Subspace Methods for Computational Relighting

Ha Q. Nguyen, Siying Liu, and Minh N. Do

Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign, Urbana, IL 61801

ABSTRACT

We propose a vector space approach for relighting a Lambertian convex object with distant light source, whose crucial task is the decomposition of the reflectance function into albedos (or reflection coefficients) and lightings based on a set of images of the same object and its 3-D model. Making use of the fact that reflectance functions are well approximated by a low-dimensional linear subspace spanned by the first few spherical harmonics, this inverse problem can be formulated as a matrix factorization, in which the basis of the subspace is encoded in the spherical harmonic matrix $S$. A necessary and sufficient condition on $S$ for unique factorization is derived with an introduction to a new notion of matrix rank called nonseparable full rank. An SVD-based algorithm for exact factorization in the noiseless case is introduced. In the presence of noise, the algorithm is slightly modified by incorporating the positivity of albedos into a convex optimization problem. Implementations of the proposed algorithms are done on a set of synthetic data.

Keywords: relighting, Lambertian surfaces, inverse rendering, reflectance function, spherical convolution, spherical harmonics, matrix factorization, singular value decomposition, convex optimization

1. INTRODUCTION

Relighting is one of the core problems in photorealism, the field of creating, or technically, rendering images that look indistinguishably from realistic ones. In relighting problems, we are given a set of images of the same scene under various lighting conditions, and the geometry of the scene as well. Our task is to synthesize another image as if it was obtained by shining the scene with a novel lighting.

Images of an object can be viewed as a reflectance function that quantifies the amount of light reflected to the camera from each point on the surface of the object. In general, reflectance function is the result of the interaction between lighting, reflection characteristic of the object (often described by the BRDF - Bidirectional Reflectance Distribution Function) and texture of the object. Consequently, the relighting problem actually consists of two sub-problems: inverse rendering and forward rendering. In the inverse rendering phase, the reflectance function is decomposed into lighting, BRDF and texture; and in forward rendering phase, the recovered BRDF and texture, together with a novel lighting are combined into a novel reflectance function, resulting in a novel image. Both forward and inverse problems have been researched extensively in the past few decades with a large body of work (see\textsuperscript{1,2} for detailed discussions). Most previous techniques (see, for example,\textsuperscript{3}) can deal only with highly controlled lighting conditions in which a single point source is usually actively positioned. These methods certainly cannot work in outdoor conditions when the lighting can be arbitrarily complex, coming from various sources of continuous distributions such as the skylight.

The difficulties of rendering under general, or uncontrolled lighting are due to the lack of a discretized framework that can efficiently describe reflectance function which has been previously interpreted as an integral. Ramamoorthi and Hanrahan in their series of work\textsuperscript{4–6} introduced a breakthrough signal processing framework for both forward and inverse rendering. In this framework, reflectance function is treated simply as a spherical convolution (sphere counterpart of circular convolution) of lighting and reflection kernel (BRDF multiplied by a half-cosine). This allows us to relate reflectance function to lighting and reflection kernel in terms of their spherical harmonic expansions (sphere counterpart of Fourier series). Transforming from space-domain to frequency-domain yields two great advantages: (1) integrals are mapped to products of coefficients; and (2) reflectance function can be well-approximated by a few low-frequency terms (because reflection kernel often varies...
Based on this framework, Ramamoorthi and Hanrahan developed efficient algorithms for both forward and inverse rendering. However, they only considered homogeneous objects with no texture, and thus ignored the local scalings of the reflection.

Basri and Jacobs in an independent work\(^7\) discovered a similar result for the special case of Lambertian surfaces whose BRDF is a constant. Here, reflectance function is the spherical convolution of lighting and half-cosine, then scaled by albedos. As a result, it has been shown in\(^7\) that any reflectance function (of a Lambertian convex object with distant light source\(^*\)) can be well-approximated by its first nine spherical harmonic coefficients. In other words, reflectance functions live close to a 9-dimensional linear subspace spanned by first 9 spherical harmonics. Basri \textit{et. al.} in the subsequent work\(^8\) matricized this important observation and mapped the forward rendering to a matrix multiplication, and inverse rendering to a matrix factorization. In this setup, the main task of relighting problem is to factorize the image matrix formed by known images into a diagonal albedo matrix, the known spherical harmonic matrix, and a lighting matrix. A heuristic algorithm was proposed in\(^8\) to solve the factorization problem by simply picking the matrix of right-singular vectors scaled by square roots of singular values of the image matrix as a lighting matrix. However, no rigorous justification for the method has been done and the well-posedness of the factorization has not been addressed.

