

Photorealistic rendering of mixed reality scenes

Joel Kronander^{1†}, Francesco Banterle², Andrew Gardner¹, Ehsan Miandji¹, and Jonas Unger¹

¹C-Research, Linköping University, Sweden

²Visual Computing Laboratory, ISTI-CNR, Italy

Abstract

Photo-realistic rendering of virtual objects into real scenes is one of the most important research problems in computer graphics. Methods for capture and rendering of mixed reality scenes are driven by a large number of applications, ranging from augmented reality to visual effects and product visualization. Recent developments in computer graphics, computer vision, and imaging technology have enabled a wide range of new mixed reality techniques including methods for advanced image based lighting, capturing spatially varying lighting conditions, and algorithms for seamlessly rendering virtual objects directly into photographs without explicit measurements of the scene lighting. This report gives an overview of the state-of-the-art in this field, and presents a categorization and comparison of current methods. Our in-depth survey provides a tool for understanding the advantages and disadvantages of each method, and gives an overview of which technique is best suited to a specific problem.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Illumination Estimation, Image-Based Lighting, Reflectance and Shading

1. Introduction

Synthesizing realistic images and seamlessly merging virtual objects into real world scenes is one of the long-standing goals of computer graphics. The production of such photo-realistic mixed reality renderings is becoming increasingly important in many application areas such as visual effects, product visualisation, and augmented reality. This has led to the development of a large number of methods for capturing the lighting conditions in real world scenes and methods for inserting virtual objects into legacy footage. Recent developments in high dynamic range imaging and computer vision have, over the last few years, enabled a wide range of new mixed reality capture and rendering methods. This paper is intended to provide a comprehensive reference and thorough comparison of the state of the art in mixed reality capture and rendering methods.

Early work focused on accurate calibration and registration to achieve geometric consistency, omitting the effects of light transport between real and virtual objects. However, it soon became apparent that ignoring the effects of illumination and cast shadows between real and virtual objects did not produce high-quality results. Indeed, one of the

key components necessary for convincing rendering of virtual objects in real scenes is perceptually consistent illumination [SUC95, KPvD*07, LMSSG10, KK12]. This entails not only illuminating virtual objects with captured or estimated lighting conditions in the real scene, but also simulating cast shadows and local light interaction (*common illumination*) between virtual and real objects. In this report we give a survey and classification of methods for achieving these goals. Apart from direct application of these methods for photorealistic rendering in mixed reality scenes, knowledge of their limitations also have applications in image forgery detection [Far09].

We present a survey and classification of methods for capturing, estimating and rendering with consistent illumination that includes the following topics:

- Capture of the lighting in a real scene using invasive measurements, such as omnidirectional HDR images.
- Estimation of the illumination environment directly from images and video.
- Accounting for the radiometric interaction between virtual and real objects, such as shadows cast from virtual to real objects.
- Efficient rendering of mixed reality scenes using Monte Carlo raytracing, precomputed radiance transport, and single-pass differential rendering methods.

† joel.kronander@liu.se

To limit the scope, we assume that the geometric calibration and registration is solved using existing methods, see for example books [HZ03, SH13] or surveys [VCP10, RU13] on the topic. We do not cover related topics such as re-lighting of real objects [JL06], or work on inverse rendering [PP03, PP05]. In comparison to previous related, but dated, surveys [Deb02, JL06] and books [RHD^{*}10], we include recent advances and the appropriate state-of-the-art methods from computer graphics, computer vision, and augmented reality.

Organization of the report - The next section describes how our survey of state-of-the-art methods and algorithms in capture and rendering of mixed reality scenes is carried out, and how we categorize and compare the different methods. To introduce readers who are unfamiliar with the topic, Section 3 then gives an overview of the most important theoretical concepts such as the rendering equation and differential rendering for modelling interactions between real and virtual objects, and discuss specific challenges and common assumptions for mixed reality scenes.

The specific methods being reviewed are described and discussed in Section 4 and Section 5 respectively. Section 6 then presents an overview of rendering methods designed for captured/estimated real world illumination, and highlights some techniques for efficient rendering of shadows and inter-reflections among real and virtual objects. Finally, in Section 7, we summarize the categorization and evaluation of the methods, and discuss open problems in the field and where future work is needed.

2. Classification and overview

In this section, we describe how we have performed the review of state-of-the-art methods and algorithms in capture and rendering of mixed reality scenes. Our review is based on a classification of the different methods into two main classes based on their intent. Within each class, we then evaluate each technique based on a set of features important to both future research in the field as well as practical user scenarios.

The main classes in our categorization are derived based on what is the intent, or goal, of a specific method and ultimately what input-data they require. Existing state-of-the-art methods can be divided into two classes:

- **Measured lighting conditions** - The first class contains methods where the intent is to capture a physically accurate model of the lighting conditions in a scene. This information is then used to produce renderings of virtual objects so that they can be seamlessly composited into e.g. backdrop image or video sequence. Although different methods rely on approximations of different accuracy, the main goal is to measure the information required to generate physically accurate renderings in a very robust and general way, e.g. rendering from novel vantage points.

- **Estimated lighting conditions** - The second class contains methods where the intent is to render virtual objects so that they can be seamlessly placed into a backdrop image or video sequence *without explicit measurements* of the scene lighting. The input is, in many cases, already existing, legacy, footage. The goal of these methods is to *estimate the lighting conditions* directly from the input image or video sequence using semi-automatic approaches. These estimations are, in many cases, based on exploiting flaws and features of the human visual system.

Within the two categories, we compare the different methods based on *capture effort, user processing effort, robustness and generality, physical accuracy, and perceptual plausibility*. The results from the categorization and evaluation are systematically reported in Table 1 in Section 7.

3. Light Transport in Mixed Reality Scenes

Realistic rendering of mixed reality scenes requires accurate simulation of both the appearance of the virtual objects as illuminated by the captured real scene, and how the appearance of the real scene is affected by the virtual objects (e.g. shadows cast from virtual objects onto real objects and color bleeding).

The appearance of the objects (both virtual and real) can be computed by solving the rendering equation [Kaj86]. Ignoring volume scattering, the outgoing radiance from a surface point x in direction $\vec{\omega}_o$ is given by

$$L_o(\mathbf{x}, \vec{\omega}_o) = L_e(\mathbf{x}, \vec{\omega}_o) + \int_{\Omega^+} L_i(\mathbf{x}, \vec{\omega}_i) f_r(\mathbf{x}, \vec{\omega}_i, \vec{\omega}_o) \langle \vec{n}, \vec{\omega}_i \rangle d\vec{\omega}_i$$

where Ω^+ is the upper hemisphere oriented around the surface normal \vec{n} at x , $\langle \cdot, \cdot \rangle$ denotes the dot product, f_r is the bi-directional reflectance function (BRDF), $L_e(\mathbf{x}, \vec{\omega}_o)$ is the emitted radiance from the surface and $L_i(\mathbf{x}, \vec{\omega}_i)$ is the incident radiance at the point from direction $\vec{\omega}_i$.

To get a better understanding of light transport in mixed reality scenes, it is constructive to separate the incident radiance into several parts,

$$L_i(\mathbf{x}, \vec{\omega}_i) = L_i^r(\mathbf{x}, \vec{\omega}_i) + L_i^v(\mathbf{x}, \vec{\omega}_i) + L_i^{v,r}(\mathbf{x}, \vec{\omega}_i)$$

The first term, $L_i^r(\mathbf{x}, \vec{\omega}_i)$, represents incident illumination originating from the real scene which has not been reflected at surfaces of the virtual objects. The second term, $L_i^v(\mathbf{x}, \vec{\omega}_i)$, is the incident illumination that has been emitted from the real or virtual scene and reflected one or more times on surfaces in the virtual scene. Finally, the third term, $L_i^{v,r}(\mathbf{x}, \vec{\omega}_i)$ represents the incident light that has interacted with both real and virtual objects in the scene.

To accurately compute the outgoing radiance at a point on a virtual object, all three terms should ideally be accounted for. The first term is described by measurements captured

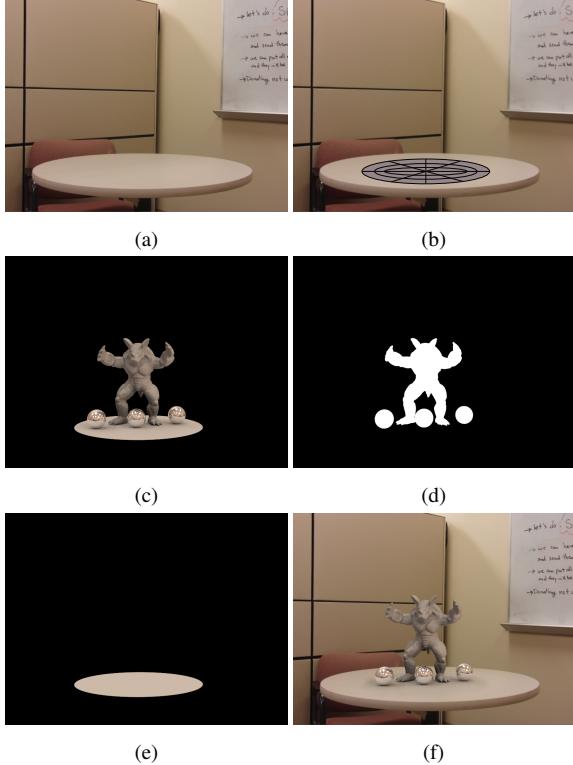


Figure 1: Differential rendering [Deb98] is a standard technique for simulating the light transport in mixed reality rendering. (a) The background image. (b) First, a model of the geometry and reflectance of the local real scene is estimated. (c) Then, a image with both the virtual objects and the modeled local real scene is rendered, I_{rpv} . (d) An alpha mask for the virtual object is also created, α (e) The model of the local real scene is rendered separately, I_r (f) The final composite is then produced by updating the background image with the rendered virtual object and the difference in the real scene, see eq 1, note that the model of the local real scene is not visible.

or estimated from the real scene, and can be sampled directly. The second and third terms need to be recursively computed using global illumination algorithms [PH10]. In order to compute the third term, it is necessary to simulate how the virtual objects affect the real scene. This requires a model describing both the geometry and reflectance properties of the real surfaces that are affected by the virtual objects. It is therefore common to divide the real scene into two parts: a *distant scene* which is not affected by the virtual objects, and a *local scene* which is affected. The local scene is modeled to enable lighting simulation.

A standard method for computing the interaction between virtual and real objects is to use a technique known as *differential rendering* [FGR93, Deb98]. Given a background im-

age, (fig 1a) the local part of the real scene is modelled, including its geometry and reflectance (fig 1b). Then an image, I_{rpv} , is rendered that includes both virtual objects and the modelled real scene objects (fig 1c). Additionally, an alpha mask, α , for the first rendered image is created that is 1 for pixels that overlap virtual objects and 0 for real objects, (fig 1d). Then a second image, I_r , that only includes the modelled real objects is rendered, (fig 1e). Now the intuition is that if I_{rpv} is the same as I_r , there is no shadowing or inter-reflections among real and virtual objects present. However, if I_{rpv} has darker pixel values there are shadows, and if I_{rpv} is brighter, inter-reflections are present. To apply the difference in reflected radiance to the background image or video R , we can update its pixel values by the difference between I_{rpv} and I_r as

$$R = \alpha * I_{rpv} + (1 - \alpha)(R + I_{rpv} - I_r) \quad (1)$$

This implies that pixels that correspond to virtual objects are taken from the first image I_{rpv} and that pixels for the background image are computed by adding the difference of the light transport with and without the virtual objects, see figure 1f.

4. Measured lighting conditions

This section describes methods based on explicit measurements of the lighting conditions in the scene into which virtual objects will be placed. We divide these methods into three different categories. The first category, see Section 4.1, is commonly referred to as image based lighting, and relies on a single environment map to capture the lighting in the scene. The second category, see Section 4.2, has extended image based lighting into the temporal domain to capture dynamically varying environment maps. The third category, see Section 4.3, captures spatial variations in the scene lighting.

4.1. Image based lighting (IBL)

Traditional IBL techniques represent the incident illumination in the scene using a single omnidirectional image, or *environment map*, capturing the angular distribution of the incident illumination at a single point. By using this captured panoramic image, also referred to as a *light probe*, during rendering, the physical lighting conditions of the real scene can be recreated in the virtual scene. Figure 2 illustrates how the panoramic HDR image is used as an estimate of the scene lighting during rendering. Each pixel in the HDR environment map can be thought of as a measurement of the lighting incident over the solid angle subtended by that pixel. By aligning the coordinate system of the virtual objects to that of the environment map, each ray which does not intersect a virtual object is used to sample the lighting captured in the image.

In the seminal work by Blinn and Nevel [BN76]

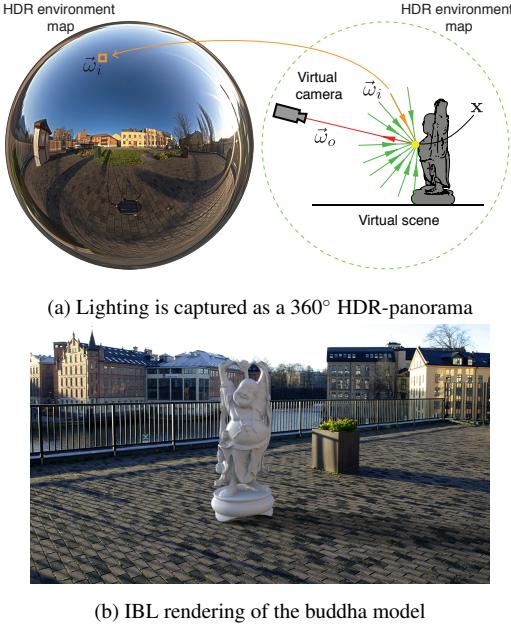


Figure 2: (a) The scene lighting is captured as a panoramic HDR image (left) and used as source of illumination during rendering (right). (b) IBL produces highly realistic renderings.

panoramic images were used to simulate perfect specular reflections and refractions. For specular scattering events, the reflected or refracted vector at a surface point can be used as a direct look-up into the environment map. While efficient, the method has inherent limitations, as only specular effects can be handled. Miller and Hoffman [MH86] and Green [Gre86] filter the environment map with a low-pass kernel in a preprocess, and look up reflected radiance by using the surface normals.

