

Optical Engineering

SPIDigitalLibrary.org/oe

Scene surface light field representation for realistic lighting

Tao Yu
Zhong Zhou
Jian Hu
QinPing Zhao



Scene surface light field representation for realistic lighting

Tao Yu

Zhong Zhou

Jian Hu

QinPing Zhao

Beihang University

State Key Laboratory of Virtual Reality Technology
and Systems

Beijing 100191, China

E-mail: zhouzhong2011@gmail.com

Abstract. Existing light field-based lighting follows the representation on plane, such as the commonly used light slab. A novel light field representation on scene surface is proposed. The illumination is captured by a panorama camera on translation stage with a nonuniform sampling and calibration strategy. Sampled light rays are reprojected on the reconstructed scene to generate the scene surface light field representations for both direct light source and indirect light. This representation is more consistent with the actual emission and reflection of light transport than the existing ones. An accurate light source estimation and a practical indirect light resampling approach are presented based on the representation. The rendering results illustrate the realistic lighting effect, especially in spatially varying illumination. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.OE.52.12.123103](https://doi.org/10.1117/1.OE.52.12.123103)]

Subject terms: image-based lighting; surface light field; light field representation; realistic lighting.

Paper 131083 received Jul. 18, 2013; revised manuscript received Oct. 29, 2013; accepted for publication Nov. 14, 2013; published online Dec. 16, 2013.

1 Introduction

With the development of the video game and film industry, realistic lighting has become an attractive hotspot. Many approaches are proposed for producing realistic lighting effects. The plenoptic function $P(x, y, z, \varphi, \theta, \lambda, t)$ proposed in Ref. 1 is a seven-dimensional function that defines the flux of a light ray with the wavelength λ through the space location (x, y, z) toward the direction (φ, θ) at time t . It can be reduced to a five-dimensional (5-D) function $P(x, y, z, \varphi, \theta)$ if we assume that the scene is static and that only three spectral bands (R, G, B) are considered. A further assumption is that the radiance of a light ray through a space free of internal occlusion remains constant. In that case, the 5-D function can be reduced to a four-dimensional (4-D) function $P(x, y, \varphi, \theta)$ with one spatial dimension eliminated. Contemporary light field representations are based on the 4-D function.

Light field describes the light spreading in every position and direction in space. McMillan and Bishop² proposed a 5-D light field representation by taking a set of panoramic images at different positions in a static scene. Levoy and Hanrahan³ presented the most commonly used 4-D light field representation light slab. As illustrated in Fig. 1, a light ray can only be determined by a pair of points on the two parallel planes at arbitrary positions. The point P_1 is parameterized by (u, v) while point P_2 by (s, t) . The line $L(u, v, s, t)$ connecting the two points indicates a light ray. A lot of work on light field is based on the 4-D representation, such as synthetic aperture imaging,⁴ light field photography,^{5,6} and integral imaging.⁷⁻⁹

Light field-based lighting is one form of image-based lighting. Traditional image-based lighting can only simulate infinitely distant light with RGB values by an environment mapping.¹⁰ Debevec¹¹ captured a light probe that was a high-dynamic range (HDR) spherical panoramic image to sample

omnidirectional light intensity at its capturing position. Several panoramic images with different exposure times were assembled into a light probe for displaying more levels of light intensity.¹² In addition, a complicated device light stage¹³ was set up by Debevec's team at the University of Southern California that has been used for producing realistic lighting effects in many famous movies, e.g., Avatar. However, a light probe just records the intensity of light rays incident into a single capturing position, so it can only represent spatially invariant illumination. Incident light field (ILF) was proposed by Unger et al.¹⁴⁻¹⁶ to generate spatially varying lighting effects. The illumination incident into a free space is captured and represented as a 4-D ILF. The ILF follows the plane-based representation and provides a realistic lighting effect in a relatively small area. But it is difficult to extend to a wider region because the complexity and data volume increase dramatically with spatially varying illumination. Although Unger discussed the possibility of extending ILF-based lighting to full room size in a Siggraph talk,¹⁷ it was too brief to show any technical details or results. Some simplifications of light field representation are valid by simplifying the data structure and data volume. Mury et al.^{18,19} simplified the measurement and reconstruction of light field structure in a finite three-dimensional (3-D) space. Hu et al.²⁰ had tried a nonuniform illumination sampling and representation to improve light field capturing and rendering efficiency. Light source plays an important role in realistic lighting, with much work concentrated on light source estimation. Lehmann and Palm²¹ introduced an approach named color line search to estimate the color of a single illuminant. Han and Sohn²² introduced a method to estimate the light source direction for face relighting. Shim²³ proposed to use a face in an image as a natural light probe¹¹ for relighting. Corsini et al.²⁴ presented a method of estimating the position, direction, and intensity of a real light source by capturing illumination with two

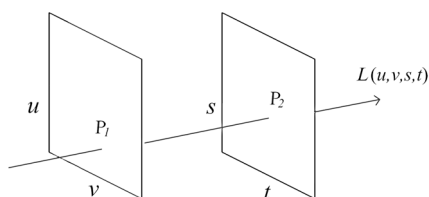


Fig. 1 The representation of light slab. $L(u, v, s, t)$ intersects with two parallel planes (uv and st) at two points (P_1 and P_2) to identify a light ray.

mirror balls. Goesele et al.²⁵ employed near-field photometry to measure and reproduce real light source in rendering.

In this paper, a novel scene surface light field representation is proposed to support realistic lighting with spatial variation. In comparison with the existing work, three contributions are made. First, a nonuniform illumination capturing and calibration strategy are employed to extend light field sampling and rendering scope to full indoor space. Second, owing to the scene surface-based representation, an accurate light source estimation of position, shape, direction, color, and intensity is presented. It is greatly beneficial to improve rendering effect and efficiency. Finally, a practical indirect light resampling method based on the scene surface is proposed for enhancing a realistic lighting effect.

