Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-based Graphics with Global Illumination and High Dynamic Range Photography

Paul Debevec

University of California at Berkeley[1]

ABSTRACT

We present a method that uses measured scene radiance and global illumination in order to add new objects to light-based models with correct lighting. The method uses a high dynamic range image-based model of the scene, rather than synthetic light sources, to illuminate the new objects. To compute the illumination, the scene is considered as three components: the distant scene, the local scene, and the synthetic objects. The distant scene is assumed to be photometrically unaffected by the objects, obviating the need for reflectance model information. The local scene is endowed with estimated reflectance model information so that it can catch shadows and receive reflected light from the new objects. Renderings are created with a standard global illumination method by simulating the interaction of light amongst the three components. A differential rendering technique allows for good results to be obtained when only an estimate of the local scene reflectance properties is known.

We apply the general method to the problem of rendering synthetic objects into real scenes. The light-based model is constructed from an approximate geometric model of the scene and by using a light probe to measure the incident illumination at the location of the synthetic objects. The global illumination solution is then composited into a photograph of the scene using the differential rendering technique. We conclude by discussing the relevance of the technique to recovering surface reflectance properties in uncontrolled lighting situations. Applications of the method include visual effects, interior design, and architectural visualization.

CR Descriptors: I.2.10 [**Artificial Intelligence**]: Vision and Scene Understanding - *Intensity, color, photometry and thresholding*; I.3.7 [**Computer Graphics**]: Three-Dimensional Graphics and Realism - *Color, shading, shadowing, and texture*; I.3.7 [**Computer Graphics**]: Three-Dimensional Graphics and Realism - *Radiosity*; I.4.1 [**Image Processing**]: Digitization - *Scanning*; I.4.8 [**Image Processing**]: Scene Analysis - *Photometry, Sensor Fusion*.

[1]Computer Science Division, University of California at Berkeley, Berkeley, CA 94720-1776. Email: debevec@cs.berkeley.edu. More information and additional results may be found at: http://www.cs.berkeley.edu/~debevec/Research

1 Introduction

Rendering synthetic objects into real-world scenes is an important application of computer graphics, particularly in architectural and visual effects domains. Oftentimes, a piece of furniture, a prop, or a digital creature or actor needs to be rendered seamlessly into a real scene. This difficult task requires that the objects be lit consistently with the surfaces in their vicinity, and that the interplay of light between the objects and their surroundings be properly simulated. Specifically, the objects should cast shadows, appear in reflections, and refract, focus, and emit light just as real objects would.

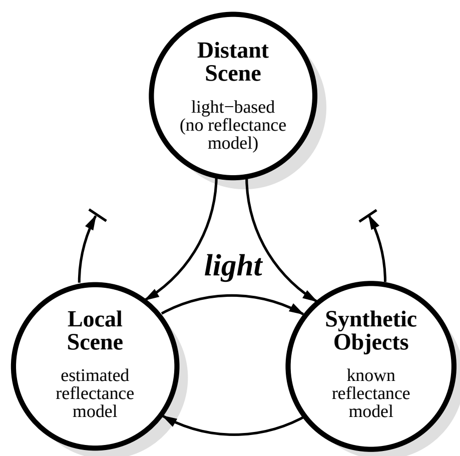


Figure 1: **The General Method** *In our method for adding synthetic objects into light-based scenes, the scene is partitioned into three components: the distant scene, the local scene, and the synthetic objects. Global illumination is used to simulate the interplay of light amongst all three components, except that light reflected back at the distant scene is ignored. As a result, BRDF information for the distant scene is unnecessary. Estimates of the geometry and material properties of the local scene are used to simulate the interaction of light between it and the synthetic objects.*

Currently available techniques for realistically rendering synthetic objects into scenes are labor intensive and not always successful. A common technique is to manually survey the positions of the light sources, and to instantiate a virtual light of equal color and intensity for each real light to illuminate the synthetic objects. Another technique is to photograph a reference object (such as a gray sphere) in the scene where the new object is to be rendered, and use its appearance as a qualitative guide in manually configuring the lighting environment. Lastly, the technique of reflection mapping is useful for mirror-like reflections. These methods typically require considerable hand-refinement and none of them easily simulates the effects of indirect illumination from the environment.

Accurately simulating the effects of both direct and indirect lighting has been the subject of research in global illumination. With a global illumination algorithm, if the entire scene were modeled with its full geometric and reflectance (BRDF) characteristics, one could correctly render a synthetic object into the scene simply by adding it to the model and recomputing the global illumination solution. Unfortunately, obtaining a full geometric and reflectance model of a large environment is extremeley difficult. Furthermore, global illumination solutions for large complex environments are extremely computationally intensive.

Moreover, it seems that having a full reflectance model of the large-scale scene should be unnecessary: under most circumstances, a new object will have no significant effect on the appearance of most of the of the distant scene. Thus, for such distant areas, knowing just its radiance (under the desired lighting conditions) should suffice.

Recently, [9] introduced a high dynamic range photographic technique that allows accurate measurements of scene radiance to be derived from a set of differently exposed photographs. This technique allows both low levels of indirect radiance from surfaces and high levels of direct radiance from light sources to be accurately recorded. When combined with image-based modeling techniques (e.g. [22, 24, 4, 10, 23, 17, 29]), and possibly active techniques for measuring geometry (e.g. [35, 30, 7, 27]) these derived radiance maps can be used to construct spatial representations of scene radiance.

We will use the term **light-based model** to refer to a representation of a scene that consists of radiance information, possibly with specific reference to light leaving surfaces, but not necessarily containing material property (BRDF) information. A light-based model can be used to evaluate the 5D plenoptic function [1] $P(\theta, \phi, V_x, V_y, V_z)$ for a given virtual or real subset of space[1]. A material-based model is converted to a light-based model by computing an illumination solution for it. A light-based model is differentiated from an image-based model in that its light values are actual measures of radiance[2], whereas image-based models may contain pixel values already transformed and truncated by the response function of an image acquisition or synthesis process.

In this paper, we present a general method for using accurate measurements of scene radiance in conjunction with global illumination to realistically add new objects to light-based models. The synthetic objects may have arbitrary material properties and can be rendered with appropriate illumination in arbitrary lighting environments. Furthermore, the objects can correctly interact with the environment around them: they cast the appropriate shadows, they are properly reflected, they can reflect and focus light, and they exhibit appropriate diffuse interreflection. The method can be carried out with commonly available equipment and software.

In this method (see Fig. 1), the scene is partitioned into three components. The first is the distant scene, which is the visible part of the environment too remote to be perceptibly affected by the synthetic object. The second is the local scene, which is the part of the environment which will be significantly affected by the presence of the objects. The third component is the synthetic objects. Our approach uses global illumination to correctly simulate the interaction of light amongst these three elements, with the exception that light radiated toward the distant environment will not be considered in the calculation. As a result, the BRDF of the distant environment need not be known — the technique uses BRDF information only for the local scene and the synthetic objects. We discuss the challenges in estimating the BRDF of the local scene, and methods for obtaining usable approximations. We also present a differential rendering technique that produces perceptually accurate results even when the estimated BRDF is somewhat inaccurate.

