Eigentransport for Efficient and Accurate All-Frequency Relighting

Abstract

We present a method for creating a geometry-dependent basis for precomputed radiance transfer. Unlike previous PRT bases, ours is derived from principal component analysis of the sampled transport functions at each vertex. It allows for efficient evaluation of shading, has low memory requirements and produces accurate results with few coefficients. We are able to capture all-frequency effects from both distant and near-field dynamic lighting in real-time and present a simple rotation scheme. Reconstruction of the final shading becomes a low-order dot product and is performed on the GPU.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Color, shading, shadowing, and texture

Keywords: Precomputed Radiance Transfer, Realistic Image Synthesis, Model Fitting, Image Error Analysis

1 Introduction

Soft and hard shadows combine to give an observer subtle cues regarding the relative position of objects as well as the lighting surrounding them. Recent precomputed radiance transfer (PRT) techniques allow for shadows to be reconstructed under dynamic, distant, low [Sloan et al. 2002] and all-frequency [Ng et al. 2003] realistic lighting environments. For diffuse materials, the shading integrand can be split into the product of two functions: lighting and cosine-weighted visibility, often called transport. We focus on diffuse relighting, however our method may be extended to glossy BRDFs.

Projecting the spherical lighting and transport functions into an alternative basis permits evaluation of the shading integral in real-time by avoiding the prohibitively slow sampling of offline cubature techniques. Originally, the Spherical Harmonics (SH) basis was chosen for its ability to capture low-frequency area lighting effects with few coefficients [Sloan et al. 2002; Kautz et al. 2002]. PRT SH reconstructions result in smooth and plausible renderings, despite the blurring of the external lighting and transport functions. Moreover, closed form rotation matrices allow SH projected lighting to be rotated without having to resample and project from the original signal.

Ng et al. use the Haar wavelet basis to capture all-frequency lighting effects resulting in both soft and hard shadows [Ng et al. 2003]. They note that, in order to capture the all-frequency lighting content, SH would require approximately two-orders of magnitude more coefficients than the Haar wavelet basis. Wang et al. derive approximate rotation matrices for the Haar wavelet basis [Wang et al. 2006].

We present a generalized mesh-dependent basis for all-frequency relighting of scenes under distant and locally varying illumination. Using significantly fewer coefficients than previous techniques, we can shade a scene on the GPU that is nearly indistinguishable from the ground truth rendering. Furthermore, we outline a simple algorithm for lighting rotations that performs intelligent light resampling in real-time. Figure 1 illustrates results using our technique.

Contributions We analyze transport function signals using a statistical approach. In particular, using eigenanalysis, we show that high-dimensional all-frequency transport effects can be accurately reconstructed in a lower-dimensional linear subspace. We illustrate how our mesh-dependent basis can capture all-frequency lighting effects with fewer coefficients than previous techniques. In fact, with the same computational cost current interactive applications, such as games, incur for low-frequency SH lighting, our basis clearly captures all-frequency effects. Ease of implementation, precomputation and evaluation speed, and increased accuracy are the key benefits our technique can bring to the game development community. Both distant (environment map) and local (approximate point source) lighting are compatible with our technique and we use a simple, efficient algorithm for rotating distant lights.

2 Previous Work

Many rendering techniques reduce the complexity of evaluating the shading that, in its unconstrained form, requires evaluation of potentially complicated functions at many spherical samples.

2.1 Basis Functions for PRT

Ramamoorthi and Hanrahan analyzed the irradiance of convex objects under distant illumination in the SH basis. They concluded that unshadowed irradiance is inherently low-dimensional and can be accurately captured with a low-order SH reconstruction [Ramamoorthi and Hanrahan 2001].
mamoorthi and Hanrahan 2001]. Sloan et al. extend this approach to include low-frequency shadowing (and indirect illumination) [Sloan et al. 2002]. Ng et al. project the transport and lighting into the orthonormal Haar basis defined over a cubemap domain [Ng et al. 2003]. This basis can capture all-frequency lighting effects such as hard and soft shadows. They note that capturing approximately the same frequency content using SH would require two orders of magnitude more coefficients.

Lehtinen and Kautz start with an SH formulation for glossy PRT rendering and perform a basis conversion for the final exit radiance representation [Lehtinen and Kautz 2003]. They project the low-frequency SH exit radiance function into a directional basis for more efficient viewpoint querying. Gautron et al. present a hemispherical basis function similar to SH for representing low-frequency functions [Gautron et al. 2004].

