Using Near-Field Light Sources to Separate Illumination from BRDF


Handle: http://hdl.handle.net/1942/19654
Using Near-Field Light Sources to Separate Illumination from BRDF.

Jeroen Put  
jeroen.put@uhasselt.be  
Nick Michiels  
nick.michiels@uhasselt.be  
Philippe Bekaert  
philippe.bekaert@uhasselt.be

Hasselt University - tUL - iMinds  
Expertise Centre for Digital Media  
Wetenschapspark 2  
3590 Diepenbeek, Belgium

Abstract

Simultaneous estimation of lighting and BRDF from multi-view images is an interesting problem in computer vision. It allows for exciting applications, such as flexible relighting in post-production, without recapturing the scene. Unfortunately, the estimation problem is made difficult because lighting and BRDF have closely entangled effects in the input images. This paper presents an algorithm to support both the estimation of distant and near-field illumination. Previous techniques are limited to distant lighting. We contribute by proposing an additional factorization of the lighting, while keeping the rendering efficient and additional data compactly stored in the wavelet domain. We reduce complexity by clustering the scene geometry into a few groups of important emitters and calculate the emitting powers by alternately solving for illumination and reflectance. We demonstrate our work on a synthetic and real datasets and show that a clean separation of distant and near-field illumination leads to a more accurate estimation and separation of lighting and BRDF.

1 Introduction

This paper treats the problem of extracting illumination and reflectance information from existing input images. The estimation is, in a sense, the opposite operation of photorealistic image rendering in the field of computer graphics. Performing this inverse operation is made difficult by problems such as noise in the input images, limited resolution, errors in the geometry information and illumination pattern frequencies that are confounded by the reflection behavior of materials. The estimation of lighting and BRDF (Bidirectional Reflectance Distribution Function) information of scenes has a number of important applications. One example is movie industry post-production. The ability to alter scenes after they have been filmed has the potential to greatly reduce the number of costly recapturing iterations.

Our goal in this paper is to simultaneously estimate illumination and reflectance information from a sparse set of multi-view input images. Current state-of-the-art techniques are restricted to distant lighting, i.e. light sources that are not positioned within the scene but are assumed to be at infinity. The problem with these techniques is that local effects, such as lighting parallax, are not taken into account. This forces near-field lighting effects to be
emulated by wrong terms in the distant lighting environment map, or with wrong reflectance estimates. This should not be a problem for reconstruction, as most techniques will faithfully reproduce the appearance of the scene in its original setting by offsetting wrong estimates in the lighting with wrong terms in the BRDF and vice versa. But afterwards, when the lighting or material properties are changed, these errors are revealed. Current techniques are thus reasonable for reconstruction, but are limited in their relighting flexibility. Figure 1 demonstrates the common problem, where erroneous correction terms in the BRDF prevent convincing relighting. In Figure 1(a), a scene is depicted with 2 objects and a red local light source $L$. The technique proposed in this paper estimates the materials and lighting in this situation correctly. Figure 1(b) demonstrates the manner in which current techniques model local lighting effects: with distant light sources at infinity on the environment map $\Omega$. These light sources affect all objects, so in order to prevent object 1 from being red, correction colors (shown as green) are introduced in the BRDF to reconstruct the gray ground truth color. Figure 1(c) shows that when relighting is performed, with a blue light source in a different location, the wrongly estimated materials (green) are revealed.

1.1 Contributions and Overview

This paper introduces an efficient method for modeling near-field lighting effects in a pipeline for simultaneous lighting and BRDF estimation, with little additional algorithmic complexity. We can summarize our contributions in a few key points:

1. A factorization of illumination in distant and near-field lighting. This allows us to model local lighting effects, many of which are ignored by current methods.

2. A clean separation of distant and near-field lighting allows more accurate estimation of illumination and materials in a scene and permits the manipulation of local light sources. It also allows using knowledge of the light sources in a scene. For example, light source positions can be identified by a HDR clustering procedure and help improve the estimation of other factors.