Devebec [Deb98] was the first to propose a general method for IBL incorporating arbitrary BRDFs, global illumination effects, and the interplay of virtual objects and the local parts of the real scene. In contrast to previous work Devebec proposed to use an HDR environment map to record the incident illumination from the real scene. HDR imaging enables the capture of the full range of light in the scene as linear-response measurements. Directly measuring or estimating the full dynamic range in the scene is a requirement to capture an accurate representation of real world lighting conditions. In this report we will not cover HDR image capture, instead we refer the reader to recent books on the topic [RHD*10, BADC11].

To capture omnidirectional HDR images, a simple, practical, setup was presented by Devebec [Deb98]. The method relies on placing a mirrored sphere in the scene where it's desired to capture the incident illumination, and capturing one

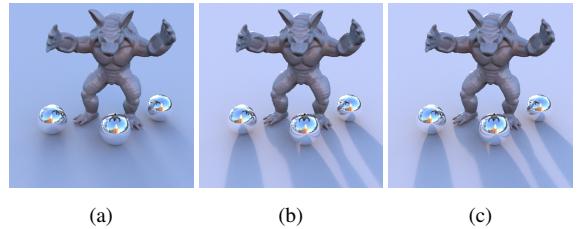


Figure 3: Using an LDR environment map during rendering gives less pronounced reflections and shadows in the scene (a), compared to a reference rendered using an HDR environment map (c). Using an inverse tone mapping operator (iTMO) [BLDC06] the dynamic range of an environment map with moderate saturation can be recovered, yielding more accurate reflections and shadows (b).

or more HDR photographs of the scene using a camera with a standard lens. A single image of the sphere covers most directions in space well except the region in front of the sphere, covered by the photographer and camera, and the region directly behind the sphere, stretched along its rim, resulting in poor resolution. To improve the resolution and cover these blind spots it's common practice to take two or more photographs of the sphere and fuse the images. For a practical tutorial see [Blo12]. One problem with this setup is that the optical system doesn't have a central projection point, i.e. the directions are measured from slightly different point in space. Although central projections can be accomplished using parabolic or hyperbolic mirrors instead [SRT*11] these are seldom used for IBL applications. Fish-eye lenses, panorama stitching [SS97, Sze06, DWH08], and specialised hardware [VR08] can also be used to capture omnidirectional HDR environment maps.

Illumination conditions in an outdoor scene often exhibit a very high dynamic range. Stumpfel *et al.* [STJ04] discuss techniques for direct HDR capture of the sky and sun using a careful selection of exposure times, aperture, and neutral density filters.

HDR images are necessary to capture light sources as they typically exceed ambient light by several orders of magnitude, but it requires recording, aligning, and assembling a range of exposures which can be time-consuming and complicates dynamic capture. Another option is to linearise and expand the dynamic range of panoramic low dynamic range images by using inverse tone mapping algorithms [Lan02, BLDC06, BDA*09]. However larger, overexposed, areas are difficult to reconstruct accurately, as information is missing in these regions. Figure 3 illustrates the difference between IBL renderings using 3a a *low dynamic range* panorama, 3b a panorama where dynamic range in the scene has been estimated using inverse tone mapping, and 3c a rendering using an HDR panorama.

Reinhard *et al.* [RHD*05] showed that if there is just one

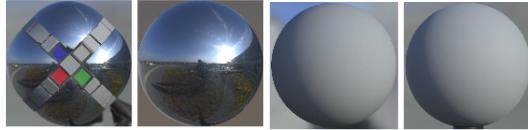


Figure 4: From left to right : The single shot light probe of Debevec *et al.* [DGBB12], a recovered specular mirror probe, a virtual diffuse sphere lit by the recovered HDR light, a real diffuse sphere in recorded light. (Images courtesy of Debevec *et al.* [DGBB12])

bright light in the scene, its intensity can be determined from an image of a diffuse grey sphere placed into the scene, with the remaining illumination imaged accurately in the mirrored sphere. Even so, two images are required, and only one saturated light source is estimated.

Alldrin *et al.* [AK06] proposed to use a multi-layered planar surface that effectively simulates a set of spatially separated BRDFs. Given the known BRDFs of the surface, and assuming uniform distant illumination, the incident illumination in the upper hemisphere can be recovered by treating each BRDF as a basis function. This probe can measure higher frequencies than a diffuse light probe without the need for HDR imaging. However, it's not sufficient for accurate illumination of highly glossy or specular objects.

Using a novel light probe design with diffuse strips between mirrored spherical quadrants Debevec *et al.* [DGBB12] demonstrated how the full dynamic range of the scene can be recovered from a single exposure, see figure 4. Based on the single shot light probe image, the intensity of multiple saturated light sources can be estimated by solving a simple linear system.

Another approach is to extract saturated regions of a LDR light probe and set the intensity manually. For editing purposes, light probes are commonly converted into a sphere of directional lights as this representation can be more intuitive to work with (in Sec 6.1.2 we discuss methods for converting light probes into a finite set of directional lights). Other approaches let the user sketch strokes on lighting features in the rendered image, with a small set of editing operations to quickly adjust the selected features, and adjust the environment map accordingly to produce the desired changes [Pel10].

4.2. Temporally varying IBL

Synthetic objects composited into real world video footage also benefit from temporally consistent and accurate illumination. Capturing a light probe at video frame rates, however, poses a different set of challenges to capturing light probes of static scenes. A variety of different methods have been proposed to address these challenges. From real-

time high dynamic range video cameras with standard light probes attached [Mys08, TKTS11, KGBU13, KGB^{*}14], to standard video cameras with special optical filters [Wae03], to custom light probes viewed by standard video cameras [CMNK13], many different approaches have had success.

Waese *et al.* [Wae03] modified a faceted lens (commonly used to create kaleidoscope effects) with increasing values of neutral density gel to the facets of the filter. This modified filter effectively produces a single image that is divided into five identical regions, with the center region capturing a direct view and the four outer regions stopped down to their respective exposure values. The filter was placed on a standard video camera which filmed a mirror ball. The exposures were then combined in real-time to compute a HDR light probe for each frame.

Havran *et al.* [HSK^{*}05] proposed to capture the incident illumination using a fish-eye camera with a logarithmic response to illumination. This provides perceptually plausible results for specular reflections, but yields less accurate reflections in diffuse and multiple-bounce reflections. When the illumination from the real scene is distant, and mainly directional without significant parallax effects, a reasonable setup for temporally varying IBL is to record an HDR video light probe slightly offset from the background footage of the real scene [Blo12, UKL^{*}13b]. Other approaches locate a light probe in the scene using a fiducial marker [KY04, HSMF05].

Aittala [Ait10] captured smoothly-varying illumination using a diffuse light probe or a rotated planar marker. Using L1-regularized least squares minimization, dominant light sources can be estimated and used in rendering. Yao *et al.* proposed a method for estimating shading of virtual objects from a detected hand, or another diffuse object with known geometry, in a RGB-D sensor image [YKK12][YKK13].

An important consideration when rendering with temporally varying illumination is temporal consistency. Depending on the input method, light probe sequences are subject to a large degree of visual noise. This noise can be remedied by temporal filtering of noisy light probe sequences [HSK^{*}05, UKL^{*}13b] or using specially designed rendering methods, see section 6.

4.3. Spatially varying illumination

A key limitation inherent to all traditional IBL techniques presented in the previous section is that they cannot capture spatially varying illumination, i.e. how the light varies from one location to another in the real scene. Spatially varying illumination, such as cast shadows and shafts of light, play a very important role in the design of lighting in visually-interesting scenes such as those generated by product visualization and cinematography. Figure 5 shows an example



(a) Traditional IBL with a single light probe.



(b) IBL including spatially varying illumination

Figure 5: shows a comparison between: (a) a traditional IBL rendering using a single high resolution omni-directional light measurement captured at a single location in the scene, and (b) a rendering with spatially varying real world illumination captured using the methods described in [UKL*13a]. The traditional IBL rendering in (a) lacks detailed lighting effects such as shadows and reflections found in (b).

of the difference between: 5a traditional IBL rendering using a single omni-directional light measurement, and 5b a rendering where the spatial variations in the scene illumination have also been captured. Although the rendering in 5a looks highly realistic, it is evident that the traditional IBL techniques based on a *single* HDR environment map fail to capture important details in the scene lighting.

Measurement and representation of spatially varying illumination, $L(\mathbf{x}, \vec{\omega}_i)$, requires the angular distribution of the scene lighting to be captured at several locations in the scene, and/or a capture of a geometric model describing the scene's structure (depth, parallax, etc.). The methods described in this section are based on different assumptions and achieve this in different ways. The number of spatial light mea-



Figure 6: shows an example rendering using a set of HDR environment maps images densely captured along a 1D path using an HDR-video camera (1 HDR image per millimeter). (Image courtesy of Unger et al. [UGY07])

surements varies from a single HDR environment map to hundreds of thousands of spatial samples, and the accuracy of the recovered geometric scene model varies from crude proxy geometry to detailed 3D reconstructions of the real scene obtained using e.g. *structure from motion* (SfM) techniques [SSS06, MGV10] or laser scanning. There is a clear trade-off between the complexity of the technique (processing time, user interaction), and the accuracy of the result. It is, however, important to note that the choice between the more accurate and less complicated methods is dependent on the application and the amount of work the user is willing to spend on the problem. Most techniques produce perceptually-plausible results, with the main difference being that the less involved methods may result in a more crude approximation of the lighting environment $L(\mathbf{x}, \vec{\omega}_i)$, and fail to include some lighting effects. Following this discussion, we divide the techniques for spatially varying IBL into three different categories:

- *Dense spatial light sampling* with little or no geometry - Techniques which use a large number of spatial and angular radiance samples and a very crude or no geometric scene representation.
- *Sparse spatial light sampling* with rough geometry - Techniques which assume that the surfaces in the real scene are lambertian, and use only a single or a small number of omni-directional HDR environment maps and a rough geometric scene model.
- *Explicit geometry* - Techniques that rely on a detailed reconstruction of the scene geometry, often recovered using computer vision methods, laser scanning, or even hand modelling. The lighting information is captured using only a few HDR images, or up to hundreds of thousands of HDR environment maps, and is represented as 2D textures or 4D surface light fields projected onto the geometric model.

The following subsections give an overview of the techniques for spatially varying IBL within each category.

4.3.1. Dense spatial light sampling

The goal of the methods described in this section is to capture and represent a slice, or subset, of the *incident light field* (ILF) at the region in the scene where virtual object will be placed during rendering. The concept of an ILF is closely related to *light fields* for photography as introduced in computer graphics by Gortler *et al.* [GGSC96] and Levoy and Hanrahan [LH96]. The goal in the context of photography is to capture and process the outgoing light field (reflected or emitted radiance) from the usually small part of the scene being photographed. This enables applications such as post capture refocusing, depth estimation, and small view-point transformations [WJV*05]. The goal of an ILF is, in contrast to light fields, to capture the illumination incident onto a region in space in a way so that the full dynamic range of the spatial and angular variations in the lighting, $L(\mathbf{x}, \vec{\omega}_i)$, can be estimated by interpolating between nearby sampling points. A comprehensive overview of ILF capture and rendering can be found in [Ung09].

Spatial variations along 1D paths - Unger *et al.* [UGY07, UGY06] used an experimental HDR-video camera [UGOJ04] attached to a mechanical arm to enable dense capture of HDR environment maps along 1D paths in space. For scenes where the lighting mainly varies along one direction, this proves to be a good approximation producing plausible results, as illustrated in Figure 6.

Spatial variations in 2D - Unger *et al.* [UWH*03] measured and parameterized the ILF incident onto a planar 2D surface in real world scenes. This was performed using a camera with a 180° field-of-view fish-eye lens attached to a motorized xy -translation stage, as displayed in Figure 7a. An example data set from [UWH*03] with 30×30 regularly-distributed spatial samples capturing the angular variations at each sample location is displayed in Figure 7b. During rendering, a very crude *auxiliary volume*, e.g. a sphere or bounding box, is placed around the scene. The captured ILF data is then sampled by projecting the sample rays backwards from the auxiliary geometry onto the capture region, where the radiance contribution is estimated using bilinear interpolation in both the spatial and angular domains. Figure 7c and 7d show a comparison between a real photograph of a scene and an ILF rendering of a virtual version of the same scene. Ihrke *et al.* [ISG*08] presented a technique for increasing the spatial resolution by capturing the ILF using imaging a moving mirror. By rotating the mirror to cover the angular domain and tracking its motion, a dense spatial and sparse angular sampling was achieved. This leads to less visible artifacts in the spatial domain.

