Free Form Incident Light Fields

J. Unger[†], S. Gustavson, P. Larsson and A. Ynnerman

VITA, Linköping University, Sweden

Abstract

This paper presents methods for photo-realistic rendering using strongly spatially variant illumination captured from real scenes. The illumination is captured along arbitrary paths in space using a high dynamic range, HDR, video camera system with position tracking. Light samples are rearranged into 4-D incident light fields (ILF) suitable for direct use as illumination in renderings. Analysis of the captured data allows for estimation of the shape, position and spatial and angular properties of light sources in the scene. The estimated light sources can be extracted from the large 4D data set and handled separately to render scenes more efficiently and with higher quality. The ILF lighting can also be edited for detailed artistic control.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism; I4.1 [Image Processing and Computer Vision]: Digitization and Image Capture;

1. Introduction

Traditional image based lighting (IBL) employs a 2D angular map of the environment, a *light probe*, to determine the incident intensity from any direction. IBL using a single light probe can only represent an angular variation and inherently assumes spatially invariant lighting. Extensions to this low-dimensional approximation have been presented in the literature, as referenced in Section 2, but due to hardware limitations, overwhelming amounts of data and lack of efficient rendering methods they have not been successfully applied in production. Recent developments in HDR imaging and computer hardware have made it feasible to capture and process more of the dimensionality of the plenoptic function. At the heart of the presented work lies dense volumetric real-time sampling of 3D sets of thousands of light probes together with efficient methods for data representations supporting data reduction, editing and fast rendering.

2. Previous work

The work presented in this paper builds on the idea of capturing omni-directional HDR images of the illumination in a real world scene, *light probe images*, and using

Figure 1: *Three synthetic objects on a real table in a real scene. Each point (x, y, z) in the rendered scene has a different environment map, interpolated from more than 40,000 light probe images.*

an estimate of the *plenoptic function* to improve the realism and visual interest in renderings. The plenoptic function $P(\lambda, x, y, z, \phi, \theta, t)$, defined by [AB91], describes the flux of light with wavelength λ through a point (x, y, z) in space from the direction (ϕ, θ) at time t. This is often reduced to a 5D function, $P(x, y, z, \phi, \theta)$, where time is fixed and only three spectral bands (R, G, B) are considered.

Traditional *image based lighting* is based on environment mapping [Bli76, MH84]. Omnidirectional HDR images are

---

† jonas.unger@itn.liu.se

captured at a single point in space as samples of the plenoptic function, $P(\phi, \theta)$, and used as illumination in renderings. [Deb98] and [SSI99] demonstrated how such radiance maps, *light probes*, could be used to photo-realistically render synthetic objects into real scenes. HDR imaging and traditional IBL are now well researched fields where many of the results have been employed in industry applications. For an overview see [RWPD06]. As an extension to traditional IBL, [UGY07] sampled the plenoptic function along linear 1D paths using an HDR video camera aimed at a mirror sphere [UG07]. In this paper we use a similar approach, but extend the measurements to irregularly spaced samples in 3D within a volume in space.

A *light field*, as introduced by [LH96, GGSC96], is a 4D ray space based on a set of 2D images of a scene captured at different viewpoints. Such 4D ray data can be analyzed such that novel views of the scene can be computed with a variety of image enhancements and effects such as active refocusing and synthetic aperture control [WJV*05]. In this work we project the captured light ray samples to a set of 4D HDR light fields enclosing the scene. During rendering, the incident light fields can be sampled such that each point in the scene has a different environment map, thus capturing both the spatial and angular variations in the illumination. A similar technique was proposed by [UWH*03], who extended IBL to use a 4D version of the plenoptic function, $P(x, y, \phi, \theta)$, by regularly sampling the illumination incident from the hemisphere above a plane in space. Such 4D HDR illumination data were named *incident light fields* (ILF) to emphasize that the idea was to capture the illumination as incident onto a region in space, as opposed to light fields for direct viewing. Although this work showed the potential of using HDR light fields as illumination in rendering, the method was of limited practical use due to long capture time and limited resolution. [MPDW03] used a similar light field technique to illuminate measured reflectance fields, introduced by [DHT*00], using captured spatially varying illumination. Taking another approach, [ISG*08] used a tilting planar mirror to obtain a dense spatial but sparse angular sampling of a 4D ILF, and showed how such data could be used for GPU hardware rendering. In contrast to [ISG*08] our sampling is dense in the angular domain and more sparse, but adaptive, in the spatial domain.

To improve rendering efficiency and allow for artistic editing of the captured illumination we extract the high intensity areas in the scene, such that they can be individually represented in their own 4D data set. The spatial and angular characteristics of a light source can be captured by measuring the exitant light field near the light source. This technique, called *near field photometry*, was presented by [Ash93]. To efficiently render scenes with light sources exhibiting complex spatial and angular variations, [HKSS98] proposed a 4D light field representation called *canned light sources*. Later [GGHS03] described a technique where the near fields of real world light sources were measured and reproduced
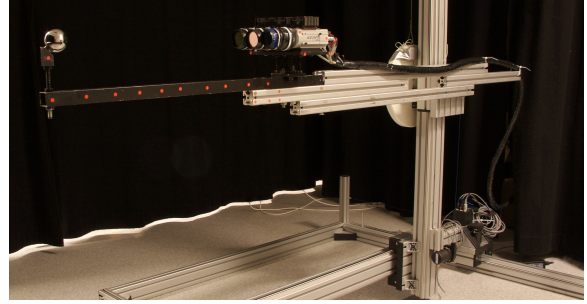

Figure 2: *The real time light probe on an (x, y, z) translation stage with motion tracking.*

in renderings using a similar representation which also supported importance sampling for efficient rendering. In our approach, a spatial and angular sampling of each light source is acquired in the region of interest in the scene and reprojected to a 4D near field representation. The extracted illuminants can also be edited. Using synthetic aperture imaging methods [WJV*05], light field data can be depth resolved and occluding objects can be removed. Similar results using synthetic aperture and confocal methods were presented by [LCV*04]. Here, we demonstrate how a blocker can be removed and how the modulation filter in a projector light can be enhanced or changed by editing of subsets of rays.

