Inverse Rendering of Lambertian Surfaces Using Subspace Methods

Ha Q. Nguyen and Minh N. Do, *Fellow, IEEE*

***Abstract*— We propose a vector space approach for inverse rendering of a Lambertian convex object with distant light sources. In this problem, the texture of the object and arbitrary lightings are both to be recovered from multiple images of the object and its 3D model. Our work is motivated by the observation that all possible images of a Lambertian object lie around a low-dimensional linear subspace spanned by the first few spherical harmonics. The inverse rendering can therefore be formulated as a matrix factorization, in which the basis of the subspace is encoded in a spherical harmonic matrix S associated with the object's geometry. 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. A singular value decomposition-based algorithm for exact factorization in the noiseless case is introduced. In the presence of noise, two algorithms, namely, alternating and optimization based are proposed to deal with two different types of noise. A random sample consensus-based algorithm is introduced to reduce the size of the optimization problem, which is equal to the number of pixels in each image. Implementations of the proposed algorithms are done on a real data set.**

***Index Terms*— Computational relighting, Lambertian surfaces, inverse rendering, reflectance function, spherical convolution, spherical harmonics, matrix factorization, singular value decomposition, convex optimization.**

I. Introduction

COMPUTATIONAL relighting is the problem of rendering images under virtual and novel lighting conditions. While it has been widely used on synthetic scenes constructed from computer graphics models or on real scenes captured under *controlled lighting* conditions, efficient relighting on general real scenes remain a research challenge due to its complexity involving both the geometric and photometric aspects in the image formation process. Existing relighting methods (see [1], [2] and the references therein) often consist in estimation, or *inverse rendering* of the scene geometry and texture, the BRDF (*Bi-directional Reflectance Distribution Function*) that defines the reflectance characteristics of the scene, and the lighting parameters. Given only the image data, the estimation of all these factors is ill-posed, and can only be achieved by strict assumptions on the scene geometry and/or light sources, or by controlled experiments (see [3]). The introduction of depth cameras and recent development of depth sensor fusion algorithms such as *KinectFusion* [4] open up new opportunities for efficient and reliable acquisition of 3D scene geometry in real-time. In this work, we explore the incorporation of geometry information in an inverse rendering problem under general, arbitrary (or *uncontrolled*) distant illumination.

Our work focuses on inverse rendering of *convex* Lambertian surfaces with *known geometry* given several images under various *distant* lighting conditions. In this model, the problem can be formulated, with the aid of a signal processing framework [5], [6], as a matrix factorization, in which an image matrix has to be decomposed into a product of an *albedo* diagonal matrix with a known *spherical harmonic* (SH) matrix in the middle and a lighting matrix. To the best of our knowledge, our work is the first one to study the problem of recovering both *nonuniform* albedos and uncontrolled lightings at the same time based on images and scene geometry. This is, to some extent, similar to an *array signal processing* problem [7], where both the DOAs (direction of arrivals) and sources are to be estimated from multiple snapshots of the array.

The main contribution of this paper is to solve the special matrix factorization mentioned above in both noiseless and noisy cases using the power of subspace methods. In the noiseless case, when both the lighting model and measurement are accurate, we gives a necessary and sufficient condition for unique factorization with an introduction to a new notion of matrix full rank, called *nonseparable full rank* (NSFR). We also solve for an exact solution using *singular value decomposition* (SVD) when the factorization is unique. In the noisy case we deal with the approximation error and measurement noise separately. When the approximation error is present, we propose an SVD-based algorithm to find the least squares solution of the factorization. This algorithm can be slightly modified to become a convex optimization that incorporates the positivity constraint on albedos. Another algorithm based on *RANdom SAmple Consensus* (RANSAC) [8] is also proposed to reduce the large size of the optimization problem which is equal to the number of pixels in each image. On the other hand, when the measurement noise is present, we propose an alternating least squares algorithm to find a local optimizer for the factorization.

