays an important role for object illumination. Moreover, the lighting conditions often change dramatically due to cloud motion. These increase the difficulty to recover lighting conditions of the scene.

Nakamae et al. [15] contributed an early work on illumination consistency for merging virtual objects into images. In general, there have been three types of approaches to acquire light sources, including measurement-based approaches, physical model-based approaches and image estimation-based approaches. Also, inverse rendering and image-based relighting largely address a similar problem. However, those methods either request that a 3D model

of the environment to be known in advance or can only deal with small indoor environments such as labs.

In this paper, we focus on estimating the light source of static outdoor scenes for online videos captured at a fixed viewpoint. Our key idea is to derive lighting conditions by learning from a set sample images of the same view under different weather and light conditions. A novel framework is proposed to estimate the lighting conditions of outdoor scenes, i.e. the intensity of the sunlight and the skylight. Our main contributions include: (1) a separation model of light and basis images for outdoor scenes to recover the lighting conditions under various weather conditions; (2) extension of the model to account for sun position variation; and (3) a reconstruction framework to acquire the lighting conditions of a live image.

Live videos that we deal with at this paper are about static outdoor scenes containing buildings, trees and sky. The videos are captured with a camera mounted on a tripod. Our method consists of two stages: an off-line learning stage and an online registration stage. In the off-line stage, we construct an image database and obtain basis images with respect to the different sun positions. During the online stage, for each input frame, we calculate its light parameters based on the knowledge learned in the off-line stage. In this paper, foggy, raining, and snowing weather conditions are not considered.

The rest of this paper is organized as follows. Section 2 introduces most related work to our studies here. Section 3 presents our model and method to estimate basis images and light source from a set of images with the same sun position. Section 4 describes an extension to our method to account for the deviation of sampled sun positions. Section 5 explains the implementation of our method. Section 6 presents some experiment results and discussions. Finally, in Sect. 7, we conclude the paper and point out some future work directions.

2 Related work

The appearance of an image regarding a static view varies greatly with the environment's lighting conditions. The methods of light source estimation are studied by researchers in several fields, such as inverse rendering, shape from shading, and intrinsic image analysis.

In inverse rendering [3], direct measurement method of the light sources can be employed to reduce the computation complexity [13, 29]. However, these approaches are obviously impractical for outdoor scenes. Alternatively, Debevec et al. [4] used the HDR environment map to capture omnidirectional incident light in a complex environment. This image-based method has been widely used to facilitate acquisition of lighting conditions, e.g. [1, 6, 19]. However, HDR is not suited for outdoor scenes because the sunlight in general is too bright, leading to a saturated image even when the HDR image is captured using a very fast shutter speed. Moreover, if the sky is clouded or the clouds drift in the sky, there will inevitable be some movement in the low dynamic images used to compile the HDR image, making resultant HDR images worthless.

The shading of an object in a photograph is jointly dependent on the BRDF and geometry of the object as well as the light source. In principle, the light source can be solved if the geometry and surface BRDF of the illuminated object are available. Such physical model-based solutions observe the cues of shading [2], shadows [20, 27], critical points [30]. Other inverse rendering methods, such as the ones proposed by Seitz et al. [21] and Nayar et al. [17], separate the illumination in a scene into its direct and global components using controlled lighting. Many methods which aim at recovering shapes from shading also estimate light sources, e.g. Sharma [22] and Hara [7]. Most of the physical model-based approaches work effectively only for indoor scenes such as lab environments since they usually require the 3D geometries of the scene to be known. For more illumination estimation approaches in mixed reality, please refer to [8].

Recently intrinsic images analysis have drawn a lot of research attentions, which separate an image into reflectance and illumination effects. Intrinsic image can be drawn from a single image [26] or image sequences [28] by adding some constraints. Matsushita et al. [14] proposed an image-based decomposing method to analyze the intrinsic component of image sequences for any daytime sequences. However, existing intrinsic image decomposition approaches are not robust enough to deal with natural images captured under little control.

Most recently, people have actively studied image sequences captured of outdoor scenes, e.g. constructing useful database [16], segmenting outdoor images [9, 12], recovering camera properties from outdoor images [10, 11]. Sunkavalli et al. [23] decomposed a video sequence to three images, two basis curves, and a compressed representation for shadows. Their method can recover any images in the sequence given these representations. However, this method is only for sunshine day, and this restriction is later removed by using a more accurate model in [24]. Although these methods analyzed the component of the scenes, no explicit lighting conditions are drawn and transferred to virtual objects. Moreover, since both approaches are performed for spatial-temporal data, they are only applicable for postprocessing outdoor videos and not suitable for solving our problem.

To solve for the illumination of outdoor scenes, two major challenges exist. First, geometries of 3D objects in nature scenes such as large scale buildings, various trees are usually difficult to acquired. Second, outdoor scenes are usually involved with complex illumination situations, including both

sunlight and skylight. Solving a global illumination equation of outdoor scenes is time consuming, which does not meet the real-time requirement of augmented reality applications.

3 Ideal light source recovering

To ensure the illumination consistency between the virtual objects and the background images, we first extract useful information from a number of sampling images of the same scene captured at different time.