The main goal of this research is to fix the shortcomings of.\(^8\) By means of subspace methods, we can give answers to the two main questions: (1) when does the factorization have unique solution; and (2) if unique, what is the exact solution of the factorization. We want to emphasize that our method, although also based on the \textit{singular value decomposition} (SVD), totally differs from that of.\(^8\) This paper is a condensed version of.\(^9\) All proofs are omitted for brevity.

The remainder of the paper is structured as follows: Sec. 2 discusses basic background of reflection equation and spherical harmonic expansion; Sec. 3 formulates the inverse rendering problem as a matrix factorization; Sec. 4 studies the uniqueness of the factorization in terms of the spherical harmonic matrix; Sec. 5 provides factorization algorithms based on SVD and convex optimization; Sec. 6 presents some numerical experiments on the proposed algorithms; and Sec. 7 draws some concluding remarks.

2. PRELIMINARIES

This section summarizes the signal-processing framework that relates reflectance function to lighting and reflection kernel in space-domain as a spherical convolution. The relation can be transformed into frequency-domain via spherical harmonic expansions.

2.1 Reflection as Convolution

Throughout this paper, we impose the following commonly used assumptions in interactive graphics and computer vision.

A1 \textit{Curved Surfaces:} characterized by surface normals.

A2 \textit{Lambertian:} same reflection for every viewing angle.

A3 \textit{Convex Objects:} no shadowing or inter-reflection.

A4 \textit{Distant Illumination:} illumination only depends on incident angle.

Consider a point \(p\) on the object surface with the normal \(n_p = [1, \alpha(p), \beta(p)]\) in some \textit{global spherical coordinate} system (illustrated in Fig. 1). Let \(Y(p)\) and \(\rho(p)\) be respectively the reflected radiance (or reflectance function) and the albedo at location \(p\). Let \(L(\theta, \phi)\) be the illuminating intensity (or lighting function) at incident angle \((\theta, \phi)\). Let also \(K(\theta, \phi) = \max(\cos(\theta), 0)\) be the \textit{Lambertian kernel} (or \textit{half cosine} function) at incident angle \((\theta, \phi)\). It was shown in\(^7\) that the reflectance function (of a Lambertian convex object with distant light

\(^*\)We will elaborate on these assumptions later on.
source) is simply a spherical convolution of the lighting function and the Lambertian kernel scaled locally by the albedos. In particular,

\[ Y(p) = \rho(p)(L \ast_s K)(\alpha(p), \beta(p)), \]

where the spherical convolution \( \ast_s \) is defined as

\[ (L \ast_s K)(\alpha, \beta) \triangleq \int_0^\pi \int_0^{2\pi} L(\theta, \varphi)K(R_{\alpha,\beta}^{-1}(\theta, \varphi)) \sin \theta \, d\theta \, d\varphi. \]

In this definition, \( R_{\alpha,\beta}^{-1} = R_y(\alpha)R_z(\beta) \) is the rotation about \( z \)-axis by \( -\beta \) and then about \( y \)-axis by \( -\alpha \), and \( R_{\alpha,\beta}^{-1}(\theta, \varphi) \) is understood as the angular coordinates of the rotation acted on the unit vector \([1, \theta, \varphi] \). This spherical convolution naturally defines a linear rotation-invariant (LRI) system with impulse response \( K(\theta, \varphi) \).

### 2.2 Spherical Harmonics as Basis

Now we want to look at the reflection equation (1) in frequency-domain because the energy of the Lambertian kernel \( K \) is compressed in low frequencies. As a spherical counterpart of Fourier basis on the circle, the spherical harmonics \( \{S_{m,n}\}_{m \geq 0, |n| \leq m} \) form an orthonormal basis for functions on the unit sphere

\[ S_{m,n}(\theta, \varphi) = N_{m,n} \cdot P_{m,n}(\cos \theta) e^{jn\varphi}, \]

where \( j = \sqrt{-1} \), \( N_{m,n} \) is the normalized term, and \( P_{m,n}(\cdot) \) is the associated Legendre function.