Manuscript received May 18, 2014; revised August 19, 2014 and October 24, 2014; accepted October 24, 2014. Date of publication October 30, 2014; date of current version November 17, 2014. This work was supported by the National Science Foundation under Grant CCF-1218682 and Grant IIS-1116012. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Christine Guillemot.

H. Q. Nguyen was with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA. He is now with the Biomedical Imaging Group, Swiss Federal Institute of Technology Lausanne, Lausanne CH-1015, Switzerland (e-mail: ha.nguyen@epfl.ch).

M. N. Do is with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: minhdo@illinois.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2014.2366297

The remainder of the paper is structured as follows. Sec. II reviews some of the related work. Sec. III discusses basic background of the reflection equation and spherical harmonic expansions. Sec. IV formulates the inverse rendering as a special matrix factorization problem. Sec. V studies the uniqueness and exact solution of the factorization. Sec. VI provides various algorithms for the matrix factorization in the presence of noise based on least squares and convex optimization. Sec. VII discusses a computational issue on the large size of the optimization problem and a heuristic way to overcome it. Sec. VIII presents some experimental results on a real dataset using the proposed algorithms. Finally, Sec. IX makes some concluding remarks.

II. Related Work

The difficulties of inverse rendering under uncontrolled lighting are due to the lack of a discretized framework that can efficiently describe the *reflectance function* which has been often interpreted as a spherical integral. Ramamoorthi and Hanrahan in their series of work [5], [9], [10] introduced a signal processing framework for inverse rendering, in which a reflectance function is treated simply as a *spherical convolution* of a lighting function with the *reflection kernel* (BRDF multiplied by the *half-cosine*) and scaled by the albedos. This allows us to relate a reflectance function of an object to the lighting and the reflection kernel in terms of their SH 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 kernels often vary slowly). Based on this framework, the authors developed various algorithms for inverse rendering, most of which are under the form of an alternating optimization in which no global solution is guaranteed. Furthermore, they only considered homogeneous objects with no texture, and thus ignoring the local scalings by the albedos of the reflection. Our method deals with nonuniform albedos and can give exact solutions in the noiseless case.

Basri and Jacobs in an independent work [6] discovered a similar result for the special case of Lambertian (or *diffuse*) surfaces whose reflection kernels are simply the half-cosine function which contains mostly low-frequency components. As a result, it was shown in [6] that any reflectance function of a Lambertian *convex* homogeneous object with distant light sources can be well-approximated by its first 9 terms of the SH series. In other words, all the images of such an object live close to a 9-dimensional linear subspace spanned by the first 9 SHs associated with the geometry of the object. Basri *et al*., in the subsequent work [11] on *photometric stereo* (PS), matricized this key observation to formulate the inverse rendering as a matrix factorization which is similar to ours. In the PS problem, however, the albedos are often assumed to be uniform, and the scene geometry is unknown and to be reconstructed. In contrast, our work assumes known geometry and the nonuniform albedos are to be recovered. The authors of [11] proposed a matrix factorization based on the SVD followed by a low-rank approximation to estimate the shape of the object. This method is very similar to Tomasi–Kanade factorization [12], that has been widely used in *structure from motion*, a classical problem in computer vision. However, this type of factorization is only within a linear ambiguity that cannot be fully resolved without any further prior knowledge. In particular, because the set of all possible images of an object, called its *illumination cone*, is invariant to any 3×3 linear transformations of the object's geometry [13], the 9-D subspace of SHs, and hence the factorization is approximately the same for different objects related by a linear transformation. If using only the first 4 SHs in the approximation, the linear ambiguity of the factorization can be reduced to a Lorentz transformation [14] with 7 degrees of freedom, by taking into account the unit norm of surface normals. Del Bue *et al*. [15] presented a bilinear factorization method to solve the PS problem, where one of the factors is constrained to lie on a specific manifold. The formulation decouples the core bilinear aspect from the manifold specificity and can be solved using Augmented Lagrange Multipliers. However, the manifold projector based on unit norm of surface normals only permits the use of first order SHs.