3. We formulate the problem so that the wavelet coefficients of the distant lighting environment map and the emitting powers of the near-field lighting can be solved together.
Any standard quadratic programming framework can be used to solve this problem robustly and efficiently.

4. We propose the use of clustering to group the most significant emitting triangles into dominant light sources. This grouping constrains the solver to more plausible solutions. We make the assumption that the distribution of light is roughly uniform along the emitting surface.

The input of our algorithm is a sparse set of (usually 5-16) different views of a scene with unknown reflectance, captured under unknown illumination. This makes our method very broadly applicable. Scene geometry is acquired using stereo methods [5] or other techniques such as structured light based on gray codes [18], or Kinect Fusion [8]. Although no currently available technique gives perfect geometry, this is not the main focus of this paper and a good approximation is already sufficient for decent results. The outputs of our algorithm are the estimated wavelet coefficients of the distant lighting environment map, the estimated emitting powers for every emitting cluster in the scene and a set of per-vertex material weights that determine a mixture of BRDF models from the MERL database [16]. Figure 2 gives a summary of our algorithmic pipeline.

![Figure 2: Schematic overview of our method. Best viewed digitally.](image)

2 Related Work

Early attempts at illumination estimation used a light probe to capture lighting and add synthetic objects in a scene [2, 13]. Schoeneman et al. [20] proposed a method for identifying intensities of a set of fixed light sources in a synthetic scene with known geometry and reflectance properties. An alternative approach is estimating the directional distribution of the incident light from a single photograph [11]. Here, a linear least squares system is used to estimate lighting effects and to perform relighting. Yu et al. [26] constructed an inverse global illumination model for recovering reflectance models of real scenes from photographs. Sato et al. [19] recover the illumination distribution of a scene from image brightness, inside shadows cast by an object of known shape in the scene. Finally, Mei et al. [12] have developed a representation for environment lighting that combines low-frequency spherical harmonics with sparse directional light sources. The factorization does not take any local emitters into account and is therefore limited to distant lighting.

The methods described above focus mainly on recovering either the illumination or the reflectance factor. We cannot, however, assume knowledge of either in the scene, so it is essential that a simultaneous estimation of BRDFs together with the illumination is made. Ramamoorthi et al. [17] provided a theoretical framework based on factorization. It predicts
under which conditions good results can be achieved and studies the well-posedness of the combined estimation problem. More recent work focuses on simultaneous estimation of these factors [10, 25].

Haber et al. [6] are able to recover high frequencies, building further on the development of efficient triple product wavelet integrals by Ng et al. [15]. Using wavelets, it is possible to represent BRDFs and illumination data with a small amount of coefficients. Previous techniques had relied on spherical harmonics [4], which lack the ability to store high frequency data without exponentially increasing the number of coefficients.

Our technique is related to Precomputed Radiance Transfer techniques (PRT) [7, 14, 21, 23], in that we factorize rendering and represent the estimated data in efficient bases. Also related is the recovery of intrinsic images [4], however instead of trying to separate lighting and BRDF we also try to recover these factors from multi-view images.

Some attempts at modeling near-field lighting effects in PRT methods have been made [22]. They are, however, restricted to planar light sources and to the scaling and translation of these lights. These techniques are also limited to rendering, but do not attempt to estimate lighting or materials. In contrast, our method supports the estimation and rendering of a broad range of light sources and supports scaling, translation and rotation. Lastly, the work of Yu et al. [27] consider the estimation of indirect lighting. However, their scenes are recorded under calibrated illumination and specular parameters are assumed to be constant across surfaces. Their technique was not designed to work in unconstrained environments.

