TEMPORALLY AND SPATIALLY VARYING IMAGE BASED LIGHTING USING HDR-VIDEO

Jonas Unger, Joel Kronander, Per Larsson, Stefan Gustavson, Anders Ynnerman

Media and Information Technology, Linköping University, Sweden

ABSTRACT

In this paper we present novel algorithms and data structures for capturing, processing and rendering with real world lighting conditions based on high dynamic range video sequences. Based on the captured HDR video data we show how traditional image based lighting can be extended to include illumination variations in both the temporal as well as the spatial domain. This enables highly realistic renderings where traditional IBL techniques using a single light probe fail to capture important details in the real world lighting environment. To demonstrate the usefulness of our approach, we show examples of both off-line and real-time rendering applications.

***Index Terms*—** HDR video, image based lighting

1. INTRODUCTION

The production of photo-realistic computer graphics renderings, seamlessly merging synthetic objects into partly or entirely virtual environments, is becoming increasingly important in many application areas. Accurately modeling synthetic lighting to match that of the real environment is a difficult and time consuming challenge. This has motivated the development of *image based lighting* (IBL) techniques [1], where the light in real world scenes is captured using *high dynamic range* (HDR) imaging and used as a source of illumination in renderings. The ease of use and high quality results have now made IBL a standard tool in most production workflows.

A fundamental limitation of IBL, however, is that it captures the lighting as a spherical distribution at a single point in the scene at a single instant in time. This means that it cannot capture how light varies from location to location or over time. The reason for this is that IBL is traditionally based on still HDR image input. In this paper, we present methods for capture and rendering with real world lighting exhibiting detailed variations in the temporal or spatial domain including: formulations of IBL taking into account the temporal and spatial domains in a formalized way, an interactive approach for scene modeling and extraction of light sources for spatial IBL, and an overview of a systems pipeline for HDR-video based IBL. No previously described method is capable of capturing and reproducing the angular, spatial and temporal variation in the scene illumination in comparable detail.

2. BACKGROUND

The key idea behind IBL [1] is to capture panoramic HDR images in real scenes and use this information as source of illumination during rendering. This is illustrated in Figure 1a. Each pixel in the panoramic HDR light probe image can be thought of as a a measurement of the radiance incident from a direction $\vec{\omega}_i$ in the real scene. For simplicity we consider only direct illumination here. The appearance of each point $\mathbf{x}$ on the virtual objects in the scene can be described as:

$$L(\mathbf{x},\vec{\omega}_o) = \int_{\Omega} L_{IBL}(\vec{\omega}_i)\rho(\mathbf{x},\vec{\omega}_i,\vec{\omega}_o)V(\mathbf{x},\vec{\omega}_i)\,\langle\vec{n},\vec{\omega}_i\rangle\,d\vec{\omega}_i \quad (1)$$

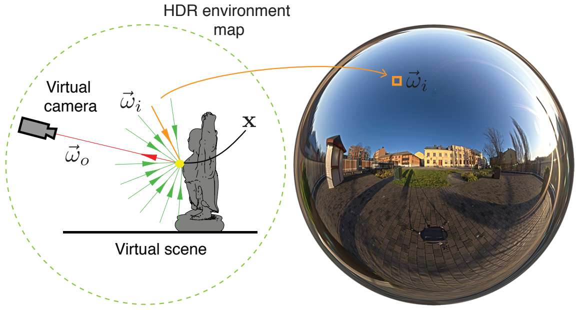
where $\vec{\omega}_o$ is the outgoing direction from the point $\mathbf{x}$ i.e. towards the camera, $L_{IBL}(\vec{\omega}_i)$ is the spherical radiance distribution captured in the light probe over the spherical domain Ω, and V is the visibility function (if the IBL environment is visible in the direction $\vec{\omega}_i$ from a point $\mathbf{x}$, then $V = 1$ otherwise $V = 0$). The BRDF, ρ, describes how radiance incident from a direction $\vec{\omega}_i$ is scattered by the material towards the outgoing direction $\vec{\omega}_o$, and $\langle\vec{n},\vec{\omega}_i\rangle$ is the falloff as a function of the angle between $\vec{\omega}_i$ and the surface normal $\vec{n}$.