Kristensen et al. use a light-simplification procedure coupled with a PCA basis to represent local lighting response for PRT geometry relighting [Kristensen et al. 2005]. Direct illumination is calculated with shadow maps and the PCA basis is used to reconstruct only the indirect lighting contributions. Since indirect illumination is much smoother than direct illumination, only a handful of coefficients are required to reconstruct an approximate local lighting response.

Tsai and Shih use spherical radial basis functions (SRBFs) for all-frequency geometry relighting [Tsai and Shih 2006]. Their datasets are compressed using clustered tensor approximation and the SRBF basis can be easily rotated. The precomputation procedure of fitting SRBFs to the transport, factored BRDF and lighting can be very time consuming and, typically, hundreds of SRBFs (each requiring six values for storage) are required to capture all-frequency effects.

Green et al. use a hybrid technique for PRT rendering of glossy materials [Green et al. 2006]; SH is used to capture the low-frequency effects, and view dependent glossy effects are represented using Gaussian terms with non-linear optimization.

Xu et al. define piece-wise constant basis functions over the sphere for reconstructing all-frequency PRT effects [Xu et al. 2007]. Approximating the multi-product shading integral as a product of the individual component integrals and using summed-addition tables results in real-time rendering performance. The mathematical simplification allows the technique to be coupled with shadow fields [Zhou et al. 2005] for rendering dynamic all-frequency glossy scenes.

Ma et al. use a spherical wavelet basis to reconstruct the triple-product shading integral for arbitrary BRDFs defined using a local parameterization [Ma et al. 2006]. As with Haar wavelets defined over cubemaps [Ng et al. 2003], a simple rotation algorithm for this basis does not exist. A significant amount of storage is required to capture all-frequency effects and results are illustrated using 80 and 320 coefficient reconstructions. Lessig proposes an orthogonal spherical wavelet basis that can be rotated in closed form, but this basis has not been used for PRT [Lessig 2007].

In general, our technique requires fewer coefficients to capture hard and soft shadowing effects, final shading is a low-order dot product calculated on the GPU, and no time consuming optimization is necessary during precomputation. We instead rely on statistical algorithms with known behavior: the singular value decomposition of the original transport dataset yields the transformation used to project the data onto a lower-dimensional linear subspace for efficient reconstruction.

### 2.2 Dimensionality of Light Transport

The most relevant previous work is that of Mahajan et al. [2007]. Analyzing the dimensionality of light transport within blocks of pixels or clusters of vertices, a relationship between the dimensionality and the block/cluster size is observed and formalized. Mahajan et al. apply their theoretical framework to determine optimal cluster sizes for CPCA geometry relighting and image relighting, whereas we consider light transport at vertices of a scene and determine a compact and efficient, global basis for reconstruction.

Sloan et al. [2003] perform clustered PCA, over vertex clusters, on spherical harmonics projected relighting transfer matrices used for low-frequency relighting of glossy materials. Liu et al. [2004] perform an SVD factorization of factored glossy BRDF representations coupled with the clustered PCA projection of the Haar projection (as opposed to the CPCA over SH transfer matrices of [Sloan et al. 2003]) of the non-diffuse transport function. These approaches perform dimensionality reduction over pre-projected transfer functions, whereas we analyze the dimensionality of the original transport function. Our approach can be extended to clusters over vertices, however our main contribution is an insight into the dimensionality of cosine-weighted visibility over an entire mesh, with an immediate application to high-fidelity, all-frequency diffuse relighting. Our approach is easy to implement and can be readily integrated into content generation pipelines used in games.

### 2.3 Rotating Projected Lighting

One of the advantages of SH PRT approaches is the ability to rotate the environment lighting in real-time. Recently, a technique was proposed for performing approximate lighting rotations for Haar compressed lighting environments without requiring the expensive light resampling procedure [Wang et al. 2006]. The representation used for the approximate Haar rotations is time consuming to compute and requires significant storage and computation during run-time evaluation, yielding only interactive relighting of arbitrary BRDF scenes without shadowing effects.

Closed form SH rotations require \( n^2 \times n^2 \) matrices for an \( n^{th} \) order SH expansion. Some work has been recently proposed to approximate SH rotation matrices for higher-order expansions [Krivánek et al. 2006], however this technique requires almost the same amount of calculation and storage as the full rotation formulation. Ma et al. did not propose a rotation scheme for their spherical Haar basis [Ma et al. 2006], although Lessig does for his basis [Lessig 2007].

Our rotation scheme does not require the explicit resampling of the lighting and, for common sampling patterns (latitude/longitude or cubemap), a simple and efficient algorithm is proposed.