3.1 Illumination model of outdoor scenes

The light source of outdoor scenes consists of the sunlight, skylight, and environment light. The sunlight is a directional light source with small solid angle, which is usually treated as a parallel light. In sunny days, the sunlight accounts for about 80% energy of all the light sources. The skylight, distributed in a hemisphere of the sky dome, is the dominate light source in cloudy days. The standard skylight distribution can be modeled following the method proposed in [25]. In this work, we use the following rendering equation for a natural environment without inter-reflection:

$$\begin{aligned} &I(x,\theta_{\text{out}},\lambda)\\ &\quad= I_{\text{sun}}(x,\lambda)+I_{\text{sky}}(x,\lambda)+I_{\text{env}}(x,\lambda)\\ &\quad= \int_{\Omega_{\text{sun}}} \rho(x,\theta_{\text{in}},\theta_{\text{out}},\lambda)S(x,\theta_{\text{in}})L_{\text{sun}}(x,\theta_{\text{in}},\lambda)\\ &\qquad\times\cos\theta(x,\theta_{\text{in}})\,d\omega\\ &\qquad+\int_{\Omega_{\text{sky}}} \rho(x,\theta_{\text{in}},\theta_{\text{out}},\lambda)S(x,\theta_{\text{in}})L_{\text{sky}}(x,\theta_{\text{in}},\lambda)\\ &\qquad\times\cos\theta(x,\theta_{\text{in}})\,d\omega\\ &\qquad+\rho(x,\lambda)L_{\text{env}}(x,\lambda), \end{aligned} \tag{1}$$

where x is a 3D point in the scene, λ denotes the wavelength, $\rho(x,\theta_{\text{in}},\theta_{\text{out}},\lambda)$ is the bi-directional reflectance distribution function (BRDF), $S(x,\theta_{\text{in}})$ is the occlusion coefficient of x, $L_{\text{sun}}(x,\theta_{\text{in}},\lambda)$ is the sunlight irradiance arriving at x, θ_{in} and θ_{out} are the incidence and reflectance angles respectively, and ω is a solid angle. Note that the incident direction of the sunlight is related to the time and date, and the longitude and latitude of the camera position [18], $L_{\text{sky}}(x,\theta_{\text{in}},\lambda)$ is the irradiance of the skylight along the direction ω, which is modeled according to the method described at [25], Ω_{sun} and Ω_{sky} are the distribution solid angle areas of the sunlight and skylight respectively.

3.2 Basis image model

Since the sun is a distant light source, it is reasonable to assume that L_{sun} is a constant for all the x in the scenes and $\cos\theta(x,\theta_{\text{in}})$ can be approximated as a constant for all the θ_{in}, i.e. $L_{\text{sun}}(x,\theta_{\text{in}},\lambda)=L_{\text{sun}}(\lambda)$, $\cos\theta(x,\theta_{\text{in}})=\cos\theta(x)$. If the scenes are static and the viewpoint is fixed, θ_{out} would only depend on x. Therefore we ignore the parameter θ_{out} in (1) and have:

$$I_{\text{sun}}(x,\lambda)=L_{\text{sun}}(\lambda)\int_{\Omega_{\text{sun}}}\rho(x,\theta_{\text{in}},\lambda)S(x,\theta_{\text{in}})\cos\theta_{\text{in}}(x)\,d\omega. \tag{2}$$

On the contrary, Ω_{sky} is the hemisphere above x, $L_{\text{env}}(\lambda)$ is the environment light. Although in sunny days the skylight is distributed asymmetrically, its percentage with respect to the whole illumination is relatively small. In the cloudy days, the skylight is dominant, and usually is more or less uniformly distributed. Hence it is reasonable to assume that the skylight is distributed uniformly, which leads to:

$$I_{\text{sky}}(x,\lambda)=L_{\text{sky}}(\lambda)\int_{\Omega_{\text{sky}}}\rho(x,\theta_{\text{in}},\lambda)S(x,\theta_{\text{in}})\cos\theta_{\text{in}}(x)\,d\omega, \tag{3}$$

where $L_{\text{sky}}(\lambda)$ is a constant for all x. We define:

$$\begin{cases} C_{\text{sun}}(x,\lambda)=\int_{\Omega_{\text{sun}}}\rho(x,\theta_{\text{in}},\lambda)S(x,\theta_{\text{in}})\cos\theta_{\text{in}}(x)\,d\omega,\\ C_{\text{sky}}(x,\lambda)=\int_{\Omega_{\text{sky}}}\rho(x,\theta_{\text{in}},\lambda)S(x,\theta_{\text{in}})\cos\theta_{\text{in}}(x)\,d\omega. \end{cases}$$

Obviously, these two terms reflect the shadows, geometry and reflectance of scene objects at each pixel. Since the distribution of environment light is very similar to that of the skylight, it can be modeled in the same form of skylight and jointly accounted for with the skylight, then any image can be approximately decomposed into a linear composition of the following two terms:

$$I(x,\lambda)=L_{\text{sun}}(\lambda)C_{\text{sun}}(x,\lambda)+L_{\text{sky}}(\lambda)C_{\text{sky}}(x,\lambda). \tag{4}$$

We notice that both $L_{\text{sun}}(\lambda)$ and $L_{\text{sky}}(\lambda)$ are constant for all x in (4), and $C_{\text{sun}}(x,\lambda)$ and $C_{\text{sky}}(x,\lambda)$ are fixed illumination maps of the same scene under the sunlight or skylight with respect to a specified sun position. We call $C_{\text{sun}}(x,\lambda)$ and $C_{\text{sky}}(x,\lambda)$ the basis images. Therefore the illumination of any images of the same scene with the same sun position can be represented as the composition of these two basis images (see Fig. 1).