More recently, Nießner, Zollhöfer, Wu *et al*. [16]–[18] introduced several methods for real-time inverse rendering from color plus depth data, in which the geometry is obtained from depth maps and some geometry refinement procedure. Unlike our work, these papers focus more on the practical side of the inverse rendering including geometry acquisition. However, similar to PS literature, these methods assume uniform albedos while estimating the lightings. In contrast, our paper puts more emphasis on the theoretical side of the matrix factorization problem where the lightings and albedos are estimated simultaneously using subspace methods.

For glossy objects when the Lambertian assumption is no longer hold, BRDF recovery has also been widely investigated in the computer graphics community [19]–[22]. In a recent work, Goldman *et al*. [23] estimated both geometry and BRDF from a small number of photographs. However, all of these methods for BRDF acquisition heavily rely on a controlled lighting setup in which the light can be easily calibrated. In this paper, we assume Lambertian objects with constant BRDF. This strict assumption allows us to rigorously tackle the inverse rendering under arbitrary general lighting conditions.

III. Preliminaries

This section summarizes the signal-processing framework [1], [6] that relates a reflectance function to the lighting and reflection kernel as a spherical convolution in space-domain. The relation can be transformed into frequency-domain via spherical harmonic expansions.

A. Reflection as Convolution

Throughout this paper, we restrict our interest to the illumination of Lambertian convex objects with distant light sources. The convexity assumption makes sure that there are no shadowing and inter-reflection which are very hard to analyze. Likewise, the Lambertian assumption implies that

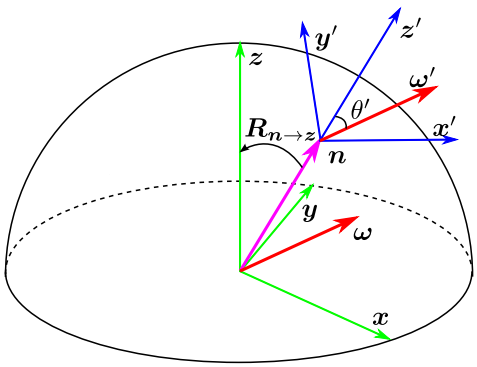


Fig. 1. Global and local coordinate systems. Green vectors define the global Cartesian coordinate system $\boldsymbol{xyz}$. Blue vectors define a local coordinate system $\boldsymbol{x'y'z'}$ associated with the normal $\boldsymbol{n}$ such that $\boldsymbol{z'} = \boldsymbol{n}$. The light vector in red is labeled $\boldsymbol{\omega}$ w.r.t. global coordinates and $\boldsymbol{\omega'}$ w.r.t. local coordinates; θ' is the polar angle of $\boldsymbol{\omega'}$. The lighting function $L(\boldsymbol{\omega})$ is given in global, whereas the reflection kernel $K(\boldsymbol{\omega'})$ is given in local coordinates. The two coordinates are related by the rotation that transfers the normal $\boldsymbol{n}$ (or $\boldsymbol{z'}$) to the polar vector $\boldsymbol{z}$: $\boldsymbol{\omega'} = \boldsymbol{R_{n\to z}\omega}$.

the reflection does not depend on the viewing angle, and so significantly simplifies our lighting model. We want to remark that, although assumed to be distant, the light source can be arbitrarily complex but not necessarily a point source.