3. Illumination Capture

The rendering equation introduced by [Kaj86] is a general foundation for global illumination algorithms. It relates the radiance, $B(\mathbf{x}, \vec{\omega}_o)$, of a scene point, $\mathbf{x}$, with normal, $\mathbf{n}$, in a viewing angle, $\vec{\omega}_o$, to incident illumination, $L(\mathbf{x}, \vec{\omega}_i)$, at that point, $\mathbf{x}$, from an angle of incidence, $\vec{\omega}_i$, and a transfer function, $T(\mathbf{x}, \vec{\omega}_i \rightarrow \vec{\omega}_o)$, which is a spatially variant BRDF denoting the surface reflectance properties at $\mathbf{x}$, for incident and exitant angles $\vec{\omega}_i$ and $\vec{\omega}_o$:

$$B(\mathbf{x}, \vec{\omega}_o) = \int_{\forall \vec{\omega}_i} L(\mathbf{x}, \vec{\omega}_i) T(\mathbf{x}, \vec{\omega}_i \rightarrow \vec{\omega}_o)(\mathbf{n} \cdot \vec{\omega}_i) d\vec{\omega}_i \quad (1)$$

Traditional IBL assumes that light sources are distant, whereby $L(\mathbf{x}, \vec{\omega}_i)$ does not depend on position and can be written $L(\vec{\omega}_i)$. This simplification, however, ignores all spatial variation such as parallax, lighting non-uniformity and cast shadows. The incident illumination, $L(\mathbf{x}, \vec{\omega}_i)$, is formally the 5D version of the plenoptic function over three spatial dimensions and two angular dimensions. By assuming that the scene is stationary during the capture time we can consider $P_m = P(x, y, z, \phi, \theta)$ as a 5D function independent of t and so:

$$L(\mathbf{x}, \vec{\omega}_i) = P_m \quad (2)$$

For the remainder of this article, we will denote $L(\mathbf{x}, \vec{\omega}_i)$

as simply $L(\mathbf{x}, \vec{\omega})$, because we will only be concerned with the illumination and not the spatially variant BRDF. The sampled version of $L(\mathbf{x}, \vec{\omega})$ will be written $L_{ij} = L(\mathbf{x}_i, \vec{\omega}_j)$ where i is the index into the set of spatial samples and j is the index into the set of angular samples.

3.1. Experimental setup

The experimental setup for light field capture consists of an HDR video camera system on a translation stage with motion tracking, see Figure 2. The camera is a Ranger C55 from Sick IVP, and the capture algorithm is described in detail in [UG07]. The imaging system captures color HDR images of 640×512 pixels with a dynamic range of 10,000,000:1 at 25 FPS. The camera is aimed at a mirror sphere, forming a real time light probe. Each captured panoramic image exhibits a field of view of approximately $356°$, taking into account the dead angle directly behind the mirror sphere. The final resolution of a light probe image after cropping is close to 512×512 pixels. The translation stage uses mechanical parts from AluFlex Systems. Tracking is achieved using JX-EP opto-mechanical encoders from Herbertek and an Arduino microcontroller. The range of motion of the translation stage is $1.5 \times 1.5 \times 1.5\ m$, with a sub-millimeter accuracy in the position tracking.

The tracked real time light probe enables rapid capture of HDR light probe data from anywhere within a bounded volume. Sampling is performed by moving the camera across a scene during continuous capture to cover the region of interest. The plenoptic function is sampled in a dependent 6D form $P_m = P(x(t), y(t), z(t), \phi, \theta, t)$. By synchronizing the motion tracking with the frame rate of the camera, the measured $L(\vec{\omega}, t)$ can be combined with the tracking data, $\mathbf{x}(t)$, to yield $L(\mathbf{x}(t), \vec{\omega}, t)$. Assuming the scene is constant during capture, this reduces to $L(\mathbf{x}, \vec{\omega})$.

The capture path is traced by manual motion in a free-form manner with a direct feedback from the tracking system displaying where illumination samples have been captured. Manual motion allows for operator-directed variations in sampling density to reflect regions of particular interest in the scene. It also yields a non-uniform sampling of the light field. Instead of resampling the data to a regular grid, we maintain the data in its original non-uniform distribution. This reduces reconstruction errors, and makes it more straightforward to perform adaptive sampling and decimation. Each light probe provides a dense sampling of the angular variation. In our setup, the camera orientation is kept fixed during capture and, therefore, the same set of angles will be captured by each light probe image. The angular variation can be reconstructed at all points without any significant error.
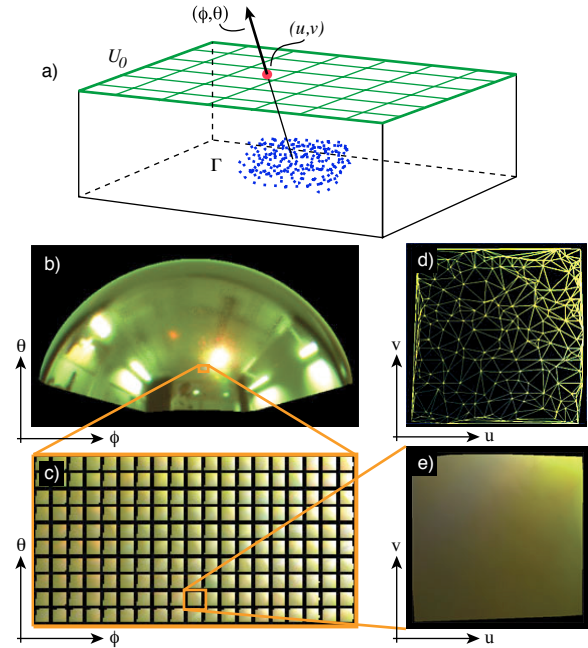