3 Estimation Process

A key part in our estimation process is the simulation of light propagation in a scene based on the reflectance equation over the hemisphere $\Omega$. We follow the approach of Ng et al. [15], whereby the precalculated visibility term $V$, the environment map $\tilde{L}$ and the BRDF slice $\rho$ that corresponds to the viewing direction $\omega_o$ are expressed in the Haar wavelet basis $\Psi$. The rendering equation can be evaluated as a wavelet triple-product integral at each point in the scene:

$$B(v, \omega_o) = \sum_k \sum_l \sum_mC_{klm}\rho_k V_l \tilde{L}_m$$

where $C_{klm} = \int_\Omega \Psi_k \Psi_l \Psi_m d\omega_i$ are called the binding or tripling coefficients of the three Haar wavelet bases. An important observation to be made here is that current state-of-the-art algorithms fail to model near-field lighting, because they assume rotated copies of an identical $\tilde{L}$ for every $v$ in Equation 1. An obvious solution would be to let every vertex of the scene have a different environment map. This, however, causes a sharp increase in coefficients to be estimated. It is also clear that such a system would be massively underconstrained.

In this paper we describe how we can include near-field lighting in our estimation pipeline with similar complexity compared to previous methods. In addition, we avoid adding too many unknowns to the system, i.e. underconstraining the system further. To solve this problem we propose factorizing illumination in distant and near-field lighting.

3.1 Factorization

We propose factorizing illumination in distant and near-field lighting components. The distant illumination is expressed in wavelet coefficients [15], which are estimated directly. Then, we further factorize the near-field lighting term at every vertex. Figure 3 gives a
Figure 3: Distant lighting and near-field lighting are split in global and local components. The global distant environment map is rotated in the local frame of every vertex. This distant lighting is added to the sum of irradiance maps of every cluster, multiplied with the global estimated emitting power of that cluster.

schematic overview of our factorization. The first factor consists of precomputed irradiance maps. The irradiance maps are defined at every vertex in the scene and are oriented around the vertex normal. They contain the influences of every visible emitting cluster on the hemisphere around the vertex. In addition to the factorization proposed by Mei et al. [12], we introduce near-field lighting originating from emitting geometry as well. Similar to BRDF, lighting and visibility, the irradiance maps are stored compactly in the wavelet domain. The second factor of near-field lighting consists of global per-cluster emitting powers (clustering will be addressed in Section 3.3). These powers are the unknowns of the near-field lighting and are solved during the lighting estimation stage. The hemispherical irradiance maps at each vertex are precalculated using raytracing. This process is analogous to the computation of the visibility term, but instead of simply marking areas as occluded we also record the color of the intersected cluster, modulated by an attenuation and cosine falloff factor. The construction of the irradiance maps is demonstrated in Figure 4. The added storage cost for this information is very reasonable in practice, as all the data is projected into the compact Haar wavelet basis. In order to keep the storage requirements and the number of unknowns low, clustering is used to group the most important emitting triangles. Rendering with the factorized representation is equivalent to adding an extra double product integral of the BRDF and the irradiance maps to Equation 1. Because Haar wavelets are an orthogonal basis, this double product integral term reduces to a simple dot product calculation [15]:

\[
B(v, \omega_0) = \sum_k \sum_l \sum_m C_{klm} \rho_k V_l \tilde{L}_m + \sum_c \sum_k \rho_k \cdot (I_c P_c)
\]  

(2)

For every cluster \(c\), visible over the hemisphere of vertex \(v\), its irradiance map \(I_c\) is multiplied by the estimated power \(P_c\). Finally, it is multiplied by the BRDF \(\rho\) at \(v\).

### 3.2 Optimization Details

By combining the visibility map with the tripling coefficients \(T_{v,lm} = \sum_l C_{klm} V_l\) and using matrix notation, we can formulate the product integrals (Equation 2) as an equation for each
Figure 4: For every vertex $v_1 \ldots v_n$ in the scene, $c$ irradiance maps are calculated which contain the influences of individual clusters of emitting triangles. Where $n$ is the total number of vertices in the scene and $c$ is the total number of clusters. (a) Irradiance map with the influence of cluster 1 impending on vertex $v$. (b) Irradiance map for cluster 2 on vertex $v$.

sample $y_i \in Y$ of the scene, with $i \in \{1 \ldots \#\text{pixels}\}$:

$$b_i = \rho_v^T(y_i)T_v(y_i)\tilde{L}_v(y_i) + \sum_c \rho_v(y_i)I_v,c(y_i)P_c$$