The rest of the paper is organized as follows: Sec. 2 describes the nonuniform illumination capturing and calibration strategy. In Sec. 3, the scene surface light field representations of both direct light source and indirect light are presented. In Sec. 4, some rendering results are illustrated. Section 5 summarizes the work presented in this paper.

2 Illumination Capturing and Calibration

In this section, an illumination capturing and calibration strategy are proposed. The illumination sample data light probes are captured in a real scene and used to record the omnidirectional light intensity at their capturing positions. After the reconstruction of the scene, the light probes calibration is applied on the reconstructed 3-D model.

2.1 Capture Setup

As shown in Fig. 2, the capture setup consists of a panorama camera placed on a translation stage driven by a numerical control device. The panorama camera is Ladybug3 from



Fig. 2 The illumination capture setups.

point gray, and it contains six 2 Mpixels cameras that enable us to take panoramas from $>80\%$ of the full 360 deg sphere. The missing view is toward to the bottom of the camera which we are less concerned with. The camera is able to capture 12 Mpixels fused panoramas by streaming them to disk through a fire wire at 15 fps. If we lower the resolution of the captured panorama, the frame rate could be higher. We choose panoramas with the resolution of 500×500 pixels to shorten the capturing time.

The translation stage is driven by the numerical control equipment with sub-millimeters accuracy in movement. It is 1-m long with a shortest step of 2 mm. At most, 500 panoramas on the stage can be captured at different positions.

The panorama camera and translation stage are moved inside a room to take panoramas as raw illumination sample data. In order to capture the illumination variation in room size with a smaller data volume at a shorter time, we employ a nonuniform capturing strategy. The illumination distribution is prejudged for the manual selection of capturing locations. More panoramas are captured at the area with more illumination variations, like the edge of a shadow. In addition, the selected locations should cover most of the scope of the room. This capturing strategy could improve sampling efficiency and resampling precision because the most illumination variations are recorded by a certain amount of panoramas which are used to extract light rays incident into their capturing positions. On account of the nonuniform capturing strategy, the capturing positions of panoramas are unknown, so a calibration procedure is required to calculate their relative positions in the room.

2.2 Light Probes Composition

We capture 20×8 panoramas on a translation stage as a group. It means 20 different capturing positions with 5-cm spacing, and eight panoramas with different exposure times from 1/256 to 1 s at each position. Every eight panoramas are composited into the light probe. Then 50 capturing locations are selected following the nonuniform capturing strategy to capture 50 groups around the room. In conclusion, this amounts to 8000 captured panoramas when all is done. They are all automatically captured by control of numerical control equipment. What we need to do is move the translation stage and start the capturing program.

Light probe is an image that records the illumination incident into its capturing position. There are two essential features of the light probe, HDR and panoramic. Light probe samples light rays by its pixels, and each pixel indicates the direction and intensity of a light ray. A regular RGB picture is inefficient to represent the luminance range in real scenes. White area will appear on the picture when overexposed, and black areas when underexposed. HDR composition is introduced to solve this problem. An HDR image is generally composed of a plurality of different exposed pictures from the same scene. As shown in Fig. 3, eight panoramas are captured at the same location with different exposure times from 1/256 to 1 s. Then the eight panoramas are composited with the light probe by the method proposed in Ref. 12.

Panoramas are captured to compose light probe for sampling light rays in all directions. A commonly used method is to take photos aiming at a mirror sphere.^{11,15-17} The view at the back of sphere cannot be seen from the picture and

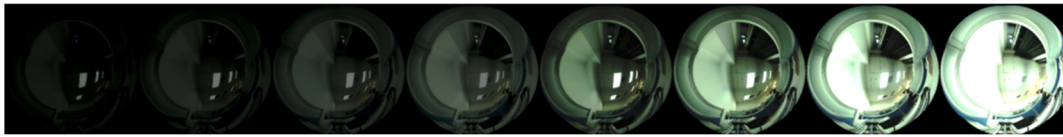


Fig. 3 An example of eight panoramas with different exposure times, they are used for light probe composition.

the camera photographs itself on the mirror sphere surface. It means that neither the illumination at the back of sphere and camera can be captured due to the missing view and occlusion. In contrast, we choose a panorama camera to capture the light probe, so as to sample the illumination that failed in the mirror sphere photos.

Figure 4(a) is the light probe example we captured. As is shown, it is an HDR spherical panoramic image. Every pixel in the image except the black area indicates a light ray sample. The positions of pixels determine the directions of light ray samples, and the radiances of these light rays are represented by their relevant pixel values. We can find that most views are captured in the scene except from the bottom of the camera. It does not matter because there are few lights emitting from the bottom in most cases.

The imaging model of the panorama camera is a classical spherical projection model. Suppose that there is a unit sphere located at the optical center of the panorama camera. The Z-axis points upward and XY are the horizontal planes across the camera optical center. Then the captured image is able to be mapped onto the surface of the unit sphere. As shown in Fig. 4(b), pixel q on the unit sphere surface indicates the light ray incident from point P into the sphere

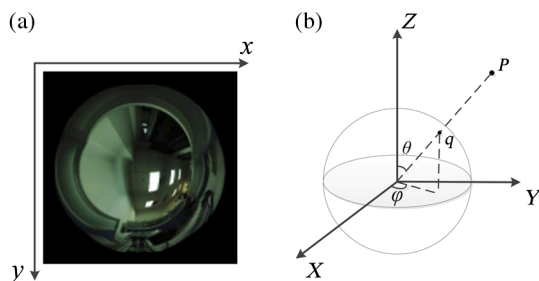


Fig. 4 (a) A light probe example captured by the panorama camera. (b) The spherical projection model. A light ray incident from P into the sphere center intersects with the sphere surface at pixel q . φ is the azimuthal angle and θ is the polar angle.

center. The light ray can only be determined by the angle pair (φ, θ) and its radiance is set as the pixel value of q .