### 3 Transport Dimensionality Reduction

Our work explicitly investigates the dimensionality of transport as a function of geometry. For a given scene, we perform per-vertex ray-tracing in order to sample the transport function

\[
T(\bar{x}, \omega) = V(\bar{x}, \omega) \max(0, n_{\hat{x}} \cdot \omega), \quad \omega \in \Omega_{n_{\hat{x}}}
\]

where \( \bar{x} \) is a vertex in the scene, \( n_{\hat{x}} \) is the normal at the vertex, \( V \) is the binary visibility function, and \( \omega \) is the sampling direction in a set of spherical sampling directions, \( \Omega_{n_{\hat{x}}} \). The choice of sampling scheme determines this set and can affect the final shading reconstruction as well as the ability to easily perform rotations. Our analysis, however, is robust to this choice. We have carried all experiments out using low-discrepancy Halton, uniform latitude/longitude and cubemap sampling schemes (see section 5.) The initial dataset is composed of \( v \) \( s \)-dimensional vectors, where \( v \) is the number of vertices and \( s \) is the number of samples used.
Performing eigenanalysis on this dataset, we observe that a large proportion of the variance can be captured using a small fraction of the eigenvectors. This subset of \(d\) eigenvectors, which we call the eigentransports, form an orthonormal basis that will be used for shading. Figure 2 illustrates how transport can be well represented in a lower-dimensional linear subspace; the first 100 eigenvalues out of tens of thousands are shown.

**Figure 2:** Captured variance: 90\% (red) and 95\% (blue) of the variance is typically captured using few coefficients. The first 100 eigenvalues (inset) of the eigentransports for the Teapot (left) and Lucy (right) meshes.

We perform principal component analysis (PCA) on the set of sampled per-vertex transport functions, keeping only the number of eigentransports required, \(d\), for our desired shading quality. The PCA decomposition yields

\[
\mathbf{T} \approx \mathbf{T}_d \cdot \mathbf{B} + \mathbf{T}^\prime
\]  

where \(\mathbf{T}\) is the full-dimensional matrix of the per-vertex sampled cosine-weighted visibility functions, \(\mathbf{T}_d\) is the matrix of per-vertex projected transport coefficients, \(\mathbf{B}\) is the matrix of eigentransport basis functions and \(\mathbf{T}^\prime\) is a matrix containing the mean transport function, \(\bar{T}\), repeated \(v\) times.

\(\mathbf{T}^\prime\) must be subtracted from the original dataset in order to perform PCA, and is added back during reconstruction. When \(d = s\), \(\mathbf{T}\) is perfectly reconstructed. Section 4.3 discusses the evaluation of the shading integral completely in the projected space, without having to explicitly add the full-dimensional mean transport function or sample any of the \(s\) spherical samples per-vertex.

4 Projection and Reconstruction

Both the per-vertex transport and lighting functions must be projected into our mesh-dependent basis before being used to evaluate the shading integral at each vertex. We will illustrate our basis’ ability to efficiently capture and reconstruct transport and lighting functions using few coefficients.

4.1 Transport Function

The eigentransport basis functions define an orthonormal basis for modeling the transport functions. One of the main benefits of a customized basis over previous PRT bases is the ability to reconstruct the relevant components of the shading integral more efficiently. The eigentransports are designed to capture the variance of the transport functions for a given scene and are able to reconstruct transport with greater accuracy per-coefficient than the most common PRT bases: SH and Haar. Using fewer coefficients, eigentransports reconstruct transport with significantly reduced numerical and visual error than SH and Haar.

4.2 Lighting Function

Although spherical external lighting environments are not a part of the dataset used to generate a basis, we project them into our basis in order to reconstruct the shading integral. The projection, in essence, only captures the important components of the lighting for a specific scene. This eliminates the storage of unnecessary lighting since only the relevant regions of the lighting (from the transport’s perspective) are retained. A projection of the environment lighting, \(L_{env}\), into our basis yields the projected lighting coefficients, \(\hat{L}_{env}\).

\[
\hat{L}_{env} = (L_{env} - \bar{T}) \cdot B^{-1} = (L_{env} - \bar{T}) \cdot B^T
\]  

This is a strictly different idea than the general basis projection of lighting used in previous PRT approaches. The projection of a lighting environment is tailored to the geometry of the scene, and thus, its reconstruction is not meant to match the original lighting but rather to capture the regions of lighting that are most relevant for a particular scene. For example, Figure 4 illustrates a scene with very few vertices sampling the bottom hemisphere (in 4(f) white is high sampling rate, black is no sampling.) Our customized basis is tailored to reconstruct the part of the signal that matters most: in this case, the upper-hemisphere.