Masselus *et al.* [MPDW03] used a similar ILF representation for relighting 6D reflectance fields, "light stage data sets", with synthetically generated ILF data. Based on the idea of 4D light fields, Goesele *et al.* [GGHS03] captured near-field light sources and showed how effects from the

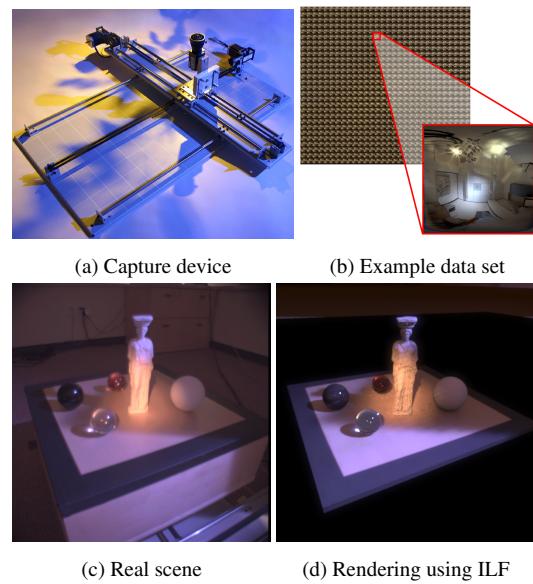
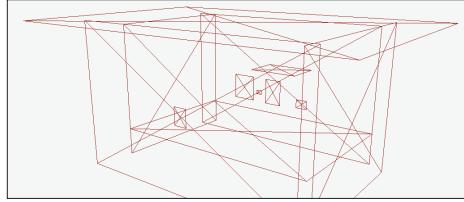


Figure 7: shows examples from ILF capture and rendering as described in [UWH*03]: (a) the apparatus used for capturing the illumination incident over the hemisphere at a large number of positions on a plane, (b) a data set with 30×30 spatial sampling positions, (c) a photograph of a real scene, and (d) a rendering of a computer graphics scene designed to match the real scene. (Images courtesy of Unger *et al.* [UWH*03])

front cover glass and/or lenses could be represented and efficiently used for rendering.

Spatial variations in 3D - The idea of ILF measurement and rendering has also been extended to capture spatial illumination variations in 3D. Unger *et al.* [UGLY08] presented a capture and rendering pipeline where a custom-built HDR video camera [UG07] was used to capture tens of thousands of irregularly spaced HDR environment maps in 3D. By tracking the movement of the camera within a cube of $1.5 \times 1.5 \times 1.5$ m, the captured light samples are projected onto a crude proxy geometry describing the scene and stored as 4D surface light fields. The system, also described in [Ung09], allows for compression of the light field data, estimation of the position, orientation and size of light sources, and editing of the recovered light sources. Figure 8a shows example proxy geometry representing the captured scene, and figure 8b shows a rendering produced using the pipeline.

Mury *et al.* [MPK09] presents a hardware and software system for measuring and analyzing the structure of ILFs in 3D. Using a custom-made device, called a Plenopter, they demonstrate measurements of local light fields represented as second order spherical harmonics. Assuming that the low-order components of light fields in natural scenes typically vary slowly and rather systematically, they show



(a) Geometric proxy model



(b) Rendering with light field captured in 3D

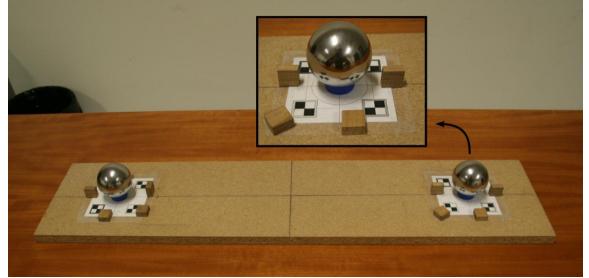
Figure 8: shows an example from the method presented by Unger *et al.* [UGLY08]: (a) the extracted proxy geometry onto which the captured scene lighting is reprojected, and (b) a final photo-realistic rendering produced using their approach. It should be noted that the proxy geometry shown in (a) is used to store 4D light fields encoding the parallax. (Images courtesy of Unger *et al.* [UGLY08])

that the second-order approximation of the radiance distribution function can be estimated reasonably well for all points in the scene using interpolation between a limited number of observations. Using a similar representation, Löw *et al.* [LYLU09] resamples densely captured irregular HDR environment maps into a regular 3D grid for efficient analysis and rendering. By representing the angular distribution at each voxel in the 3D volume using spherical harmonics projections, they demonstrate how this representation can be used for real-time rendering of virtual objects illuminated by real world lighting with spatial variations in 3D.

4.3.2. Sparse spatial light sampling

Various situations exist where an exhaustive capture of spatially varying illumination measurements is not feasible. Film sets, for example, are rapidly changing environments that must be captured quickly before the lights are struck and the set redressed. The methods described in this section are designed to capture spatially varying illumination information with very few HDR environment map samples, typically only one or two. Compared to the methods presented in Sections 4.3.1 and 4.3.3, these techniques allow for very fast and inexpensive lighting capture which makes them valuable in dynamic environments.

Many of these methods exploit computer vision and ge-



(a)



(b)

(c)

Figure 9: An example of SLP by Corsini *et al.*'s method [CCC08]: (a) the acquisition setup. (b) a photograph of a scene. (c) a rendering using SLP. (Images courtesy of Corsini *et al.* [CCC08])

ometric relationships in order to recover geometric information about the scene and enable spatial variations in the lighting. Many also assume their subjects are made up of Lambertian surfaces, and although they reach a reasonable quality in terms of capturing spatially varying effects, some of their assumptions render them incapable of matching the immersive realism of the more advanced methods described in Section 4.3.3.

Sato *et al.* [SSI99] proposed one of the first methods. In their work, two omni-directional cameras are used to generate the spatial radiance distribution of the environment using stereo matching. This resulted in a 3D mesh with lighting information which was then used as an area light source for providing spatially varying illumination in a renderer. Following this work, Corsini *et al.* [CCC08] proposed *stereo light probes* (SLP), where two HDR environment maps are captured in a computer vision setup. Exploiting spherical stereo, area light sources are extracted from HDR environment maps and used for rendering, see Figure 9. A similar and concurrent work was proposed by Korn *et al.* [KSAB08]. Happa *et al.* [HBRDC11] proposed a method for improving the lighting in 3D modeled or 3D scanned environments for cultural heritage. A few HDR environment maps of the modeled or scanned site are acquired, then manually aligned to the 3D mesh of the environment. Finally, from each HDR environment map, light is emitted in a fashion similar to photon mapping [Jen01] or instant radiosity [Kel97]. A similar technique (HDR photographs and 3D model alignment) was proposed by Kölzer *et al.* [KNG11]; but their approach does not exploit spatial information, and

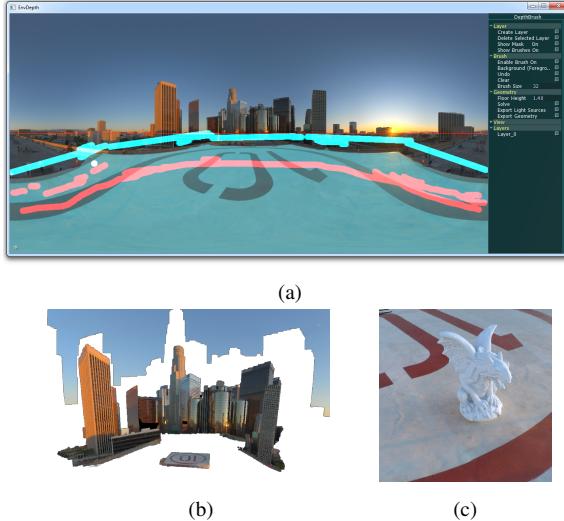


Figure 10: An example of EnvyDepth by Banterle et al.’s method [BCD*13]: (a) EnvyDepth user interface for modeling the environment. (b) a visualization of the virtual point light sources. (c) a rendered virtual object using the VPLs visualized in (b). (Images courtesy of Banterle et al. [BCD*13])

yields an environment map used to render with standard IBL. From a modeling perspective, Grosch [Gro05b] proposed to reconstruct a 3d scene using a modeling interface [SHS98] for adding rectangles and boxes starting from a single HDR environment map as input. The generated 3d scene with associated radiance is then used for rendering and augmented reality. A more general system, *EnvDepth*, was introduced by Banterle et al. [BCD*13]. This system allows users to paint depth onto an HDR environment map using geometric constraints for creating primitives such as planes, curved planes, and domes. EnvDepth outputs a depth map from which virtual point light sources are generated and used for generating rendering in a straightforward way, see Figure 10.

4.3.3. Explicit geometry

Most spatial variations in scene illumination, such as sharp shadows, shafts of light, and parallax effects are difficult to capture and represent accurately without an estimate of the scene’s 3D geometry. However, if an accurate model of the scene is recovered, many of these effects come for free.

Debevec et al. [DTG*04, TSE*04, Deb05] presented a system that used time-of-flight laser scanning to capture scenes in outdoor environments (the Parthenon in Greece is used as the example). Based on the recovered geometry, captured textures, and measurements of the lighting, the system projects the textures onto the 3D model and estimates the material properties at the surfaces in the scene us-

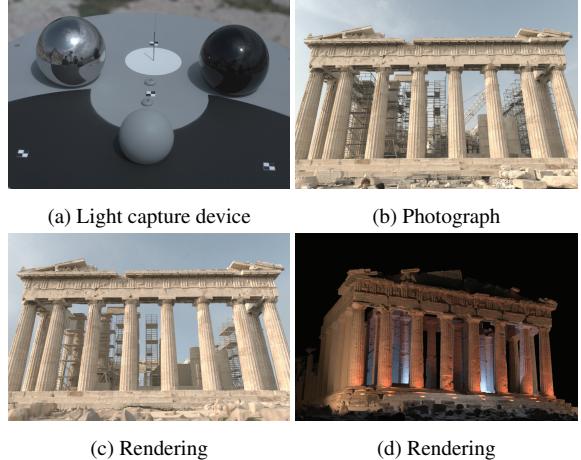


Figure 11: shows examples from the reconstruction of Parthenon described in [DTG*04]: (a) shows the light capture device developed within the project, (b) a photograph of the real scene, (c) a synthetic rendering from the same vantage point as the photograph, and (d) a rendering using a synthetically generated lighting setup. (Images courtesy of Debevec et al. [DTG*04])

ing inverse global illumination techniques in similar fashion as [YDMH99]. In order to estimate the lighting conditions in the real scene they developed a device for accurate illumination capture, see Figure 11a. The report demonstrates how the recovered scene model can be used to generate highly realistic renderings with full global illumination using natural illumination [STJ04], see Figure 11.

Unger et al. [UKL*13a, UKL*13b] described a systems pipeline for capture, processing and rendering of *Virtual Photo Sets* (VPS). Although laser scanning or other active depth sensors can be used [NDI*11], the capture pipeline is purely image-based and relies on SfM methods [SSS06, MGV10] with dense geometry estimation [Fur10], and a set of interactive tools for estimation and semi-automatic adjustment of the recovered scene geometry. The VPS model consists of 3D geometry onto which the lighting information captured in the HDR-video sequences is projected. The captured lighting information is stored as either 2D textures or 4D light fields on the surfaces. The paper also describes tools for estimating the position, size, and orientation of the light sources in the scene, and an approach for estimating the BRDF on densely sampled surfaces in the recovered scene model. The recovered VPS model is intended to be used as lighting information to produce photo-realistic renderings of virtual objects composited into high-quality backdrop images. An example rendering is displayed in Figure 5.

Meillandet et al. [MBC13] proposed a system based on dense real-time 3D tracking and mapping with a RGB-D

camera (e.g. a Kinect) to recover a rough geometric scene. Camera pose and dense scene structure are estimated simultaneously with the observed dynamic range to fuse LDR exposures into HDR light fields on surfaces in the scene.

Scene geometry recovered using laser scanning, by surveying landmarks, and/or hand-modelling is commonly used in visual effects production. In recent years the use of captured or painted HDR textures has become an increasingly important tool in the production of realistic content. Bloch [Blo12] presents a nice overview with many practical tutorials describing how these techniques are carried out in practice, using commercial hardware and software systems.

5. Estimated lighting conditions

The requirement of physical measurements in the real scene, for example introducing light probes or other measurement devices, is a tedious and time consuming process. For legacy photos or videos, such physical measurements in the real scene are not feasible. To avoid physical measurements in the real scene, a large body of previous work has focused on extracting approximate illumination information from images or video sequences directly. The computation of illumination directly from regular images is, in the general case, an ill-posed problem with many possible solutions leading to the same observed image. As a result, assumptions about the environment must be made, such as known scene geometry, Lambertian material reflectance, or by enforcing priors on the illumination distribution.

We categorize the methods in this class according to how they recover the incident illumination from the real scene. In section 5.1 we discuss methods that assume there are real objects in the scene with known or trivial geometry and reflectance properties, from which the incident illumination can be estimated. Another set of methods exploit properties of the human visual system to estimate illumination models that are perceived as realistic, though they may not be physically meaningful. These methods are discussed in section 5.2. While these methods often produce realistic results, they tend to be less reliable and produce illumination estimates that can be hard to manually edit. In section 5.3 we discuss methods that are designed for outdoor scenes where priors on the sky distribution can be used to produce high frequency environment maps.

In this context, it is worth mentioning that there are also methods that produce photorealistic composites of virtual objects placed into existing photographs by querying large image-based object libraries [LHE^{*}07, GCZ^{*}12]. However these approaches cannot handle inserting a specific model, but rather proposes a set of images of objects in a specific category that roughly match illumination conditions in the desired photograph.



Figure 12: Results from Gruber *et al.* [GLS^{*}14] showing real-time Augmented Reality with estimated illumination from a RGB-D video stream. (Images courtesy of Gruber *et al.* [GLS^{*}14])

5.1. Estimating illumination from objects with known geometry and reflectance

In many scenes, common objects with known or trivial geometry and reflectance properties can be used estimate the incident illumination.

Tsumura *et al.* [TDMM03] observed that that eyes could serve as natural light probes. Based on this observation, Nishino *et al.* [NN04] proposed a robust framework for estimating the incident illumination in the scene by detecting eyes and estimating the illumination by observing the reflected scene radiance from the cornea of the eye, see Figure 17.

Rammamorthi and Hanrahan [RH01a, RH01b] investigated estimating the incoming radiance from irradiance measurements, e.g. the estimation of the lighting from images of a homogeneous convex curved Lambertian surface of known geometry under distant illumination.

Recently, Knorr and Kurz [KK14] proposed a framework for estimating the real-world lighting conditions based on the captured appearance of a human face. The method is based on learning a face-appearance model from an offline dataset of faces under known illumination. At run-time they then recover the most plausible real-world lighting conditions in a spherical harmonics bases for the captured face appearance.