Light probe is an angular map and the image coordinate origin is set at the center of the map. The relationship between the image coordinate (x, y) in Fig. 4(a) and direction (φ, θ) in Fig. 4(b) is given by

$$\begin{cases} \varphi = \arctan y/x \\ \theta = \pi(x^2 + y^2)^{1/2}, \end{cases} \quad (1)$$

where $x \in [-1, 1]$, $y \in [-1, 1]$, $\varphi \in [0, 2\pi)$, and $\theta \in [0, \pi]$.

2.3 Scene Reconstruction from Kinect Fusion

We employ an open source library called the point cloud library²⁶ which includes Kinect fusion algorithm²⁷ to reconstruct the room. Kinect fusion is a 3-D reconstruction algorithm based on Microsoft Kinect. There are an infrared transmitter, an infrared receiver, and a color camera in Kinect. It is able to capture color image and depth image at the same time. Figure 5 shows an example of the color image and depth image captured by Kinect. Since the infrared sensors and color camera are precalibrated, the pixels of color image and depth image are in a one-to-one correspondence. With the depth information and the pose of camera, which is estimated by the corresponding color image sequence, the object geometry can be reconstructed.

The truncated signed distance function is used to be a volumetric representation of the scene from the depth image, and the iterative closest point algorithm is used to estimate the pose of Kinect from two sequential color frames.²⁸ In conjunction with the two approaches, a fused dense point cloud of the scene is provided. After the point cloud is triangle meshed, the scene surface in the view is generated. As Kinect fusion is unable to reconstruct an unbounded extended area, an incremental surface reconstruction algorithm²⁹ is required to reconstruct the whole scene geometry. The reconstructed area is shifted with the movement of Kinect, so that we can continuously expand the reconstructed surface in an incremental update.

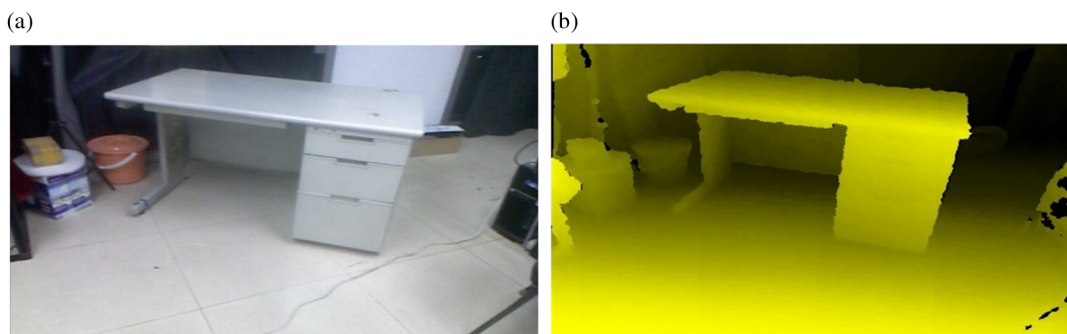


Fig. 5 The color image and depth image captured by Kinect at the same time. (a) The color image captured in the scene. (b) The relevant depth image.

The whole scene is reconstructed after the area has been completely scanned.

2.4 Light Probes Calibration

The light rays sampled by light probes are reprojected to the reconstructed scene model to set up the scene surface light field representation. Therefore, the relative positions of the light probes in the scene are required to be known in advance. Two calibration processes are needed to get the relationship. One is extrinsic camera parameters' calibration and the other is the transforming relationship between the camera coordinate system and the world coordinate system.

The extrinsic camera parameters calibration is, in fact, to find out the relative positions of all the light probes. With our setup, light probes are captured on the translation stage with precise movements controlled by numerically control equipment. It is easy to calculate the relative position on the stage

$$P_k = P_s + \frac{k}{n}(P_e - P_s). \quad (2)$$

In Eq. (2), P_s is the start position on the stage, P_e is the end position, and n is the number of panoramic images captured on the stage. Then, we can obtain the position of the k 'th panoramic image P_k in the capturing sequence.

We move the camera with the translation stage in the room to capture light probes. The subsequent problem is to calibrate all the light probes. As the internal relationship has been given in Eq. (2), only the start position P_s and the end position P_e are needed to be calculated. As illustrated in Fig. 6, a panorama camera at different positions always has overlapping fields of view, so structure from a motion (SFM) algorithm is appropriate for estimating the extrinsic camera parameters. We use Bundler,^{30,31} which is an SFM toolkit, to calibrate the rotation and translation of the panorama camera. Bundler takes a set of images as input and outputs the extrinsic camera parameters in a camera coordinate system. One image is selected and its capturing position is set as the coordinate origin. The rotation and translation of other images' capturing positions are calculated by Bundler in the selected camera coordinates. Because the amount of light probes are captured around the room and required to be calibrated into the camera coordinates, we finish the calibration work incrementally in groups. Groups of images are calibrated into the

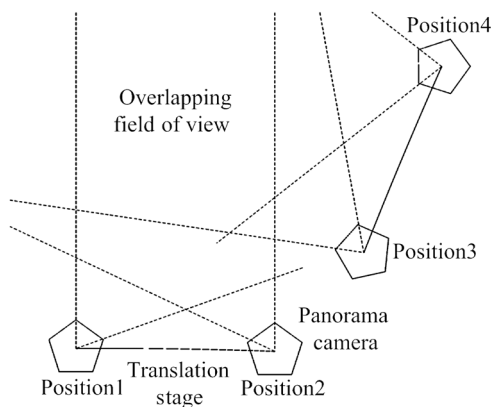


Fig. 6 The view of the panorama camera. Due to its omnidirectional vision, they always have overlapping field of view at different positions.

camera coordinate system sequentially, and they share the common origin, so that they are in the same coordinate system and their translation vector can be seen as their spatial coordinates.