**Figure 4:** Scene geometry biasing: the basis reconstructs detail in the more visible regions of the sphere.

Many of our meshes have added ground plane geometry that biases the sampling density. This additional geometry biasing actually decreases the effectiveness of our basis; however, we choose to include it in most of our scenes for illustrative purposes. All ground planes are coarsely sampled, sometimes introducing shadow interpolation artifacts across vertices. Figure 6 illustrates the sampling bias. Figure 7 illustrates our basis’ ability to tailor the lighting to a particular scene. The sampling densities in Figure 6 match the reconstruction patterns in Figures 7(i) to 7(p).

4.2.1 Local Lighting Approximation

Since the full dimensional dataset we use to derive our mesh-dependent basis is based on a sub-sampling of the spherical domain, we can determine local lighting projections into our basis in a very straightforward manner.

Near-field lighting, unlike distant environmental lighting environments, depends on vertex location. To determine the projection of
a local light with position $L_p$ at a vertex $\bar{x}$, we determine the $k$-nearest directions from $\Omega_{nx}$ to $\|L_p - \bar{x}\|$ (which can be done in closed form for cubemap sampling). For most scenes $k = 1$ is sufficient. A per-vertex full-dimensional lighting vector, $L_{FD}$, can be composed, after finding the $k$-nearest directions, by setting the appropriate direction indices in the vector to a constant value dependent on the original sampling, $\alpha$ (see equation 10). This vector is sparse (often only a single non-zero element) and the evaluation of equation 3 would then result in a simple scaling of rows in the inverse eigentransport matrix (after the projection of the full-dimensional mean transport function):

$$L^c = (L_{FD} - \hat{T}) \cdot B^T = \alpha \sum_{i=1}^{k} B^T(i, \cdot) - \hat{T} \cdot B^T$$

where $\hat{T}$ is the row of $\hat{T}$ belonging to the current vertex, and $\hat{T}_p$ is the projection of the mean transport function into the basis. Figure 5 illustrates our basis’ support for local lighting. As with all the shading results generated with our system, this scene was rendered at approximately 200 FPS.

### 4.3 Shading Reconstruction

Our goal is to evaluate the diffuse shading integral at each vertex without having to explicitly reconstruct the transport or lighting functions from their projected forms. The shading integral is

$$B(\bar{x}) = \frac{\rho}{\pi} \int_{\Omega_{\text{ns}}} L(\bar{x}, \omega) T(\bar{x}, \omega) d\omega$$

where $B(\bar{x})$ is the shade at $\bar{x}$, $\rho$ is the diffuse reflectivity of the surface and $L$ may be independent of $\bar{x}$ in the case of distant-only lighting. Individually, the transport and lighting functions can be reconstructed from their projection coefficients and the mean transport function

$$L(\bar{x}, \omega) \approx L^c \cdot B + \hat{T} \quad \text{or} \quad L(\omega) \approx L^c_{env} \cdot B + \hat{T}$$

and we eliminate the addition of the full-dimensional mean transport vector by projecting it into the basis, leaving us with the following mean-compensated lighting and transport coefficients required
Figure 7: Original environment maps (7(a), 7(e)), SH (7(b), 7(g)), Haar (7(c), 7(d), 7(h)) and our basis reconstructions (7(i) to 7(p)), and sampling density of the teapot scene (7(f)). Number of coefficients are in parentheses and our bases are named according to their associated geometry.

Figure 6: Meshes with (6(a), 6(b), 6(d), 6(e)) and without (6(c), 6(f)) added ground planes.

Figure 8: Fewer than 200 coefficients (8(b)) almost perfectly reconstructs the ground truth (8(a)). At 36 coefficients (8(d)) our technique captures significantly more detail than an equivalent SH reconstruction (8(f)).

5 Handling Lighting Rotation

To rotate the lighting (or scene) normally involves rotation of the original lighting dataset (cubemap) and reprojecting into our basis. The brute-force resampling of the cubemap data is the computational bottleneck of this operation, since projecting into our basis is a simple matrix vector multiplication. Depending on the original number of samples used during precomputation, the brute-force resampling and projection can take anywhere between 0.02 seconds (for approximately 1000 samples) to 2 seconds (for approximately 30000 samples.)

We present a simple algorithm that uses the original sampled lighting (in canonical orientation), $L_{env}^e$, to determine the rotated full-dimensional lighting, $L_{env}^R$. Equation 3 is applied to $L_{env}^R$ to obtain the coefficients used for shading.