Some works use RGBD cameras to dynamically approximate and update the geometry in the scene, using, for example, a Kinect sensor [NDI^{*}11]. Based on a Lambertian scene assumption they can recover low frequency incident illumination reflecting temporal variations [GRTS12, GLS^{*}14].

Using a guided video capture, Jachnik *et al.* [JND12] reconstruct a light field for a simple planar surface, such as a glossy book cover, and factorize the light field into a diffuse and a specular part. The specular part is then used to reconstruct an environment map describing the incident illumination.

Assuming a Lambertian scene the illumination can be estimated using shadows cast from objects with known geometry [SSI03, HDH03, WS03, OSS04]. Mei *et al.* [MLJ09] estimate illumination from shadows by assuming that the illumination can be represented by a sparse set of directional

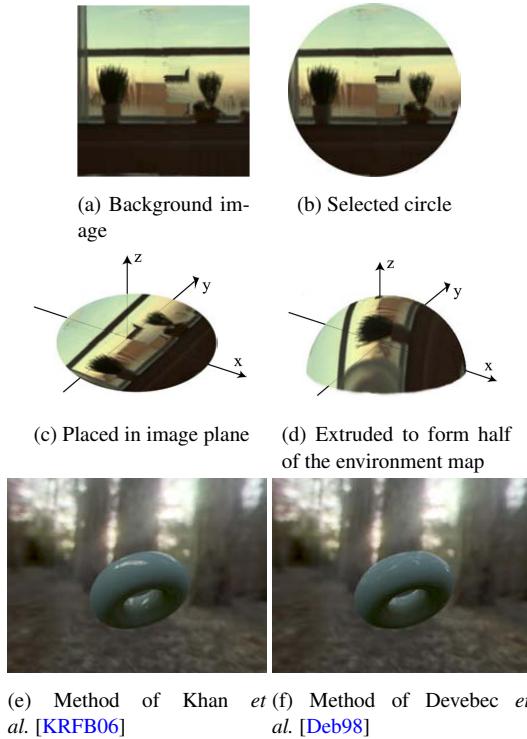


Figure 13: Khan *et al.* [KRB06] propose a simple method to create a perceptually plausible environment map from a single legacy photo. From the background image (a) a circular selection is made (b). The pixels are then mapped onto a 3D plane (c) which is then extruded forward and backward in space (d) to create a spherical environment map. (e) Show a rendering using an extracted environment map, the results are perceptually plausible and similar to that of traditional IBL (f). (Images courtesy of Khan *et al.* [KRB06])

lights, and solve for the illumination parameters using l1-regularised least squares.

Kholgade *et al.* [KSES14] propose a semi-automatic approach to fit publicly available 3D models to targeted objects in a photograph. Assuming a Lambertian reflection model, the illumination conditions are estimated by optimizing a regularized cost function given the observed pixel values of the fitted 3D model. The illumination is represented by an environment map and a sparse illumination prior based on L1-regularized von Mises-Fisher [Fis53] kernel coefficients.

5.2. Perceptually plausible illumination

User studies have shown that humans cannot distinguish between a range of widely different illumination conditions [RF07, LMSSG10]. Furthermore, it has been shown that local illumination consistency is more important than globally consistent illumination [OCS05, KPvD*07].

This enables methods to exploit properties of the human visual system to produce not necessarily physically correct, but perceptually plausible illumination models that are perceived as realistic.

Reinhard *et al.* [RAC*04] proposed a very simple and easily implemented method for inserting virtual objects into legacy photographs and videos. Their method is based on transferring color statistics [RAGS01] from the legacy photograph/video of insertion to the rendered virtual object. Although, this method allows very quick results and is fully automatic, the quality is not high and it may fail with specular material and shadow reproduction because only color statistics are transferred.

Khan *et al.* [KRB06] approximates the incident illumination from a single HDR image for performing image based material editing. In a first step, the object to be edited is removed from the image. Assuming that the precise configuration of the illumination is less important than the overall image statistics, the missing pixels are then filled with an in-painting algorithm that tries to match the color distribution and spectral slope of the rest of the image. To acquire a fully spherical representation of the illumination, a circular region of the unpainted image is then extruded forwards and backwards from the image plane into 3D. While this distorts the illumination, it provides local consistency, which has been shown to be important for the correct perception of objects [OCS05], see Figure 13.

Lopez-Moreno *et al.* [LMHRG10] propose a semi-automatic system for estimating a set of discrete directional light sources illuminating a real scene. The user identifies a object silhouette, which is then used to extract illumination information. Based on the assumption that the normals of the object lie in the image plane along the silhouette, the number of light sources and the azimuthal angle of the light sources can be estimated. Using the assumption that the objects are globally convex, they then estimate the zenith angles of the light sources by looking for maximums of the shading along azimuthal direction on the object surface.

With a small amount of user annotation in a single photograph, Karsch *et al.* [KF11] recovers a rough model of geometry and the position, shape, and intensity of light sources in the scene, see Figure 14. The first step is a semi-automatic reconstruction of a simple geometric model of the scene, based on vanishing points and image-annotations by the user for larger planar surfaces, such as tables. The user then annotates visible area light sources in the image and possibly also models external light sources using a 3D modeling software. The extracted light sources are then adjusted to match the observed direct illumination component provided by an intrinsic image extraction. Light shafts, produced by distant sources such as the sun, are recovered by letting the user draw a bounding box around the visible light shafts in the scene and the source of the illumination (e.g. windows). Using a shadow detection algorithm and the recovered scene

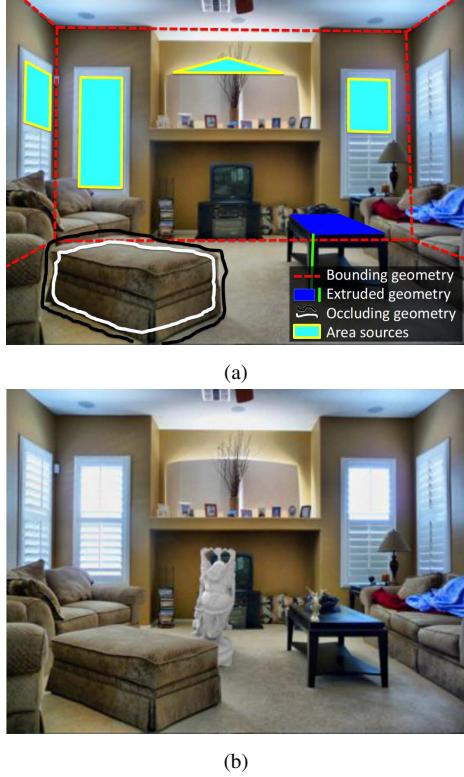


Figure 14: With a small amount of user annotation (an example is shown in (a)), Karsch *et al.* [KF11] recovers a rough model of the scene. The annotated scene properties are then further optimized to correspond to observed properties in the image. Using the estimated scene geometry and illumination, realistic renderings of virtual objects placed in real scenes are produced (b). (Images courtesy of Karsch *et al.* [KF11])

geometry, the direction and extent of the light shafts can then be estimated. Based on a user study with several indoor scenes, they demonstrate that synthetic images produced by the method are confusable with real scenes, and that the method is competitive to traditional IBL techniques using a single light probe.

Karsch *et al.* [KSH*14] proposed a completely automatic system for inserting non-specular objects into single photographs of indoor scenes assuming Lambertian scene reflectance and diffuse illumination. First, a rough depth map is extracted from the input image using an improved version of the non-parametric depth transfer method described in Karsh *et al.* [KLK14]. Depth values are estimated by matching the input image to a database of RGB-D images, where the most similar images in the RGB domain are used to infer the depth at each pixel in the input image. The camera parameters are then estimated from vanishing points [HZ03]. Spatially varying diffuse albedo is then es-

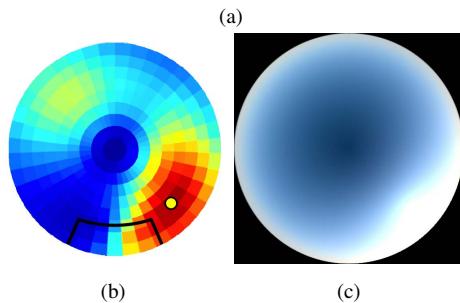


Figure 15: Using the automatic framework proposed by Karsch *et al.* [KSH*14] a user can simply drag and drop a 3D model into a real photograph to produce highly realistic image compositions. (Images courtesy of Karsch *et al.* [KSH*14])

timated using Color Retinex [GJAF09]. The scene illumination is estimated in two steps. First, visible light sources in the image are detected using a classifier trained on features extracted from superpixels [ASS*12]. To estimate the out-of-view light sources (not visible in the image), a data-driven approach is used, in which a dataset of panoramic images is used to find the panoramas that provide a similar illumination to the visible image. Finally, the estimated intensity of each recovered light source is refined using an optimization procedure where the renderings of the estimated scene (geometry and albedo) is matched to the input photograph. The method produces high quality results for many scenes, see figure 15. However, the method often fails to recover the scene and illumination models in scenes where the diffuse assumption does not hold, e.g. outdoor scenes. In scenes where the depth map cannot be accurately estimated, the method produces inaccurate shadows cast from the virtual objects onto the real scene.

5.3. Recovering natural illumination in outdoor scenes

A body of previous work has focused on estimating the illumination in outdoor scenes using priors on the sky appearance. Examples include detecting the sun, fitting physically based parametric sky models, and exploiting statistical properties of natural illumination [DWA04].



(c)

Figure 16: Lalonde *et al.* estimate a synthetic sky model (c) by using information from the sky, shading, shadows, and visible pedestrians to infer a distribution of likely sun positions (b). Based on the estimated sky model synthetic objects can be rendered realistically in real scenes. (Images courtesy of Lalonde *et al.* [LEN12])

The pioneering work by Nakamae et al. [NHIN86] considered composing photographs with virtual elements, and was one of the first works to point out the importance of using a radiometric model to improve the image composition. Input photographs are calibrated and a very simple geometric model of the real scene is extracted. The sun is positioned within the system according to the time and date when the picture was taken. The sun intensity and an ambient term are estimated from two polygons in the image. The estimation of the illumination and geometry, however, is very rough and the results are therefore limited in accuracy.

Madsen and Nielsen [MN08] proposed a system for outdoor illumination estimation by analyzing shadows and using the time of the capture to infer the sun position.

A simple illumination model of outdoor scenes is adopted by Liu *et al.* [LQX*09], which assumes that the sun is a directional light source and the sky is an uniform area light source. The illumination can therefore be estimated with the illumination related statistical parameters or basis learned from sample images. Liu *et al.* [LG12] adopts the same assumptions, but takes spatial and temporal coherence of the illumination as a prior to recover illumination from videos captured with moving cameras. Nevertheless, the simple



Figure 17: Xing *et al.* [XZP13] recently proposed a framework for inserting virtual objects into outdoor scenes that assumes a parametric sky model and accounts for surface radiosity in the scene. (Images courtesy of Xing *et al.* [XZP13])

light source assumption of outdoor scenes prevent these approaches from achieving highly realistic rendering results.

Lalonde *et al.* [LEN09a] uses Perez's sky model [PSM93] as prior information to estimate outdoor illumination using a time-lapse image sequence as input. In the sequence, the scene is assumed to remain constant while the illumination conditions vary over time. In each frame, they recover an environment map describing the illumination in the scene in two steps. First, the sun's position and a parametric sky model are fitted to the part of the sky visible in the images. To synthesize the part of the environment map not covered by the sky the perceptually based method described in Khan *et al.* [KRFB06] described above is used.

In a series of publications, Lalonde and collaborators have developed methods for estimating natural illumination environments from a single image [LEN09b][LEN12][LE10][Lal11]. To estimate the illumination in outdoor scenes, the visible sky is combined in a probabilistic way with other scene features, such as cast shadows and shading on vertical surfaces and convex objects, as well as with illumination priors from large image collections.

Recently Xing *et al.* [XZP13] proposed a framework for compositing virtual objects into a outdoor scene. First, the user annotates the scene to reconstruct scene geometry. Then, a parametric sky model is fitted to the parts of the sky visible in the image. The sun is represented as a directional light, and the sky distribution is modelled and estimated according to the Perez sky model [PSM93] which is similar to Lalonde *et al.* [LE10]. The material properties in the recovered scene are then estimated by solving a linear model with only six free parameters. The illumination contribution from the environment is assumed to be diffuse, and its dimension is reduced by exploiting a spherical harmonics expansion. The environment map is finally computed by solving a lin-

ear problem utilizing the color of the skylight and sunlight as constraints.

6. Rendering

In this section we describe methods for efficient rendering of mixed reality scenes. We divide previous work into two main categories, general algorithms for rendering with static and dynamic environment maps, see section 6.1, and methods for interactive differential rendering specifically developed, but not limited to, interactive Augmented Reality applications 6.2.

We also note that recent, alternative approaches have considered sidestepping the simulation of light transport by using a *shading probe* to directly capture diffuse global illumination in the real scene (and not incident illumination). Using the measured shading response of the probe, virtual objects can be effectively shaded without expensive light transport simulations [CMNK13]. However these methods are still limited to diffuse shading.

6.1. Rendering with environment maps

A large body of work has proposed efficient methods to render scenes with distant real-world illumination represented as an environment map. The most general methods use Monte Carlo sampling techniques to approximate light transport in the scene, described in section 6.1.1. While Monte Carlo methods sample the environment map on the fly for each shading point, a different category of methods have been proposed for converting the environment map to a set of directional point light sources in a pre-process. These methods are discussed in section 6.1.2. For real-time rendering, an efficient alternative is precomputed radiance transfer (PRT), discussed in section 6.1.3.