The scene is reconstructed by Kinect fusion algorithm, while the world coordinate system is set up by the pose of Kinect when it captures the first frame. The starting capturing position of Kinect is the origin of the world coordinate system. In order to calculate the relative positions of light probes in the scene, it is required to align the world coordinate system with the camera coordinate system. We add the first frame of Kinect into Bundler to calculate the rotation R_1 and translation T_1 from the selected origin light probe to Kinect start position. This matrix M_1 is, in fact, the transformation matrix of two coordinates. To calculate the real position of the light probe in the scene, it would be transformed from the camera coordinates to the world coordinates by M_1 . Each light probe is marked by an extrinsic parameters matrix M_2 , records the rotation R_2 , and translation T_2 in the camera coordinates. Equation (3) is the transforming equation to transform light probes into the world coordinates. Matrix M is the real pose of the light probes in the world coordinates

$$\begin{aligned} M &= M_1^{-1} \cdot M_2 = \begin{bmatrix} R_1 & T_1 \\ 0 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} R_2 & T_2 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} R_1^{-1}R_2 & R_1^{-1}T_2 - R_1^{-1}T_1 \\ 0 & 1 \end{bmatrix}. \end{aligned} \quad (3)$$

3 Scene Surface Light Field Representation

In this section, the scene surface light field representation is presented including direct light source and indirect light. It is more consistent with the actual emission and reflection of light transport than the previous ones, thus it provides accurate direct light source estimation and indirect light resampling.

3.1 Direct Light Source Representation

3.1.1 Light source extraction

As shown in Fig. 7(a), light sources only occupy a little area on the light probe, but they contribute most of the illumination in the scene. It is easy to see that extracting the light source from the scene and making a dense representation will greatly improve the rendering efficiency and effect.

Figure 7(a) is the light probe sample in latitude–longitude image format. As shown in this image, the intensity of the pixels in the light sources' areas is much higher than the other areas, so we employ Otsu algorithm³² to evaluate a binarization threshold according to the image's pixel intensity. Figure 7(b) is the image segmentation result based on the binarization threshold, and highlighted areas are extracted as light source areas. In addition, we can find that not only the lamps but also some other highlighted areas, e.g., the mirror reflection of windows, are regard as light sources. Such extractions are processed on each light probe.

The relative positions between light probes and recovered scenes have been solved. Consequently, the real emitting locations of the light sources can be extracted from the scene through reprojected of the white areas in Fig. 7(b).

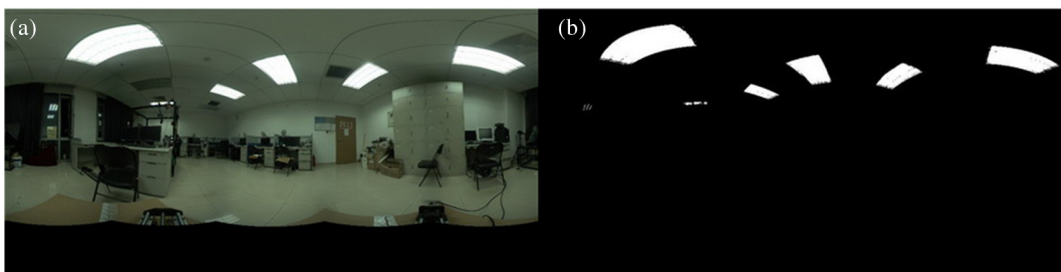


Fig. 7 (a) A light probe captured in our laboratory. (b) Image binarization result. White areas are extracted as light sources.

As illustrated in Fig. 8, the steps of light sources extraction are as below:

1. The light rays are reprojected from the extracted white areas on the light probes to the reconstructed scene, along the direction (φ, θ) by spherical projection which has been described in Sec. 2.2. As the scene is meshed by triangular patches, the reprojected light rays will intersect with some patches at some points.
2. Step 1 is applied to each light probe.
3. Only a small part of patches contain intersections. These patches except those contain few intersections are regard as light sources areas. A threshold is determined by OTSU algorithm according to the number of intersections each patch contains. The patches that contain intersections less than the threshold value are abandoned for their supposed accuracy errors or discontinuity.
4. The other patches are picked out to construct light sources.

3.1.2 Anisotropic point light sources

Both direct light source and indirect light representations are set up on scene surface. The basis representation of light ray sample is shown as follows:

$$L(u, v, \varphi, \theta) \rightarrow \text{RGB}. \tag{4}$$

In Eq. (4), (u, v) defines a parameterization of geometry surface and (φ, θ) represents the direction of a light ray.

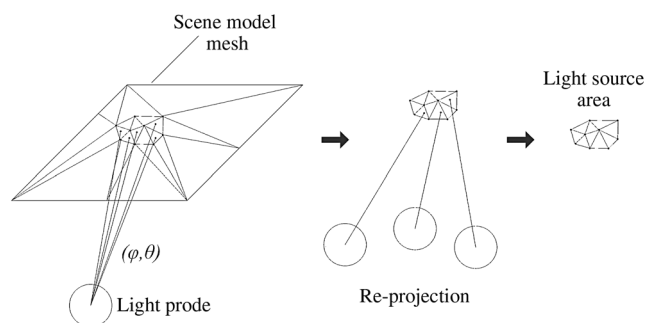


Fig. 8 The steps of light sources extraction. The light rays sampled by the highlighted areas on the light probes are reprojected to the scene. They intersect with the scene surface at some patches. These patch sets are regard as light sources.