In traditional IBL, the captured radiance distribution L_{IBL} is independent of $\mathbf{x}$ and cannot recreate spatial variations in the real scene illumination. The main reason for this is that HDR image capture has been difficult and time consuming. Most HDR capture techniques are based on fusing different exposures captured with e.g. different exposure time [2]. This requires both the scene and the camera to be stationary over the duration of the entire capture. One of the key contributions in this paper is to formalize the transition from using still HDR panoramas to the use of HDR-video input.

The temporal IBL described in Section 4 is, similarly to [3], using a single or possibly a few filtered panoramic HDR frames at each time step. In contrast to previous work, however, we use high quality HDR-video captured in real scenes as input instead of simulated video. Using techniques adopted from pre-computed radiance transfer (PRT) [4], we also extend this to real-time processing and rendering. The spatial IBL described in Section 5 relies on reconstructing a

---

This work was funded by the Swedish Foundation for Strategic Research (SSF) through grant IIS11-0081, Linköping University Center for Industrial Information Technology (CENIIT), COST Action IC1005 on HDR video, and the Swedish Research Council through the Linnaeus Environment CADICS.



(a) Lighting is captured as a 360° HDR panorama and used during rendering

(b) IBL rendering of the buddha model

Fig. 1. **a)** The scene lighting is captured as a panoramic HDR image (right) and used as source of illumination during rendering (left). **b)** IBL produces highly realistic renderings.

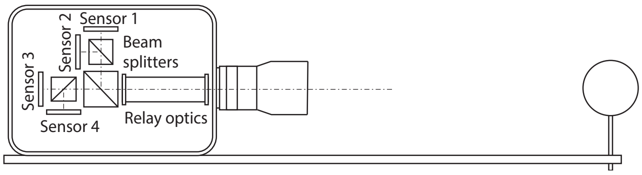
geometric model of the scene onto which the radiance information captured in the HDR-video is reprojected. For this, we combine automatic scene reconstruction methods based on structure from motion [5, 6] with a visualization and interaction scheme inspired by synthetic aperture imaging [7]. This enables full control over the reconstruction process and allows the user to rapidly edit or model parts of the scene. Our approach is inspired by incident light fields [8, 9], but extends these ideas with tools for rapidly building fully general and significantly more detailed representations of the spatial and angular variations in the scene lighting.

3. HDR-VIDEO CAPTURE

HDR-video is a rapidly emerging field and a number of different capture methods have been proposed [10]. Although there are sensors, e.g. with logarithmic response, that can capture the full dynamic range in most scenes, the currently best camera systems in terms of resolution, noise characteristics, and overall image quality are based on setups with multiple, synchronized sensors that simultaneously capture different exposures of the scene [11, 12]. Such high resolution HDR-video is what enables the capture of the lighting information required for temporal and spatial IBL.

The HDR-video sequences used in the experiments described in this paper were captured using two different camera setups, see Figure 2a. The first setup, described in [13, 14], is implemented around the Ranger C55 camera platform from SICK-IVP. The camera micro-code was modified to support sequential HDR-capture with per-pixel minimal time disparity between the exposures. This rolling shutter camera exhibits a dynamic range of 23 f-stops at 25 fps with a resolution of 960×512 pixels . The second setup, described in [12], is a multi-sensor setup with four high quality, global shutter Kodak KAI-04050 CCD sensors with a resolution of 2336×1752 pixels. A common front lens projects the incident light distribution via a custom built relay-lens onto a beam-splitter arrangement that distributes the light onto the four synchronized sensors as illustrated in Figure 2b. The raw sensor data is streamed to a host computer where it is processed on the GPU into final HDR images and written to disk. The 12 bits linear A/D conversion of the sensors in combination with the different ND-filters yield a dynamic range of 24 f-stops. The dynamic range can be extended further by running the sensors with different integration times. This system allows for capture of 4 Mpixel HDR-video at up to 32 fps. To capture HDR panoramas we use a mirror sphere setup, depicted in Figure 2b, with a near 360° image of the environment.

