Interactive Relighting of Virtual Objects under Environment Lighting

Nick Michiels, Jeroen Put and Philippe Bekaert
Hasselt University - tU - iMinds, Expertise Centre for Digital Media
Wetenschapspark 2, 3590 Diepenbeek, Belgium
{nick.michiels, jeroen.put, philippe.bekaert}@uhasselt.be

Keywords: rendering; spherical Gaussians; real-time; triple product; gpu

Abstract: Current relighting applications often constrain one or several factors of the rendering equation to keep the rendering speed real-time. For example, visibility is often precalculated and animations are not allowed, changes in lighting are limited to simple rotation or the lighting is not very detailed. Other techniques compromise on quality and often coarsely tabulate BRDF functions. In order to solve these problems, some techniques have started to use spherical radial basis functions. However, solving the triple product integral does not guarantee interactivity. In order to dynamically change lighting conditions or alter scene geometry and materials, these three factors need to be converted to the SRBF representation in a fast manner. This paper presents a method to perform the SRBF data construction and rendering in real-time. To support dynamic high-frequency lighting, a multiscale residual transformation algorithm is applied. Area lights are detected through a peak detection algorithm. By using voxel cone tracing and a subsampling scheme, animated geometry casts soft shadows dynamically. We demonstrate the effectiveness of our method with a real-time application. Users can shine with multiple light sources onto a camera and the animated virtual scene is relit accordingly.

1 INTRODUCTION

Interaction between the real environment and virtual scenes is crucial in real-time relighting applications. Realistic integration of computer generated objects in real environments is used in movie post-production and in augmented reality applications. Fully interactive applications allow immediate changes of three important factors. Firstly, they allow changes in the lighting in the scene. These changes can affect both distant and near-field changes in illumination. Second, the geometry of the objects in the scene can change with animations or with user input. Lastly, the materials that constitute the objects can change. However, the evaluation of the required rendering equation in real-time has so far remained impossible without constraining at least one of these three factors. Besides, it is also important that the representable content can contain high-frequency details. There are several problems to solve before reaching real-time framerates. Our approach combines the advantages of previous approaches and allows real-time relighting of dynamic virtual scenes with real-time captured environment lighting. The data is represented using spherical radial basis functions (SRBFs). This paper will explain how to construct such a representation efficiently. Each factor needs to be dynamic and requires a fast transformation to the SRBF representation. For environment lighting, we show how a multiscale residual algorithm can be used for SRBF fitting. The visibility needs to be sampled in real-time and we show how voxel cone tracing together with an efficient sampling scheme allow for accurate soft
Table 1: Comparison to other techniques that use precomputed radiance transfer for triple product rendering. All compared techniques use spherical harmonics wavelets or spherical Gaussians to solve the rendering equation. However, they all constrain at least one of the factors. Our approach allow for a real-time update for as well lighting, materials and visibility.

<table>
<thead>
<tr>
<th>Year</th>
<th>All-Frequency</th>
<th>Dynamic Lighting</th>
<th>Dynamic Geometry</th>
<th>Dynamic Reflectance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sloan et al., 2003</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Ng et al., 2004</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Tsai and Shih, 2006</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Haber et al., 2009</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Wang et al., 2009</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Lam et al., 2010</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Meunier et al., 2010</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Iwasaki et al., 2012</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 2: Related Work

Our technique is related to Precomputed Radiance Transfer techniques (PRT) (Sloan et al., 2002; Ng et al., 2003; Sun et al., 2007; Imber et al., 2014), in that we factorize rendering and represent the estimated data in efficient bases.

Ng et al. (Ng et al., 2004) represent the visibility term, the environment map and the BRDF slice that corresponds to the viewing direction $\omega_o$ in the Haar wavelet basis $\Psi$. This basis is defined over the hemisphere using the hemi-octahedral parameterization as introduced by Praun and Hoppe (Praun and Hoppe, 2003). Later this approach is extended to work with high-order wavelets (Michiels et al., 2014). Wavelet representations are good for compression and can be used efficiently in the triple product integration of PRT rendering. However, the factors themselves cannot be transformed easily to the wavelet representation and require offline processing.

Haber et al. (Haber et al., 2009) use the aforementioned approach to efficiently compress and render with the three factors. However, none of the three factors are dynamic. For example, all BRDF slices are first sampled in pixel domain and are then transformed to wavelets and the visibility is preprocessed with ray tracing.

A new basis to represent the three factors of the rendering are spherical radial basis functions (SRBFs). The advantage of SRBFs is that they have an efficient rotation operator. Tsai et al. (Tsai and Shih, 2006) proposed a method for rendering with all-frequency lighting using SRBFs. The lighting Gaussians are fitted using an optimization process which results in high compression ratios, but can take up to several minutes to fit the full lighting factor. Furthermore, the visibility is precomputed and compressed using Clustered Tensor Approximation (CTA) and is therefore limited to static scenes only.