$L(u, v, \varphi, \theta)$ indicates the radiance of a certain light ray and is valued by RGB color in HDR.

The direct light source representation can be seen as a collection of anisotropic point light sources located on the scene surface. Because the scene is reconstructed by triangular mesh and the areas of light sources have been extracted, we can easily create such anisotropic point light sources at the vertexes of triangular patches in light source areas.

Figure 9 is an example of an anisotropic point light source. All the rays from a certain vertex of triangular patch in the light source area are emitting to the capturing positions of the light probes, and form an anisotropic point light source. Each light ray intersects with the light probe at one pixel and the pixel value is set as the radiance of the light ray. Another key issue is to solve the directions of these light rays. Since the positions of the vertexes and the light probes have all been calibrated into the world coordinate system, the directions of their connecting rays can be calculated by

$$\begin{cases} \varphi = \arccos \frac{\vec{v}_{xy} \cdot \vec{x}}{|\vec{v}_{xy}|}, \\ \theta = \arccos \frac{\vec{v} \cdot \vec{z}}{|\vec{v}|} \end{cases} \tag{5}$$

where the definition of light ray direction (φ, θ) is the same as that in Fig. 4(b), φ is the azimuthal angle and θ is the polar angle, \vec{v} is the direction vector of the light rays, \vec{v}_{xy} is the projecting vector of \vec{v} in XY -plane, and \vec{x} and \vec{z} are the unit vectors in X -axis and Z -axis.

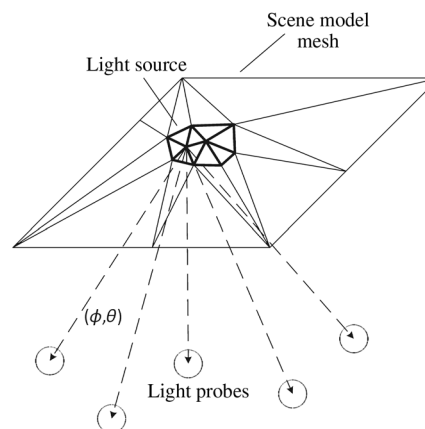


Fig. 9 Anisotropic point light sources are created at each mesh vertex in extracted light source areas. The light rays are sampled by emitting to all the light probes.

3.1.3 Light source resampling

The number of light rays that each anisotropic point light source emits is equal to the number of light probe samples. It is too sparse to represent direct light source, so a resampling process is needed. As illustrated in Fig. 10, the anisotropic point light source is resampled to be continuous in angular dimension. Such a resampling process is executed to all the created point light sources.

The steps of an anisotropic point light source resampling are:

1. Recording the direction (φ, θ) and radiance (relevant color value on the light probe) of each emitting light ray.
2. Mapping the radiances of all the light rays to an angular map. As the discretization of light probes are captured, the map shows a discrete points distribution.
3. A Delaunay triangulation is applied to generate a triangular mesh.
4. Make the radiance of light rays continuous in angular distribution by using the barycentric interpolation in triangle.

After resampling in angular dimension, the point light source is resampled in spatial dimension to get a complete direct light source representation. This resampling is similar to that in angular dimension. Assume that a light ray L_x emits from the light source and the start position is not at a mesh vertex, then it must start from the edge or inside of a triangular patch. As the three vertexes of the relevant triangular patch have been resampled to be continuous in angular dimension, the radiances of three light rays emitting from the three vertexes and toward the same direction with L_x are known. Then the radiance of L_x can be solved by triangular barycentric interpolation with the three known light rays. After finishing the resampling of both angular and spatial dimensions, a dense direct light source representation is established.

Since the direct light source representation is created on the reconstructed scene surface, the position and shape of light sources are accurate. The emission and reflection of light ray samples from the light sources are more consistent with the real situation of light transport than the existing planar parameterization. It improves the resampling accuracy of the direction and intensity of light rays.

3.2 Indirect Light Representation

Indirect light represents the illumination that does not emit from the light source, such as light reflected from walls, furniture, windows, etc. To illuminate something,

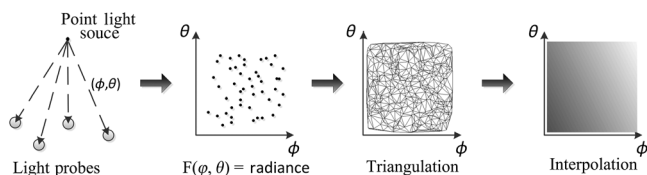


Fig. 10 The process of point light source resampling. The sampled light rays are mapped to an angular map, with a Delaunay triangulation and barycentric interpolation, they are resampled to be continuous.

the contribution of indirect light is much smaller than that of the light source, but it is an essential part of realistic lighting, especially to render the details in a dark area. Most light in space is indirect light, so it is infeasible to represent the indirect light like the direct light source due to the large data volume. A sparse data representation and a fast resampling method for the indirect light are fit for balancing rendering effect and time consumption. The indirect light representation is based on the light probes sequence retrieval. A kd-tree structure is used to index light probes sampled in the scene for efficient searching. During rendering, the nearest light probes beside the objects to be rendered are quickly found to resample light rays.