6.1.1. Monte Carlo rendering

Monte Carlo rendering has a long history in computer graphics [CPC84, Kaj86] and is one of the most general methods for solving the rendering integral (eq. 1). These methods rely on averaging a large number of random samples of light transport in the scene. The methods are thus stochastic in nature and a large body of work has focused on deriving estimators with good convergence rates for both offline [PH10] and real-time rendering [RDGK12]. Importance sampling techniques reduce the variance of estimators by taking information about the integral into account to guide the sampling. For scenes with high frequency illumination, importance sampling of the environment map is essential for efficient rendering of diffuse surfaces [PH10]. To render extremely glossy surfaces in diffuse illumination, a better choice is to sample proportional to the BRDF of the surface. Veach and Guibas [VG95] proposed Multiple Importance Sampling for deriving robust and efficient estimators in cases when either the illumination or

BRDF is complex. Using resampling, Burke *et al.* [BGH05] and Talbot *et al.* [TCE05], samples proportional to the product of the illumination and the BRDF, providing better estimators in scenes with high frequency illumination and glossy materials. Approaches sampling directly from an approximated product distribution have also been proposed [CJAMJ05, CETC06, CAM08, RCL*08, JCJ09]. Other approaches pre-filter the environment map to reduce aliasing [KC08]. Lu *et al.* [LPG13] proposed a method for efficient sampling of dynamic environment maps. Using Sequential Monte Carlo samplers, Ghosh *et al.* [GD06] proposed a method for efficient rendering with dynamic environment maps that propagated a product distribution between frames. While this method is efficient for CPU implementations, recent work has shown that the efficiency of GPU implementations of the method are limited [KDJ*14]. Bashford-Rogers *et al.* [BRDC14] proposed a method for efficient global illumination rendering with environment maps.

6.1.2. Conversion to directional light sources

Another approach is to use a pre-processing step to transform the environment map to a set of finite directional light sources. By placing directional light sources so that the resulting angular illumination distribution is close to that represented by the environment map, these light sources can be used in any general rendering algorithm designed for directional light sources, including real-time rendering [AMHH08, ESAW11]. The downside with these methods is that they introduce additional correlation, as they use the same light samples for all shaded points. The potential artifacts, such as banded shadows and reflections, appear when a few directional light sources are used for the approximation, however these artifacts may sometimes be less objectionable than severe Monte Carlo noise.

Several methods for computing the intensity and distribution of the directional light sources have been proposed. Structured importance sampling proposed by Agarwal *et al.* [ARB03] first decomposes the environment map into strata based on the illumination intensity and expected variance due to visibility in real scenes. All pixels in each strata are then pre-integrated and approximated by a single directional light in the centre of the strata. Kollig and Keller propose to generate light sources randomly on the environment map and then stratify them using Loyds relaxation method [Llo82]. Ostromoukhov *et al.* [ODJ04] places a light source at each vertex of a Penrose tiling which is hierarchically subdivided and thresholded with respect to the luminance value. Using pre-computed correction vectors, the spectral characteristics of the distribution are then improved by trying to enforce a blue-noise Fourier spectrum. Debevec [Deb05] proposed a simple algorithm in which the environment map is hierarchically decomposed into a two-dimensional tree that recursively subdivides the area into regions of equal luminance until there are as many regions as

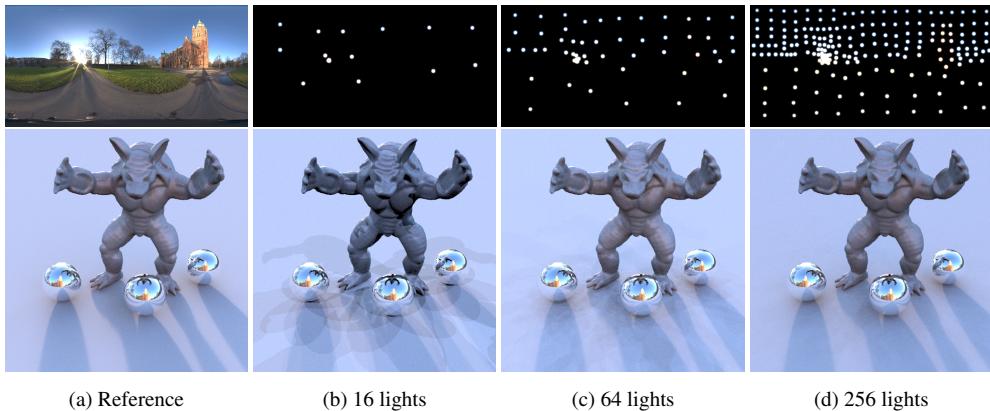


Figure 18: A common practice is to convert an environment map into a finite set of directional lights before rendering. (a) Reference Monte Carlo rendering. (b) However, this can result in visible banding artifacts in reflections and shadows (here using [Deb05]), these artefacts can be reduced by using more lights (c,d).

light sources requested. Light sources are then placed into the weighted centre of each region. Viriyothai and Debevec [VD09] proposed a modified version of the same algorithm where regions are subdivided so that the variances are minimised in the subregions. Wanet *et al.* [Wan05] proposed a spherical q^2 -tree, a hierarchical structure that subdivides the environment map into equal quadrilaterals proportional to solid angles. For dynamic environment maps, the given frames q^2 -tree is constructed based on the q^2 -tree of the previous frame, to improve temporal consistency. Havran *et al.* [HSK^{*}05] proposed a post-processing step to improve temporally consistency of animated sequences, in which extracted light sources intensities and positions are temporally filtered to reduce flickering artifacts. In a later work Wan *et al.* [WMWL11] proposed a global approach in which the dynamic environment map is treated as a spatiotemporal volume which is then sampled by adaptively stratifying the volume.

6.1.3. Precomputed Radiance Transfer

An efficient method for real-time and interactive rendering with environment maps is Precomputed Radiance Transfer (PRT) [SKS02, Ram09]. Given an environment map representing incoming radiance, the main idea of this method is to precompute transfer functions on the surface of an object. These functions locally map the incoming radiance to the outgoing radiance and are computationally expensive. Both the environment map and the transfer functions are projected onto an orthogonal basis. A large body of research has been devoted for finding a suitable basis. A few examples are: spherical harmonics (SH) [SKS02, Slo08], wavelets [NRH04], radial basis functions [TS06], principal components [NSKF07] and spherical piecewise constant basis functions [JFHT08]. Nowrouzezahrai *et al.* [NGM^{*}11] proposed to factorize the spherical harmonic projection of an environment map into two separate terms, a directional

term, and a global term. This factorization enables rendering dynamic scenes with both hard and soft shadows. Grosset *et al.* [GEM07] propose to use a grid of spherical harmonics for real-time rendering with spatially varying illumination in indoor scenes, considering daylight reaching the room through windows and indirect light in the room. A unified and comprehensive formulation of PRT is presented by Lehtinen in [Leh07].

PRT methods can also be extended to handle glossy BRDFs [SZC^{*}07], local deformation [SLS05], subsurface scattering [WTL05], dynamic scenes with rigid motion [IDYN07], etc. Using GPUs, real-time computation of spherical harmonics from an HDR video light probe is also possible [HKMU12]. For the case of glossy BRDFs, the number of coefficients is typically high, making real-time performance unachievable. To tackle this problem, a number of methods have been proposed for compression of coefficients with minimal degradation of visual quality. Principal Component Analysis (PCA) [LK03], clustered-PCA [SHHS03], Biclustering [SHR^{*}11] and clustered tensor approximation [TS06] are a few examples. Despite the efforts in this direction, a PRT method for interactive global illumination of fully dynamic and deformable scenes with arbitrary materials has not yet been proposed.

6.2. Interactive differential rendering

Standard differential rendering requires that the scene is rendered two times, first with only the local real scene model, and then with both real and virtual objects. However, many regions remain unchanged and the same work is done twice without any visual effect. A more efficient approach is to use a single pass where changes in lighting introduced by a virtual object are directly simulated. Grosh *et al.* [Gro05a] modified photon mapping [Jen01] by using a differential photon map to render interactions between virtual and real

objects in a single pass. A set of photons for each pixel are shot in a direction perpendicular to the bounding sphere of the environment map. The photon distribution is carried out as in [Jen01], but if a photon hits a virtual object, a negative flux is stored at the next intersection with a real surface. Rendering is carried out for diffuse and reflective/refractive objects separately.

A method for real-time global illumination of indoor scenes with diffuse materials, lit by environment lighting was proposed in [GEM07]. The direct light from outside is used to update a near-field representation of the indirect light in the room by a dynamic form of the irradiance volume. Importance sampling and shadow mapping is used for direct lighting. Knecht and collaborators [KTM^{*}10, KTMW12] have proposed methods to combine instant radiosity [Kel97][RGK^{*}08] with differential rendering requiring only a single render pass to achieve real-time performance for diffuse and glossy objects. They extended this method to handle reflective and refractive objects, taking into account caustic effects [KTWW13]. In addition, they assume that the geometry for real objects is given and is static.

Kan *et al.* [KK12] propose a method for interactive global illumination of a mixed scene using photon mapping, which enables caustics and reflective/refractive materials. Later, the same authors proposed a real-time single pass differential rendering approach using irradiance caching [WRC88]. By analyzing different ray types and intersection scenarios, the irradiance is separated for real and virtual objects, which can be used for computing differential irradiance. The irradiance cache record will then store the real and differential irradiances, which are used for irradiance cache splatting on the GPU. While their method produces plausible results for multiple-bounce global illumination, even under the depth-of-field effect, it is limited to diffuse materials and requires a precomputation stage. Lensing *et al.* [LB12] utilized reflective shadow mapping [DS05] for a single-bounce diffuse indirect illumination without a precomputation stage. The proposed method is purely image-based and uses guided image filtering [HST10] to overcome depth image errors.

7. Conclusion

We conclude this report by summarizing and comparing a selection of the methods described in Section 4 and Section 5. We also give an overview of what are the open problems that are not solved with current methods, and present important venues for future work.

7.1. Comparison

This section presents a summary and comparison of state-of-the-art methods for capture and rendering of mixed reality scenes. Section 4 and Section 5 review a large number of methods. For methods that have been iteratively ex-

tended and described in more than one paper we have, in most cases, selected the most recent and general version. For cases, where there are differences between seemingly similar methods, we have included both. The methods are classified into the two categories described in Section 2: *measured lighting conditions* and *estimated lighting conditions*.

In order to describe each method in the context of the progress of the field, we have compared them according to a number of criteria. Many of the methods are radically different in intent, robustness, and accuracy, and thus difficult to compare. The comparisons and scores are therefore based on both information from the original papers, and the subjective judgement of the authors (by consensus). We believe, however, that the comparisons give important indications about the usability and performance of the selected methods. The following criteria have been used:

1. *Required Data* - Different methods require different input data. In this column we list the data that is required for measuring or estimating the incident illumination from the real scene, e.g. HDR or LDR light probes, laser scans, images or video.
2. *Illumination Assumption* - The methods assume different models for the captured illumination. Important considerations are if the illumination is assumed to be distant (often represented using an environment map) or spatially varying and if high-frequency illumination can be estimated.
3. *Capture (time,effort)* - This criterion gives an indication of how much time and effort is required during capture. A high score is given to methods that require less effort.
4. *Processing (time,effort)* - This criterion indicates how much time and effort is required after the initial capture has been performed. This includes necessary user annotations and how easy it is to reconstruct, represent and edit the recovered scene illumination. A high score is given to methods that require less effort.
5. *Robustness and Generality* - Different methods assume different limitations. This criterion is higher if the method is robust during rendering within the approximations made in the *illumination assumption* (2). Robustness reflects change of viewpoint, virtual object geometry, and material appearance.
6. *Physical accuracy* - This criterion reflects how physically accurate the reconstructed scene illumination and the final renderings are.
7. *Perceptual plausibility* - This criterion is related to the quality of the results in the context of perceived artifacts and plausibility. Note that this is different from physical accuracy.

Note that we have not included criteria related to image synthesis such as rendering time and memory requirements. This is motivated by the fact that most methods included in this survey can be used with a range of different rendering methods, and that the choice of rendering algorithm is heav-

ily application dependent. For example, an HDR environment map can be used to produce images in both offline renderings using Monte Carlo methods, as well as in interactive applications using e.g. precomputed radiance transfer techniques. However, the illumination assumptions criteria (2) loosely reflects how heavy or lightweight the requirements are during rendering, e.g. in general, a single light probe exhibits a smaller memory footprint and enables computationally efficient representations compared to full reconstruction of geometry and lighting in the scene.

An overview of the surveyed methods comparing them according to the listed criteria is given in Table 1.

7.2. Open problems and future work

Although photorealistic rendering of mixed reality scenes has been a focus area in research and industry for more than a decade, there are still many open problems to be solved.

Current measurement based methods can capture spatial variations in the scene illumination if the scene is static. Spatially varying illumination usually requires significant effort, both during capture and processing. We believe that the next important step towards more general scene capture is the development of techniques for robust estimation of material properties on the surfaces in general scenes. This would lead the way towards new methods for measuring and estimating spatially varying illumination. Future improvements in computer vision and capturing devices could greatly improve this field. For example, devices such as the ones proposed in project Tango by Google [Goo14] look very promising. New and more robust inverse global illumination techniques would also allow for more accurate simulation of how virtual objects affect real scenes.

Current methods for photorealistic object insertion in legacy images are limited in their generality. Complex illumination environments and specular objects are often problematic for these methods. Future work should therefore continue to improve the robustness and physical accuracy of the estimated illumination environments.

Another very interesting problem that has not been solved is photorealistic object insertion in legacy video. The key challenge in this context is to ensure consistency in the estimation of material properties (color and reflectance) and lighting throughout the video. This will require the development of new models and new optimization schemes taking into account the temporal coherency inferred by the scene. A significant challenge is to extend these methods to allow for lighting conditions which vary both spatially within the scene as well as over time.

Most methods that consider photorealistic composition of virtual objects into legacy images and video are limited by the LDR nature of the content. The increased availability of HDR content (images and video) opens new possibilities for

developing more general and robust methods. Another interesting venue for future work in this area is to use methods from inverse tone mapping to recover HDR illumination from LDR media.