In the following analysis, we refer to Fig. 1 that illustrates the *global* and *local* coordinate systems associated with a location on the surface of the object. Let $\boldsymbol{p}$ be a point on the surface with normal vector $\boldsymbol{n}$, and $L(\boldsymbol{\omega})$ be the lighting function that defines the incident radiance at angle $\boldsymbol{\omega}$ w.r.t. some global coordinate system $\boldsymbol{xyz}$. The reflected radiance (or *reflectance function*) $Y(\boldsymbol{p}, \boldsymbol{n})$ at location $\boldsymbol{p}$ and normal $\boldsymbol{n}$, is then given by integrating the reflections over all possible light directions. In particular,

$$Y(\boldsymbol{p}, \boldsymbol{n}) = \rho(\boldsymbol{p}) \int_{\Omega} L(\boldsymbol{\omega}) \cos(\theta') d\boldsymbol{\omega}', \tag{1}$$

where $\rho(\boldsymbol{p})$ is the albedo at location $\boldsymbol{p}$; $\boldsymbol{\omega'}$ is the light direction w.r.t. the local coordinate system $\boldsymbol{x'y'z'}$ obtained by rotating $\boldsymbol{xyz}$ so that $\boldsymbol{z'} = \boldsymbol{n}$; θ' is the corresponding polar angle of $\boldsymbol{\omega'}$; and the integrating region Ω is the upper-hemisphere of the unit sphere. The integral in (1) can be extended to the whole unit sphere S as

$$Y(\boldsymbol{p}, \boldsymbol{n}) = \rho(\boldsymbol{p}) \int_{S} L(\boldsymbol{\omega}) K(\boldsymbol{\omega'}) d\boldsymbol{\omega}', \tag{2}$$

where $K(\boldsymbol{\omega'}) \triangleq \max(\cos\theta', 0)$ is a *half-cosine* function of the local polar angle θ'. We will later refer to $K(\boldsymbol{\omega'})$ as the *Lambertian kernel*. Note that, although written as $K(\boldsymbol{\omega'})$ for general purpose, this kernel is actually a 1-D function that depends only on the polar angle of the incident direction in local coordinates.

By definition, the global and local coordinates are related by a 3D rotation: $\boldsymbol{\omega'} = \boldsymbol{R_{n\to z}\omega}$, where $\boldsymbol{R_{n\to z}}$ is the rotation matrix that brings $\boldsymbol{n} = \boldsymbol{z'}$ back to $\boldsymbol{z}$. Substituting this relation into (2), we arrive at

$$Y(\boldsymbol{p}, \boldsymbol{n}) = \rho(\boldsymbol{p}) \int_{S} L(\boldsymbol{\omega}) K(\boldsymbol{R_{n\to z}\omega}) d\boldsymbol{\omega}, \tag{3}$$

We can easily recognize the integral in (3) as a *spherical convolution*, in which the translation operator in a classical circular convolution is replaced with a rotation operator. We adopt the notation $\circledast_s$ for spherical convolution and write

$$(L \circledast_s K)(\boldsymbol{n}) \triangleq \int_{S} L(\boldsymbol{\omega}) K(\boldsymbol{R_{n\to z}\omega})\, d\boldsymbol{\omega}. \tag{4}$$

Similar to circular convolutions, the above spherical convolution defines a *linear rotation-invariant* (LRI) system with impulse response $K(\boldsymbol{\omega})$.

To summarize the above analysis, the reflectance function is now a spherical convolution of the lighting and Lambertian kernel scaled by albedos, i.e.,

$$Y(\boldsymbol{p}, \boldsymbol{n}) = \rho(\boldsymbol{p})(L \circledast_s K)(\boldsymbol{n}). \tag{5}$$

B. Spherical Harmonics as Basis

Now we want to look at the reflection equation (5) in frequency-domain because the energy of the Lambertian kernel $K(\boldsymbol{\omega'})$ is compacted 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, whose formulas are given by

$$S_{m,n}(\boldsymbol{\omega}) = N_{m,n} \cdot P_{m,n}(\cos\theta)\, e^{jn\varphi}, \tag{6}$$

where $j = \sqrt{-1}$; θ and φ are respectively the polar and azimuthal angles of $\boldsymbol{\omega}$: $\theta = \arccos(z)$, $\varphi = \arctan(y/x)$; $P_{m,n}(\cdot)$ is the *associated Legendre function*; and $N_{m,n}$ is the normalized term.[1]