We demonstrate the general method for the specific case of rendering synthetic objects into particular views of a scene (such as background plates) rather than into a general image-based model. In this method, a light probe is used to acquire a high dynamic range panoramic radiance map near the location where the object will be rendered. A simple example of a light probe is a camera aimed at a mirrored sphere, a configuration commonly used for acquiring environment maps. An approximate geometric model of the scene is created (via surveying, photogrammetry, or 3D scanning) and mapped with radiance values measured with the light probe. The distant scene, local scene, and synthetic objects are rendered with global illumination from the same point of view as the background plate, and the results are composited into the background plate with a differential rendering technique.

1.1 Overview

The rest of this paper is organized as follows. In the next section we discuss work related to this paper. Section 3 introduces the basic technique of using acquired maps of scene radiance to illuminate synthetic objects. Section 4 presents the general method we will use to render synthetic objects into real scenes. Section 5 describes a practical technique based on this method using a *light probe* to measure incident illumination. Section 6 presents a differential rendering technique for rendering the local environment with only an approximate description of its reflectance. Section 7 presents a simple method to approximately recover the diffuse reflectance characteristics of the local environment. Section 8 presents results obtained with the technique. Section 9 discusses future directions for this work, and we conclude in Section 10.

2 Background and Related Work

The practice of adding new objects to photographs dates to the early days of photography in the simple form of pasting a cut-out from one picture onto another. While the technique conveys the idea of the new object being in the scene, it usually fails to produce an image that as a whole is a believable photograph. Attaining such realism requires a number of aspects of the two images to match. First, the camera projections should be consistent, otherwise the object may seem too foreshortened or skewed relative to the rest of the picture. Second, the patterns of film grain and film response should match. Third, the lighting on the object needs to be consistent with other objects in the environment. Lastly, the object needs to cast realistic shadows and reflections on the scene. Skilled artists found that by giving these considerations due attention, synthetic objects could be painted into still photographs convincingly.

In optical film compositing, the use of object mattes to prevent particular sections of film from being exposed made the same sort of cut-and-paste compositing possible for moving images. However, the increased demands of realism imposed by the dynamic nature of film made matching camera positions and lighting even more critical. As a result, care was taken to light the objects appropriately for the scene into which they were to be composited. This would still not account for the objects casting shadows onto the scene, so often these were painted in by an artist frame by frame [13, 2, 28]. Digital film scanning and compositing [26] helped make this process far more efficient.

Work in global illumination [16, 19] has recently produced algorithms (e.g. [31]) and software (e.g. [33]) to realistically simulate lighting in synthetic scenes, including indirect lighting with both specular and diffuse reflections. We leverage this work in order to create realistic renderings.

Some work has been done on the specific problem of compositing objects into photography. [25] presented a procedure for ren-

[1] Time and wavelength dependence can be included to represent the general 7D plenoptic function as appropriate.

[2] In practice, the measures of radiance are with respect to a discrete set of spectral distributions such as the standard tristimulus model.

dering architecture into background photographs using knowledge of the sun position and measurements or approximations of the local ambient light. For diffuse buildings in diffuse scenes, the technique is effective. The technique of *reflection mapping* (also called *environment mapping*) [3, 18] produces realistic results for mirror-like objects. In reflection mapping, a panoramic image is rendered or photographed from the location of the object. Then, the surface normals of the object are used to index into the panoramic image by reflecting rays from the desired viewpoint. As a result, the shiny object appears to properly reflect the desired environment[3]. However, the technique is limited to mirror-like reflection and does not account for objects casting light or shadows on the environment.

A common visual effects technique for having synthetic objects cast shadows on an existing environment is to create an approximate geometric model of the environment local to the object, and then compute the shadows from the various light sources. The shadows can then be subtracted from the background image. In the hands of professional artists this technique can produce excellent results, but it requires knowing the position, size, shape, color, and intensity of each of the scene's light sources. Furthermore, it does not account for diffuse reflection from the scene, and light reflected by the objects onto the scene must be handled specially.

To properly model the interaction of light between the objects and the local scene, we pose the compositing problem as a global illumination computation as in [14] and [12]. As in this work, we apply the effect of the synthetic objects in the lighting solution as a differential update to the original appearance of the scene. In the previous work an approximate model of the entire scene and its original light sources is constructed; the positions and sizes of the light sources are measured manually. Rough methods are used to estimate diffuse-only reflectance characteristics of the scene, which are then used to estimate the intensities of the light sources. [12] additionally presents a method for performing fast updates of the illumination solution in the case of moving objects. As in the previous work, we leverage the basic result from incremental radiosity [6, 5] that making a small change to a scene does not require recomputing the entire solution.

3 Illuminating synthetic objects with real light

In this section we propose that computer-generated objects be lit by actual recordings of light from the scene, using global illumination. Performing the lighting in this manner provides a unified and physically accurate alternative to manually attempting to replicate incident illumination conditions.

Accurately recording light in a scene is difficult because of the high dynamic range that scenes typically exhibit; this wide range of brightness is the result of light sources being relatively concentrated. As a result, the intensity of a source is often two to six orders of magnitude larger than the intensity of the non-emissive parts of an environment. However, it is necessary to accurately record both the large areas of indirect light from the environment and the concentrated areas of direct light from the sources since both are significant parts of the illumination solution.

Using the technique introduced in [9], we can acquire correct measures of scene radiance using conventional imaging equipment. The images, called *radiance maps*, are derived from a series of images with different sensor integration times and a technique for computing and accounting for the imaging system response function f. We can use these measures to illuminate synthetic objects exhibiting arbitrary material properties.

Fig. 2 shows a high-dynamic range lighting environment with electric, natural, and indirect lighting. This environment was recorded by taking a full dynamic range photograph of a mirrored ball on a table (see Section 5). A digital camera was used to acquire a series images in one-stop exposure increments from $\frac{1}{4}$ to $\frac{1}{10000}$ second. The images were fused using the technique in [9].

The environment is displayed at three exposure levels (-0, -3.5, and -7.0 stops) to show its full dynamic range. Recovered RGB radiance values for several points in the scene and on the two major light sources are indicated; the color difference between the tungsten lamp and the sky is evident. A single low-dynamic range photograph would be unable to record the correct colors and intensities over the entire scene.

Fig. 3(a-e) shows the results of using this panoramic radiance map to synthetically light a variety of materials using the RADIANCE global illumination algorithm [33]. The materials are: (a) perfectly reflective, (b) rough gold, (c) perfectly diffuse gray material, (d) shiny green plastic, and (e) dull orange plastic. Since we are computing a full illumination solution, the objects exhibit self-reflection and shadows from the light sources as appropriate. Note that in (c) the protrusions produce two noticeable shadows of slightly different colors, one corresponding to the ceiling light and a softer shadow corresponding to the window.

The shiny plastic object in (d) has a 4 percent specular component with a Gaussian roughness of 0.04 [32]. Since the object's surface both blurs and attenuates the light with its rough specular component, the reflections fall within the dynamic range of our display device and the different colors of the light sources can be seen. In (e) the rough plastic diffuses the incident light over a much larger area.