(a) HDR-video cameras



(b) Light probe setup and multi-sensor layout

Fig. 2. **a)** The two HDR-video cameras used in the experiments. **b)** In our multi-sensor setup, the light incident onto the front lens is distributed onto four sensors with different ND-filters. HDR panoramas are captured through the reflection in a mirror sphere.

4. TEMPORAL IMAGE BASED LIGHTING

To extend IBL to include the temporal domain, we capture panoramic HDR-video sequences and switch the lighting environment for each rendered frame. If the HDR-camera moves along a path in the scene, the rendered objects will appear as they are following the same path. Extending equation 1, this can be described as:

$$L(\mathbf{x}, \vec{\omega}_o, t) = \int_{\Omega} L_T(\vec{\omega}_i, t) \rho(\mathbf{x}, \vec{\omega}_i, \vec{\omega}_o) V(\mathbf{x}, \vec{\omega}_i) \langle \vec{n}, \vec{\omega}_i \rangle \, d\vec{\omega}_i \quad (2)$$

where $L_T(\vec{\omega}_i, t)$ describes the angular radiance distribution at each time step t. The temporal sampling at time steps

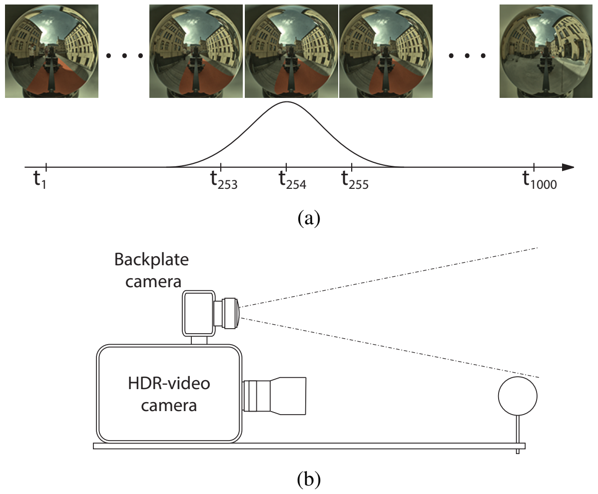


Fig. 3. **a)** An example panoramic HDR-Video sequence. Each rendered frame use and individual lighting measurement. To avoid flickering it is useful to apply temporal filtering over a few frames. **b)** For each frame, our setup synchronously captures both a panoramic lighting measurement and a high resolution backplate.

$t = t_0, t_1, ...t_n$ is discretized corresponding to the frame rate of the HDR-camera as illustrated in Figure 3a. Our capture setup for temporal IBL, shown in Figure 3b, consists of two synchronized cameras: one HDR-video camera capturing the spherical lighting distribution, L_T, and one camera capturing high resolution backplates.

High frequency changes in the time domain, e.g. rapid motion in an out of shadows, may lead to flickering artifacts in the rendered images. This is because we, at each time step, use only a single angular radiance distribution for the entire virtual scene. This is similar to traditional IBL. It is therefore useful to apply filtering to smooth out rapid changes in the temporal domain as illustrated in Figure 3b. In our pipeline, we normally use a gaussian or triangular filter over 3-5 frames. Figure 4 displays three frames from a sequence using temporal IBL rendering. The dynamic HDR environment used to illuminate the virtual helicopter is shown in the upper left corner in each frame. For efficient sampling of the lighting environment during rendering, we use a sequential Monte-Carlo approach [3], taking into account the coherency between frames in the HDR-video sequences.

To enable real-time rendering we use a GPU based algorithm for fast *spherical harmonics* (SH) projections of the captured light probe frames [15]. By pre-computing SH projections of the material and local visibility on the virtual objects we use techniques from pre-computed radiance transfer, see [4], to enable real-time rendering with streaming high resolution (1700 × 1700 pixels) HDR-video input.