(3)

This format is used to build the bilinear system necessary to solve for both lighting and BRDF parameters. For every observed point $y_i$ in the input images, a least squares bilinear equation $E_{y_i}$ between that data sample and the rendered sample $b_i$ with the current lighting and material estimates is added to the system:

$$E_{y_i} = (y_i - b_i)^2$$

(4)

The BRDF $\rho$, the distant lighting $\tilde{L}$ and the emitting powers $P$ that minimize this objective function need to be found. The full bilinear system can be rewritten for input into any quadratic programming solver:

$$E = x^TQx - 2x^Tc + Y^TY$$

(5)

where $Q = B^TB$ and $c = B^TY$. Optimizing for lighting is the equivalent of keeping the BRDF fixed and the lighting properties $x \in \mathbb{R}^{(#\text{coeffs} + \#\text{clusters}) \times 1}$ unknown. To avoid alternating between distant and near-field illumination we estimate distant lighting coefficients and emitting powers all at once. We achieve this by combining the influences $D$ and $N$ in one large matrix $B$ (see Figure 5):

$$B = [D \mid N_1 \ldots N_{\#\text{clusters}}]$$

(6)

with $D$ the influences of distant light wavelet coefficients on each pixel of the rendered image and $N_c (c \in \{1 \ldots \#\text{clusters}\})$ the influences of all emitting clusters.

Every iteration in the bilinear optimization process solves for either illumination or BRDF. For BRDF optimization, we follow the same procedure as Weistroffer et al. [24]. It should be noted that such a quadratic programming problem is convex and leads to an exact optimum [1]. However, because the optimization process cannot solve for both factors at once, the solution might still end up in a local mimimum. This was a significant problem for previous techniques, because the calculated estimation was final and could not be altered. In our framework, we can incorporate extra knowledge of the near-field scene illumination to disambiguate the lighting situation. This significantly increases the accuracy of the estimation.
3.3 Clustering

The optimization, as described in the previous section, is too underconstrained and computationally intensive if every triangle of the scene is considered as a potential emitter. This would slow down rendering and make the combined matrix $B$ from Equation 6 unwieldy, since it would introduce extra columns per triangle in the scene. For complex scenes, this matrix might be too large to keep in main memory. We conclude that there is a need for some kind of clustering of triangles into a few groups of important emitting primitives. There are many ways to perform this clustering operation. Using Debevec’s [3] method for calculating the radiance map of an image, we have a good guess of the radiance at each pixel. We use connected components to group high-radiance pixels together and are able to identify the light sources in the scene. The geometry of the scene can then be clustered into groups of $N$ emitting primitives, with $N$ the number of large connected components. Clustering based on radiance maps requires the light sources to be visible in at least one input image. Areas with high radiance but without intersecting geometry are assigned to distant lighting.

4 Results on Synthetic Data

Our renderer is used to create several viewpoints of a scene with two objects: a red Lambertian plane and a densely tessellated metallic sphere with a Lafortune BRDF. The two objects
are lit by a photographically captured environment map around the scene. For this synthetic data set, the exact geometry is given to the algorithm as input. This eliminates errors in the estimation of lighting and materials due to geometric reconstruction errors. The input images are used to extract color information per viewpoint at every vertex. This information is interpolated over the pixels of each triangle. All the input data are then fed into a bilinear system, as described in previous sections. Next, the system is solved and the outputs are the combination of weights of basis BRDFs from the MERL database and an estimation of the environment illumination. The estimated environment map is shown in Figure 6. We have chosen to compare with the technique of Haber et al. [6] as it also represents data in the wavelet domain and therefore provides a fair comparison. Our estimated environment map is closer to the ground truth, both numerically and visually. The figure also shows that the reconstruction of the scene with our technique is closer to the ground truth. We now rerender the scene, but with the near-field plane repositioned. Our method successfully incorporates the plane at the new position in the reflection on the sphere. However, the method of Haber et al. [6] assigned the red reflection to the distant environment map. As a result, when the plane is moved, their method fails to model the near-field scene transformation. The last part of the figure compares the rendered results when the distant environment map is replaced with another one. As explained earlier on, the technique of Haber et al. [6] assigns erroneous correction colors to the BRDF, which are revealed as severe artifacts when the illumination is changed. With our solution, these artifacts are strongly mitigated. There are, however, some residual correction colors that remain visible at certain edges in the images. These artifacts arise from discretization of the per-vertex irradiance maps.