Given a desired rotation for the environment lighting, we apply the
inverse rotation to the $s$ points of the sampling lattice. Now, instead of resampling the environment based on the directions given by the new point locations, we compute the sampling coordinates of the new point positions and bilinearly interpolate the lighting based on the previously sampled values, yielding the rotated full-dimensional lighting, $L_{\text{rot}}^t$. This recomputation and interpolation can be performed efficiently in real-time if a uniform latitude/longitude or (as in our case) cubemap parametrization is used.

Figure 9 compares a brute-force rotated and sampled lighting environment to the lighting obtained using our algorithm. After projection into our basis, the differences become even smaller.

![Figure 9](image)

**Figure 9:** Brute-force resampled lighting (9(a),9(c)) and the lighting obtained with our algorithm (9(b),9(d)).

## 6 Error Analysis

We measure numerical error using the $L_2$ norm over the reconstructed transport functions. Figure 10 illustrates how our basis reconstructs transport with lower error than Haar and SH. Haar coefficients are chosen to minimize the $L_2$ error, which means that for an $n$ coefficient Haar reconstruction, potentially many more than $n$ coefficients must be stored. For scenes with biasing geometry, Haar typically outperforms our technique only after a few hundred coefficients.

![Figure 10](image)

**Figure 10:** Error for transport reconstruction.

Our basis, especially for the first 100 coefficient reconstructions, outperforms Haar and SH. For interactive applications like games, low-order reconstructions are preferred since they require less run-time computation to shade. Furthermore, GPU constraints limit the number of constant variables that can be used in a shader.

## 7 Results

Figure 11 compares our rendering results to ground truth renderings. All-frequency shadows are captured under dynamic lighting with real-time performance and low error.

### 7.1 Memory Requirements

If the lighting is pre-projected (in its canonical orientation) for use with our rotation algorithm, then only the per-vertex coefficients and global lighting coefficients need to be stored for the run-time. This cost, in bytes, is $4d(v + 1)$ for single precision floating point values. The Asian Dragon (39372 vertices), Happy Buddha (18978 vertices) and Lucy (41293 vertices) meshes require 7.2MB, 3.4MB and 7.5MB respectively for a 48 coefficient eigen-transport reconstruction. Local lighting and brute-force lighting rotation/reprojection both require storage of the projection matrix, adding $4ds$ bytes to the previous requirements.

### 7.2 Precomputation Times

We use a standard PCA decomposition implemented in MATLAB to determine our geometry dependent basis. PCA computation times vary depending on the number of vertices of the mesh and the number of original sampling directions in the dataset. In general, the precomputation times are reasonable considering the amount of time required to sample the original transport, which is required for PRT precomputations with any other basis. For the meshes in our dataset, PCA took between 4 and 22 minutes. In contrast, generating the sampled transport function took anywhere between 2 and 13 minutes.

## 8 Limitations and Future Work

Since eigentransports are derived from the scene geometry, sampling biasing due to added geometry can negatively affect the effectiveness of the basis. In our experiments, the scenes without added ground planes required significantly fewer reconstruction coefficients to capture more of the transport variance.

All approximated local lights are only accurate outside the convex hull of the scene since the self-occlusions used to calculate transport are with respect to global visibility.

As it stands, assuming diffuse surfaces simplifies our experiments and allows us to analyze the dimensionality of transport without considering the potentially non-linear effects of arbitrary BRDFs. Performing a similar investigation of dimensionality reduction while including the effects of arbitrary BRDFs is an interesting area of future work we plan on exploring. Furthermore, our precomputation system can be extended, in a manner similar to that presented in [Sloan et al. 2002], to include indirect illumination effects at no extra run-time cost and without any changes to the mathematical analysis we presented.

Our mesh-dependent basis can be extended to vertex-cluster dependent bases over a mesh. The dimensionality of transport within clusters of vertices on a mesh is bounded by the global dimensionality and is likely to be much smaller. Investigating the trade-offs of cluster-size and required reconstruction dimensionality, in a similar fashion as Mahajan et al. [2007], is left to future work.

## 9 Conclusion

We present a novel, scene-dependent basis for geometry relighting under dynamic global and local lighting. We use PCA to minimize
the least-squares error of per-vertex transport over a mesh. The resulting PCA transformation is used to project lighting into our basis.

We achieve lower numerical error than SH and Haar bases and require fewer basis functions to shade scenes with near ground-truth detail. The storage requirements of our basis are modest and pre-computation times are low.

**References**