3.3 Camera response curve

A camera usually transfers the received intensity of the scene to RGB value adhering to a response curve. With the known exposure time, an image can be converted to a radiance map. We recover the camera response curve using the technique described in [5]. The illumination estimation is

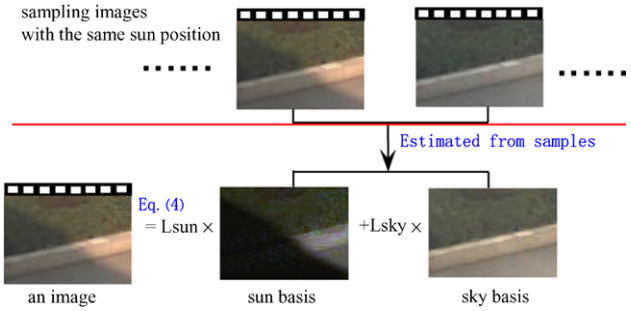


Fig. 1 An image can be expressed by the basis images

then performed in the radiance space. Since there is an unknown scaling factor of every channel of color images, we can only recover the relative radiance of background image. Correspondingly, the sunlight and the skylight is also estimated in a relative manner.

3.4 Estimation of basis images from samples

For simplicity, in the following, we ignore the symbol λ and represent color with RGB components. For a pixel x_j, $j = 1, \ldots, m$ in all the n registered images, we have:

$$\begin{cases} I_1(x_j) = L_{\text{sun},1} C_{\text{sun}}(x_j) + L_{\text{sky},1} C_{\text{sky}}(x_j) \\ I_2(x_j) = L_{\text{sun},2} C_{\text{sun}}(x_j) + L_{\text{sky},2} C_{\text{sky}}(x_j) \\ \quad \vdots \\ I_n(x_j) = L_{\text{sun},n} C_{\text{sun}}(x_j) + L_{\text{sky},n} C_{\text{sky}}(x_j) \end{cases} \tag{5}$$

for every pixel x_j in an image I_i, in case of the $L_{\text{sun},i}$ and $L_{\text{sky},i}$ are available, $I_i(x_j)$ involves only two unknowns, thus we need two registered images under different weather conditions to solve for these two unknowns. If more than two images are available, we use the least-squared method to derive an optimal solution of the equations. We denote the solution as $\hat{C}_{\text{sun}}(x_j)$ and $\hat{C}_{\text{sky}}(x_j)$.

For all pixels x_j $(j = 1, \ldots, m)$ in an image I where I can be one of I_i's $(i = 1, \ldots, n)$ or captured under other lighting conditions as long as it is acquired under a similar sun position as $I_i, i = 1, \ldots, n$, we have:

$$\begin{cases} I(x_1) = L_{\text{sun}} \hat{C}_{\text{sun}}(x_1) + L_{\text{sky}} \hat{C}_{\text{sky}}(x_1) \\ I(x_2) = L_{\text{sun}} \hat{C}_{\text{sun}}(x_2) + L_{\text{sky}} \hat{C}_{\text{sky}}(x_2) \\ \quad \vdots \\ I(x_m) = L_{\text{sun}} \hat{C}_{\text{sun}}(x_m) + L_{\text{sky}} \hat{C}_{\text{sky}}(x_m). \end{cases} \tag{6}$$

Also we use least squared method to compute L_{sun} and L_{sky}. Denote the solution as $\hat{L}_{\text{sun}}$, and $\hat{L}_{\text{sky}}$, then we can reconstruct the image I_i from the basis images as:

$$\begin{cases} \hat{I}(x_1) = \hat{L}_{\text{sun}} \hat{C}_{\text{sun}}(x_1) + \hat{L}_{\text{sky}} \hat{C}_{\text{sky}}(x_1) \\ \hat{I}(x_2) = \hat{L}_{\text{sun}} \hat{C}_{\text{sun}}(x_2) + \hat{L}_{\text{sky}} \hat{C}_{\text{sky}}(x_2) \\ \quad \vdots \\ \hat{I}(x_m) = \hat{L}_{\text{sun}} \hat{C}_{\text{sun}}(x_n) + \hat{L}_{\text{sky}} \hat{C}_{\text{sky}}(x_n). \end{cases} \tag{7}$$

We call the difference of the reconstructed image and the original image as an error map, i.e. $E(x_j) = \|\hat{I}(x_j) - I(x_j)\|$, for all $j = 1, 2, \ldots, m$.