Wang et al. (Wang et al., 2009) also presented rendering with SRBFs for all-frequency lighting. They lack the possibility to render dynamic scenes since the visibility is precalculated using a spherical signed distance function (SSDF). Moreover, SSDFs are not able to accurately render detailed shadows. To represent the lighting factor, they use the same elaborate fitting process as Tsai et al. (Tsai and Shih, 2006).

Iwasaki et al. (Iwasaki et al., 2012) developed a real-time renderer for dynamic scenes which is able to render under all-frequency lighting using spherical Gaussians. The lighting can be rotated, but cannot be updated dynamically. The visibility is dynamically estimated by projecting bounding volumes onto a grid of patches of the hemisphere. Occluded patches are represented with spherical Gaussians. However, updating the bounding volume tree is strongly dependent on the complexity of the scene. Besides, their tree traversal uses suboptimal branching operation on shadows.
GPUs.

Our method also uses spherical radial basis functions to represent all three factors. In contrast to previous techniques, we are able to construct and update all three factors in real-time. We show how a residual transformation technique can be used to efficiently transform the lighting information to spherical Gaussians. Cone tracing together with peak-detection of powerful and high-frequency lights allow for interactive rendering of soft shadows. Table 1 gives an overview of all the discussed techniques. Note that only our technique allows real-time adjustment of the lighting. Other techniques support only static lighting or rotation of the environment map. Furthermore, the performance of our real-time visibility tracing is not dependent on the amount of polygons.

3 TRIPLE PRODUCT INTEGRATION

A key part in our framework is the simulation of light propagation in a scene based on the reflectance equation over the hemisphere $\Omega$ (Kajiya, 1986; Ng et al., 2004):

$$B(v, \omega_0) = \int_{\Omega} p(v, \omega_0, \omega_i) V(v, \omega_0) L(v, \omega_i) d\omega_0$$

where $p$ is the BRDF, incorporating the cosine term for the incident illumination from $\omega_0$, at position $v$ and the outgoing direction $\omega_i$, $V$ is the hemispherical visibility term evaluated around $v$ and $L$ is the incident illumination rotated in the local frame of $v$. The components affect the appearance of the rendered result. Now the rendering equation can be evaluated as a triple-product integral at each point in the scene:

$$B(v, \omega_0) = \sum_{k} \sum_{l} \sum_{m} C_{klm} p_k V_l \tilde{L}_m$$

where $C_{klm} = \int_{\Omega} \Psi_k \Psi_l \Psi_m d\omega_0$ are called the binding or tripling coefficients of the three bases. The efficiency of this calculation is dependent on the choice of basis to represent the factors. Calculations in the pixel domain will be inefficient, as pixels cannot sparsely represent the information of the factors. Other techniques using spherical harmonics (Sloan et al., 2002), Haar wavelets or eigenbases are often able to more compactly represent the content. These bases, however, place their own restrictions on the data or the operations that can be performed on it. For example, spherical harmonics are not able to sparsely represent high-frequency signals. On the other hand, wavelets are very good at compression but lack flexibility. Wavelets are constrained to static scenes, since the visibility needs to be precomputed. Moreover, rotation is a complex operation in the wavelet domain (Wang et al., 2006). We choose to represent all three factors in the triple product integral with hierarchical spherical Gaussians.

A spherical Gaussian $G$ is defined as

$$G(\nu \cdot p, \lambda, \mu) = \Psi(\nu, \lambda, \mu)$$

where $\nu \cdot p$ is the dot product between a vector $\nu$ on the sphere and the center of the Gaussian, $\lambda$ the Gaussian bandwidth and $\mu$ the amplitude. Spherical Gaussians have several advantages:

- A hierarchy of spherical Gaussians allows the efficient representation of both low- and high-frequency information.
- Spherical Gaussians are rotationally symmetric and therefore allow the efficient rotation of the environment map to the local vertex coordinate system by rotating the center of the Gaussian.
- BRDF functions can be well approximated with spherical Gaussians, allowing run-time evaluation of reflectance.
- The mixing of Gaussians is additive and allows for a less complex implementation, compared to spherical harmonics or wavelet theory.