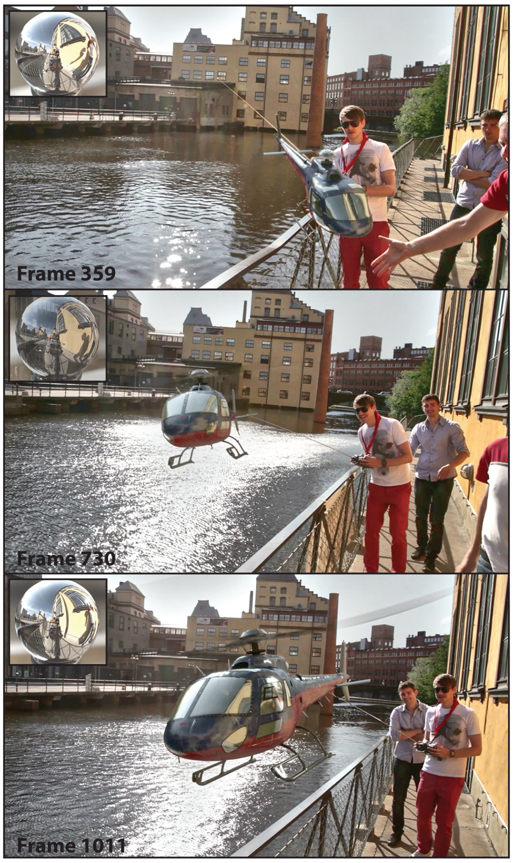
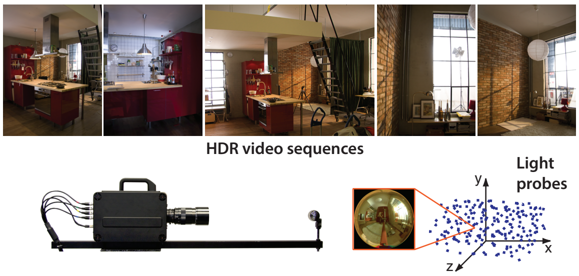


Fig. 4. Three frames from a video sequence of a virtual helicopter rendered using temporal IBL. The lighting measurement used in each frame is shown i the left upper corner.

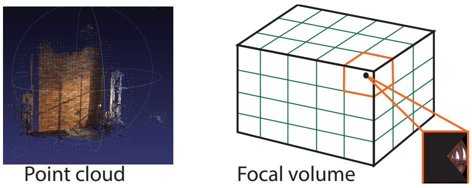
5. SPATIALLY VARYING IMAGE BASED LIGHTING

To extend IBL to include detailed spatial variations in the illumination, we assume that the scene is stationary during the duration of the capture. The goal is to reconstruct a geometric model of the scene onto which the radiance information captured in the panoramic HDR-video sequences can be reprojected and stored as HDR-textures representing both the surfaces in the scene as well as the direct light sources. A scene is, as illustrated in Figure 5a, built from a combination of both panoramic HDR-video sequences captured using the mirror sphere setup and high resolution HDR-video frames captured without the mirror sphere. The position and orientation of each image is estimated using external tracking systems. For tracking we use both a $1.5 \times 1.5 \times 1.5$ m translation stage with an accuracy in the order of 0.1 mm as well as optical tracking based on external cameras and tracking markers.

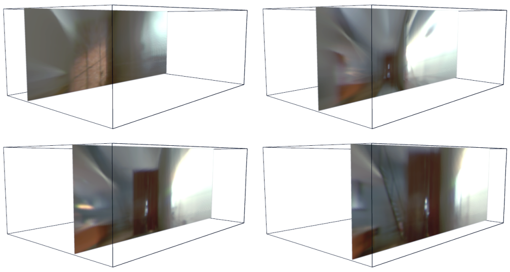
Based on the image data we compute two different data sets, see Figure 5b. The first is a dense point cloud that describes the scene geometry. The point cloud is estimated using state-of-the-art structure from motion techniques [5, 6] based on the "ordinary" HDR-video frames. The second data