As illustrated in Fig. 11, the bunny is a virtual object rendered by indirect light. $L_1, L_2,$ and L_3 are quickly found out by a nearest search approach based on kd-tree with the maximum distance d . Light ray r is a random sample in ray tracing. Different from a traditional ray tracing algorithm, the whole scene is regarded as a light source. After r is projected to the scene surface at the intersection P_1 , ray $r_1, r_2,$ and r_3 from P_1 to the nearest neighbors $L_1, L_2,$ and L_3 are sampled to calculate the radiance of r .

The radiance of ray sample in ray tracing is solved by

$$\begin{cases} R = \sum_{i=0}^n R_i \frac{IDW L_i}{\sum_{i=0}^n IDW L_i} \\ IDW L_i = \left(\frac{d - (P - P_{L_i})}{d(P - P_{L_i})} \right)^2 \end{cases} \quad (6)$$

where IDW stands for inverse distance weighted algorithm³³ which defines the weight of the near light rays' radiance contribution. P and P_{L_i} are the positions of the virtual objects and the near light probes, while d is the threshold of max distance. R is the radiance of the light ray sample in ray tracing, such as r in Fig. 11. R_i is the radiance of the light rays captured by nearby light probes, such as $r_1, r_2,$ and r_3 .

4 Rendering Results

In this section, some rendering results are illustrated. Custom mental-ray shaders for the direct light source and indirect light are implemented, respectively, and then used in global illumination rendering.

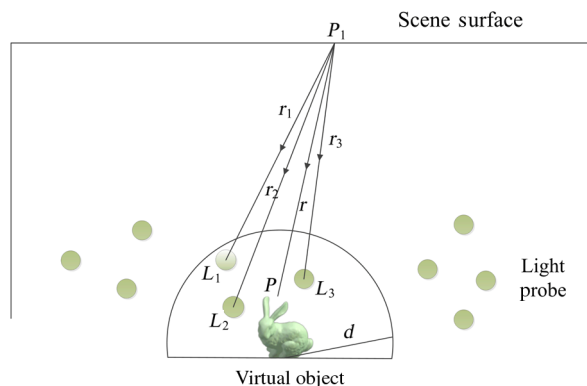


Fig. 11 The sketch of indirect light resampling. The nearest light probe neighbors $L_1, L_2,$ and L_3 are found out by kd-tree index and used in virtual object lighting. The random light ray sample r in ray tracing is estimated by $r_1, r_2,$ and r_3 with inverse distance weighted algorithm.

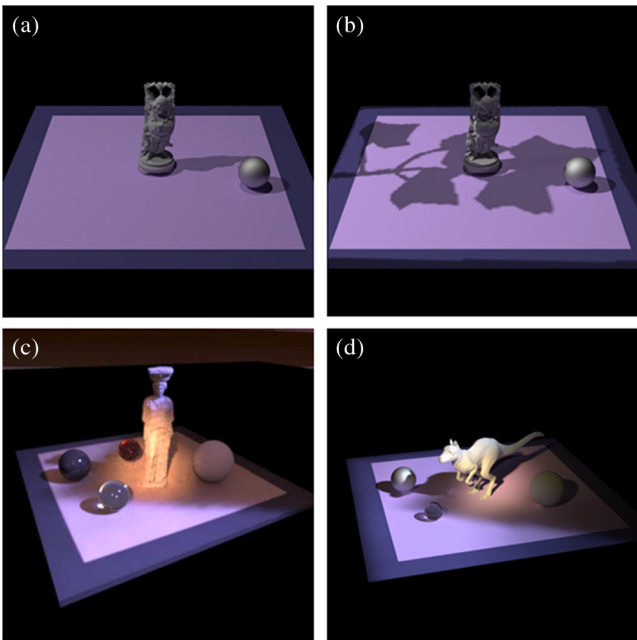


Fig. 12 Renderings based on different light field representations. (a) Environment mapping. (b) Our result with a spatially varying illumination captured. (c) Incident light field rendering with a uniform sampling of illumination. (d) Our result based on nonuniform sampling strategy.

Figure 12 is a comparison of rendering results based on the different light field representations. Figure 12(a) is the traditional image based lighting result illuminated by an environment map. Since the illumination is measured by a single point sample, infinitely distant light can only be simulated to light virtual objects. It just fits for performing a uniform illumination due to the missing spatial variation. By contrast, we show a spatially varying lighting effect in Fig. 12(b). The complicated illumination that is occluded by a bonsai is captured and represented in our light field representation, then used in rendering. As shown, the complicated shadow reveals that our light field representation works well in spatially varying lighting.

Figure 12(c) is rendered in Ref. 14 by ILF, which is parameterized on a plane. Also its illumination is captured with a uniform spatially sampling resolution. About 30×30 light probes are captured to compose an ILF. It shows a realistic lighting effect. Figure 12(d) is our result based on a nonuniform capturing strategy. With prejudging of illumination distribution, 160 light probes are chosen to be sampled at the regions diverse in illumination, e.g., the edge of shadows and spotlighting areas. Also only 40 light probes are sampled at the other regions. A realistic lighting effect almost the same as Fig. 12(c) is produced with less than one-fourth of the light probes sampled. Moreover, smoother and softer shadow edges are displayed in Fig. 12(d), since more light probes are sampled here than that in Fig. 12(c). This comparison figure

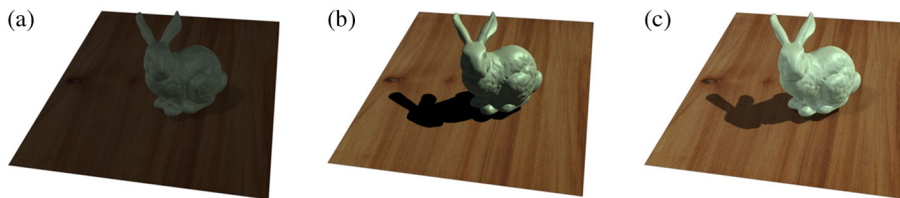


Fig. 13 The rendering results of indirect light and direct light source. (a) The lighting result of indirect light only. (b) The lighting result of direct light only. (c) The lighting result of complete light field.