To illustrate the importance of using high dynamic range radiance maps, the same renderings were produced using just one of the original photographs as the lighting environment. In this single image, similar in appearance to Fig. 2(a), the brightest regions had been truncated to approximately 2 percent of their true values. The rendering of the mirrored surface (f) appears similar to (a) since it is displayed in low-dynamic range printed form. Significant errors are noticeable in (g-j) since these materials blur the incident light. In (g), the blurring of the rough material darkens the light sources, whereas in (b) they remain saturated. Renderings (h-j) are very dark due to the missed light; thus we have brightened by a factor of eight on the right in order to make qualitative comparisons to (c-e) possible. In each it can be seen that the low-dynamic range image of the lighting environment fails to capture the information necessary to simulate correct color balance, shadows, and highlights.

Fig. 4 shows a collection of objects with different material properties illuminated by two different environments. A wide variety of light interaction between the objects and the environment can be seen. The (synthetic) mirrored ball reflects both the synthetic objects as well as the environment. The floating diffuse ball shows a subtle color shift along its right edge as it shadows itself from the windows and is lit primarily by the incandescent lamp in Fig. 4(a). The reflection of the environment in the black ball (which has a specular intensity of 0.04) shows the colors of the light sources, which are too bright to be seen in the mirrored ball. A variety of shadows, reflections, and focused light can be observed on the resting surface.

The next section describes how the technique of using radiance maps to illuminate synthetic objects can be extended to compute the proper photometric interaction of the objects with the scene. It also describes how high dynamic range photography and image-based modeling combine in a natural manner to allow the simulation of arbitrary (non-infinite) lighting environments.

4 The General Method

This section explains our method for adding new objects to light-based scene representations. As in Fig. 1, we partition our scene into three parts: the distant scene, the local scene, and the synthetic

[3]Using the surface normal indexing method, the object will not reflect itself. Correct self-reflection can be obtained through ray tracing.

Figure 2: **An omnidirectional radiance map** This full dynamic range lighting environment was acquired by photographing a mirrored ball balanced on the cap of a pen sitting on a table. The environment contains natural, electric, and indirect light. The three views of this image adjusted to (a) +0 stops, (b) -3.5 stops, and (c) -7.0 stops show that the full dynamic range of the scene has been captured without saturation. As a result, the image usefully records the direction, color, and intensity of all forms of incident light.

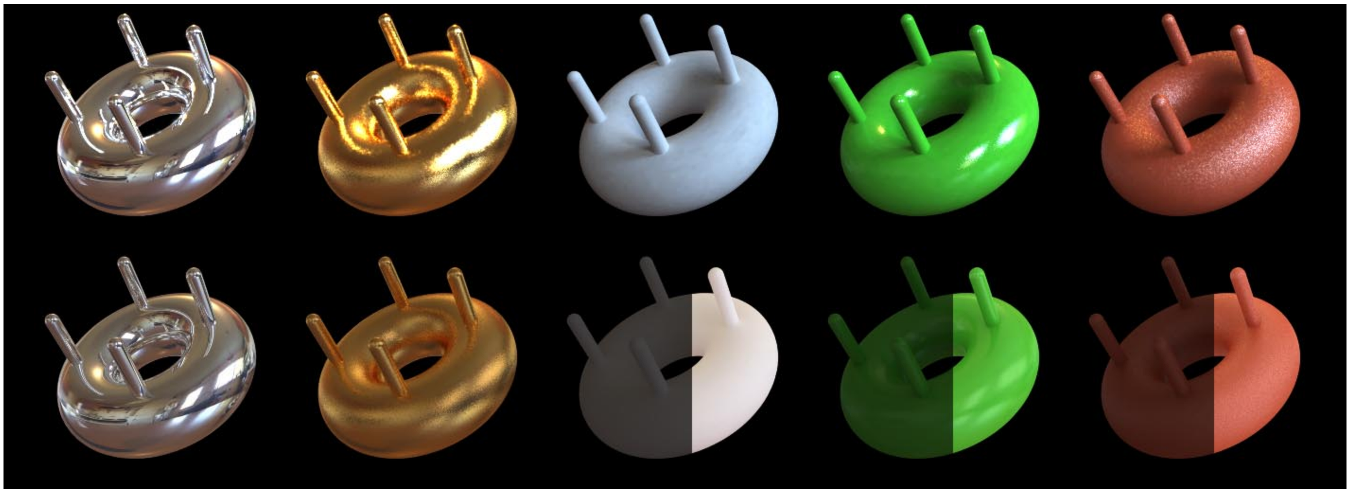


Figure 3: **Illuminating synthetic objects with real light (Top row: a,b,c,d,e)** *With full dynamic range measurements of scene radiance from Fig. 2.* **(Bottom row: f,g,h,i,j)** *With low dynamic range information from a single photograph of the ball. The right sides of images (h,i,j) have been brightened by a factor of six to allow qualitative comparison to (c,d,e). The high dynamic range measurements of scene radiance are necessary to produce proper lighting on the objects.*

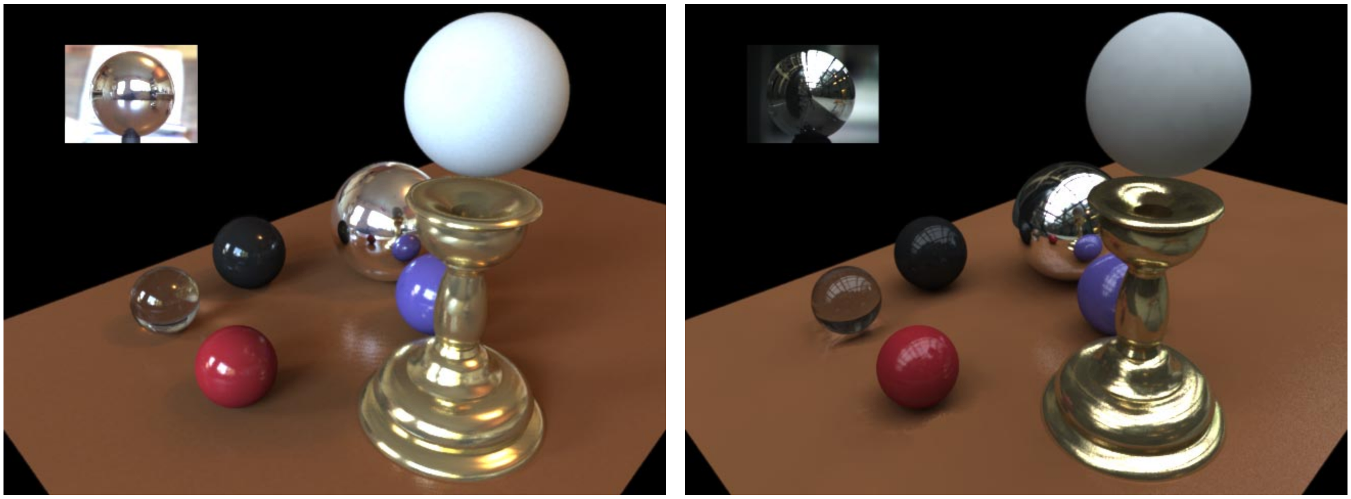


Figure 4: **Synthetic objects lit by two different environments (a)** *A collection of objects is illuminated by the radiance information in 2. The objects exhibit appropriate interreflection.* **(b)** *The same objects are illuminated by different radiance information obtained in an outdoor urban environment on an overcast day. The radiance map used for the illumination is shown in the upper left of each image. Candle holder model courtesy of Gregory Ward Larson.*

objects. We describe the geometric and photometric requirements for each of these components.