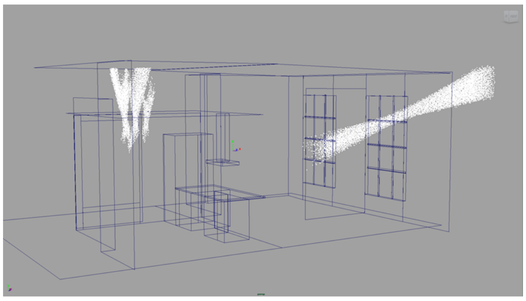
(a) A scene is captured using HDR-video sequences and panoramas



(b) Point cloud and focal volume



(c) Illustration of a focal surface swept through the scene



(d) Interactive light source extraction

Fig. 5. Overview of the different steps in our capture and processing pipeline for spatially varying IBL

set is a volumetric representation of how the scene radiance varies in the scene computed from the panoramic HDR-video sequences. We refer to this data set as a *focal volume*. Each pixel in each HDR panorama corresponds to a radiance sample l captured at a certain (known) position $\mathbf{x}_l$ and direction $\vec{\omega}_l$ in the scene. The focal volume is constructed by tracing each captured radiance sample, l, through the volume from its capture position, $\mathbf{x}_l$, along its direction $\vec{\omega}_l$. At each voxel in the volume we compute the mean value of all radiance samples that intersects it.

A focal volume can be thought of as a large synthetic aperture projection using a 3D aperture of the size and shape of the capture region. This 3D data set allows the user to rapidly improve the estimate from the point cloud, or model parts of the scene by hand. A surface placed inside the volume can be thought of as a *focal surface*. Points on a focal surface that coincide with a surface in the real scene will be in focus, but will rapidly go out of focus with increasing distance to the real surface. Due to the HDR nature of the data, light sources will be represented by high intensity voxels that can be easily thresholded to segment out the direct light sources in the scene. Figures 5c displays examples captured during interactive modeling of an example scene. Due to the large synthetic aperture, the depth of field is very narrow, and objects in the scene cause strong and robust local focus maxima.

As demonstrated by [9], it is important to extract the light sources in the scene so that they can be labeled and sampled efficiently during rendering. Figure 5d shows the final geometry recovered from the scene and how the light sources are robustly revealed as high intensity cones in the volume. This enables well-defined segmentation of the voxels corresponding to each light, and makes it possible to extract its position and spatial extent using spatial selection and thresholding. As the cone converges, the voxel intensities increase to a distinct maximum at the position of the light source. From a user-specified position on the cone and a threshold, the conic region is segmented using either region-growing or manual selection in the modeling software. The orientation of the cone is analyzed using principal component analysis, and a plane orthogonal to the strongest principal component is constructed at the cone apex, covering the extent of the light source as specified by the user.

In a final step, the captured radiance samples, l, are reprojected onto the recovered geometry where they are stored as HDR textures. During rendering of a point $\mathbf{x}$ on an object, the recovered model is used to estimate the incident radiance. This is carried out by tracing rays towards the recovered model of the real scene. The geometric and radiometric model describes both the spatial and angular variations in the lighting, and the rendering integral can be formulated as:

$$L(\mathbf{x}, \vec{\omega}_o) = \int_{\Omega} L_{\mathcal{S}}(\mathbf{x}, \vec{\omega}_i)\ \rho(\mathbf{x}, \vec{\omega}_o, \vec{\omega}_i)\ \langle \vec{n}, \vec{\omega}_i \rangle\ d\vec{\omega}_i \qquad (3)$$

where $L_{\mathcal{S}}(\mathbf{x}, \vec{\omega}_i)$ describes how the radiance distribution varies across the scene, here including both the captured environment and virtual objects. Figure 6a shows a photo-realistic rendering of virtual furniture placed into the recovered scene from Figure 5. As demonstrated in Figure 6b, it is also possible to change the appearance of the scene by e.g. replacing the back wall. The figure also shows a comparison between using our approach and traditional IBL. It is evident that spatially varying IBL enables a higher degree of visual realism.