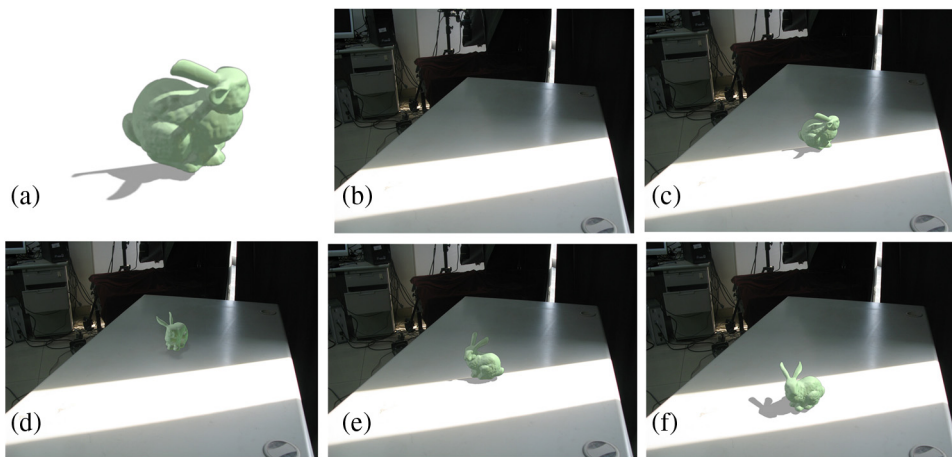


Fig. 14 The virtual bunny is placed into the scene and rendered by the captured illumination to demonstrate a consistent lighting effect at different locations. (a) The bunny lighted by the captured illumination at the shadow edge. (b) The original scene. (c) A fusion of virtual object and real scene. (d) Virtual bunny at shadowy area. (e) Virtual bunny at shadow edge. (f) Virtual bunny at bright area.

is also presented in our previous work on nonuniform light field representation.²⁰ The new representation presented in this paper retains the ability to handle illumination spatial variation.

Figure 13 illustrates the lighting contributions of a direct light source and an indirect light. Direct light sources are extracted from the scene and sunshine constitutes the main light source. Figure 13(a) is rendered only by indirect light, 13(b) is merely lighted by direct light, and Fig. 13(c) is a rendering result of the composition of a light source and indirect light. The virtual bunny in Fig. 13(b) has a brighter surface and sharper shadow than 13(a). Obviously, the lighting contribution of indirect light is much less than the direct light source. However, most details in the dark area can be seen in Fig. 13(a) rather than Fig. 13(b) due to the light source occlusion. It is indicated that the indirect light plays an important role in realistic lighting as a light source. Figure 13(c) is the final result of our light field lighting. It shows a realistic lighting effect.

Figure 14 illustrates a fusion effect of a virtual object and a real scene with consistent lighting in a spatially varying illumination environment. The illumination is captured and represented on the reconstructed laboratory surface. The virtual bunny (Stanford bunny) is lighted by the scene surface light field at different positions, and then rendered into the captured scene. The shadow and appearance of the bunny show an obvious difference at different positions, and they are just consistent with the scene illumination. The fusion results demonstrate a realistic lighting effect.

5 Conclusion

In this paper, a scene surface light field representation is proposed to generate a realistic lighting effect. A complete solution of capturing, processing, representing, and rendering on spatially varying illumination in full indoor space is presented. In contrast to the existing work, a nonuniform capturing and calibration strategy are employed to extend illumination sampling and rendering scope to full indoor space. Then, the captured scene is reconstructed by Kinect fusion, and a scene surface based light field is created, which is more consistent with the real physical property of light transport than the existing planar representation. The rendering results based on the scene surface light field illustrate the realistic lighting effect especially in spatially varying illumination.

The scene surface-based light field is represented on the reconstructed scene model, so the method is suitable for a bounded scene, such as interior space. It cannot treat with unbounded scenes whose reconstruction is not complete. Instead, regular planar representation is applicable. In addition, compared with the existing approaches, our method introduces extra scene reconstruction which will cost more time. In the future, we plan to optimize the scene reconstruction and light probes capturing process to improve the efficiency.

Acknowledgments

This paper is supported by the National High Technology Research and Development 863 Program of China under Grant (Nos. 2012AA011801 and 2012AA011803) and National Natural Science Foundation of China (No. 61300066).