1. **A light-based model of the distant scene**
   The distant scene is constructed as a light-based model. The synthetic objects will receive light from this model, so it is necessary that the model store true measures of radiance rather than low dynamic range pixel values from conventional images. The light-based model can take on any form, using very little explicit geometry [23, 17], some geometry [24], moderate geometry [10], or be a full 3D scan of an environment with view-dependent texture-mapped [11] radiance. What is important is for the model to provide accurate measures of incident illumination in the vicinity of the objects, as well as from the desired viewpoint. In the next section we will present a convenient procedure for constructing a minimal model that meets these requirements.
   In the global illumination computation, the distant scene radiates light toward the local scene and the synthetic objects, but ignores light reflected back to it. We assume that no area of the distant scene will be significantly affected by light reflecting from the synthetic objects; if that were the case, the area should instead belong to the local scene, which contains the BRDF information necessary to interact with light. In the RADIANCE [33] system, this exclusively emissive behavior can be specified with the ”glow” material property.
2. **An approximate material-based model of the local scene**
   The local scene consists of the surfaces that will photometrically interact with the synthetic objects. It is this geometry onto which the objects will cast shadows and reflect light. Since the local scene needs to fully participate in the illumination solution, both its geometry and reflectance characteristics should be known, at least approximately. If the geometry of the local scene is not readily available with sufficient accuracy from the light-based model of the distant scene, there are various techniques available for determining its geometry through active or passive methods. In the common case where the local scene is a flat surface that supports the synthetic objects, its geometry is determined easily from the camera pose. Methods for estimating the BRDF of the local scene are discussed in Section 7.
   Usually, the local scene will be the part of the scene that is geometrically close to the synthetic objects. When the local scene is mostly diffuse, the rendering equation shows that the visible effect of the objects on the local scene decreases as the inverse square of the distance between the two. Nonetheless, there is a variety of circumstances in which synthetic objects can significantly affect areas of the scene not in the immediate vicinity. Some common circumstances are:
   - If there are concentrated light sources illuminating the object, then the object can cast a significant shadow on a distant surface collinear with it and the light source.
   - If there are concentrated light sources and the object is flat and specular, it can focus a significant amount of light onto a distant part of the scene.
   - If a part of the distant scene is flat and specular (e.g. a mirror on a wall), its appearance can be significantly affected by a synthetic object.
   - If the synthetic object emits light (e.g. a synthetic laser), it can affect the appearance of the distant scene significantly.

   These situations should be considered in choosing which parts of the scene should be considered local and which parts distant. Any part of the scene that will be significantly affected in its appearance from the desired viewpoint should be included as part of the local scene.
   Since the local scene is a full BRDF model, it can be added to the global illumination problem as would any other object. The local scene may consist of any number of surfaces and objects with different material properties. For example, the local scene could consist of a patch of floor beneath the synthetic object to catch shadows as well as a mirror surface hanging on the opposite wall to catch a reflection. The local scene replaces the corresponding part of the light-based model of the distant scene.
   Since it can be difficult to determine the precise BRDF characteristics of the local scene, it is often desirable to have only the *change* in the local scene's appearance be computed with the BRDF estimate; its appearance due to illumination from the distant scene is taken from the original light-based model. This differential rendering method is presented in Section 6.
3. **Complete material-based models of the objects**
   The synthetic objects themselves may consist of any variety of shapes and materials supported by the global illumination software, including plastics, metals, emitters, and dielectrics such as glass and water. They should be placed in their desired geometric correspondence to the local scene.

Once the distant scene, local scene, and synthetic objects are properly modeled and positioned, the global illumination software can be used in the normal fashion to produce renderings from the desired viewpoints.

5 Compositing using a light probe

This section presents a particular technique for constructing a light-based model of a real scene suitable for adding synthetic objects at a particular location. This technique is useful for compositing objects into actual photography of a scene.

In Section 4, we mentioned that the light-based model of the distant scene needs to appear correctly in the vicinity of the synthetic objects as well as from the desired viewpoints. This latter requirement can be satisfied if it is possible to directly acquire radiance maps of the scene from the desired viewpoints. The former requirement, that the appear photometrically correct in all directions in the vicinity of the synthetic objects, arises because this information comprises the incident light which will illuminate the objects.

To obtain this part of the light-based model, we acquire a full dynamic range omnidirectional radiance map near the location of the synthetic object or objects. One technique for acquiring this radiance map is to photograph a spherical first-surface mirror, such as a polished steel ball, placed at or near the desired location of the synthetic object[4]. This procedure is illustrated in Fig. 7(a). An actual radiance map obtained using this method is shown in Fig. 2.

The radiance measurements observed in the ball are mapped onto the geometry of the distant scene. In many circumstances this model can be very simple. In particular, if the objects are small and resting on a flat surface, one can model the scene as a horizontal plane for the resting surface and a large dome for the rest of the environment. Fig. 7(c) illustrates the ball image being mapped onto a table surface and the walls and ceiling of a finite room; 5 shows the resulting light-based model.

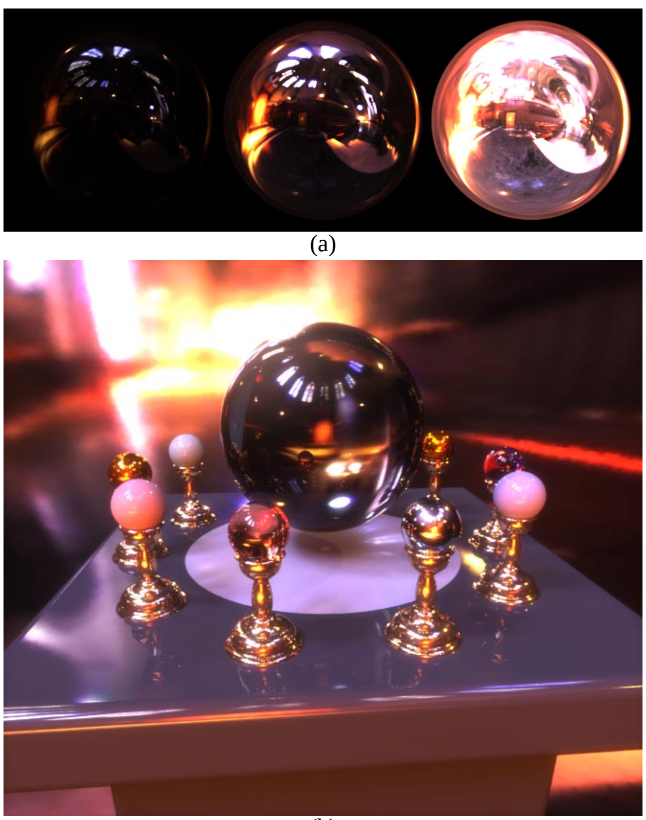
5.1 Mapping from the probe to the scene model

To precisely determine the mapping between coordinates on the ball and rays in the world, one needs to record the position of the ball

[4]Parabolic mirrors combined with telecentric lenses [34] can be used to obtain hemispherical fields of view with a consistent principal point, if so desired.

relative to the camera, the size of the ball, and the camera parameters such as its location in the scene and focal length. With this information, it is straightforward to trace rays from the camera center through the pixels of the image, and reflect rays off the ball into the environment. Often a good approximation results from assuming the ball is small relative to the environment and that the camera's view is orthographic.