Every function \( F(\theta, \varphi) \) on the unit sphere has a spherical harmonic expansion

\[ F(\theta, \varphi) = \sum_{m=0}^{\infty} \sum_{n=-m}^{m} f_{m,n} S_{m,n}(\theta, \varphi), \]

where the harmonic coefficients \( f_{m,n} \) are given by

\[ f_{m,n} = \int_0^\pi \int_0^{2\pi} F(\theta, \varphi) S_{m,n}^*(\theta, \varphi) \sin \theta \, d\theta \, d\varphi. \]

The feature that makes spherical harmonics similar to Fourier basis on the circle is that they are eigensignals of LRI systems. This key property maps the spherical convolution in space-domain to a multiplication in frequency-domain.

---

1. See [7] for specific formulae of \( P_{m,n}(\cdot) \) and \( N_{m,n} \).
2. See a special case of the Funk-Hecke theorem \([7, \text{Thm. 1}]\).
domain. Particularly, let the harmonic coefficients of $L$, $K$ and $L \ast K$ be $\{\ell_{m,n}\}_{m \geq 0, |n| \leq m}$, $\{k_{m,0}\}_{m \geq 0}$, and $\{y_{m,n}\}_{m \geq 0, |n| \leq m}$, respectively, then it was shown in\textsuperscript{5,7} that

$$y_{m,n} = \sqrt{\frac{4\pi}{2m+1}} k_{m,0} \ell_{m,n}. \quad (3)$$

Moreover, it was also shown in\textsuperscript{7} that the convolution $L \ast K$ can be well-approximated using first few terms in its spherical harmonic expansion, namely

$$(L \ast K)(\alpha, \beta) \approx \sum_{m=0}^{M'} \sum_{n=-m}^{m} y_{m,n} S_{m,n}(\alpha, \beta),$$

for small $M'$ (for instance, using $M' = 3$ preserves roughly 97.96\% the energy of $L \ast K$.) As a result, the reflectance function can be approximated by

$$Y(p) \approx \rho(p) \sum_{m=1}^{M} y_{m} S_{m}(\alpha(p), \beta(p)), \quad (4)$$

where $M = M'^2$, and the double index $(m, n)$ was converted to the single index $m'$ using the relation $(m, n) \leftrightarrow m' = m^2 + m + n + 1$ (to keep the notations simple, we replace $m'$ back by $m$.) In words, this approximation says that, ignoring the albedos, reflectance functions live around a low-dimensional linear subspace spanned by the first few spherical harmonics.

3. PROBLEM STATEMENT

3.1 From Reflection to Images

Suppose we are given $J$ images taken at the same viewpoint of the same object under different lighting conditions. Let $N$ be the number of pixels in each image, and $p_i$ be the location on the surface corresponding to pixel $i$, for $i = 1, \ldots, N$. From (4), the intensity at pixel $i$ of image $j$ is approximated by

$$Y_j(p_i) \approx \rho(p_i) \sum_{m=1}^{M} y_{m} S_{m}(\alpha(p_i), \beta(p_i)), \quad \text{for } 1 \leq i \leq N, 1 \leq j \leq J. \quad (5)$$

We can put all equations in (5) into a single matrix form by defining matrices $Y \in \mathbb{R}^{N \times J}, \Phi \in \mathbb{R}^{N \times N}$ (diagonal matrix of size $N$); $S \in \mathbb{R}^{N \times M}$; and $X \in \mathbb{R}^{M \times J}$ as follows

$$Y_{ij} = Y_j(p_i), \quad \Phi_{ii} = \rho(p_i), \quad S_{im} = S_m(\alpha(p_i), \beta(p_i)), \quad X_{mj} = y_m^j, \quad (6)$$

for $1 \leq i \leq N, 1 \leq j \leq J, 1 \leq m \leq M,$ and $M_{ij}$ denotes the entry of matrix $M$ at row $i$ and column $j$. With these notations, (5) becomes

$$Y \approx \Phi S X, \quad (7)$$

where $Y$, $\Phi$, $S$ and $X$ will be respectively referred to as image, albedo, spherical harmonic and lighting matrices.