Realistic real-time rendering with captured or estimated illumination is an ongoing effort, and is likely to continue into the future. General interactions between real and virtual objects, dynamic scenes, glossy materials, and high-frequency illumination environments are all open research problems. Recent developments in interactive global illumination algorithms [RDGK12] are likely to also have a large impact on photorealistic rendering of mixed reality scenes. Developing rendering algorithms specifically for mobile devices is another important aspect of future work that will enable widespread adoption of photorealistic augment reality.

This comparison concludes our attempt to survey and classify this rapidly evolving research area. It is now possible to generate highly realistic renderings of virtual objects in real scene using state-of-the-art methods. However, as indicated above, many problems still remain. We are optimistic that future research in the area will provide answers to many of these questions.

Acknowledgment. We thank the authors of the original publications for providing images and data: Paul Debevec, Kevin Karsch, Erik Reinhard, Jean-François Lalonde, Lukas Gruber, and Guanyu Xing. This project was funded by the Swedish Foundation for Strategic Research through grant IIS11- 0081, Linkping University Center for Industrial Information Technology, and the Linnaeus research environment CADICS.

Class/Method	Required data	Illumination assumption	Capture (time,effort)	Processing (time,effort)	Robustness & generality	Physical accuracy	Perceptual plausibility
Measured lighting							
<i>IBL</i>							
Debevec <i>et al.</i> [Deb98]	HDR LP	Dist.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Banterle <i>et al.</i> [BDA*09]	LDR LP	Dist./Smooth	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Debevec <i>et al.</i> [DGBB12]	LDR SHLP	Dist.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Calian <i>et al.</i> [CMNK13]	LDR SP	Dist./Diffuse	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
<i>Temporal IBL</i>							
Waese <i>et al.</i> [Wae03]	HDR Vid. LP	Dist.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Havran <i>et al.</i> [HSK*05]	Log-HDR Vid.	Dist./Diffuse	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Grosch <i>et al.</i> [GEM07]	HDR Vid.	Dist./Diffuse	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Yao <i>et al.</i> [YKK13]	RGB-D Vid.	Dist./Diffuse	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Unger <i>et al.</i> [UKL*13b]	HDR Vid. LP	Dist.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
<i>Sparse capture</i>							
Sato <i>et al.</i> [SSI99]	Stereo LPs	Spat.Var.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Grosch [Gro05b]	HDR LP	Diff. Spat.Var.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Corsini <i>et al.</i> [CCC08]	Stereo LPs	Spat.Var.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Banterle <i>et al.</i> [BCD*13]	HDR LP	Diff. Spat.Var.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
<i>Dense capture</i>							
Unger <i>et al.</i> [UGY07]	HDR Vid. LP	1D Spat.Var.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Unger <i>et al.</i> [UWH*03]	ILF LP	2D Spat.Var.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Ihrke <i>et al.</i> [ISG*08]	moving mirror	2D Spat.Var.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Unger <i>et al.</i> [UGLY08]	HDR Vid. LP	3D Spat.Var.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Mury <i>et al.</i> [MPK09]	Plenopter	Diff. Spat.Var.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
<i>Explicit Geometry</i>							
Debevec <i>et al.</i> [DTG*04]	Model	3D Spat.Var.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Meilland <i>et al.</i> [MBC13]	RGB-D Vid.	3D Spat.Var.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Unger <i>et al.</i> [UKL*13a]	LP Vid./Laser	3D Spat.Var.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Estimated lighting							
<i>Implicit light probes</i>							
Nishino <i>et al.</i> [NN04]	Img. w. eye	Dist./Low-freq.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Mei <i>et al.</i> [MLJ09]	Img.	Sparse	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Jachnik <i>et al.</i> [JND12]	Vid.	Dist.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Gruber <i>et al.</i> [GLS*14]	RGB-D Vid.	Dist./Low-freq.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Kholgade <i>et al.</i> [KSES14]	Img. Model	Dist.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
<i>Perceptual</i>							
Khan <i>et al.</i> [KRFB06]	Img.	Percep./Symm.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Lopez-Moreno [LMHRG10]	img.	Percep.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Karsch <i>et al.</i> [KF11]	Img.	User Model	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Karsch <i>et al.</i> [KSH*14]	Img.	Percep.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
<i>Outdoor illumination</i>							
Lalonde <i>et al.</i> [LEN09a]	T-lapse w. sky	Sky/Symm.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Liu <i>et al.</i> [LG12]	Vid. w. sky	Sky	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Lalonde <i>et al.</i> [LEN12]	Img. w. sky	Sky/Symm.	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Xing <i>et al.</i> [XZP13]	Img. w. sky	Sky/Diffuse	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★

Table 1: Overview of surveyed methods for illumination-consistent photorealistic rendering of synthetic objects placed into real scenes. Please see Sec. 7.1 for a description of the properties/scores. Abbreviations used : (LP) Light Probe (SHLP) Single Shot Light Probe (SP) Shading Probe (Spat.Var.) Spatially varying (Img.) Image, (Vid.) video, (T-lapse) Time-lapse image sequence, (RGB-D) RGB image/video with depth channel, e.g. obtained by a kinect sensor. (Dist.) Distant (Percep.) Perceptually Plausible (low-freq.) Low-Frequency (Symm.)

References

- [Ait10] AITTALA M.: Inverse lighting and photorealistic rendering for augmented reality. *The Visual Computer* 26, 6-8 (2010), 669–678. [5](#)
- [AK06] ALLDRIN N., KRIEGMAN D.: A Planar Light Probe. In *CVPR* (2006), vol. 2, pp. 2324–2330. [5](#)
- [AMHH08] AKENINE-MÖLLER T., HAINES E., HOFFMAN N.: *Real-time rendering*, third ed. A K Peters, 2008. [14](#)
- [ARBJ03] AGARWAL S., RAMAMOORTHI R., BELONGIE S., JENSEN H. W.: Structured importance sampling of environment maps. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2003)* 22, 3 (2003), 605–612. [14](#)
- [ASS*12] ACHANTA R., SHAJI A., SMITH K., LUCCHI A., FUÀ P., SUSSTRUNK S.: Slic superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34, 11 (2012), 2274–2282. [12](#)
- [BADC11] BANTERLE F., ARTUSI A., DEBATTISTA K., CHALMERS A.: *Advanced high dynamic range imaging: theory and practice*. CRC Press, 2011. [4](#)
- [BCD*13] BANTERLE F., CALLIERI M., DELLEPIANE M., CORSINI M., PELLACINI F., SCOPIGNO R.: EnvyDepth: An Interface for Recovering Local Natural Illumination from Environment Maps. *Computer Graphics Forum (Proc. of Pacific Graphics 2013)* 32, 7 (2013), 411–420. [9, 18](#)
- [BDA*09] BANTERLE F., DEBATTISTA K., ARTUSI A., PAT-TANAIK S., MYSZKOWSKI K., LEDDA P., CHALMERS A.: High dynamic range imaging and low dynamic range expansion for generating hdr content. *Computer Graphics Forum* 28, 8 (2009), 2343–2367. [4, 18](#)
- [BGH05] BURKE D., GHOSH A., HEIDRICH W.: Bidirectional importance sampling for direct illumination. In *Rendering Techniques (Proc. of EGSR)* (2005), pp. 147–156. [14](#)
- [BLDC06] BANTERLE F., LEDDA P., DEBATTISTA K., CHALMERS A.: Inverse tone mapping. In *GRAPHITE* (2006), pp. 349–356. [4](#)
- [Blo12] BLOCH C.: *The HDRI Handbook 2.0*. Rocky Nook, 2012. [4, 5, 10](#)
- [BN76] BLINN J., NEWELL M.: Texture and reflection in computer generated images. *Communications of the ACM* 19, 10 (1976), 542–547. [3](#)
- [BRDC14] BASHFORD-ROGERS T., DEBATTISTA K., CHALMERS A.: Importance Driven Environment Map Sampling. *IEEE Transactions on Visualization and Computer Graphics* 20, 6 (2014), 907–918. [14](#)
- [CAM08] CLARBERG P., AKENINE-MÖLLERY T.: Practical product importance sampling for direct illumination. *Computer Graphics Forum (Proc. of Eurographics 2008)* 27, 2 (2008), 681–690. [14](#)
- [CCC08] CORSINI M., CALLIERI M., CIGNONI P.: Stereo Light Probe. *Computer Graphics Forum (Proc. of Eurographics 2008)* 27, 2 (2008), 291–300. [8, 18](#)
- [CETC06] CLARBERG P., EGBERT P. K., TALBOT J. F., CARDON D. L.: Two stage importance sampling for direct lighting. In *Rendering Techniques (Proc. of EGSR)* (2006), pp. 103–113. [14](#)
- [CJAMJ05] CLARBERG P., JAROSZ W., AKENINE-MÖLLER T., JENSEN H. W.: Wavelet importance sampling: efficiently evaluating products of complex functions. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2005)* 24, 3 (2005), 1166–1175. [14](#)
- [CMNK13] CALIAN D. A., MITCHELL K., NOWROUZEZHRAI D., KAUTZ J.: The shading probe: fast appearance acquisition for mobile AR. In *SIGGRAPH Asia Technical Briefs* (2013). [5, 14, 18](#)
- [CPC84] COOK R. L., PORTER T., CARPENTER L.: Distributed ray tracing. *Computer Graphics (Proc. of SIGGRAPH 1984)* 18, 3 (1984), 137–145. [14](#)
- [Deb98] DEBEVEC P.: Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *SIGGRAPH* (1998), pp. 1–10. [3, 4, 11, 18](#)
- [Deb02] DEBEVEC P.: Image-based lighting. *IEEE Computer Graphics and Applications* 22, 2 (2002), 26–34. [2](#)
- [Deb05] DEBEVEC P.: Making the parthenon. In *International Symposium on Virtual Reality, Archaeology, and Cultural Heritage* (2005). [9, 14, 15](#)
- [DGBB12] DEBEVEC P., GRAHAM P., BUSCH J., BOLAS M.: A single-shot light probe. In *SIGGRAPH Talks* (2012), pp. 10:1–10:1. [5, 18](#)
- [DS05] DACHSBACHER C., STAMMINGER M.: Reflective shadow maps. In *Symposium on Interactive 3D Graphics and Games (I3D)* (2005), pp. 203–231. [16](#)
- [DTG*04] DEBEVEC P., TCHOU C., GARDNER A., HAWKINS T., POULLIS C., STUMPFEL J., JONES A., YUN N., EINARSSON P., LUNDGREEN T., FAJARDO M., MARTINEZ P.: *Estimating Surface Reflectance Properties of a Complex Scene under Captured Natural Illumination*. Tech. Rep. ICT-TR-06.2004, University of Southern California Institute for Creative Technologies, 2004. [9, 18](#)
- [DWA04] DROR R. O., WILLSKY A. S., ADELSON E. H.: Statistical characterization of real-world illumination. *Journal of Vision* 4, 9 (2004), 11. [12](#)
- [DWH08] DIVERDI S., WITHER J., HOLLEREI T.: Envisor: Online environment map construction for mixed reality. *IEEE Virtual Reality* (2008), 19–26. [4](#)
- [ESAW11] EISEMANN E., SCHWARZ M., ASSARSSON U., WIMMER M.: *Real-time shadows*. CRC Press, 2011. [14](#)
- [Far09] FARID H.: Image forgery detection. *Signal Processing Magazine, IEEE* 26, 2 (2009), 16–25. [1](#)
- [FGR93] FOURNIER A., GUNAWAN A., ROMANZIN C.: Common Illumination between Real and Computer Generated Scenes. In *Graphics Interface* (1993), pp. 254–254. [3](#)
- [Fis53] FISHER R.: Dispersion on a sphere. In *Proc. of the Royal Society of London* (1953), vol. 217, pp. 295–305. [11](#)
- [Fur10] FURUKAWA Y.: Accurate, dense, and robust multi-view stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32, 8 (2010), 1362–1376. [9](#)
- [GCZ*12] GOLDBERG C., CHEN T., ZHANG F.-L., SHAMIR A., HU S.-M.: Data-driven object manipulation in images. In *Computer Graphics Forum* (2012), vol. 31, pp. 265–274. [10](#)
- [GD06] GHOSH A., DOUCET A.: Sequential sampling for dynamic environment map illumination. In *Rendering Techniques (Proc. of EGSR)* (2006), pp. 115–126. [14](#)
- [GEM07] GROSCH T., EBLE T., MUELLER S.: Consistent interactive augmentation of live camera images with correct near-field illumination. In *ACM Symposium on Virtual Reality Software and Technology (VRST)* (2007), vol. 1, pp. 125–132. [15, 16, 18](#)
- [GGHS03] GOESELE M., GRANIER X., HEIDRICH W., SEIDEL H.-P.: Accurate light source acquisition and rendering. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2003)* 22, 3 (2003), 621–630. [7](#)