The data acquired from a single ball image will exhibit a number of artifacts. First, the camera (and possibly the photographer) will be visible. The ball, in observing the scene, interacts with it: the ball (and its support) can appear in reflections, cast shadows, and can reflect light back onto surfaces. Lastly, the ball will not reflect the scene directly behind it, and will poorly sample the area nearby. If care is taken in positioning the ball and camera, these effects can be minimized and will have a negligible effect on the final renderings. If the artifacts are significant, the images can be fixed manually in image editing program or by selectively combining images of the ball taken from different directions; Fig. 6 shows a relatively artifact-free enviroment constructed using the latter method. We have found that combining two images of the ball taken ninety degrees apart from each other allows us to eliminate the camera's appearance and to avoid poor sampling.



(a)

(b)

Figure 6: **Rendering with a Combined Probe Image** *The full dynamic range environment map shown at the top was assembled from two light probe images taken ninety degrees apart from each other. As a result, the only visible artifact is small amount of the probe support visible on the floor. The map is shown at -4.5, 0, and +4.5 stops. The bottom rendering was produced using this lighting information, and exhibits diffuse and specular reflections, shadows from different sources of light, reflections, and caustics.*

5.2 Creating renderings

To render the objects into the scene, a synthetic local scene model is created as described in Section 4. Images of the scene from the desired viewpoint(s) are taken (Fig. 7(a)), and their position relative to the scene is recorded through pose-instrumented cameras or (as in our work) photogrammetry. The location of the ball in the scene is also recorded at this time. The global illumination software is then run to render the objects, local scene, and distant scene from the desired viewpoint (Fig. 7(d)).

The objects and local scene are then composited onto the background image. To perform this compositing, a mask is created by rendering the objects and local scene in white and the distant scene in black. If objects in the distant scene (which may appear in front of the objects or local scene from certain viewpoints) are geometrically modeled, they will properly obscure the local scene and the objects as necessary. This compositing can be considered as a subset of the general method (Section 4) wherein the light-based model of the distant scene acts as follows: if (V_x, V_y, V_z) corresponds to an actual view of the scene, return the radiance value looking in direction (θ, ϕ). Otherwise, return the radiance value obtained by casting the ray $(\theta, \phi, V_x, V_y, V_z)$ onto the radiance-mapped distant scene model.

In the next section we describe a more robust method of compositing the local scene into the background image.

6 Improving quality with differential rendering

The method we have presented so far requires that the local scene be modeled accurately in both its geometry and its spatially varying material properties. If the model is inaccurate, the appearance of the local scene will not be consistent with the appearance of adjacent distant scene. Such a border is readily apparent in Fig. 8(c), since the local scene was modeled with a homogeneous BRDF when in reality it exhibits a patterned albedo (see [21]). In this section we describe a method for greatly reducing such effects.

Suppose that we compute a global illumination solution for the local and distant scene models without including the synthetic objects. If the BRDF and geometry of the local scene model were perfectly accurate, then one would expect the appearance of the rendered local scene to be consistent with its appearance in the light-based model of the entire scene. Let us call the appearance of the local scene from the desired viewpoint in the light-based model LS_b. In the context of the method described in Section 5, LS_b is simply the background image. We will let LS_{noobj} denote the appearance of the local scene, without the synthetic objects, as calculated by the global illumination solution. The error in the rendered local scene (without the objects) is thus: $Err_{ls} = LS_{noobj} - LS_b$. This error results from the difference between the BRDF characteristics of the actual local scene as compared to the modeled local scene.

Let LS_{obj} denote the appearance of the local environment as calculated by the global illumination solution with the synthetic objects in place. We can compensate for the error if we compute our final rendering LS_{final} as:

$$LS_{final} = LS_{obj} - Err_{ls}$$

Equivalently, we can write:

$$LS_{final} = LS_b + (LS_{obj} - LS_{noobj})$$

In this form, we see that whenever LS_{obj} and LS_{noobj} are the same (i.e. the addition of the objects to the scene had no effect on the local scene) the final rendering of the local scene is equivalent to LS_b (e.g. the background plate). When LS_{obj} is darker than LS_{noobj}, light is subtracted from the background to form shadows,

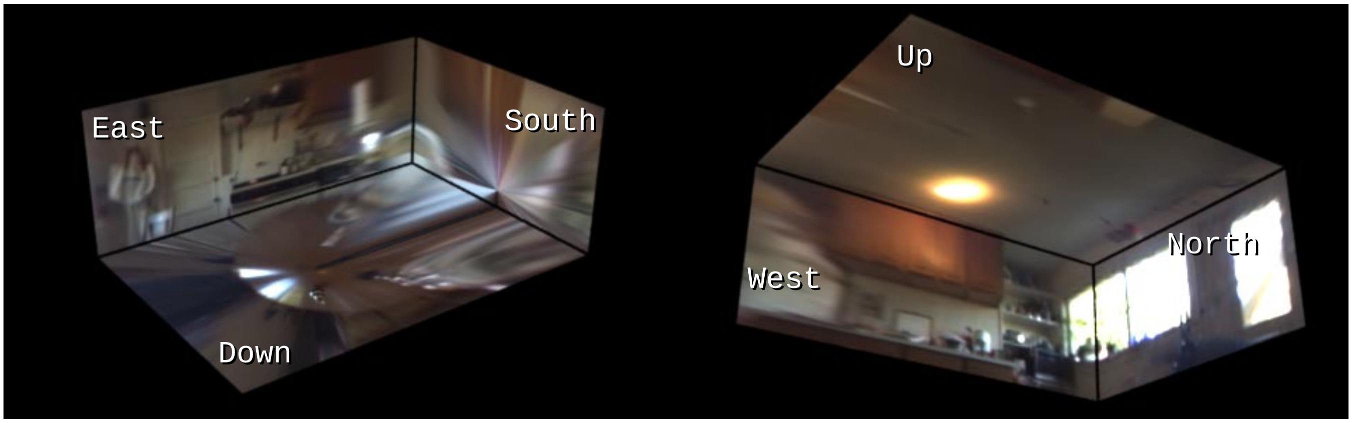


Figure 5: **A Light-Based Model** *A simple light-based model of a room is constructed by mapping the image from a light probe onto a box. The box corresponds to the upper half of the room, with the bottom face of the box being coincident with the top of the table. The model contains the full dynamic range of the original scene, which is not reproduced in its entirety in this figure.*

and when LS_{obj} is lighter than LS_{noobj} light is added to the background to produce reflections and caustics.

Stated more generally, the appearance of the local scene without the objects is computed with the correct reflectance characteristics lit by the correct environment, and the change in appearance due to the presence of the synthetic objects is computed with the modeled reflectance characteristics as lit by the modeled environment. While the realism of LS_{final} still benefits from having a good model of the reflectance characteristics of the local scene, the perceptual effect of small errors in albedo or specular properties is considerably reduced. Fig. 8(g) shows a final rendering in which the local environment is computed using this differential rendering technique. The objects are composited into the image directly from the LS_{obj} solution shown in Fig. 8(c).