To make our algorithm more robust, we manually pick up some regions to obtain a rough sunlight and skylight condition using the intensity of a diffuse surface in the scene [15]. This provides an initial estimation for $L_{\text{sun},i}$'s and $L_{\text{sky},i}$'s for all the i in (5). We then iteratively compute the basis images and the light parameters using (5) and (6). Notice that with the two basis images of $C_{\cdot}(x)$, we can refine the two lighting parameters $L_{\cdot,i}$. Also if $L_{\cdot,i}$ and $C_{\cdot}(x)$ are one solution, $\mu L_{\cdot,i}$ and $1/\mu C_{\cdot}(x)$ will also be a solution. Since we have assumed that skylight is uniform, its basis image should be constant. Therefore, we can give a coefficient constraint $\frac{1}{n}\sum_{i=1}^{n} L_{\text{sky},i} = 1$ to fix the scale. We usually use more than three images to recovery basis images to achieve robust solutions. The resolved skylight basis images have errors around 3∼8 intensity grades. Since there are 255 intensity levels, such an error level is within a decent noise range.

4 Extended light source recovering

We have derived a linear model for illumination due to the sunlight and skylight for the images with the same sun position. It is found that, the orbit of the sun in the sky dome is on a plane in the same day, and gradually rotates in latitude day after day. Therefore, we need to extend this model to those images with small deviation of the sun position.

4.1 Extension problems

To insert the virtual object into the real scenes with consistent illuminance, we need to specify the relationship at least in one image, i.e. give a scale of lighting condition of an image, to make the intensity of virtual objects match that of the real scene. All others can be automatically registered by our algorithm.

In practice, we manually select some key frames to avoid formulating an ill-posed linear equations and make our algorithm more robust. Note that our model is constrained to the same sun position, which would cause errors for the images

where the sun position deviates from that of the key frames. In this case, the main error is the difference of shadow term $S(x, \theta_{\text{in}})$, and cosine term $\cos\theta_{\text{in}}(x)$ in $C_{\text{sun}}(x)$.

4.2 Compensating of cosine term

At the off-line stage, we select two or three key images in a sunny day from our database in which the sun direction denoted as $e_i, i = 1, \ldots, s$ ($s = 2$ or 3) can serve as basis vectors of a 2D or 3D space since the sun directions are in a plane in the first day. And the light conditions are also known in these selected images. Therefore, for a diffuse surface in the scene with normal vector N, the product of surface reflectance k_d and the cosine angle between N and e_i are known. We denote $\langle e_i, N\rangle$ as the cosine term of the incidence and $J_i = k_d \times \langle e_i, N\rangle$. Any sun direction can be expressed as a linear combination of these basis vectors: $v = \sum_{i=1}^{s} \kappa_i e_i$ where κ_i can be solved for a known v. Suppose the sun direction of the closest key frame is $v' = \sum_{i=1}^{s} \nu_i e_i$, and for a pixel in the images, we have:

$$\begin{aligned} I &= k_d \langle v, N\rangle = k_d \sum_{i=1}^{s} \kappa_i \langle e_i, N\rangle = \sum_{i=1}^{s} \kappa_i J_i, \\ I' &= k_d \langle v', N\rangle = k_d \sum_{i=1}^{s} \nu_i \langle e_i, N\rangle = \sum_{i=1}^{s} \nu_i J_i, \end{aligned} \tag{8}$$

where I_i is known. Consequently, we have the extended cosine term:

$$\cos\theta_{\text{in}} = \langle v, N\rangle = \langle v', N\rangle \frac{\sum_{i=1}^{s} \kappa_i J_i}{\sum_{i=1}^{s} \nu_i J_i}. \tag{9}$$

Now we can compensate the intensity due to the direction of the sunlight deviating from that of the key frame without knowing the normal vector of the scenes.

4.3 Compensating of shadow term

For any sun position, we select a sample sun position which is the closest. We denote it as k, and its corresponding basis images are $C_{\text{sun},k}(x)$ and $C_{\text{sun},k}(x)$. We adopt the compensated sunlight basis images $C'_{\text{sun},k}(x)$ as its sunlight basis image, and then recover the light conditions by using the method described in Sect. 3.4. After that we estimate the error map. If the error is more than threshold ε, we just blind these pixels out of the extended basis images, and then the extended basis images are renewed. We iterate the process until convergence. During the deletion process, if the percentage of remaining pixels is less than a threshold, we divide the strip of the sun position into small parts, select the key frames and then estimate the light conditions as described above.

With the compensation of shadow and cosine angle, the light condition of an image can be recovered precisely using the same method as described in Sect. 3.4.

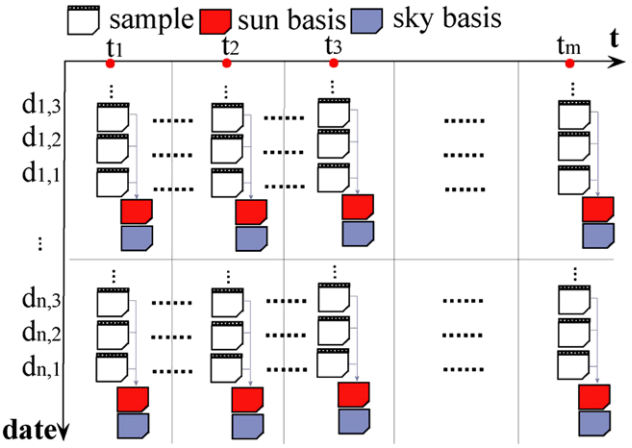


Fig. 2 Region splitting due to the course of the sun on different date, and each region has a pair of key basis images

5 System implementation

Our system consists of two stages: an off-line stage for basis image acquisition and an online stage for light condition recovery. Since basis images are related to the sun position, the sequence images captured in day time should be carefully organized.