References

1. E. H. Adelson and J. Bergen, "The plenoptic function and the elements of early vision," *Comput. Models Visual Process.* **1**(2), 4–6 (1991).
2. L. McMillan and G. Bishop, "Plenoptic modeling: an image-based rendering system," in *Proc. of ACM SIGGRAPH*, S. G. Mair, Ed., pp. 39–46, ACM, New York (1995).
3. M. Levoy and P. Hanrahan, "Light field rendering," in *Proc. of ACM SIGGRAPH*, J. Fujii, Ed., pp. 31–42, ACM, New York (1996).
4. V. Vaish et al., "Reconstructing occluded surfaces using synthetic apertures: stereo, focus and robust measures," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 2331–2338, IEEE, Piscataway, New Jersey (2006).
5. R. Ng et al., "Light field photography with a hand-held plenoptic camera," Computer Science Technical Report CSTR 2 (2005).
6. Z. Xu, J. Ke, and E. Y. Lam, "High-resolution light field photography using two masks," *Opt. Express* **20**(10), 10971–10983 (2012).
7. Y. Xu et al., "Homogeneous light field model for interactive control of viewing parameters of integral imaging displays," *Opt. Express* **20**(13), 14137–14151 (2012).
8. F. Wu et al., "High-optical-efficiency integral imaging display based on gradient-aperture pinhole array," *Opt. Eng.* **52**(5), 054002 (2013).
9. J.-H. Park, "Comparative analysis on light field reconstruction characteristics of autostereoscopic three-dimensional display technologies," *Proc. SPIE* **8556**, 855605 (2012).
10. G. S. Miller and C. R. Hoffman, "Illumination and reflection maps: simulated objects in simulated and real environments," in *Proc. of ACM SIGGRAPH 84, Advanced Computer Graphics Animation Seminar Notes*, Vol. 190, ACM, New York (1984).
11. P. Debevec, "Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography," in *Proc. of ACM SIGGRAPH*, pp. 189–198, ACM, New York (1998).
12. P. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proc. of ACM SIGGRAPH*, pp. 369–378, ACM, New York (1997).
13. A. Wenger et al., "Performance relighting and reflectance transformation with time-multiplexed illumination," in *Proc. of ACM SIGGRAPH*, M. Gross, Ed., pp. 756–764, ACM, New York (2005).
14. J. Unger et al., "Capturing and rendering with incident light fields," in *Proc. of the 14th Euro Graphics Symposium on Rendering*, pp. 141–149, Eurographics Association, Aire-la-Ville, Switzerland (2003).
15. J. Unger, S. Gustavson, and A. Ynnerman, "Spatially varying image based lighting by light probe sequences," *Visual Comput.* **23**(7), 453–465 (2007).
16. J. Unger et al., "Free form incident light fields," *Comput. Graph. Forum* **27**(4), 1293–1301 (2008).
17. J. Unger, "Next generation image based lighting using HDR video," in *Proc. of ACM SIGGRAPH Talks* (2011).
18. A. A. Mury, S. C. Pont, and J. J. Koenderink, "Light field constancy within natural scenes," *Appl. Opt.* **46**(29), 7308–7316 (2007).
19. A. A. Mury, S. C. Pont, and J. J. Koenderink, "Representing the light field in finite three-dimensional spaces from sparse discrete samples," *Appl. Opt.* **48**(3), 450–457 (2008).
20. J. Hu et al., "Non-uniform illumination representation based on HDR light probe sequences," in *Int. Conf. on Virtual Reality and Visualization*, pp. 62–68, IEEE Press, Piscataway, New Jersey (2012).
21. T. M. Lehmann and C. Palm, "Color line search for illuminant estimation in real-world scenes," *J. Opt. Soc. Am. A* **18**(11), 2679–2691 (2001).
22. H. Han and K. Sohn, "Face relighting based on virtual irradiance sphere and reflection coefficient," *Opt. Eng.* **47**(2) 027204 (2008).
23. H. Shim, "Faces as light probes for relighting," *Opt. Eng.* **51**(7) 077002 (2012).
24. M. Corsini, M. Callieri, and P. Cignoni, "Stereo light probe," *Comput. Graph. Forum* **27**(2), 291–300 (2008).
25. M. Goesele et al., "Accurate light source acquisition and rendering," in *Proc. of ACM SIGGRAPH*, A. P. Rockwood, Ed., pp. 621–630, ACM, New York (2003).
26. R. B. Rusu and S. Cousins, "3D is here: point cloud library (PCL)," in *Int. Conf. on Robotics and Automation (ICRA)*, pp. 1–4, IEEE, Piscataway, New Jersey (2011).
27. S. Izadi et al., "Kinect fusion: real-time 3D reconstruction and interaction using a moving depth camera," in *Proc. of the 24th Annual ACM Symposium on User Interface Software and Technology*, pp. 559–568 (2011).
28. R. A. Newcombe et al., "Kinect fusion: real-time dense surface mapping and tracking," in *10th IEEE Int. Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 127–136, IEEE (2011).
29. T. Whelan et al., "Kintinuous: spatially extended kinect fusion," in *3rd Robotics: Science and Systems (RSS) Workshop on RGB-D: Advanced Reasoning with Depth Cameras*, Sydney, Australia (2012).
30. N. Snavely, S. M. Seitz, and R. Szeliski, "Photo tourism: exploring photo collections in 3D," *ACM Trans. Graph.* 835–846 (2006).
31. N. Snavely, S. M. Seitz, and R. Szeliski, "Modeling the world from internet photo collections," *Int. J. Comput. Vision* **80**(2), 189–210 (2008).

32. N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man Cybern.* 9(1), 62–66 (1979).
33. D. Shepard, "A two-dimensional interpolation function for irregularly-spaced data," in *Proc. of the 1968 23rd ACM National Conf.*, pp. 517–524, ACM (1968).



Tao Yu received his BS degree from Beijing Institute of Technology and MS degree from Beijing Jiaotong University in 2008 and 2011, respectively. He is currently working toward his PhD at State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing. His research interests include light field-based lighting and HDR imaging.



Zhong Zhou received BS degree from Nanjing University and PhD degree from Beihang University in 1999 and 2005, respectively. He is an associate professor, PhD adviser, State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing, China. His main research interests include augmented virtual environment, natural phenomena simulation, distributed virtual environment, and Internet-based VR technologies. He is a member of IEEE, ACM, and CCF.



Jian Hu received his BS degree from Sichuan University and MS degree from Beihang University in 2010 and 2013, respectively. His research interests include light field rendering.



Qiping Zhao is a professor, PhD adviser, the Director of State Key Laboratory of Virtual Reality Technology and Systems, Beihang University and Director-general of Chinese Association for System Simulation. He has published more than 160 papers, 33 national authorized patents, and three books. His research interests mainly lie in virtual reality.