It is important to stress that this technique can still produce abitrarily wrong results depending on the amount of error in the estimated local scene BRDF and the inaccuracies in the light-based model of the distance scene. In fact, Err_{ls} may be larger than LS_{obj}, causing LS_{final} to be negative. An alternate approach is to compensate for the *relative* error in the appearance of the local scene: $LS_{final} = LS_b(LS_{obj}/LS_{noobj})$. Inaccuracies in the local scene BDRF will also be reflected in the objects.

In the next section we discuss techniques for estimating the BRDF of the local scene.

7 Estimating the local scene BRDF

Simulating the interaction of light between the local scene and the synthetic objects requires a model of the reflectance characteristics of the local scene. Considerable recent work [32, 20, 8, 27] has presented methods for measuring the reflectance properties of materials through observation under controlled lighting configurations. Furthermore, reflectance characteristics can also be measured with commercial radiometric devices.

It would be more convenient if the local scene reflectance could be estimated directly from observation. Since the light-based model contains information about the radiance of the local scene as well as its irradiance, it actually contains information about the local scene reflectance. If we hypothesize reflectance characteristics for the local scene, we can illuminate the local scene with its known irradiance from the light-based model. If our hypothesis is correct, then the appearance should be consistent with the measured appearance. This suggests the following iterative method for recovering the reflectance properties of the local scene:

1. Assume a reflectance model for the local scene (e.g. diffuse only, diffuse + specular, metallic, or arbitrary BRDF, including spatial variation)
2. Choose approximate initial values for the parameters of the reflectance model
3. Compute a global illumination solution for the local scene with the current parameters using the observed lighting configuration or configurations.
4. Compare the appearance of the rendered local scene to its actual appearance in one or more views.
5. If the renderings are not consistent, adjust the parameters of the reflectance model and return to step 3.

Efficient methods of performing the adjustment in step 5 that exploit the properties of particular reflectance models are left as future work. However, assuming a diffuse-only model of the local scene in step 1 makes the adjustment in step 5 straightforward. We have:

$$L_{r1}(\theta_r, \phi_r) = \int_0^{2\pi} \int_0^{\pi/2} \rho_d L_i(\theta_i, \phi_i) \cos\theta_i \, \sin\theta_i \, d\theta_i \, d\phi_i =$$

$$\rho_d \int_0^{2\pi} \int_0^{\pi/2} L_i(\theta_i, \phi_i) \cos\theta_i \, \sin\theta_i \, d\theta_i \, d\phi_i$$

If we initialize the local scene to be perfectly diffuse ($\rho_d = 1$) everywhere, we have:

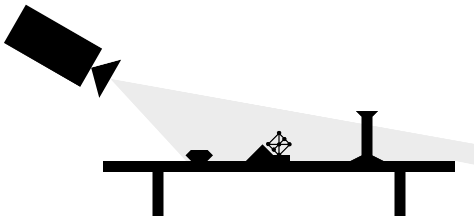
$$L_{r2}(\theta_r, \phi_r) = \int_0^{2\pi} \int_0^{\pi/2} L_i(\theta_i, \phi_i) \cos\theta_i \, \sin\theta_i \, d\theta_i \, d\phi_i$$

The updated diffuse reflectance coefficient for each part of the local scene can be computed as:

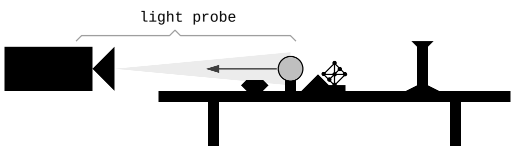
$$\rho_d' = \frac{L_{r1}(\theta_r, \phi_r)}{L_{r2}(\theta_r, \phi_r)}$$

In this manner, we use the global illumination calculation to render each patch as a perfectly diffuse reflector, and compare the resulting radiance to the observed value. Dividing the two quantities yields the next estimate of the diffuse reflection coefficient ρ_d'. If there is no interreflection within the local scene, then the ρ_d' estimates will make the renderings consistent. If there is interreflection, then the algorithm should be iterated until there is convergence.

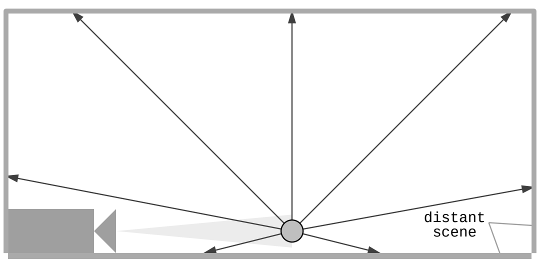
For a trichromatic image, the red, green, and blue diffuse reflectance values are computed independently. The diffuse characteristics of the background material used to produce Fig. 8(c) were



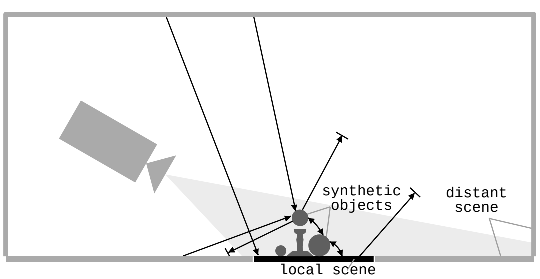
(a) Acquiring the background photograph



(b) Using the light probe



(c) Constructing the light-based model



(d) Computing the global illumination solution

Figure 7: **Using a light probe (a)** *The background plate of the scene (some objects on a table) is taken.* **(b)** *A light probe (in this case, the camera photographing a steel ball) records the incident radiance near the location of where the synthetic objects are to be placed.* **(c)** *A simplified light-based model of the distant scene is created as a planar surface for the table and a finite box to represent the rest of the room. The scene is texture-mapped in high dynamic range with the radiance map from the light probe. The objects on the table, which were not explicitly modeled, become projected onto the table.* **(d)** *Synthetic objects and a BRDF model of the local scene are added to the light-based model of the distant scene. A global illumination solution of this configuration is computed with light coming from the distant scene and interacting with the local scene and synthetic objects. Light reflected back to the distant scene is ignored. The results of this rendering are composited (possibly with differential rendering) into the background plate from (a) to achieve the final result.*

computed using this method, although it was assumed that the entire local scene had the same diffuse reflectance.

In the standard "plastic" illumination model, just two more coefficients – those for specular intensity and roughness – need to be specified. In Fig. 8, the specular coefficients for the local scene were estimated manually based on the specular reflection of the window in the table in Fig. 2.

8 Compositing Results

Fig. 5 shows a simple light-based model of a room constructed using the panoramic radiance map from Fig. 2. The room model begins at the height of the table and continues to the ceiling; its measurements and the position of the ball within it were measured manually. The table surface is visible on the bottom face. Since the room model is finite in size, the light sources are effectively local rather than infinite. The stretching on the south wall is due to the poor sampling toward the silhouette edge of the ball.

Figs. 4 and 6 show complex arrangements of synthetic objects lit entirely by a variety of light-based models. The selection and composition of the objects in the scene was chosen to exhibit a wide variety of light interactions, including diffuse and specular reflectance, multiple soft shadows, and reflected and focused light. Each rendering was produced using the RADIANCE system with two diffuse light bounces and a relatively high density of ambient sample points.