5.1 Preparing the sample image database

A database is constructed to collect sample images of the same scene under a wide variety of weather and illumination conditions. We capture images for 0.2 to 1 frames per second. It takes at least two days with different weather conditions to find out the solution of basis images. After that, more and more images are added into the data base gradually, since the orbit of the sun is related to the season, and the basis images are constructed gradually.

Our collecting process lasted one year, all the possible sun positions were sampled. More samples may increase the robustness, but not really necessary. Minimally, our system can work as long as there are sample images captured from two different weather conditions.

5.2 Database organization

As the orbit of the sun is a strip in the sky, the database is organized by regions of the sun position as shown in Fig. 2. Each region has one key sun position, which is associated with one pair of basis images, and a set of images with the same sun position and typical light conditions, which are selected from captured images. Usually we use 15 degree interval. The other samples are abandoned to keep the database compact.

The sun gradually sweeps the region of the sky dome in both longitude and latitude. When the system runs after at

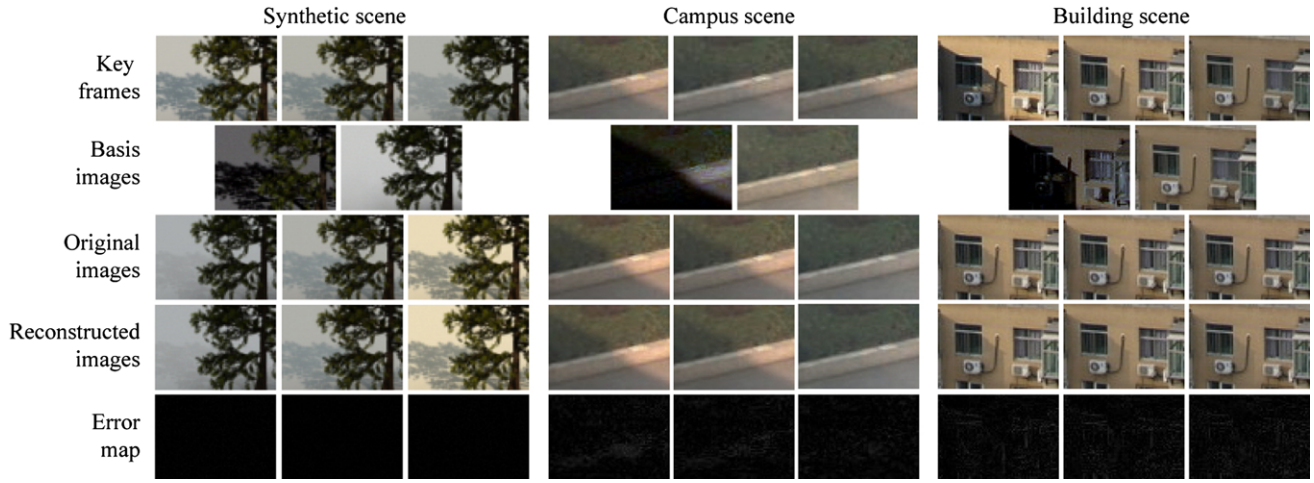


Fig. 3 The decomposition and reconstruction of images with the same sun position. Note that all the error maps have been multiplied by 3 times for better visualization

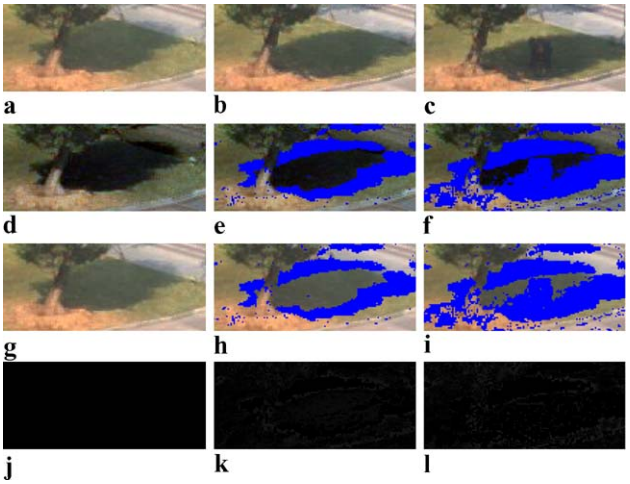


Fig. 4 Compensation of the difference of the sun position due to the time and date difference: (**a**) is a key frame, (**b**), (**c**) are two frames with 8 and 15 degree deviation w.r.t that of (**a**); (**d–f**) are the corresponding extended basis images for the sunlight with removed pixels marked by blue colors; (**g–i**) are the reconstructed images of (**a–c**); and (**j–i**) are their error maps

least two days, we accumulate the sample images into database due to the sun motion in latitude until all of the regions in the sky dome are covered.