\textsuperscript{5}Since $K(\theta, \varphi) = \max(\cos \theta, 0)$ does not depend on $\varphi$, it is orthogonal to all $S_{m,n}$ with $n \neq 0$, resulting in $k_{m,n} = 0$ for all $n \neq 0$. Thus, kernel $K$ is fully described by the coefficients $\{k_{m,0}\}_{m \geq 0}$.
3.2 The Matrix Factorization Problem

For analytical purpose, we first assume the noiseless case when the approximation in (7) is replaced with exact equation

\[ Y = \Phi S X, \]  

(8)

In light of Eq. (8), the inverse rendering problem becomes a matrix factorization in which the image matrix \( Y \in \mathbb{R}^{N \times J} \) and spherical harmonic matrix \( S \in \mathbb{R}^{N \times M} \) are known (can be computed from given images and 3-D model of the object); the albedo matrix \( \Phi \in \mathbb{R}^N \) and lighting matrix \( X \in \mathbb{R}^{M \times J} \) are to be recovered. Once the albedo and lighting matrices are reconstructed, say \( \Phi \) and \( X \), the forward rendering becomes an obvious matrix multiplication

\[ Y_{\text{new}} = \Phi S X_{\text{new}}, \]

where \( Y_{\text{new}} \) corresponds to novel images under novel lightings associated with matrix \( X_{\text{new}} \). Therefore, the rest of this paper only focuses on the (harder) inverse rendering, or matrix factorization problem. Before tackling the problem, we need a few further assumptions.

A5 \( J \geq M \): the number of images is greater than the number of spherical harmonics used in approximation.

A6 Eq. (8) always has a solution \((\Phi, X) \in \mathbb{D}_+^N \times \mathbb{R}^{M \times J}_\text{full}\).

In assumption A6, \( \mathbb{D}_+^N \) denotes the set of all \( N \times N \) diagonal matrices with nonzero entries on the diagonal; and \( \mathbb{R}^{M \times J}_\text{full} \) denotes the set of all full-rank \( M \times J \) matrices. Assuming that \( \Phi \in \mathbb{D}_+^N \), or all albedos are nonzero, does not restrict ourselves because if there are pixels corresponding to zero albedo, we can mask them out of the equations. Also, the assumption that \( X \) has full rank is reasonable and often made in classical algorithms for array signal processing such as MUSIC (MUltiple Signal Classification)\(^{10}\) and ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques).\(^{11}\) It essentially says that the images are taken under diversified lighting conditions.

4. UNIQUENESS OF THE FACTORIZATION

The very first question one always asks when dealing with an inverse problem is when the recovery is unique. This section provides a necessary and sufficient condition on the spherical harmonic matrix \( S \) such that the matrix factorization (8) is unique up to some scaling factor. It is certainly the best we can hope for since if \((\Phi, X)\) is a solution to (8) then so is \((\alpha \Phi, \frac{1}{\alpha} X)\), for any scalar \( \alpha \neq 0 \). The condition can be stated neatly by Thm. 1, the first main result of this paper, with an introduction to a new notion of matrix full rank.

**Definition 1.** A tall matrix \( S \in \mathbb{D}^N_\text{full} \) of \( N > M \) with no zero rows is said to have nonseparable full rank if there do not exist nonempty disjoint sets \( N_1 \) and \( N_2 \) such that \( N_1 \cup N_2 = N \) and \( \text{rank}(S_{N_1}) + \text{rank}(S_{N_2}) = M \).

Here, \( S_J \) denotes the submatrix of \( S \) with rows indexed by a subset \( J \) of \( N \triangleq \{1, 2, \ldots, N\} \). In words, a matrix has nonseparable full rank if it has full column rank and we can not separate its rows into two groups with ranks add up to the rank of the matrix. That justifies the term “nonseparable full rank.” Fig. 2 illustrates the notion of nonseparable full rank in comparison with regular full rank.

**Theorem 1.** Equation (8) has no solutions in \( \mathbb{D}^N \times \mathbb{R}^{M \times J}_\text{full} \) other than \((\alpha \Phi, \frac{1}{\alpha} X)\), for some scalar \( \alpha \neq 0 \), if and only if \( S \) has nonseparable full rank.