Figure 3: *A free-form ILF data structure. **a)** Upward directions (ϕ, θ) from the irregular point cloud of light probe samples in Γ are projected to a point (u, v) in the plane U_0 above the scene and stored in a 4D data structure $L(u, v, \phi, \theta)$. **b)** Each direction in the angular map stores a bucket of positioned intensity samples. **c)** A close-up reveals the bucket structure of the angular map. **d)** The sample set for each bucket is triangulated, and **e)** reconstruction is performed by bilinear interpolation.*

3.2. Viewpoint correction

The reflection from a spherical surface does not have a single viewpoint [SGN06]. In ordinary IBL this is not a problem, because the spatial variation of the lighting is neglected, but in our case the spatial variation is captured and represented in the data, and the sample distance is significantly smaller than the diameter of the reflective sphere. Therefore, individual ray samples are associated with different viewpoints. Any point along a ray can be used as its origin, e.g. the actual hit point on the sphere.

Because the HDR capture by our particular camera system is performed using multiple exposures and a rolling shutter, different pixels in the image will be exposed at different instants in time. Using knowledge of which pixel and which exposure was used for each sample in the HDR image, we can compensate in full for this by linear interpolation between adjacent tracking positions.

4. Illumination Representation

A captured light probe sequence is a sampling of light from all directions, $\vec{\omega}$, incident onto a set of irregularly spaced

points, **x**, along a path within a convex spatial region, $\mathbf{x} \in \Gamma$, as indicated in Figure 3a. The objective is to cover a convex hull of suitable volume with a dense enough sampling. To create a data structure suitable for fast ray lookup during rendering, all sampled rays are reprojected onto a bounding surface enclosing the scene. No resampling is performed in this process, it is merely a reordering and projection of individual rays from a 5D ray space to a 4D subspace. It is beneficial to make the bounding volume match the actual geometry of the environment, but such a match is by no means crucial as long as there exists a parameterization such that individual ray origins can be found rapidly during rendering. An additional feature of this data structure is that it allows for straightforward adaptive decimation of the data. Similarly to [UWH*03], we call this irregularly 4D ray representation a *free form incident light field* (ILF).

4.1. Ray Space Partitioning and Binning

Our choice of bounding surface is a set of planes, U, similar to a cube environment mapping but with unrestricted size, orientation and proportions. The near $360°$ field of view of each light probe image is partitioned across these bounding planes such that all positions for a certain direction, $\vec{\omega}_j$, are collected in the same plane with a 2D mapping **u**. The rays corresponding to a certain bounding plane are projected along their respective direction, $\vec{\omega}_j$, and stored in a planar data structure where the new position, $\mathbf{u}_i$, of each ray is restricted to a plane, U, see Figure 3a. The figure denotes the projected ILF as:

$$L(\mathbf{u}_i, \vec{\omega}_j) = L(u, v, \phi, \theta) \quad (3)$$

The projection can be performed without error or loss of generality, provided that the sampled volume Γ is free from internalt occlusion. In this process we also perform a binning of the rays depending on their direction $\vec{\omega}_j$ such that all rays in a certain direction are stored in the same angular bucket. This means that the rays are sorted on angle $\vec{\omega}_j$, with the regular and dense sampling from the light probe images being kept for the two $\vec{\omega}$ dimensions. Within each $\vec{\omega}_j$ bucket, the positions, $\mathbf{u}_i$, are stored in an irregular point cloud. To reconstruct the signal at any point in the plane, we use a Delaunay triangulation with barycentric interpolation between sample positions. The spatial samples are indexed in a balanced kd-tree structure for efficient lookup. This data structure is presented in Figure 3b-e.

The number of angular buckets used in the binning determines the resolution of the environment map seen from each point in the scene. In our experiments, the number of bins is determined by the resolution of the input light probe images. The exact number of bins used in the example renderings is discussed further in Section 6.
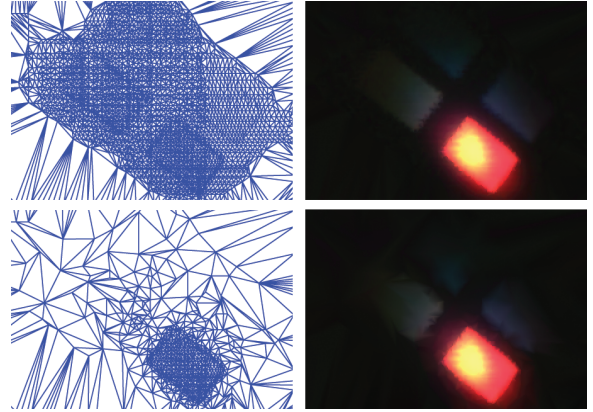

Figure 4: *Adaptive decimation. From 26,000 original samples, the global decimation (top) reduces the data set to 5,000 samples, and local decimation (bottom) reduces this particular high contrast bucket to 1,500 points. The reconstructed signal (right) has an RMS error of 0.3% and a total energy error of 0.1%.*

4.2. Adaptive Decimation

Real world lighting exhibits local regions with slow spatial variations. Therefore, the free-form ILF representation can be made more compact using adaptive non-uniform sampling. The raw data for each bucket, $\vec{\omega}_j$, contains one $\mathbf{u}_i$ for each light probe position, $\mathbf{x}_i$, in the captured sequence, but many samples are redundant and can be removed. An initial global decimation is based on inter-image differences between adjacent light probes. An entire image is considered redundant if it is very similar to both the previous and the next image, and its captured position is close to either of those images. The similarity is measured in terms of energy in the absolute difference between images. The decimation is performed in a conservative manner, such that it only removes closely spaced points in regions where there is very little variation in the incident lighting.