5.3 Online recovery of the lighting conditions

During online process, for each frames, we can find out its basis images, and then recover the L_{sun} and L_{sky} by using the pixels in basis images. Although the basis images may have masked pixels, it does not affect our solution. Our method requires only two pixels with different illumination conditions for obtaining the solution, which is very easy to be satisfied. Of course, increasing the sampling points will increase the robustness for recovering the light conditions.

We employ the following procedure to recover the light conditions online. For any online image:

- Select key position with the sun position closest to that of this image;
- Fetch the basis images from the database;
- Use these basis images to recover the light condition by compensating the incidence difference;
- Reconstruct this image by recovering light condition and basis images;
- Remove pixels with large error from the basis images;
- Repeat the above process until no pixels are removed;
- Use the updated basis images to recover light parameters according to (6).

6 Results and discussion

The proposed approach for light recovery is examined with a synthetic scene and three real scenes. The sample images of the database were taken by a Cannon Eos5D, SX110, Dragon Fly camera on a tripod. The virtual objects were rendered by our software. Image registration is pre-processed due to occasional move of the tripod or the jitter of shutter. The algorithm is implemented on PC with 2.2 G CPU.

The key images of a synthetic and two real scenes are demonstrated in Fig. 3. We select a set of original images with the same sun position (sampled within 3 to 5 days) but different weather conditions, and solve the basis images for the sunlight and skylight. Note that the shadow area of the sunlight is black in the sunlight basis image. For any other image taken under the same sun position, it can be repre-

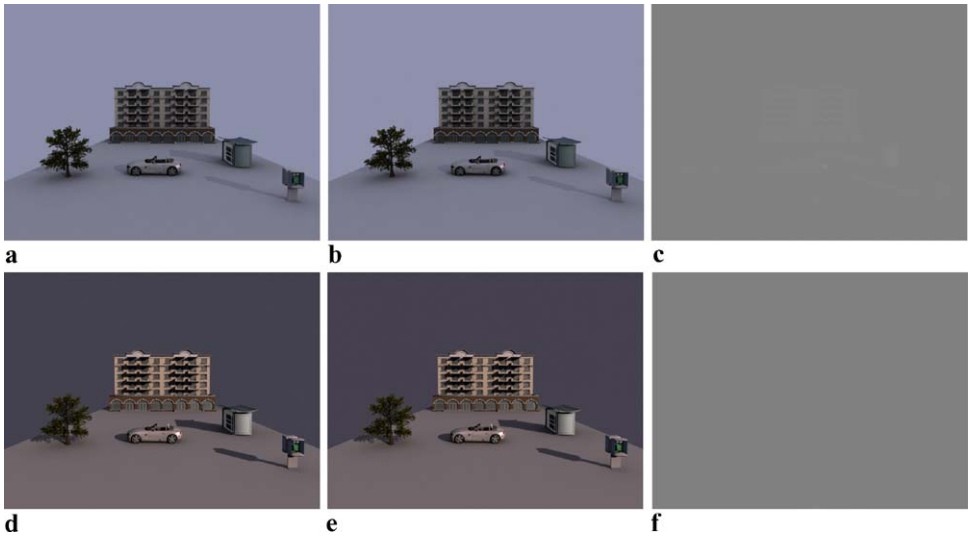


Fig. 5 The comparison between the recovered light condition and ground truth for the synthesis scene. (**a, d**) Original images; (**b, e**) reconstructed images; (**c, f**) error maps, added by 128 intensity values for better visualization

Table 1 The remaining pixels after removing large error caused by the position deviation of the sun

Deviation angle	Synthetic scene, $\varepsilon = 20$		Campus scene, $\varepsilon = 25$	
	Percentage of remaining pixels	Percentage of pixels error less than 10	Percentage of remaining pixels	Percentage of pixels error less than 10
5°	97.08%	92.38%	94.01%	86.72%
13°	90.77%	83.25%	89.23%	79.1%
21°	78.43%	71.91%	75%	68.2%

Fig. 6 Selected images in typical weather conditions from sequences after being composed with virtual objects: (**a**) campus scenes; (**b**) another view of the campus scene; (**c**) building scenes

sented as a linear combination of these basis images. After obtaining its light parameters using (6), we can reconstruct it according to (7). From the figure, we can see the reconstructed images are very similar to the original ones and the reconstruction errors, multiplied by 3 times for better visualization are very small. For campus scene shown in Fig. 6, there is no complete cloudy images for solving basis images. However, the cloudy image can still be correctly recombined with very small error.

The effect of our compensation method is shown in Fig. 4. We adopt ε around 20 out of 255 intensity levels as a threshold to reject the deviated shadow pixels. While tolerating certain noise, such a threshold is effective for rejecting deviated shadow pixels since the intensity difference of a suspected shadow pixel between two frames is usually larger than this value. Table 1 shows the percent of remaining pixels after rejection for synthetic scene and campus scene. It is found that there are still enough pixels to perform the calculation. From Fig. 4(d–f), we notice that the more deviation there is, the more pixels are removed by our algorithm. Also, in Fig. 4(c), there is a new object intruded into the shadow, and it is also removed by the algorithm. The reconstructed images in Fig. 4(g–i) are satisfactory. From the error maps shown in Fig. 4(k–m), it can be seen that the error increases when the deviation angle increases, but the result looks still satisfactory. In our experiment, $\pm 10^\circ$ deviation is still acceptable for our algorithm.