Additionally, Gaussians have an extra interesting characteristic. The product and convolution of Gaus-
Gaussian lobes can be solved analytically (Tsai and Shih, 2006):
\[
\int_{S^2} G_1(v \cdot p_1, \lambda_1, \mu_1) G_2(v \cdot p_2, \lambda_2, \mu_2) G_3(v \cdot p_3, \lambda_3, \mu_3) \, dv = \\
\frac{4\pi \lambda_1 \lambda_2 \lambda_3 \sinh(\|r\|)}{c^{\lambda_1 + \lambda_2 + \lambda_3} \|r\|}
\]  
(4)

where \( r = \|\lambda_1 p_1 + \lambda_2 p_2 + \lambda_3 p_3\| \). The binding coefficients out of Equation 4 are now reduced to a simple formula. However, solving the triple product integral efficiently is not enough for real-time relighting applications. In order to interactively change lighting conditions or alter scene geometry and materials, these three factors need to be converted to the SRBF representation in a fast manner. In the next sections, we will demonstrate how we can make the visibility, materials and lighting factors dynamic by using spherical radial basis functions.

4 DYNAMIC MATERIALS

There are a few characteristics a good material representation must have:
1. The representation should be compact.
2. Easy rotation of the BRDF into the global lighting and visibility frame.

SRBF functions can approximate almost any BRDF. They are rotationally symmetric and easily rotatable. They also allow for a compact all-frequency representation, due to their multiscale nature. It is known that BRDF functions expressed in terms of a normal distribution function can be approximated with Gaussian lobes (Ngan et al., 2005; Wang et al., 2009). These Gaussians are then integrated in the triple product rendering integral.

5 DYNAMIC VISIBILITY

To evaluate the complex visibility term in the rendering equation in real-time, a fast and dynamic approximation is required. Previous techniques have relied on precomputation, but such a solution is no longer feasible when changes in the geometry occur frequently. In our framework, we combine PRT techniques with voxel cone tracing (Crassin et al., 2011), to evaluate the visibility in real-time entirely on the GPU.

The first step in our processing pipeline is to voxelize the geometry to a 3D texture. This is done by implementing a one-pass voxelization algorithm (Crassin and Green, 2012). We use a geometry shader to project all polygon fragments to their dominant axis. Voxelization is a real-time process and only needs to be partially updated every frame for dynamic geometry. For this we use a dirty flag mechanism that leaves the static geometry untouched.

The second step is to perform cone tracing in the voxel volume. Each visible pixel corresponds to a 3D position in voxel space. From this position, new cones can be constructed and traced throughout the voxels. Current voxel cone tracing is often used to perform ambient occlusion or low frequency global illumination effects. However in our application, we shoot cones to evaluate the visibility factor of the triple product integral. For this to work, a mapping from cones to spherical Gaussians is essential. Suppose we have a Gaussian with size \( \lambda \). Let us fit a cone that coincides with the distance \( \alpha \) where the Gaussian evaluates 0.5:
\[
e^{\lambda \cos(\alpha) - 1} = 0.5
\]  
(5)

Solving for \( \alpha \):
\[
\alpha = \cos^{-1}\left(\frac{\log(0.5)}{\lambda} + 1.0\right)
\]  
(6)

where \( \alpha \) is used as the solid angle for the cone to trace.

The last step is to select the appropriate cones to trace. A straightforward method would trace one visibility cone for each BRDF lobe. However, important high-frequency detail might be missed. For BRDF lobes with a large solid angle, all visibility details are integrated over a large area of the hemisphere. A better approach is to subsample the BRDF lobe.
Figure 4: Adaptive subsampling of visibility cones. Left: Gaussian of a Lambertian BRDF lobe (orange) approximated with a number of visibility cones (green). Middle: Gaussian of a glossy BRDF lobe (orange) approximated with the same number of visibility cones (green), rotated and scaled to enclose the BRDF. Right: for a Gaussian of a specular BRDF lobe (orange) it suffices to only trace a single visibility cone (green) that coincides the lobe.

Figure 6: Peak detection. The input HDR environment map is thresholded to keep only the high-intensity values. These peaks are then detected using connected components. Each connected component can now be represented with a number of Gaussians. These Gaussians are used to accurately render soft shadow for high-frequency lighting.

with smaller visibility cones. The subsampling density depends on the characteristics of the BRDF. A circle packing algorithm is used to determine the center position and radius. Figure 4 shows our subsampling scheme. For diffuse Lambertian BRDF lobes, a small number of uniformly traced visibility cones suffice. For more glossy BRDF lobes, the cones are scaled and rotated into the lobe. The number of cones used for subsampling determines the desired level of quality. Because the size of a BRDF lobe dictates the recoverable frequencies (Ramamoorthi and Hanrahan, 2001), our subdivision of scaled cones within the BRDF lobe retains approximately the same level of quality for all lobe sizes. In our implementation we obtained good quality renderings using a subdivision of 7 cones.