Fig. 8(a) is a background plate image into which the synthetic objects will be rendered. In 8(b) a calibration grid was placed on the table in order to determine the camera pose relative to the scene and to the mirrored ball, which can also be seen. The poses were determined using the photogrammetric method in [10]. In 8(c), a model of the local scene as well as the synthetic objects is geometrically matched and composited onto the background image. Note that the local scene, while the same average color as the table, is readily distinguishable at its edges and because it lacks the correct variations in albedo.

Fig. 8(d) shows the results of lighting the local scene model with the light-based model of the room, without the objects. This image will be compared to 8(c) in order to determine the effect the synthetic objects have on the local scene. Fig. 8(e) is a mask image in which the white areas indicate the location of the synthetic objects. If the distant or local scene were to occlude the objects, such regions would be dark in this image.

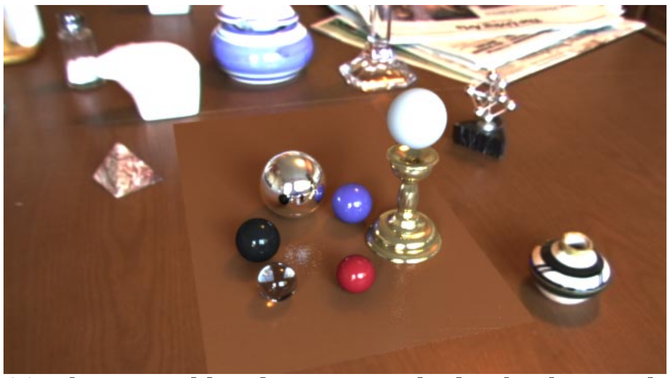
Fig. 8(f) shows the difference between the appearance of the local scene rendered with (8(c)) and without (8(d)) the objects. For illustration purposes, the difference in radiance values have been offset so that zero difference is shown in gray. The objects have been masked out using image 8(e). This difference image encodes both the shadowing (dark areas) and reflected and focussed light (light areas) imposed on the local scene by the addition of the synthetic objects.

Fig. 8(g) shows the final result using the differential rendering method described in Section 6. The synthetic objects are copied directly from the global illumination solution 8(c) using the object mask 8(e). The effects the objects have on the local scene are included by adding the difference image 8(f) (without offset) to the background image. The remainder of the scene is copied directly from the background image 8(a). Note that in the mirror ball's reflection, the modeled local scene can be observed without the effects of differential rendering — a limitation of the compositing technique.

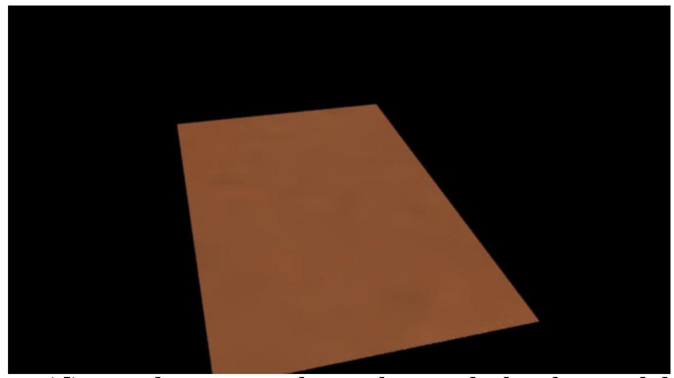
In this final rendering, the synthetic objects exhibit a consistent appearance with the real objects present in the background image 8(a) in both their diffuse and specular shading, as well as the direction and coloration of their shadows. The somewhat speckled nature of the object reflections seen in the table surface is due to

(a) Background photograph

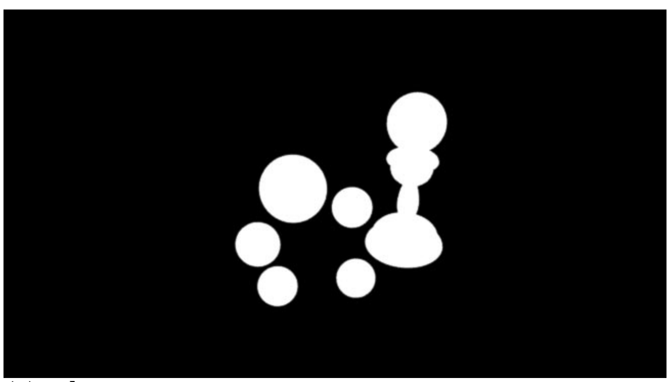
(b) Camera calibration grid and light probe



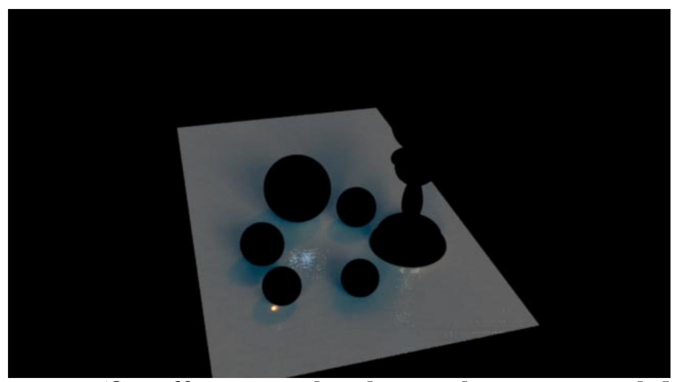
(c) Objects and local scene matched to background



(d) Local scene, without objects, lit by the model



(e) Object matte



(f) Difference in local scene between **c** and **d**

(g) Final result with differential rendering

Figure 8: Compositing synthetic objects into a real scene using a light probe and differential rendering

the stochastic nature of the particular global illumination algorithm used.

The differential rendering technique successfully eliminates the border between the local scene and the background image seen in 8(c). Note that the albedo texture of the table in the local scene area is preserved, and that a specular reflection of a background object on the table (appearing just to the left of the floating sphere) is correctly preserved in the final rendering. The local scene also exhibits reflections from the synthetic objects. A caustic from the glass ball focusing the light of the ceiling lamp onto the table is evident.

9 Future work

The method proposed here suggests a number of areas for future work. One area is to investigate methods of automatically recovering more general reflectance models for the local scene geometry, as proposed in Section 7. With such information available, the program might also also be able to suggest which areas of the scene should be considered as part of the local scene and which can safely be considered distant, given the position and reflectance characteristics of the desired synthetic objects.

Some additional work could be done to allow the global illumination algorithm to compute the ilumination solution more efficiently. One technique would be to have an algorithm automatically locate and identify concentrated light sources in the light-based model of the scene. With such knowledge, the algorithm could compute most of the direct illumination in a forward manner, which could dramatically increase the efficiency with which an accurate solution could be calculated. To the same end, use of the method presented in [15] to expedite the solution could be investigated. For the case of compositing moving objects into scenes, greatly increased efficiency could be obtained by adapting incremental radiosity methods to the current framework.

10 Conclusion

We have presented a general framework for adding new objects to light-based models with correct illumination. The method leverages a technique of using high dynamic range images of real scene radiance to synthetically illuminate new objects with arbitrary reflectance characteristics. We leverage this technique in a general method to simulate interplay of light between synthetic objects and the light-based environment, including shadows, reflections, and caustics. The method can be implemented with standard global illumination techniques.