A subsequent local decimation algorithm operates on each angular bucket separately, where each vertex, $\mathbf{u}_i$, in the Delaunay triangulation is considered for removal. The method is based on [Sch97], but uses a different optimization criterion. The local error, E_{L_i}, introduced by removing a vertex, $\mathbf{u}_i$, from its local neighborhood in bucket $\vec{\omega}_j$ is computed as the difference in energy between the original triangle set, $\triangle_n$, and decimated triangle set, $\triangle_m$:

$$E_{Li} = \sum_n \int_{\triangle_n} L(u, v, \vec{\omega}_j) du dv - \sum_m \int_{\triangle_m} L(u, v, \vec{\omega}_j) du dv \quad (4)$$

Vertex $\mathbf{u}_i$ is removed if $|E_{Li}| < T_L$, i.e. if the introduced local error is less than a threshold, T_L. To avoid systematic errors we also set a bound for the accumulated global error within each bucket. Together these criteria preserve local and global energy and local contrast in the ILF, while efficiently removing vertices in areas with a largely constant
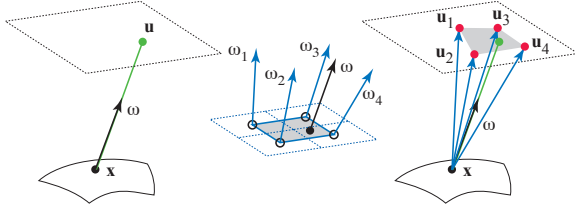
Figure 5: *ILF ray sampling. The value at sample position **u** in direction $\vec{\omega}$ is adjusted by reprojection along the four nearest sampled directions in the angular domain, followed by interpolation in the spatial domain.*

gradient. The resulting data set after decimation is a 4D indexed structure, L_{ij}, where a set of angle buckets, $\vec{\omega}_j$, contain a variable number of irregularly spaced sample points, $\mathbf{u}_i$. Figure 4 shows the triangles and the reconstructed (u, v) spatial intensity variation for one bucket before and after decimation. The local threshold T_L in this example was set to 0.1% of the average radiant energy per square millimeter in the entire light field. The resulting full ILF was binned in 256×256 angular buckets of $26,000$ samples each. Most of the energy was in the top plane, and its original data set comprised around 250M rays. After decimation, it could be represented using only 8M rays concentrated to high contrast regions, a data reduction of 97%. The RMS error of a typical high-intensity bucket, see Figure 4, was 0.3%, and the total relative error was less than 0.1%. Low intensity buckets exhibit larger relative errors, but similar absolute errors, and are decimated significantly more than buckets with high intensity and high contrast content.

5. Rendering with Incident Light Fields

Support for ILF illumination can be implemented in a standard global illumination framework which evaluates light contributions along explicit sample rays. Renderings in this article were produced using a custom plug-in for *mental-ray* from Mental Images. The 4D ILF is represented as a generalized area light source, where the intensity varies according to a shader function of both position, $\mathbf{u}$, and angle, $\vec{\omega}$. Each sample ray, $(\mathbf{x}, \vec{\omega})$, is projected to the ILF $\mathbf{u}$ plane, where $L(\mathbf{u}, \vec{\omega})$ is evaluated by interpolation. The interdependency of $\mathbf{u}$ and $\vec{\omega}$ is handled similarly to the *depth correction* method suggested by [GGSC96] and [UWH*03]. The ILF ray sampling is illustrated in Figure 5. A sample ray, $(\mathbf{x}, \vec{\omega})$, has an actual hit point, $\mathbf{u}$, on the SLF plane, but typically neither the exact direction $\vec{\omega}$ nor the exact hit point $\mathbf{x}$ have been sampled directly, hence interpolation is required in both the angular and spatial dimensions. This 4D interpolation is performed in two 2D steps. First, the four closest angular samples, $\vec{\omega}_{1...4}$, are picked and four rays are re-projected from the point, $\mathbf{x}$, in those directions to four adjusted hit points, $\mathbf{u}_{1...4}$. The values for L_{ij} at those points are evaluated by the previously described Delaunay interpolation, and the final

value is computed by bilinear interpolation between those four points.

The sample volume typically does not cover the entire scene, but a smaller region, Γ, of special interest where highly detailed rendering is desired. For points $\mathbf{x} \notin \Gamma$ data is available only for directions, $\vec{\omega}_i$, that intersect Γ. To enable rendering of objects outside Γ, extrapolation is employed. For directions, $\vec{\omega}$, not intersecting Γ, a clamp-to-edge extrapolation finds the closest point on the boundary and assumes it is representative for all $\mathbf{x} \notin \Gamma$ in that direction. This assumption removes the spatial variations outside Γ, but maintains a variation with direction to the sample volume and gives a predictable behavior even outside the sampled region.

An example rendering with a photographic reference is shown in Figure 6b,c. The illumination was captured in a test scene with spatially varying lighting, containing an area light with a perforated blocker and a light projector with a gobo pattern, see Figure 6a. Six high resolution ILF planes were the only sources of illumination. While this kind of direct rendering using an ILF data set from the full environment yields a high level of realism, it is impractical. The large amount of data and the large number of ray samples required causes long rendering times. As for image based lighting in general, a blind stochastic sampling has difficulties in reliably hitting the small angular regions where the light sources are located.