- [GGSC96] GORTLER S. J., GRZESZCZUK R., SZELISKI R., COHEN M. F.: The lumigraph. In *SIGGRAPH* (1996), pp. 43–54. 7
- [GJAF09] GROSSE R., JOHNSON M. K., ADELSON E. H., FREEMAN W. T.: Ground truth dataset and baseline evaluations for intrinsic image algorithms. In *ICCV* (2009), pp. 2335–2342. 12
- [GLS^{*}14] GRUBER L., LANGLOTZ T., SEN P., HOHERER T., SCHMALSTIEG D.: Efficient and robust radiance transfer for probeless photorealistic augmented reality. *IEEE Virtual Reality* (2014), 15–20. 10, 18
- [Goo14] GOOGLE INC.: Project tango. <https://www.google.com/atap/projecttango/>, 2014. 17
- [Gre86] GREENE N.: Environment mapping and other applications of world projections. *Computer Graphics and Applications* 6, 11 (1986), 21–29. 4
- [Gro05a] GROSCH T.: Differential photon mapping: Consistent augmentation of photographs with correction of all light paths. *Eurographics Short Presentations* (2005), 53–56. 15
- [Gro05b] GROSCH T.: Panoar: Interactive augmentation of omnidirectional images with consistent lighting. In *Mirage 2005, Computer Vision / Computer Graphics Collaboration Techniques and Applications* (2005), pp. 25–34. 9, 18
- [GRTS12] GRUBER L., RICHTER-TRUMMER T., SCHMALSTIEG D.: Real-time photometric registration from arbitrary geometry. In *ISMAR* (2012), pp. 119–128. 10
- [HBRDC11] HAPPA J., BASHFORD-ROGERS T., DEBATTISTA K., CHALMERS A.: Rendering interior cultural heritage scenes using image-based shooting. In *Eurographics, Heritage Area Papers* (2011). 8
- [HDH03] HALLER M., DRAB S., HARTMANN W.: A real-time shadow approach for an augmented reality application using shadow volumes. In *ACM Symposium on Virtual Reality Software and Technology (VRST)* (2003), pp. 56–65. 10
- [HKMU12] HAJISHARIF S., KRONANDER J., MIANDJI E., UNGER J.: Real-time image based lighting with streaming HDR-light probe sequences. In *SIGRAD* (2012). 15
- [HSK^{*}05] HAVRAN V., SMYK M., KRAWCZYK G., MYSZKOWSKI K., SEIDEL H.-P.: Interactive system for dynamic scene lighting using captured video environment maps. In *Rendering Techniques (Proc. of EGSR)* (2005), pp. 31–42. 5, 15, 18
- [HSMF05] HEYMANN S., SMOLIC A., MULLER K., FROEHLICH B.: Illumination reconstruction from real-time video for interactive augmented reality. In *International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS)* (2005). 5
- [HST10] HE K., SUN J., TANG X.: Guided image filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35, 6 (2010), 1397–1409. 16
- [HZ03] HARTLEY R., ZISSERMAN A.: *Multiple view geometry in computer vision*. Cambridge university press, 2003. 2, 12
- [IDYN07] IWASAKI K., DOBASHI Y., YOSHIMOTO F., NISHITA T.: Precomputed radiance transfer for dynamic scenes taking into account light interreflection. In *Rendering Techniques (Proc. of EGSR)* (2007), pp. 35–44. 15
- [ISG^{*}08] IHRKE I., STICH T., GOTTSCHLICH H., MAGNOR M., SEIDEL H.-P.: Fast incident light field acquisition and rendering. *Journal of WSCG* 16, 1-3 (2008), 25–32. 7, 18
- [JCJ09] JAROSZ W., CARR N. A., JENSEN H. W.: Importance sampling spherical harmonics. *Computer Graphics Forum (Proc. of Eurographics 2009)* 28, 2 (2009), 577–586. 14
- [Jen01] JENSEN H. W.: *Realistic image synthesis using photon mapping*. AK Peters, Ltd., 2001. 8, 15, 16
- [JFHT08] JIA Y.-T., FU H., HU S., TAI C.-L.: Spherical piecewise constant basis functions for all-frequency precomputed radiance transfer. *IEEE Transactions on Visualization and Computer Graphics* 14, 2 (2008), 454–467. 15
- [JL06] JACOBS K., LOSCOS C.: Classification of Illumination Methods for Mixed Reality. *Computer Graphics Forum* 25, 1 (2006), 29–51. 2
- [JND12] JACHNIK J., NEWCOMBE R. A., DAVISON A. J.: Real-time surface light-field capture for augmentation of planar specular surfaces. In *ISMAR* (2012), IEEE, pp. 91–97. 10, 18
- [Kaj86] KAJIYA J. T.: The rendering equation. *Computer Graphics (Proceedings of SIGGRAPH '86)* 20, 4 (1986), 143–150. 2, 14
- [KC08] KŘIVÁNEK J., COLBERT M.: Real-time shading with filtered importance sampling. In *Computer Graphics Forum (Proc. of EGSR)* (2008), vol. 27, Wiley Online Library, pp. 1147–1154. 14
- [KDJ^{*}14] KRONANDER J., DAHLIN J., JÖNSSON D., KOK M., B. SCHÖN T., UNGER J.: Real-time video based lighting using gpu ray-tracing. In *EUSPICO* (2014). 14
- [Kel97] KELLER A.: Instant radiosity. In *Proc. of SIGGRAPH* (1997), pp. 49–56. 8, 16
- [KF11] KARSCH K., FORSYTH D.: Rendering Synthetic Objects into Legacy Photographs. *ACM Transactions on Graphics (Proc. of SIGGRAPH Asia 2011)* 30, 6 (2011), 157:1–157:12. 11, 12, 18
- [KGB^{*}14] KRONANDER J., GUSTAVSON S., BONNET G., YNERNER A., UNGER J.: A unified framework for multi-sensor HDR video reconstruction. *Signal Processing: Image Communication* 2, 19 (2014), 203–215. 5
- [KGBU13] KRONANDER J., GUSTAVSON S., BONNET G., UNGER J.: Unified HDR reconstruction from raw CFA data. In *IEEE International Conference on Computational Photography (ICCP), 2013* (2013). 5
- [KK12] KÁN P., KAUFMANN H.: High-quality reflections, refractions, and caustics in augmented reality and their contribution to visual coherence. In *ISMAR* (2012), pp. 99–108. 1, 16
- [KK14] KNORN S., KURZ D.: Real-time illumination estimation from faces for coherent rendering. In *ISMAR* (2014), pp. 113–122. 10
- [KLK14] KARSCH K., LIU C., KANG S. B.: Depthtransfer: Depth extraction from video using non-parametric sampling. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2014). 12
- [KNG11] KÖLZER K., NAGL F., GRIMM P.: Synthesizing Relative Radiance for Realistic Rendering of Virtual Objects in 3D Photo Collections. *Eurographics, High Dynamic Range Area Paper* (2011), 71–72. 8
- [KPvD^{*}07] KOENDERINK J. J., PONT S. C., VAN DOORN A. J., KAPPERS A. M., TODD J. T.: The visual light field. *Perception* 36, 11 (2007), 1595. 1, 11
- [KRFB06] KHAN E. A., REINHARD E., FLEMING R. W., BÜLTHOFF H. H.: Image-based material editing. *ACM Transactions on Graphics (Proc. of Siggraph 2006)* 25, 3 (2006). 11, 13, 18
- [KSAB08] KORN M., STANGE M., ARB A. V., BLUM L.: Interactive augmentation of live images using a hdr stereo camera. *Journal of Virtual Reality and Broadcasting* 4, 12 (2008). 8

- [KSES14] KHOGLADE N., SIMON T., EFROS A., SHEIKH Y.: 3d object manipulation in a single photograph using stock 3d models. *ACM Transactions on Graphics* 33, 4 (2014), 127:1–127:12. [11, 18](#)
- [KSH*14] KARSCH K., SUNKAVALLI K., HADAP S., CARR N., JIN H., FONTE R., SITTIG M., FORSYTH D.: Automatic scene inference for 3d object compositing. *ACM Transactions on Graphics* 33, 3 (June 2014). [12, 18](#)
- [KTM*10] KNECHT M., TRAXLER C., MATTAUTSCH O., PURGATHOFER W., WIMMER M.: Differential instant radiosity for mixed reality. In *ISMAR* (2010), pp. 99–107. [16](#)
- [KTMW12] KNECHT M., TRAXLER C., MATTAUTSCH O., WIMMER M.: Reciprocal shading for mixed reality. *Computers & Graphics* 36, 7 (2012), 846–856. [16](#)
- [KTWW13] KNECHT M., TRAXLER C., WINKLHOFER C., WIMMER M.: Reflective and refractive objects for mixed reality. *IEEE Transactions Visualization and Computer Graphics* 19, 4 (2013), 576–582. [16](#)
- [KY04] KANBARA M., YOKOYA N.: Real-time estimation of light source environment for photorealistic augmented reality. In *ICPR* (2004), vol. 2, pp. 911–914. [5](#)
- [Lal11] LALONDE J.: *Understanding and recreating visual appearance under natural illumination*. PhD thesis, Carnegie Mellon University, 2011. [13](#)
- [Lan02] LANDIS H.: Production-ready global illumination. *Siggraph course notes* 16, 2002 (2002), 11. [4](#)
- [LB12] LENZING P., BROLL W.: Instant indirect illumination for dynamic mixed reality scenes. In *ISMAR* (2012), pp. 109–118. [16](#)
- [LE10] LALONDE J., EFROS A.: *Synthesizing environment maps from a single image*. Tech. rep., Carnegie Mellon University, 2010. [13](#)
- [Leh07] LEHTINEN J.: A framework for precomputed and captured light transport. *ACM Transactions on Graphics* 26, 4 (2007). [15](#)
- [LEN09a] LALONDE J., EFROS A., NARASIMHAN S.: Webcam clip art: Appearance and illuminant transfer from time-lapse sequences. *ACM Transactions on Graphics (Proc. of SIGGRAPH Asia 2009)* 28, 5 (2009), 131:1–131:10. [13, 18](#)
- [LEN09b] LALONDE J.-F., EFROS A. A., NARASIMHAN S. G.: Estimating natural illumination from a single outdoor image. *ICCV* (2009), 183–190. [13](#)
- [LEN12] LALONDE J.-F., EFROS A. A., NARASIMHAN S. G.: Estimating the Natural Illumination Conditions from a Single Outdoor Image. *International Journal of Computer Vision* 98, 2 (2012), 123–145. [13, 18](#)
- [LG12] LIU Y., GRANIER X.: Online tracking of outdoor lighting variations for augmented reality with moving cameras. *IEEE Transactions on Visualization and Computer Graphics* 18, 4 (2012), 573–80. [13, 18](#)
- [LH96] LEVOY M., HANRAHAN P.: Light field rendering. In *SIGGRAPH* (1996), pp. 31–42. [7](#)
- [LHE*07] LALONDE J.-F., HOIEM D., EFROS A. A., ROTHER C., WINN J., CRIMINISI A.: Photo clip art. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2007)* 26, 3 (2007). [10](#)
- [LK03] LEHTINEN J., KAUTZ J.: Matrix radiance transfer. In *Symposium on Interactive 3D Graphics and Games (I3D)* (2003), pp. 59–64. [15](#)
- [Llo82] LLOYD S.: Least squares quantization in pcm. *IEEE Transactions on Information Theory* 28, 2 (1982), 129–137. [14](#)
- [LMHRG10] LOPEZ-MORENO J., HADAP S., REINHARD E., GUTIERREZ D.: Compositing images through light source detection. *Computers & Graphics* 34, 6 (2010), 698–707. [11, 18](#)
- [LMSSG10] LOPEZ-MORENO J., SUNDSTEDT V., SANGORRIN F., GUTIERREZ D.: Measuring the perception of light inconsistencies. In *Symposium on Applied Perception in Graphics and Visualization* (2010), pp. 25–32. [1, 11](#)
- [LPG13] LU H., PACANOWSKI R., GRANIER X.: Real-Time Importance Sampling of Dynamic Environment Maps. *Eurographics Short Papers* (2013). [14](#)
- [LQX*09] LIU Y., QIN X., XU S., NAKAMEE E., PENG Q.: Light source estimation of outdoor scenes for mixed reality. *The Visual Computer* 25, 5–7 (2009), 637–646. [13](#)
- [LYLU09] LÖW J., YNNERMAN A., LARSSON P., UNGER J.: HDR Light Probe Sequence Resampling for Realtime Incident Light Field Rendering. In *Spring Conference on Computer Graphics* (2009), pp. 43–50. [8](#)
- [MBC13] MEILLAND M., BARAT C., COMPORT A.: 3D High Dynamic Range dense visual SLAM and its application to real-time object re-lighting. In *ISMAR* (2013), pp. 143–152. [9, 18](#)
- [MGV10] MOONS T., GOOL L. V., VERGAUWEN M.: 3D Reconstruction from Multiple Images Part 1: Principles. *Foundations and Trends in Computer Graphics and Vision* 4, 4 (2010), 287–404. [6, 9](#)
- [MH86] MILLER G., HOFFMAN C. R.: Illumination and reflection maps: Simulated objects in simulated and real environments. In *Proc. of SIGGRAPH 86* (1986). [4](#)
- [MLJ09] MEI X., LING H., JACOBS D.: Sparse representation of cast shadows via 1l-regularized least squares. In *ICCV* (2009), pp. 583–590. [10, 18](#)
- [MN08] MADSEN C. B., NIELSEN M.: Towards probe-less augmented reality: a position paper. In *International Conference on Graphics Theory and Applications* (2008), pp. 255–261. [13](#)
- [MPDW03] MASSELUS V., PEERS P., DUTRE P., WILLEMS Y. D.: Relighting with 4d incident light fields. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2003)* 22, 3 (2003), 613–620. [7](#)
- [MPK09] MURY A. A., PONT S. C., KOENDERINK J. J.: Representing the light field in finite three-dimensional spaces from sparse discrete samples. *Applied optics* 48, 3 (2009), 450–7. [7, 18](#)
- [Mys08] MYSZKOWSKI K.: *High Dynamic Range Video*. Morgan and Claypool Publishers, 2008. [5](#)
- [NDI*11] NEWCOMBE R. A., DAVISON A. J., IZADI S., KOHLI P., HILLIGES O., SHOTTON J., MOLYNEAUX D., HODGES S., KIM D., FITZGIBBON A.: Kinectfusion: Real-time dense surface mapping and tracking. In *ISMAR* (2011), pp. 127–136. [9, 10](#)
- [NGM*11] NOWROUZEZHRAI D., GEIGER S., MITCHELL K., SUMNER R., JAROSZ W., GROSS M.: Light factorization for mixed-frequency shadows in augmented reality. *ISMAR* (2011), 173–179. [15](#)
- [NHIN86] NAKAMEE E., HARADA K., ISHIZAKI T., NISHITA T.: A montage method: The overlaying of the computer generated images onto a background photograph. *Computer Graphics (Proceedings of SIGGRAPH '86)* 20, 4 (1986), 207–214. [13](#)
- [NN04] NISHINO K., NAYAR S.: Eyes for relighting. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2004)* 23, 3 (2004), 704–711. [10, 18](#)
- [NRH04] NG R., RAMAMOORTHI R., HANRAHAN P.: Triple product wavelet integrals for all-frequency relighting. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2004)* 23, 3 (2004), 477–487. [15](#)