The recovery of the light condition is first illustrated by the synthesized scene with two light conditions shown in Fig. 5(a, d). The reconstructed images are shown in Fig. 5(b, e) with recovered light parameters. Figure 5(c, f) show the error maps, the reconstruction error is very small.

Our algorithm is also examined by 3 real video sequences as shown in Fig. 6. To check the quality of the recovering of the light source, some virtual objects are inserted into the real scenes. In the videos, the sunlight sometimes disappears, or decreases in the real scenes, even so the shadows and the color of the virtual objects match the background very well. The virtual objects in these three scenes from top to bottom are a budda and box, two bonsai trees and a box, a red pavilion with a gray plane on the top of a building, respectively. Please see the supplementary video to this paper for more details.

The computing time for online light condition recovery is about 21 fps with our not yet optimized code when the number of pixels used for calculation is within 60×30, which is a quite large number for our application. From these examples, we can see that the quality and the speed provided can meet the requirement of online application. Although the examples are all limited to the fixed view, our method is also applicable for camera panning by registering views in advance.

7 Conclusion and future work

We have proposed a light source estimation method for outdoor scene images. In the method, we first derive the basis images from sample images of database, the light condition of an online image can then be estimated with these basis images. To make our approach applicable to images under deviated sun positions, we develop a scheme to compensate the incidence term of sunlight and then remove the pixels whose shadow term changes with respect to the basis images. Unlike traditional methods in literature, our method requires no geometry information of the scene, which otherwise would be difficult to acquire in practice.

Although in our method the shadow areas of the sunlight are not required to be known in advance, the existence of both shadowed and lit areas of the sunlight are necessary to make the linear equations used in our algorithm well-posed.

No wonder that outdoor illumination estimation is a very challenging task. In this paper, we overlook the case of the virtual objects being in shadow regions and this will be one of our future work. We also plan to extend our work to cases with moving viewpoints. Also, more complex weather conditions should be considered, such as local cloud shadows and fogs.

Acknowledgements This paper is supported by 863 Program of China (No. 2007AA01Z326), 973 Program of China (No. 2009CB320802). The second author is also supported by the Open Project Program of the State Key Lab of CAD&CG (Grant No. A0809), Zhejiang University.

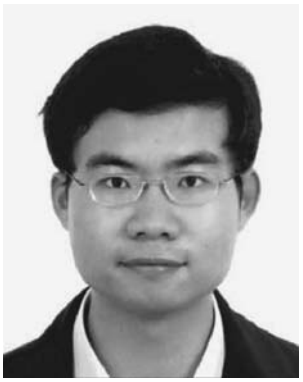
References

1. Agusanto, K., Li, L., Chuangui, Z., Sing, N.W.: Photorealistic rendering for augmented reality using environment illumination. In: Proc. IEEE/ACM International Symposium on Augmented and Mixed Reality, pp. 208–216 (2003)
2. Andersen, M.S., Jensen, T., Madsen, C.B.: Estimation of dynamic light changes in outdoor scenes without the use of calibration objects. In: ICPR, vol. 4, pp. 91–94 (2006)
3. Choudhury, B., Chandran, S.: A survey of image-based relighting techniques. In: GRAPP 2006: Proceedings of the First International Conference on Computer Graphics Theory and Applications, pp. 176–183 (2006)
4. 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, pp. 189–198 (1998)
5. Debevec, P.E., Malik, J.: Recovering high dynamic range radiance maps from photographs. In: Proc. SIGGRAPH '97, pp. 369–378 (1997)
6. Gibson, S., Cook, J., Howard, T., Hubbold, R.: Rapid shadow generation in real-world lighting environments. In: EGSR, pp. 219–229 (2003)