5.1. Illuminant Extraction for Directed Sampling

An ILF can describe arbitrarily complex lighting in a scene. However, as described, it also presents a considerable sampling problem during rendering. Importance sampling of a single 2D light probe is straightforward, but for the 4D ILFs it is intractable to predetermine good sampling distributions for all possible cases. Since the planar ILF geometry generally does not match the actual scene geometry, tabulation of important high energy areas in the scene might differ significantly between each angular bucket. By matching the ILF geometry to that of the high intensity areas in the scene, compact high energy clusters can be located in 3D space, and sampling can be efficiently concentrated during rendering. Therefore, we take a scene-based approach and separate the illumination into several terms such that:

$$L(\mathbf{u}, \vec{\omega}) = L_0(\mathbf{u}_{k_0}, \vec{\omega}_{k_0}) + \sum_{n=1}^{N} L_n(\mathbf{u}_{k_n}, \vec{\omega}_{k_n}) \qquad (5)$$

where k is the joint index over the entire (i, j) sample set and k_n represent extracted subsets. L_0 is the residual background ILF which covers the environment. It has low intensity and predominantly low frequency content, while $L_{1...N}$ have high intensity, high frequency content and strong localization. These high intensity areas represent illuminants which are extracted and re-projected to local *source light fields* (SLFs), which are explicitly sampled more densely during rendering, similarly to a traditional area light source.
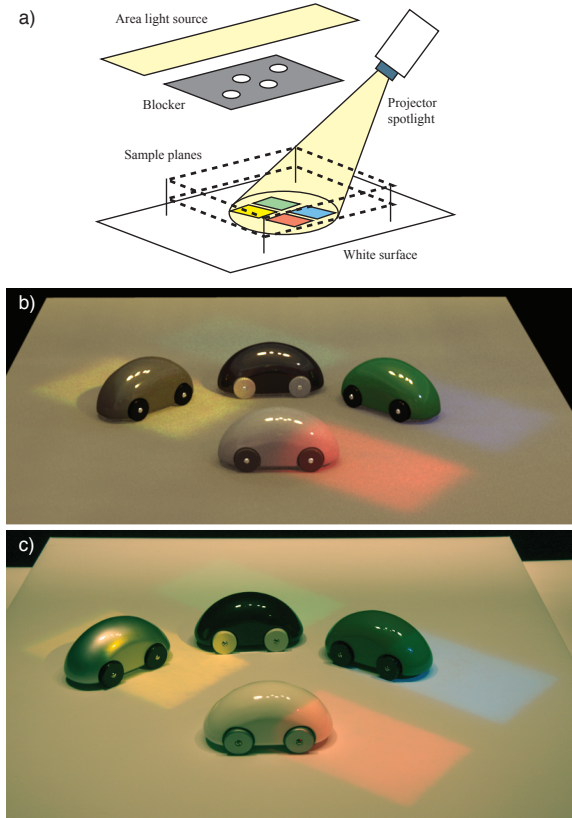
Figure 6: *a) Test scene setup. Direct illumination comes from an occluded area light from above and a light projector from the upper right. b) A rendering with an ILF based on 20,000 light probes from the scene. c) A reference photograph from the scene. The rendering used 6,000 light samples at each diffuse point. The rendering time was several hours, but the resulting image still has noticeable noise.*

Extracting $L_{1...N}$ from the scene is a 4D clustering problem. To separate light sources, both intensity and color as well as spatial and angular discrimination can be employed. The extraction is implemented as a semi-automated method where a user selects an angular region in a sum of the ILF across all positions, $L_j = \sum_i L(\mathbf{u}_i, \vec{\omega}_j)$, or in a max or mean image of the angular buckets, see Figure 7a,b. This selection is followed by a thresholding such that a set of representative rays, $L(\mathbf{u}_{k_n}, \vec{\omega}_{k_n})$, for the illuminant are selected. The subset of rays are then projected to points, $\tilde{\mathbf{u}}$, in a plane, U_n, with the same orientation as the ILF but a variable offset, z. The optimal z value is found by minimizing the projected cluster size in $\tilde{\mathbf{u}}$:

$$z_n = \arg\min_z \left(A(\tilde{\mathbf{u}}_{k_n}(z)) \right) \qquad (6)$$

where $A(\tilde{\mathbf{u}}_{k_n}(z))$ is a suitable measure of the cluster size of $\tilde{\mathbf{u}}_{k_n}(z)$, e.g. the norm of its vector standard deviation, and
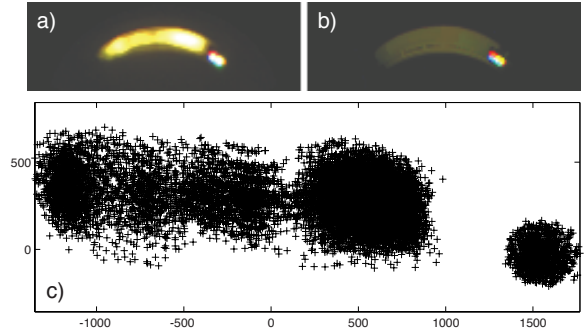

Figure 7: *a) Mean intensity and b) maximum intensity for each (ϕ, θ) bucket of one ILF. c) When high intensity rays are projected to a suitably chosen plane, light sources which are separate in the scene will be separated in the plane, which allows for unambiguous extraction. Here, rays from the partly blocked area light are to the left and rays from the light projector to the right. Units are in millimeters.*

$\tilde{\mathbf{u}}_{k_n}(z)$ is the projection of $\mathbf{u}_{k_n}$ along $\vec{\omega}_{k_n}$ to the plane U_n at offset z.