For the particular case of rendering synthetic objects into real scenes (rather than general light-based models), we have presented a practical instance of the method that uses a light probe to record incident illumination in the vicinity of the synthetic objects. In addition, we have described a differential rendering technique that can convincingly render the interplay of light between objects and the local scene when only approximate reflectance information for the local scene is available. Lastly, we presented an iterative approach for determining reflectance characteristics of the local scene based on measured geometry and observed radiance in uncontrolled lighting conditions. It is our hope that the techniques presented here will be useful in practice as well as comprise a useful framework for combining material-based and light-based graphics.

Acknowledgments

The author wishes to thank Chris Bregler, David Forsyth, Jianbo Shi, Charles Ying, Steve Chenney, and Andrean Kalemis for the various forms of help and advice they provided. Special gratitude is also due to Jitendra Malik for helping make this work possible. Discussions with Michael Naimark and Steve Saunders helped motivate this work. Tim Hawkins provided extensive assistance on improving and revising this paper and provided invaluable assistance with image acquisition. Gregory Ward Larson deserves great thanks for the RADIANCE lighting simulation system and his invaluable assistance and advice in using RADIANCE in this research, for assisting with reflectance measurements, and for very helpful comments and suggestions on the paper. This research was supported by a Multidisciplinary University Research Initiative on three dimensional direct visualization from ONR and BMDO, grant FDN00014-96-1-1200.

References

[1] ADELSON, E. H., AND BERGEN, J. R. *Computational Models of Visual Processing*. MIT Press, Cambridge, Mass., 1991, ch. 1. The Plenoptic Function and the Elements of Early Vision.

[2] AZARMI, M. *Optical Effects Cinematography: Its Development, Methods, and Techniques*. University Microfilms International, Ann Arbor, Michigan, 1973.

[3] BLINN, J. F. Texture and reflection in computer generated images. *Communications of the ACM 19*, 10 (October 1976), 542–547.

[4] CHEN, E. QuickTime VR - an image-based approach to virtual environment navigation. In *SIGGRAPH '95* (1995).

[5] CHEN, S. E. Incremental radiosity: An extension of progressive radiosity to an interactive synthesis system. In *SIGGRAPH '90* (1990), pp. 135–144.

[6] COHEN, M. F., CHEN, S. E., WALLACE, J. R., AND GREENBERG, D. P. A progressive refinement approach to fast radiosity image generation. In *SIGGRAPH '88* (1988), pp. 75–84.

[7] CURLESS, B., AND LEVOY, M. A volumetric method for building complex models from range images. In *SIGGRAPH '96* (1996), pp. 303–312.

[8] DANA, K. J., GINNEKEN, B., NAYAR, S. K., AND KOENDERINK, J. J. Reflectance and texture of real-world surfaces. In *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.* (1997), pp. 151–157.

[9] DEBEVEC, P. E., AND MALIK, J. Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH '97* (August 1997), pp. 369–378.

[10] DEBEVEC, P. E., TAYLOR, C. J., AND MALIK, J. Modeling and rendering architecture from photographs: A hybrid geometry- and image-based approach. In *SIGGRAPH '96* (August 1996), pp. 11–20.

[11] DEBEVEC, P. E., YU, Y., AND BORSHUKOV, G. D. Efficient view-dependent image-based rendering with projective texture-mapping. Tech. Rep. UCB//CSD-98-1003, University of California at Berkeley, 1998.

[12] DRETTAKIS, G., ROBERT, L., AND BOUGNOUX, S. Interactive common illumination for computer augmented reality. In *8th Eurographics workshop on Rendering, St. Etienne, France* (May 1997), J. Dorsey and P. Slusallek, Eds., pp. 45–57.

[13] FIELDING, R. *The Technique of Special Effects Cinematography*. Hastings House, New York, 1968.

[14] FOURNIER, A., GUNAWAN, A., AND ROMANZIN, C. Common illumination between real and computer generated scenes. In *Graphics Interface* (May 1993), pp. 254–262.

[15] GERSHBEIN, R., SCHRODER, P., AND HANRAHAN, P. Textures and radiosity: Controlling emission and reflection with texture maps. In *SIGGRAPH '94* (1994).

[16] GORAL, C. M., TORRANCE, K. E., GREENBERG, D. P., AND BATTAILE, B. Modeling the interaction of light between diffuse surfaces. In *SIGGRAPH '84* (1984), pp. 213–222.

[17] GORTLER, S. J., GRZESZCZUK, R., SZELISKI, R., AND COHEN, M. F. The Lumigraph. In *SIGGRAPH '96* (1996), pp. 43–54.

[18] HECKBERT, P. S. Survey of texture mapping. *IEEE Computer Graphics and Applications 6*, 11 (November 1986), 56–67.

[19] KAJIYA, J. The rendering equation. In *SIGGRAPH '86* (1986), pp. 143–150.

[20] KARNER, K. F., MAYER, H., AND GERVAUTZ, M. An image based measurement system for anisotropic reflection. In *EUROGRAPHICS Annual Conference Proceedings* (1996).

[21] KOENDERINK, J. J., AND VAN DOORN, A. J. Illuminance texture due to surface mesostructure. *J. Opt. Soc. Am. 13*, 3 (1996).

[22] LAVEAU, S., AND FAUGERAS, O. 3-D scene representation as a collection of images. In *Proceedings of 12th International Conference on Pattern Recognition* (1994), vol. 1, pp. 689–691.

[23] LEVOY, M., AND HANRAHAN, P. Light field rendering. In *SIGGRAPH '96* (1996), pp. 31–42.

[24] MCMILLAN, L., AND BISHOP, G. Plenoptic Modeling: An image-based rendering system. In *SIGGRAPH '95* (1995).

[25] NAKAMAE, E., HARADA, K., AND ISHIZAKI, T. A montage method: The overlaying of the computer generated images onto a background photograph. In *SIGGRAPH '86* (1986), pp. 207–214.

[26] PORTER, T., AND DUFF, T. Compositing digital images. In *SIGGRAPH 84* (July 1984), pp. 253–259.

[27] SATO, Y., WHEELER, M. D., AND IKEUCHI, K. Object shape and reflectance modeling from observation. In *SIGGRAPH '97* (1997), pp. 379–387.

[28] SMITH, T. G. *Industrial Light and Magic: The Art of Special Effects*. Ballantine Books, New York, 1986.

[29] SZELISKI, R. Image mosaicing for tele-reality applications. In *IEEE Computer Graphics and Applications* (1996).

[30] TURK, G., AND LEVOY, M. Zippered polygon meshes from range images. In *SIGGRAPH '94* (1994), pp. 311–318.

[31] VEACH, E., AND GUIBAS, L. J. Metropolis light transport. In *SIGGRAPH '97* (August 1997), pp. 65–76.

[32] WARD, G. J. Measuring and modeling anisotropic reflection. In *SIGGRAPH '92* (July 1992), pp. 265–272.

[33] WARD, G. J. The radiance lighting simulation and rendering system. In *SIGGRAPH '94* (July 1994), pp. 459–472.

[34] WATANABE, M., AND NAYAR, S. K. Telecentric optics for computational vision. In *Proceedings of Image Understanding Workshop (IUW 96)* (February 1996).

[35] Y.CHEN, AND MEDIONI, G. Object modeling from multiple range images. *Image and Vision Computing 10*, 3 (April 1992), 145–155.