6 DYNAMIC LIGHTING

Real-time changes to the lighting environment can drastically alter the look of three-dimensional models. To capture the entire distant lighting sphere, we use HDR omnidirectional video frames extracted from the Ladybug3 (PointGrey, 2014) camera. To convert the lighting to an RBF representation, we perform an RBF fitting algorithm based on Ferrari et al. (Ferrari et al., 2004). We apply the algorithm on the sphere using the Healpix (Gorski et al., 2005) distribution, utilizing a multiscale algorithm. The algorithm starts from the environment map as input to the first scale. The input of each scale is approximated with a fixed grid of SRBFs. The difference of the approximated reconstruction and the input data is the residue and is passed as input to the next scale. This process continues until the residual error is below a certain threshold or the maximum level of detail is reached. The described algorithm is fast and allows for a real-time transformation. Unlike other methods, no complex optimization is used (Tsai and Shih, 2006; Lam et al., 2010). An example of a multiscale approximation is shown in Figure 5.

The subsampling scheme described in Section 5 performs well for large area light sources and low-frequency lighting. For these cases, the visibility cones can stay relatively large. In the case of very bright area lights, subsampling of visibility cones would quickly become too fine-grained to be tractable. We treat these lights as a special case. Their areas are easily identified with a peak detection algorithm. First, the incoming environment map lighting is thresholded, so that only the peaks with the most influence remain. Then, the connected component algorithm is used to group pixels into the lights they constitute. Lastly, a number of Gaussians (1-3) are fitted to match the shape of each connected compo-
Figure 5: Multiscale residual transformation of environment map into SRBF representation from a coarse scale (left) to more refined scales (right). Each level adds extra detail to the reconstruction.

Figure 6 shows a few examples of our peak detection algorithm applied on different environment maps.

7 RESULTS

We have designed a custom application to demonstrate the relighting capabilities of our technique. The application was tested on an Intel Core i7 3.4 GHz processor with 8 GB of RAM and a GeForce 770 GTX. The triple product calculation for material, visibility and lighting is done entirely with GLSL shaders on the GPU. The application can be used to place virtual models under real-world lighting conditions. Our peak detection process of Section 6 is performed on the distant illumination environment map in real-time. Afterwards, the model can be viewed under these lighting conditions and changes can be made. Unique to our approach is that all factors can be changed simultaneously and in real-time. The lighting can be altered, the model can be moved or deformed and the materials can be edited. If there are spatially-varying material parameters in the scene, they can be encoded in textures defined over the virtual model.

Figure 7 shows results with animated geometry. This geometry causes dynamic changes in the visibility factor. Our algorithm is able to cope with this in real-time by using voxel cone tracing. The sampling scheme of visibility cones, as explained in Section 5, is implemented to quickly get an integrated visibility value over a certain solid angle, defined by the BRDF properties. Extra importance sampling of visibility cones is done by finding peaks in the lighting and is a key feature to efficiently render accurate high-frequency area light sources. Figure 7 clearly shows that a change in geometry caused different visibility and that our system is able to render with soft shadows. Figure 8 illustrates the importance of peak detection for shadows. The same environment map is used to render both images, but the right rendering is done with peak detection and the left with only the subsampling scheme for visibility.

To demonstrate dynamic visibility, we placed an omnidirectional camera, such as the Ladybug3 (PointGrey, 2014) of Point Grey, near the desired position of the virtual object in the scene. The residual transformation algorithm is implemented to extract the spherical Gaussians to render with. Now, a virtual model can be relit with one or more flashlights, held before the omnidirectional camera. The results are given in Figure 9. Note the high-quality soft shadows resulting from our peak detection algorithm.

To measure the influence of scene complexity on performance, the scene was rendered with models of varying complexity and lighting conditions with a different amount of peaks. An overview of these exper-
Figure 7: Dynamic visibility\(^1\). Geometry can be changed using animations or user input. Adaptive cone tracing in the voxel volume based on the subsampling scheme and peak detection allows for real time rendering of soft shadows.

Figure 9: Dynamic lighting\(^1\). The environment lighting is captured using the Ladybug3 (PointGrey, 2014) omnidirectional camera. Our peak detection algorithm is performed in real-time and passed to the renderer. The virtual object is now rendered with the lighting conditions from the captured frames. Note how the virtual shadow positions are aligned with the real light source.

\(^1\)Animations are courtesy of Mixamo. https://www.mixamo.com/motions

\(^2\)It can be seen that geometric complexity has only a small influence on the framerate. Environmental complexity has a more noticeable influence, but scales well with the number of peaks.
Acknowledgements

This work has been made possible by the Fund for Scientific Research-Flanders (FWO-V Remixing reality) and with the help of a PhD specialization bursary from the IWT. The authors acknowledge financial support from the European Commission (FP7 IP SCENE).

References