6. CONCLUSION AND FUTURE WORK

This paper presented a set of algorithms and a framework extending traditional IBL to include the temporal and spatial do-

(a) Photo-realistic rendering using spatial IBL

(b) Comparison of spatial IBL vs. traditional IBL with single light probe

Fig. 6. **a**) Rendering of virtual furniture against a photographic backdrop. The objects are illuminated by the recovered model from the scene displayed in Figure 5a. **b)** Comparison between spatially varying IBL (left) and traditional IBL (right). The spatial variations in the illumination plays a key role in the realism and visual richness.

mains. Our pipeline includes tools for measuring the scene illumination as HDR panoramas, algorithms for processing the data and data structures for efficient scene reconstruction. We have demonstrated how our methods can be applied to both high quality off-line rendering for special effects and product visualizations, as well as for real-time rendering.

Our approach is limited in that it cannot capture both the temporal and spatial domains simultaneously. For capture of spatial lighting variations, we need to fix time by keeping the scene stationary. In future work we will adress the challenge of efficient capture and rendering with real world lighting exhibiting variations in both the spatial and temporal domains.

7. ACKNOWLEDGEMENTS

The authors would like to thank Christian Bloch for rendering of the helicopter sequence displayed in Figure 4.

8. REFERENCES

[1] Paul Debevec, "Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography," in *Proc. SIGGRAPH*, 1998, pp. 189–198.

[2] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec, *High Dynamic Range Imaging – Acquisition, Display and Image-Based Lighting*, Morgan Kaufmann, San Francisco, CA, 2006.

[3] Abhijeet Ghosh, Arnaud Doucet, and Wolfgang Heidrich, "Sequential sampling for dynamic environment map illumination," in *Proc. of EGSR*, 2006.

[4] Ravi Ramamoorthi, "Precomputation-based rendering," *Found. Trends. Comput. Graph. Vis.*, vol. 3, April 2009.

[5] S.N. Sinha, D. Steedly, and R. Szeliski, "Piecewise Planar Stereo for Image-based Rendering," in *ICCV*, 2009.

[6] Yasutaka Furukawa, Brian Curless, Steven M. Seitz, and Richard Szeliski, "Reconstrucing building interiors from images," in *Proc. of ICCV*, 2009.

[7] Vaibhav Vaish, Marc Levoy, Richard Szeliski, C. L. Zitnick, and Sing Bing Kang, "Reconstructing occluded surfaces using synthetic apertures: Stereo, focus and robust measures," in *Proc. CVPR*, 2006.

[8] Jonas Unger, Andreas Wenger, Andrew Gardner, Tim Hawkins, and Paul Debevec, "Capturing and rendering with incident light fields," in *Proc. of EGSR*, 2003.

[9] Jonas Unger, Stefan Gustavson, Per Larsson, and Anders Ynnerman, "Free form incident light fields," *Computer Graphics Forum*, vol. 27, no. 4, 2008.

[10] K. Myszkowski, R. Mantiuk, and G. Krawczyk, *High Dynamic Range Video*, Morgan & Claypool, 2008.

[11] Michael D. Tocci, Chris Kiser, Nora Tocci, and Pradeep Sen, "A Versatile HDR Video Production System," *ACM Trans. on Graph. (TOG)*, vol. 30, no. 4, 2011.

[12] Joel Kronander, Stefan Gustavson, Gerhard Bonnet, and Jonas Unger, "Unified hdr reconstruction from raw cfa data," in *Proc. of the IEEE International Conference on Computational Photography*, April 2013.

[13] Jonas Unger, Stefan Gustavson, Mark Ollila, and Mattias Johannesson, "A real time light probe," in *In Proc. of Eurographics Annual Conference*, Short Papers and Interactive Demos, 2004.

[14] Jonas Unger and Stefan Gustavson, "High-dynamic-range video for photometric measurement of illumination," in *Sensors, Cameras and Systems for Scientific/Industrial Applications X, IS&T/SPIE International Symposium on Electronic Imaging*, vol. 6501, 2007.

[15] Saghi Hajisharif, Joel Kronander, Ehsan Miandji, and Jonas Unger, "Real-time image based lighting with streaming HDR-light probe sequences," in *Proc. of SIGRAD '12*, 2012.