*Proof. See.*\(^9\) \( \Box \)
Algorithm 1 SVD-based Recovery of Albedo Matrix

**Inputs:** Image matrix $Y$, spherical harmonic matrix $S$

**Output:** Estimate $\hat{\Phi}$ of albedo matrix

1. Compute $M = (I - SS^\dagger) \odot (YY^T)$.
2. Compute the eigenvalue decomposition of $M$.
3. Let $z^*$ be the eigenvector associated with the smallest eigenvalue of $M$.
4. Return $\hat{\Phi} = \text{diag}((z_1^{-1}, \ldots, (z_N^{-1}))$.

5. **FACTORIZATION ALGORITHMS**

This section is to answer the second question of how to solve the matrix factorization problem given that it has unique solution, or $S$ has nonseparable full rank by Thm. 1. Note that since $S$ has full column rank, once $\Phi$ is known, $X$ can be uniquely recovered by

$$X = S^\dagger (\Phi^{-1} Y),$$

where $S^\dagger = (S^T S)^{-1} S^T$ is the pseudo-inverse of $S$. Therefore, we can focus on finding $\Phi$ which has a diagonal structure. In the noiseless case, an SVD-based algorithm is proposed to exactly recover $\Phi$. In the noisy case, it is modified into an optimization-based algorithm to find an estimate of $\Phi$.

### 5.1 SVD-based Algorithm

The algorithm is based on the following theorem, second main result of this paper.

**Theorem 2.** $\Phi = \text{diag}(\varphi)$ is a solution to Equation (8) if and only if $z = \varphi \odot^{-1}$ is a nontrivial solution to

$$\left( (I - SS^\dagger) \odot (YY^T) \right) z = 0,$$

where $\odot$ denotes element-wise operators.

**Proof.** See.9 □

Thm. 2 naturally gives rise to Algorithm 1, in which we solve (10) for $z$ by picking the eigenvector associated with the smallest eigenvalue of the positive definite matrix $(I - SS^\dagger) \odot (YY^T)$.

The following corollary is very useful in checking whether a matrix $S$ has nonseparable full rank. It can be deduced directly from Thm. 1 and Thm. 2 by setting $\Phi = I_N$ and $X = I_M$.

**Corollary 1.** A matrix $S \in \mathbb{R}^{N \times M}$ has nonseparable full rank if and only if

$$\text{rank}\left( (I - SS^\dagger) \odot (SS^T) \right) = N - 1.$$
Algorithm 2 Optimization-based Recovery of Albedo Matrix

| Inputs: | Image matrix $Y$, spherical harmonic matrix $S$ |
| Output: | Estimate $\hat{\Phi}$ of albedo matrix |

1. Compute $M = (I - SS^T) \odot (YY^T)$.
2. Let $c = 1_{N \times 1}$.
3. Use a convex programming method to find $z^*$ that minimizes $\|Mz\|_2$ s.t. $c^Tz = 1$, $z \geq 0$.
4. Return $\hat{\Phi} = \text{diag}(z_1^{-1}, \ldots, z_N^{-1})$.

We conclude this subsection by giving a few comments on this result. Without identifying the nonseparable full rank with the uniqueness of the corresponding matrix factorization, we can hardly see the connection between Definition 1 and Corollary 1. When fixing $M$ and growing $N$, checking if an $N \times M$ matrix has nonseparable full rank using the brute-force approach (i.e., computing the rank of every submatrix) would be exponentially complex. However, Corollary 1 provides a much more efficient indirect way to do so, in which only the rank of an $N \times N$ matrix needs to be computed, resulting in a polynomial complexity.

5.2 Optimization-based Algorithm

Since $\Phi$ is an albedo matrix, its diagonal entries must be all positive. It follows that the vector $z = (\varphi_1^{-1}, \ldots, \varphi_N^{-1})$ must be positive as well. If the proposed model is perfect then $\hat{\Phi}$ (and therefore $z$) is unique (provided that $S$ has nonseparable full rank), and we do not need to worry about the positivity of $z$. However, it can never be the case due to approximation error, model mismatch, measurement error, etc. Eq. (8) should therefore be remodeled as

$$Y + W = \Phi SX,$$

where $W$ is a white noise with standard deviation $\sigma_{\text{noise}}$. Consequently, the matrix $M$ in Algorithm 1 is modified as

$$M = (I - SS^T) \odot ((Y + W)(Y + W)^T).$$

Now the smallest eigenvalue of $M$ may be different from zero, and we can not guarantee that its corresponding eigenvector has all positive entries. Therefore the positivity of $z$ should be incorporated into the recovery.