7. Hara, K., Nishino, K., Ikeuchi, K.: Light source position and reflectance estimation from a single view without the distant illumination assumption. IEEE Trans. Pattern Anal. Mach. Intell. **27**(4), 493–505 (2005)
8. Jacobs, K., Loscos, C.: Classification of illumination methods for mixed reality. Comput. Graph. Forum **25**(1), 29–51 (2006)
9. Jacobs, N., Roman, N., Pless, R.: Consistent temporal variations in many outdoor scenes. In: CVPR, pp. 1–6 (2007)
10. Lalonde, S.G.N.J.-F., Efros, A.A.: What does the sky tell us about the camera? In: ECCV (2008)
11. Kim, F.J.P.M., S.J.: Radiometric calibration with illumination change for outdoor scene analysis. In: CVPR (2008)
12. Koppal, S.J., Narasimhan, S.G.: Clustering appearance for scene analysis. In: CVPR, vol. 2, pp. 1323–1330 (2006)
13. Loscos, C., Drettakis, G., Robert, L.: Interactive virtual relighting of real scenes. IEEE Trans. Vis. Comput. Graph. **6**(3), 289–305 (2000)
14. Matsushita, Y., Lin, S., Kang, S.B., Shum, H.-Y.: Estimating intrinsic images from image sequences with biased illumination. In: ECCV, pp. 274–286 (2004)
15. Nakamae, E., Harada, K., Ishizaki, T., Nishita, T.: A montage method: the overlaying of the computer generated images onto a background photograph. In: Proc. SIGGRAPH '86, vol. 20, pp. 207–214 (1986)
16. Narasimhan, S.G., Wang, C., Nayar, S.K.: All the images of an outdoor scene. In: ECCV, pp. 148–162 (2002)
17. Nayar, S.K., Krishnan, G., Grossberg, M.D., Raskar, R.: Fast separation of direct and global components of a scene using high frequency illumination. ACM Trans. Graph. **25**(3), 935–944 (2006)
18. Rees, W.: Physical Principles of Remote Sensing. Cambridge Univ. Press, Cambridge (1990)
19. Sato, I., Sato, Y., Ikeuchi, K.: Acquiring a radiance distribution to superimpose virtual objects onto a real scene. IEEE Trans. Vis. Comput. Graph. **5**(1), 1–12 (1999)
20. Sato, I., Sato, Y., Ikeuchi, K.: Illumination distribution from shadows. In: CVPR, pp. 306–312 (1999)
21. Seitz, S.M., Matsushita, Y., Kutulakos, K.N.: A theory of inverse light transport. In: ICCV '05: Proceedings of the Tenth IEEE International Conference on Computer Vision, pp. 1440–1447 (2005)
22. Sharma, S., Joshi, M.V.: A practical approach for simultaneous estimation of light source position, scene structure, and blind restoration using photometric observations. EURASIP J. Adv. Signal Process. **2008**(3), 12 pages (2008)
23. Sunkavalli, K., Matusik, W., Pfister, H., Rusinkiewicz, S.: Factored time-lapse video. ACM Trans. Graph. **26**(3), 101 (2007)
24. Sunkavalli, K., Romeiro, F., Matusik, W., Zickler, T., Pfister, H.: What do color changes reveal about an outdoor scene? In: CVPR (2008)
25. Tadamura, K., Nakamae, E., Kaneda, K., Baba, M., Yamashita, H., Nishita, T.: Modeling of skylight and rendering of outdoor scenes. Comput. Graph. Forum **12**(3), 189–200 (1993)
26. Tappen, M.F., Freeman, W.T., Adelson, E.H.: Recovering intrinsic images from a single image. IEEE Trans. Pattern Anal. Mach. Intell. **27**(9), 1459–1472 (2005)
27. Wang, Y., Samaras, D.: Estimation of multiple directional light sources for synthesis of augmented reality images. Graph. Models **65**(4), 185–205 (2003)
28. Weiss, Y.: Deriving intrinsic images from image sequences. In: Computer Vision, IEEE International Conference on, vol. 2, p. 68 (2001)
29. Yu, Y., Debevec, P.E., Malik, J., Hawkins, T.: Inverse global illumination: Recovering reflectance models of real scenes from photographs. In: Proc. SIGGRAPH '99, pp. 215–224 (1999)
30. Zhang, Y., Yang, Y.H.: Multiple illuminant direction detection with application to image synthesis. IEEE Trans. Pattern Anal. Mach. Intell. **23**(8), 915–920 (2001)

Yanli Liu is now a Ph.D. candidate of State Key Lab of CAD&CG in Zhejiang University. Her main research interests include computer graphics, image and video processing.

Xueying Qin received her Ph.D. from Hiroshima University of Japan in 2001, and M.Sc. and B.Sc. from Zhejiang University and Peking University in 1991 and 1988, respectively. Currently, she is Professor of Shandong University. Her main research interests are augmented reality, video-based rendering, and photo-realistic rendering.



Songhua Xu is a member of the CAD and CG State Key Lab of China at Zhejiang University and an honorary researcher in the Department of Computer Science at the University of Hong Kong. He is also pursuing computer graphics research with the Computer Science Department at Yale University. His research interests include artificial intelligence, computer graphics, and human–computer interaction. He received his Bachelor of Engineering summa cum laude in Computer Science from Zhejiang University. He is a member of the AAAI and Eurographics.

Eihachiro Nakamae is now the Chairman of Sanei and Honorary Professor of Hiroshima University. He was Researcher Associate at Hiroshima University in 1956, Professor from 1968 to 1992, and Associate Researcher at Clarkson College of Technology, Potsdam, N.Y., from 1973 to 1974. He was Professor at Hiroshima Prefectural University from 1992 to 1995 and Professor from 1996 to 1999. He received his B.E., M.E., and Ph.D. in Electrical Engineering in 1954, 1956, and 1967 from Waseda University. His research interests include computer graphics, image processing, and electronic machinery.

Qunsheng Peng graduated from Beijing Mechanical College in 1970 and received a Ph.D. from the Department of Computing Studies, University of East Anglia, UK in 1983. Now he is Professor of computer graphics at Zhejiang University. He serves currently as a member of the editorial boards of several international and Chinese journals. His research interests include realistic image synthesis, computer animation, scientific data visualization, virtual reality, bio-molecule modeling.