5 Results on Captured Data

Since existing techniques have not published a dataset with near-field light sources, we have created our own for comparison purposes. We intend to share our datasets with the community (URL omitted for reviewing).

With the statuette dataset we simulate conditions similar to the synthetic dataset. The white statuette is captured under a red light source. The goal is to estimate the material as accurate as possible. Figure 7(a) shows one of the input images. The reconstructed version of the scene is depicted in Figure 7(b). This reconstruction is achieved by optimizing the ren-

![Figure 7: A white statuette under a red light source dataset. (a) One of the input images. (b) Reconstruction after optimization. (c) The technique of Haber et al. [6] reveals compensation colors in materials when rendered under white lighting. (d) Material estimated by our technique rendered under white lighting. Best viewed digitally.](image-url)
dering factors so that the pixels of the rendered image closely match the pixels of the input image in the $L^2$-norm. Although the reconstruction looks good for the technique of Haber et al. [6], this algorithm does not take near-field lighting effects into account. Thus, while faithfully reproducing the appearance of the scene in its original setting, it offsets wrong estimates in the lighting with wrong terms in the BRDF and vice versa. Afterwards, when rendering the model under white lighting, the compensation colors in the BRDF are revealed (Figure 7(c)). Our technique includes near-field lighting effects in the optimization, resulting in an improved separation of lighting and BRDF. The result minimizes the amount of compensation colors. Figure 7(d) shows our result under white lighting. Our technique better recovers the white material properties.

The second captured dataset is constructed from 8 camera views of a table scene with 2 objects and a desk lamp light source within the scene. Geometry is acquired with structured light scanning based on gray codes. The same solving procedure is followed as with the synthetic dataset. Clustering is performed with the method as described in section 3.3, to identify the emitting primitives. The desk lamp must be visible in at least one of the input images for this to work. Figure 8 demonstrates our results of relighting on the table scene. The reconstruction demonstrates that a plausible estimation of materials and lighting has been made. Note that when we reposition the light source, correct shadows are cast from the new direction. However, in some cases a slight residual from the old shadow can be observed. This is due to inaccuracies introduced by noise in the input images that our renderer does not reproduce. We therefore cannot always filter out the old shadow completely. Most of the rendering errors in the scene arise from incorrectly reconstructed geometry, such as holes in the mesh. Refining geometric data with extra knowledge of near-field light sources is considered as future work. The focus of our method, however, is on the modeling of near-field lighting. The general appearance of the scene is correctly captured and its parameters can be freely manipulated.

6 Conclusion

This paper has presented an algorithm for simultaneous estimation of illumination and BRDF in uncontrolled images that supports both distant and near-field illumination. Previous techniques were limited to distant lighting or estimated near-field illumination in highly constrained environments. Our technique is based on a further factorization of the lighting, while keeping the rendering efficient. Complexity is reduced by clustering the scene geometry into a few groups of important emitters. Finally, the emitting powers are calculated.
by alternatingly solving for illumination and reflectance. We have demonstrated our results on simultaneous illumination and BRDF estimation with information from near-field light sources. We have shown that extra knowledge of near-field light sources leads to more accurate estimates. The synthetic dataset clearly reveals that our technique, in contrast to previous work, does not suffer from disturbing correction colors. The captured dataset indicates that our method allows relighting of interesting scenes.

7 Acknowledgements

This work was partly made possible by the Agency for Innovation by Science and Technology in Flanders (IWT).

References