Even if rays from more than one illuminant are selected, separate SLFs can still be resolved since the clusters in $\mathbf{u}_{\tilde{k}_n}$ will separate during the reprojection. Such separation is displayed in Figure 7c, where the left cluster is the partly occluded area light and the right circular cluster is the light projector from the test scene in Figure 6. At the optimal $z = z_n$, the convex hull of $\tilde{\mathbf{u}}_{k_n}(z_n), \forall i$, is used as an estimation of the spatial extent, U_n, of illuminant n. All rays from the original ILF that intersect U_n are then stored in the SLF to describe both the shape and detailed appearance of the illuminant. Rays associated with the SLF are removed from the ILF, and the holes introduced in the background are interpolated over by a re-triangulation of the Delaunay mesh. Both localized point-like light sources and area light sources can thus be extracted to SLF planes with optimal position and spatial extents, closely representing actual illuminants. This representation is similar to the canned light sources used in [HKSS98].

A suitable threshold for the extraction can be found using an energy histogram over the ray samples. The illuminants stand out very clearly from the background in the HDR data. As described above, when an angular and spatial subset is considered, the discrimination between illuminants can also be based on the projected position in the scene for better separation and increased robustness.

During rendering, an extracted SLF presents a slightly different problem from an ILF. The SLF contains only a subset of directions, $\vec{\omega}_{k_n}$, captured in Γ. Because the SLF contains only rays emitted from the illuminant, re-projected to the illuminant proxy, the SLF representation itself contains the parallax and shape information for the illuminant. For points $\mathbf{x} \notin \Gamma$ there exist directions, $\vec{\omega}$, which intersect the ex-

Figure 8: *a) A light source SLF extracted and separated from an ILF. b) The residual background ILF with missing data. c) The sum of the influence of both the SLF and the ILF.*
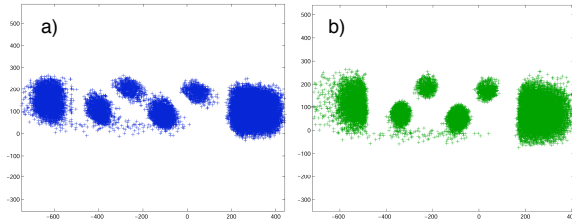

Figure 9: *Extracted high-intensity rays from the area light. a) Rays projected to the plane of the light source, showing maximum convergence. b) Rays projected to the plane of the blocker, showing crisp edges and holes for the blocker. Units are in millimeters.*

tracted illuminant but are not represented in the SLF data. Such directions are extrapolated using the closest direction represented in the SLF as a representative. If Γ is chosen to cover the significant variations in illumination, this allows for qualitatively satisfactory rendering of objects outside Γ.

An example of an extracted light source can be seen in Figure 8. In this case the sample region, Γ, was smaller in extent than the light source (a large chandelier), so there is a spot of missing data for some angles, $\vec{\omega}$, in the background. Assuming, however, that the background is locally uniform, these missing values can be interpolated and the hole can be filled. This is in effect an educated guess of missing data in the background ILF when seen from scene points outside the sample region. It will make specular reflections of the environment look significantly better from such points, and it will not change the ILF as seen from within the sample region, because the holes will be covered by the corresponding extracted SLFs.

5.2. Illuminant Analysis and Editing

An SLF data set can be manipulated like a classic light source: it can be *rotated, translated* and globally have its *intensity and color changed*, but since it contains 4D ray data, it is also possible to perform more complex editing of its spatial and angular properties on individual rays. Such *synthetic light aperture editing* includes *blocker removal, refocusing* and *enhancement or editing of projected patterns*. As examples of this, we present removal of an occluder and enhancement and editing of a projected pattern.

In order to edit individual rays, we model the SLF, $L_n(\mathbf{u}_{k_n}, \vec{\omega}_{k_n})$, as a collection of rays with equal intensity,

$I_n(\mathbf{u}_{k_n}, \vec{\omega}_{k_n})$, multiplied by a modulation function, $f_z(s,t)$, to represent filters and blockers in the ray path, where (s,t) are planar coordinates projected to a plane, U_n, at distance z from the SLF.

$$L_n(\mathbf{u}_i, \vec{\omega}_j) = I_n(\mathbf{u}_i, \vec{\omega}_j) \times f_z(s,t) \qquad (7)$$

This 2D function, $f_z(s,t)$, is not a fully general description of a 4D light field. Light sources with complex angular and spatial interdependence require the full 4D representation to be preserved, but many real world light sources are suitable for this representation. In our experiments, the diffuse area light with a planar occluder and the projector spotlight could both be well described by this simplified model.

Analysis of the modulation, f, for the area light in Figure 6 reveals that it can be accurately described as a diffuse area light source with an on-off modulation pattern $f_z(s,t)$. This distance z can be found by reprojecting the rays in I_n to the plane of maximum focus of the blocker edges. Figure 9 shows a subset of high energy rays from the SLF with z at the light source and z at the blocker, respectively. By removing all rays which hit the region of the extracted light source but do not have a high intensity, the blocked-out parts will be interpolated over during rendering, and the light source will be synthetically unoccluded. A rendering with the unoccluded area light is shown in Figure 11b,d.

The projected pattern can be edited by replacing $f_z(s,t)$ with a new function, $g_z(s,t)$. An example of this is displayed in Figure 10, where a complex filter pattern was deliberately undersampled (a). Undersampling of a detailed light pattern will result in loss of detail and introduce aliasing, (b). To enhance the data, the new modulation function, g, was created from a reference photo of the projector pattern captured in the scene (c). The photograph was rectified and manually aligned with the analyzed projector filter. Since the projector gobo is located in a plane, f could be replaced with a function g describing the planar modulation in the image. By replacing f with g, a much enhanced rendering could be created (d). Varying z shows that the depth of field properties in the SLF are correctly preserved (d, e, f).