- [NSKF07] NOWROUZEZAHRAI D., SIMARI P., KALOGERAKIS E., FIUME E.: Eigentransport for efficient and accurate all-frequency relighting. In *GRAPHITE* (2007), pp. 163–169. 15
- [OCS05] OSTROVSKY Y., CAVANAGH P., SINHA P.: Perceiving illumination inconsistencies in scenes. *Perception* 34, 11 (2005). 11
- [ODJ04] OSTROMOUKHOV V., DONOHUE C., JODOIN P.-M.: Fast hierarchical importance sampling with blue noise properties. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2004)* 23, 3 (2004), 488–495. 14
- [OSS04] OKABE T., SATO I., SATO Y.: Spherical harmonics vs. haar wavelets: Basis for recovering illumination from cast shadows. In *CVPR* (2004). 10
- [Pel10] PELLACINI F.: envylight: An interface for editing natural illumination. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2010)* 29, 4 (2010), 34:1–34:8. 5
- [PH10] PHARR M., HUMPHREYS G.: *Physically based rendering: From theory to implementation*. Morgan Kaufmann, 2010. 3, 14
- [PP03] PATOW G., PUEYO X.: A survey of inverse rendering problems. In *Computer graphics forum* (2003), vol. 22, Wiley Online Library, pp. 663–687. 2
- [PP05] PATOW G., PUEYO X.: A survey of inverse surface design from light transport behavior specification. In *Computer Graphics Forum* (2005), vol. 24, Wiley Online Library, pp. 773–789. 2
- [PSM93] PEREZ R., SEALS R., MICHALSKY J.: All-weather model for sky luminance distribution — preliminary configuration and validation. *Solar Energy* 50, 3 (1993), 235 – 245. 13
- [RAC*04] REINHARD E., AKYUZ A. O., COLBERT M., HUGHES C. E., O’CONNOR M.: Real-time color blending of rendered and captured video. In *Interservice/Industry Training, Simulation and Education Conference* (2004). 11
- [RAGS01] REINHARD E., ASHIKHMAR M., GOOCH B., SHIRLEY P.: Color transfer between images. *IEEE Computer Graphics and Applications* 21, 5 (2001), 34–41. 11
- [Ram09] RAMAMOORTHI R.: Precomputation-based rendering. *Foundations and Trends® in Computer Graphics and Vision* 3, 4 (2009), 281–369. 15
- [RCL*08] ROUSSELLE F., CLARBERG P., LEBLANC L., OSTROMOUKHOV V., POULIN P.: Efficient product sampling using hierarchical thresholding. *The Visual Computer* 24, 7-9 (2008), 465–474. 14
- [RDGK12] RITSCHEL T., DACHSBACHER C., GROSCH T., KAUTZ J.: The state of the art in interactive global illumination. *Computer Graphics Forum* 31, 1 (2012). 14, 17
- [RF07] RAMANARAYANAN G., FERWERDA J.: Visual equivalence: towards a new standard for image fidelity. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2007)* 26, 3 (2007), 76:1–76:11. 11
- [RGK*08] RITSCHEL T., GROSCH T., KIM M. H., SEIDEL H.-P., DACHSBACHER C., KAUTZ J.: Imperfect shadow maps for efficient computation of indirect illumination. In *ACM Transactions on Graphics (TOG) (Proc of SIGGRAPH Asia 2008)* (2008), vol. 27, p. 129. 16
- [RH01a] RAMAMOORTHI R., HANRAHAN P.: On the relationship between radiance and irradiance: determining the illumination from images of a convex lambertian object. *JOSA A* 18, 10 (2001), 2448–2459. 10
- [RH01b] RAMAMOORTHI R., HANRAHAN P.: A signal-processing framework for inverse rendering. In *SIGGRAPH* (2001), pp. 117–128. 10
- [RHD*05] REINHARD E., HEIDRICH W., DEBEVEC P., PAT-TANAIC S., WARD G., MYSZKOWSKI K.: *High dynamic range imaging: acquisition, display, and image-based lighting*, first ed. Morgan Kaufmann, 2005, ch. 9, pp. 396–401. 4
- [RHD*10] REINHARD E., HEIDRICH W., DEBEVEC P., PAT-TANAIC S., WARD G., MYSZKOWSKI K.: *High dynamic range imaging: acquisition, display, and image-based lighting*, second ed. Morgan Kaufmann, 2010. 2, 4
- [RU13] RABBI I., ULLAH S.: A survey on augmented reality challenges and tracking. *Acta Graphica znanstveni časopis za tiskarstvo i grafičke komunikacije* 24, 1-2 (2013), 29–46. 2
- [SH13] SCHMALSTIEG D., HOLLERER T.: *Augmented Reality: Theory and Practice*. Addison-Wesley, 2013. 2
- [SHHS03] SLOAN P.-P., HALL J., HART J., SNYDER J.: Clustered principal components for precomputed radiance transfer. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2003)* 22, 3 (2003), 382–391. 15
- [SHR*11] SUN X., HOU Q., REN Z., ZHOU K., GUO B.: Radiance transfer biclustering for real-time all-frequency biscale rendering. *IEEE Transactions on Visualization and Computer Graphics* 17, 1 (2011), 64–73. 15
- [SHS98] SHUM H.-Y., HAN M., SZELISKI R.: Interactive construction of 3d models from panoramic mosaics. In *CVPR* (1998). 9
- [SKS02] SLOAN P.-P., KAUTZ J., SNYDER J.: Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2002)* 21, 3 (2002), 527–536. 15
- [Slo08] SLOAN P.-P.: Stupid spherical harmonics (sh) tricks. In *Game developers conference* (2008), vol. 9. 15
- [SLS05] SLOAN P.-P., LUNA B., SNYDER J.: Local, deformable precomputed radiance transfer. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2005)* 24, 3 (2005), 1216–1224. 15
- [SRT*11] STURM P., RAMALINGAM S., TARDIF J.-P., GAS-PARINI S., BARRETO J.: Camera models and fundamental concepts used in geometric computer vision. *Foundations and Trends® in Computer Graphics and Vision* 6, 1–2 (2011), 1–183. 4
- [SS97] SZELISKI R., SHUM H.-Y.: Creating full view panoramic image mosaics and environment maps. In *SIGGRAPH* (1997), pp. 251–258. 4
- [SSI99] SATO I., SATO Y., IKEUCHI K.: Acquiring a Radiance Distribution to Superimpose Virtual Objects onto a Real Scene. *IEEE Transactions on Visualization and Computer Graphics* 5, 1 (1999), 1–12. 8, 18
- [SSI03] SATO I., SATO Y., IKEUCHI K.: Illumination from shadows. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25, 3 (2003), 290–300. 10
- [SSS06] SNAVELY N., SEITZ S. M., SZELISKI R.: Photo tourism: Exploring image collections in 3d. *ACM Transactions on Graphics (SIGGRAPH 2006)* (2006). 6, 9
- [STJ04] STUMPFEL J., TCHOU C., JONES A.: Direct HDR capture of the sun and sky. In *Afrigraph* (2004). 4, 9
- [SUC95] SLATER M., USOH M., CHRYSANTHOU Y.: The influence of dynamic shadows on presence in immersive virtual environments. In *Selected Papers of the Eurographics Workshops on Virtual Environments ’95* (1995), VE ’95, Springer-Verlag, pp. 8–21. 1
- [SZC*07] SUN X., ZHOU K., CHEN Y., LIN S., SHI J., GUO B.: Interactive relighting with dynamic brdfs. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2007)* 26, 3 (2007). 15

- [Sze06] SZELISKI R.: Image alignment and stitching: A tutorial. *Foundations and Trends in Computer Graphics and Vision* 2, 1 (2006), 1–104. 4
- [TCE05] TALBOT J. F., CLINE D., EGBERT P.: Importance re-sampling for global illumination. In *Rendering Techniques 2005 (Proc. of EGSR)* (2005), pp. 139–146. 14
- [TDMM03] TSUMURA N., DANG M. N., MAKINO T., MIYAKE Y.: Estimating the directions to light sources using images of eye for reconstructing 3d human face. In *Color and Imaging Conference* (2003), vol. 2003, pp. 77–81. 10
- [TKTS11] TOCCI M. D., KISER C., TOCCI N., SEN P.: A Versatile HDR Video Production System. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2011)* 30, 4 (2011). 5
- [TS06] TSAI Y.-T., SHIH Z.-C.: All-frequency precomputed radiance transfer using spherical radial basis functions and clustered tensor approximation. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2006)* 25, 3 (2006), 967–976. 15
- [TSE*04] TCHOU C., STUMPFEL J., EINARSSON P., FAJARDO M., DEBEVEC P.: Unlighting the parthenon. In *SIGGRAPH Sketches* (2004). 9
- [UG07] UNGER J., GUSTAVSON S.: High-dynamic-range video for photometric measurement of illumination. In *SPIE* (2007), vol. 6501. 7
- [UGLY08] UNGER J., GUSTAVSON S., LARSSON P., YNNERMAN A.: Free Form Incident Light Fields. *Computer Graphics Forum (Proc. of EGSR 2008)* 27, 4 (2008), 1293–1301. 7, 8, 18
- [UGOJ04] UNGER J., GUSTAVSON S., OLLILA M., JOHANNESSON M.: A real time light probe. In *Eurographics, Short Papers* (2004), pp. 17–21. 7
- [UGY06] UNGER J., GUSTAVSON S., YNNERMAN A.: Densely sampled light probe sequences for spatially variant image based lighting. In *GRAPHITE* (2006), pp. 341–347. 7
- [UGY07] UNGER J., GUSTAVSON S., YNNERMAN A.: Spatially Varying Image Based Lighting by Light Probe Sequences Capture, Processing and Rendering. *The Visual Computer* (2007), 1–13. 6, 7, 18
- [UKL*13a] UNGER J., KRONANDER J., LARSSON P., GUSTAVSON S., LÖW J., YNNERMAN A.: Spatially varying image based lighting using hdr-video. *Computers and Graphics* 37, 7 (2013). 6, 9, 18
- [UKL*13b] UNGER J., KRONANDER J., LARSSON P., GUSTAVSON S., YNNERMAN A.: Temporally and spatially varying image based lighting using hdr-video. In *EUSIPCO* (2013). 5, 9, 18
- [Ung09] UNGER J.: *Incident Light Fields*. PhD thesis, The Department of Science and Technology, Linköping University, The Institute of Technology, January 2009. Linköping Studies in Science and Technology, Dissertations, No. 1233. 7
- [UWH*03] UNGER J., WENGER A., HAWKINS T., GARDNER A., DEBEVEC P.: Capturing and rendering with incident light fields. In *Rendering Techniques 2003 (Proc. of EGSR)* (2003), pp. 1–10. 7, 18
- [VD09] VIRIYOTHAI K., DEBEVEC P.: Variance minimization light probe sampling. *SIGGRAPH Posters* (2009). 15
- [VG95] VEACH E., GUIBAS L. J.: Optimally combining sampling techniques for Monte Carlo rendering. In *SIGGRAPH* (1995), pp. 419–428. 14
- [VKP10] VAN KREVELEN D., POELMAN R.: A survey of augmented reality technologies, applications and limitations. *International Journal of Virtual Reality* 9, 2 (2010), 1. 2
- [VR08] VR S.: Spherocam hdr. www.spheron.com, 2008. 4
- [Wae03] WAENSE J.: A real time high dynamic range light probe. In *SIGGRAPH Sketches* (2003). 5, 18
- [Wan05] WAN L.: Spherical Q2-tree for Sampling Dynamic. In *Rendering Techniques 2005 (Proc. of EGSR)* (2005). 15
- [WJV*05] WILBURN B., JOSHI N., VAISH V., TALVALA E.-V., ANTUNEZ E., BARTH A., ADAMS A., HOROWITZ M., LEVOY M.: High performance imaging using large camera arrays. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2005)* 24, 3 (2005), 765–776. 7
- [WMWL11] WAN L., MAK S.-K., WONG T.-T., LEUNG C.-S.: Spatiotemporal sampling of dynamic environment sequences. *IEEE Transactions on Visualization and Computer Graphics* 17, 10 (2011), 1499–1509. 15
- [WRC88] WARD G. J., RUBINSTEIN F. M., CLEAR R. D.: A ray tracing solution for diffuse interreflection. *Computer Graphics (Proc. of SIGGRAPH 1988)* 22, 4 (1988), 85–92. 16
- [WS03] WANG Y., SAMARAS D.: Estimation of multiple directional light sources for synthesis of augmented reality images. *Graphical Models* 65, 4 (2003), 185–205. 10
- [WTL05] WANG R., TRAN J., LUEBKE D.: All-frequency interactive relighting of translucent objects with single and multiple scattering. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2005)* 24, 3 (2005), 1202–1207. 15
- [XZP13] XING G., ZHOU X., PENG Q.: Lighting Simulation of Augmented Outdoor Scene Based on a Legacy Photograph. *Computer Graphics Forum (Proc. of Pacific Graphics 2013)* 32, 7 (2013). 13, 18
- [YDMH99] YU Y., DEBEVEC P., MALIK J., HAWKINS T.: Inverse global illumination: Recovering reflectance models of real scenes from photographs. In *SIGGRAPH* (1999), pp. 215–224. 9
- [YKK12] YAO Y., KAWAMURA H., KOJIMA A.: Shading derivation from an unspecified object for augmented reality. *ICPR* (2012), 57–60. 5
- [YKK13] YAO Y., KAWAMURA H., KOJIMA A.: The Hand as a Shading Probe. *SIGGRAPH Posters* (2013). 5, 18