We first note that, finding the eigenvector corresponding to the smallest eigenvalue of $M$ is nothing but solving the optimization problem

$$\min \|Mz\|_2 \quad \text{s.t.} \quad \|z\|_2 = 1.$$  \hspace{1cm} (12)

Now we can adjust (12) by adding the positivity constraint as

$$\min \|Mz\|_2 \quad \text{s.t.} \quad \|z\|_2 = 1 \text{ and } z \geq 0.$$  \hspace{1cm} (13)

Solving the optimization problem (13) is hard due to the nonconvexity of the feasible set. To convexify the problem, one can relax the constraint as

$$\min \|Mz\|_2 \quad \text{s.t.} \quad c^Tz = 1 \text{ and } z \geq 0,$$

where $c = [1, 1, \ldots, 1]^T \in \mathbb{R}^N$. Solving (14) is now easy using some well-developed convex programming method. This optimization-based algorithm to recover $\hat{\Phi}$ is summarized in Algorithm 2.
Figure 3. Synthetic images of the same object under different lighting conditions.

(a) Ground truth (b) $\sigma_{\text{noise}} = 1$, SNR = 37.1789 dBs

(c) $\sigma_{\text{noise}} = 3$, SNR = 24.5037 dBs (d) $\sigma_{\text{noise}} = 5$, SNR = 20.8765 dBs

Figure 4. Recovered albedos with corresponding SNRs under various levels of noise.

(a) Ground truth (b) $\sigma_{\text{noise}} = 1$, SNR = 37.1789 dBs

(c) $\sigma_{\text{noise}} = 3$, SNR = 24.5037 dBs (d) $\sigma_{\text{noise}} = 5$, SNR = 20.8765 dBs
6. SIMULATIONS
Simulations are performed on a data set of \( J = 12 \) images of the same object shown in Fig. 3 using MATLAB, with the convex programming cvx provided by. These images were synthesized by first randomly generating the lighting matrix \( X \) and then forming the product \( \Phi SX \), where \( \Phi \) is the ground truth for the albedos. The spherical harmonic matrix \( S \) was computed from the 3-D model of the object using \( (6) \), with \( M = 9 \). Using Corollary 1, we can easily verify that \( S \) has nonseparable full rank. The size of each image is \( 340 \times 512 \); excluding zero-albedo pixels results in \( N = 35983 \). The albedo matrix \( \Phi \) was recovered using Algorithm 2 under various levels of noise (white Gaussian noise of different variances was added to the images.) The reconstructions of the albedos are visually shown in Fig. 4 in comparison with the ground truth. All of them were normalized to have the same norm for comparison. When the additive noise is negligible (e.g. \( \sigma_{\text{noise}} = 1 \)), the reconstruction is almost exact that conforms to our developed theory. However, when the noise increases just a little bit (e.g. \( \sigma_{\text{noise}} = 5 \)), the reconstruction performance decreases significantly.

7. CONCLUSIONS
We have studied the relighting problem of a Lambertian convex object with distant light source. Under these assumptions, the reflectance functions live close to a low-dimensional linear subspace spanned by the first few spherical harmonics. The inverse rendering phase of relighting thus becomes a factorization of the image matrix into albedo and lighting matrices, given the spherical harmonic matrix. This factorization problem is well-posed if and only if the spherical harmonic matrix has nonseparable full rank, a stronger notion of full rank. In the noiseless case, exact factorization (up to some scale) can be done via an SVD-based algorithm. When the noise is present, an optimization-based algorithm is proposed with slight modifications. Simulations are performed only on synthetic data but already suggest that the proposed algorithm is quite sensitive to noise.

ACKNOWLEDGMENTS
This work was supported in part by the US National Science Foundation under Grant CCF-1218682.

REFERENCES