6. Results

Using ILF data with 256×256 angular buckets from the scene in Figure 6a two illuminants containing 95% of the energy were extracted, each to their own SLF plane. Because the residual background was very dark, all ILF planes except the top plane were subsampled to a 64×64 bucket resolution. After processing, the ILF was represented using 13M rays, or 277 MB using 20 bytes per ray. Around 30% of the rays were concentrated in the SLFs. The physical size of the area light source was 20 cm by 120 cm, and at the optimal distance the corresponding measured rays converged well to within a rectangle of about that size, see Figure 9a. The perforated blocker was removed from the data set, thereby synthesizing an unobstructed area light. The light projector was
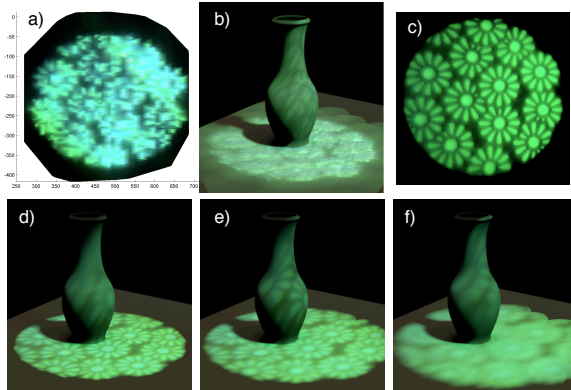
Figure 10: *a) A 2D blocker extracted from under-sampled data. b) Corresponding rendering. c) A photosourced image replacing the under-sampled pattern. d) A much improved rendering. e), f) The narrow depth of field of the real projector is preserved, and the pattern can be refocused by moving the position of the 2D blocker.*

projected similarly to its estimated area of origin, a 10 cm circle. Its actual aperture was slightly smaller than this, but glare and a limited angular resolution of the light probe images limits the precision of the re-projection for each ray. By thus separating the data set to two small SLFs which were densely sampled during rendering and a large, sparsely sampled residual environment ILF, the rendering time was orders of magnitude shorter (minutes instead of hours), while improving the visual quality, see Figure 11.

Figure 10 displays how synthetic light aperture editing can be performed on an SLF data set to enhance projected patterns. As demonstrated in the figure, the light source focus and depth of field properties are preserved.

For the rendering in Figure 1 a capture was performed in a production-like environment, a local castle. The original point cloud of light probe data was around 9G light ray samples from more than 40,000 light probes with a resolution of 500×500 pixels. This data set was subsampled to 256×256 angular buckets, resulting in 2.25G rays. The ILF was decimated and processed such that a large chandelier, a mirror, a wall-mounted lamp and three windows were extracted as displayed in Figure 12. The extracted illuminants contained 82% of the total energy in the scene, and the contrast ratio in the residual ILF, expressed as the maximum value divided by the mean value of the rays, was about 10:1 compared to 1000:1 before the extraction. The extracted SLFs and residual ILF contained only 90M rays, 4% of the unprocessed ILF, a data size of 1.8GB. The extracted illuminants contained 13M rays, 270 MB in total.

7. Conclusion and Future Work

The approach presented in this paper provides a means for capture, modeling and rendering of arbitrarily spatially vary-
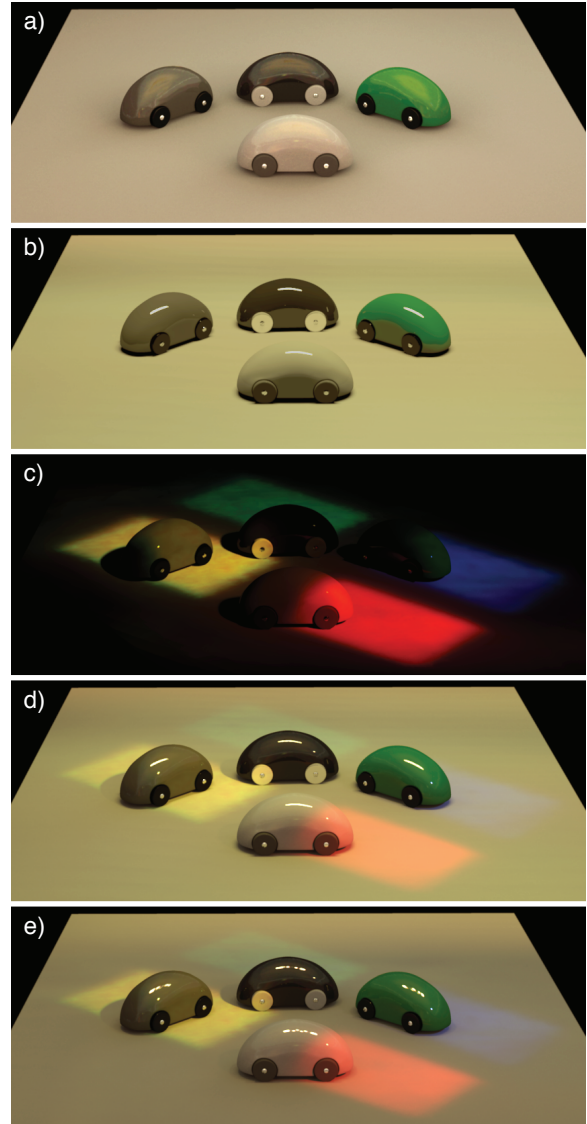

Figure 11: *Rendering with extracted SLFs. a) Influence of the residual ILF, L_0, amplified for presentation. b) Influence of the area light, L_1, with a blocker synthetically removed. c) Influence of the light projector, L_2. d) Final rendering using all lights, $L = L_0 + L_1 + L_2$. e) Rendering with the original occluded area light for direct comparison with Figure 6. Rendering time was cut to four minutes, compared to five hours for the unseparated ILF, L.*

ing illumination, which can be implemented in any global illumination framework. This is a significant improvement over existing IBL methods, as it incorporates the full spatial dimensionality of the plenoptic function. Light field illumination provides unprecedented levels of realism, and our approach is well suited to make immediate use of upcoming commercial HDR video cameras.
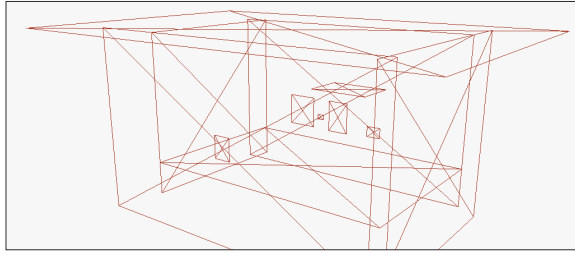
Figure 12: *The residual ILF planes (large rectangles) and seven extracted SLFs (small rectangles) used for the rendering in Figure 1, as seen in the 3D editor (Maya).*

Directions for future work include fully automated and improved extraction of illuminants, removal of the capture device and operator from the ILF, data compression beyond the decimation presented in this work, and further software development to incorporate ILF methods in existing work-flow pipelines.

Acknowledgements

References

[AB91]   ADELSON E. H., BERGEN J. R.: *Computational Models of Visual Processing*. MIT Press, Cambridge, Mass., 1991, ch. 1. The Plenoptic Function and the Elements of Early Vision.

[Ash93]   ASHDOWN I.: Near-field photometry: A new approach. *Journal of the Illuminating Engineering Society 22*, 1 (Winter 1993), 163–180.

[Bli76]   BLINN J. F.: Texture and reflection in computer generated images. *Communications of the ACM 19*, 10 (October 1976), 542–547.

[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 '98: ACM SIGGRAPH 1998 Papers* (1998), ACM, pp. 189–198.

[DHT*00] DEBEVEC P., HAWKINS T., TCHOU C., DUIKER H.-P., SAROKIN W., SAGAR M.: Acquiring the reflectance field of a human face. *SIGGRAPH '00: ACM SIGGRAPH 2000 Papers* (July 2000), 145–156.

[GGHS03] GOESELE M., GRANIER X., HEIDRICH W., SEIDEL H.-P.: Accurate light source acquisition and rendering. In *SIGGRAPH '03: ACM SIGGRAPH 2003 Papers* (2003), ACM, pp. 621–630.

[GGSC96] GORTLER S. J., GRZESZCZUK R., SZELISKI R., CO-HEN M. F.: The lumigraph. In *SIGGRAPH '96: ACM SIGGRAPH 1996 Papers* (1996), ACM, pp. 43–54.

[HKSS98] HEIDRICH W., KAUTZ J., SLUSALLEK P., SEIDEL H.-P.: Canned lightsources. In *Proceedings of the EG Rendering Workshop 1998* (1998), pp. 293–300.

[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.

[Kaj86]   KAJIYA J. T.: The rendering equation. *SIGGRAPH Comput. Graph. 20*, 4 (1986), 143–150.

[LCV*04] LEVOY M., CHEN B., VAISH V., HOROWITZ M., MCDOWALL I., BOLAS M.: Synthetic aperture confocal imaging. *ACM Trans. Graph. 23*, 3 (2004), 825–834.

[LH96]   LEVOY M., HANRAHAN P.: Light field rendering. In *SIGGRAPH '96: ACM SIGGRAPH 1996 Papers* (1996), ACM, pp. 31–42.

[MH84]   MILLER G. S., HOFFMAN C. R.: Illumination and reflection maps: Simulated objects in simulated and real environments. In *SIGGRAPH 84 Course Notes for Advanced Computer Graphics Animation* (July 1984).

[MPDW03] MASSELUS V., PEERS P., DUTRE; P., WILLEMS Y. D.: Relighting with 4d incident light fields. *ACM Trans. Graph. 22*, 3 (2003), 613–620.

[RWPD06] REINHARD E., WARD G., PATTANAIK S., DEBEVEC P.: *High Dynamic Range Imaging – Acquisition, Display and Image-Based Lighting*. Morgan Kaufmann, San Francisco, CA, 2006.

[Sch97]   SCHROEDER W. J.: A topology modifying progressive decimation algorithm. In *VIS '97: Proceedings of the 8th conference on Visualization '97* (1997), IEEE Computer Society Press, pp. 205–ff.

[SGN06]   SWAMINATHAN R., GROSSBERG M. D., NAYAR S. K.: Non-Single Viewpoint Catadioptric Cameras: Geometry and Analysis. *International Journal of Computer Vision 66*, 3 (Mar 2006), 211–229.

[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 (January-March 1999), 1–12.

[UG07]   UNGER J., GUSTAVSON S.: High-dynamic-range video for photometric measurement of illumination. In *Proceedings of Sensors, Cameras and Systems for Scientific/Industrial Applications X, IS&T/SPIE 19th International Symposium on Electronic Imaging* (Feb 2007), vol. 6501.

[UGY07]   UNGER J., GUSTAVSON S., YNNERMAN A.: Spatially varying image based lighting by light probe sequences: Capture, processing and rendering. *Vis. Comput. 23*, 7 (2007), 453–465.

[UWH*03] UNGER J., WENGER A., HAWKINS T., GARDNER A., DEBEVEC P.: Capturing and rendering with incident light fields. In *Proceedings of the EG Rendering Workshop 2003* (2003), Eurographics Association, pp. 141–149.

[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 Trans. Graph. 24*, 3 (2005), 765–776.