Similarly to Fourier series of functions on a circle, every function $F(\boldsymbol{\omega})$ on the unit sphere has a spherical harmonic (SH) expansion

$$F(\boldsymbol{\omega}) = \sum_{m=0}^{\infty} \sum_{n=-m}^{m} f_{m,n} \cdot S_{m,n}(\boldsymbol{\omega}),$$

where the SH coefficients $f_{m,n}$ are given by

$$f_{m,n} = \int_S F(\boldsymbol{\omega}) S^*_{m,n}(\boldsymbol{\omega})\, d\boldsymbol{\omega}.$$

The feature that makes spherical harmonics similar to Fourier basis on a circle is that they are eigensignals of LRI systems.[2] This key property maps a spherical convolution in space-domain to a multiplication in frequency-domain. Particularly, let the SH coefficients of L, K and $L \circledast_s K$ be $\{\ell_{m,n}\}$, $\{k_{m,n}\}$, and $\{x_{m,n}\}$, respectively, then it was shown in [5] and [6] that

$$x_{m,n} = \sqrt{\frac{4\pi}{2m+1}}\, k_{m,0}\, \ell_{m,n}, \quad m \geq 0, |n| \leq m \tag{7}$$

Moreover, it was also shown in [6] that the convolution $L \circledast_s K$ can be well-approximated using first few terms in its SH expansion, namely

$$(L \circledast_s K)(\boldsymbol{n}) \approx \sum_{m'=0}^{M'-1} \sum_{n'=-m'}^{m'} x_{m',n'} \cdot S_{m',n'}(\boldsymbol{n}),$$

[1] See [6] for specific formulas of $P_{m,n}(\cdot)$ and $N_{m,n}$.

[2] See a special case of the Funk-Hecke theorem [6, Th. 1]. Note that since $K(\theta')$ is a 1-D function of the polar angle, its SH coefficients $k_{m,n}$ are nonzero only for $n = 0$.

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

$$Y(\boldsymbol{p}, \boldsymbol{n}) \approx \rho(\boldsymbol{p}) \sum_{m=1}^{M} x_m \cdot S_m(\boldsymbol{n}), \tag{8}$$

where $M = M'^2$, and the double index (m', n') was converted to the single index $m = m'^2 + m' + n' + 1$. In words, this approximation says that, all the reflectance functions of a Lambertian textureless object live around a low-dimensional linear subspace spanned by the first few SHs evaluated at the normal vectors of the object.

It is important to note that the SHs presented above are complex-valued (for analytical simplicity). In experiments, however, we can use the real-valued SHs which are real and imaginary parts of the above functions. From now on, we assume that the SHs, $\{S_m(\boldsymbol{\omega})\}_{m \geq 1}$, are real-valued, where the specific forms for the first 9 of them at $\boldsymbol{\omega} = (x, y, z)$, for $x^2 + y^2 + z^2 = 1$, are given below [6, eq. (7)]

$$\begin{aligned} &S_1 = \frac{1}{\sqrt{4\pi}}, \quad S_2 = \sqrt{\frac{3}{4\pi}}x, \quad S_3 = \sqrt{\frac{3}{4\pi}}y, \quad S_4 = \sqrt{\frac{3}{4\pi}}z, \\ &S_5 = \sqrt{\frac{15}{4\pi}}xy, \quad S_6 = \sqrt{\frac{15}{4\pi}}yz, \quad S_7 = \sqrt{\frac{5}{16\pi}}(3z^2 - 1), \\ &S_8 = \sqrt{\frac{15}{4\pi}}zx, \quad S_9 = \sqrt{\frac{15}{16\pi}}(x^2 - y^2). \end{aligned} \tag{9}$$

IV. Problem Statement

A. 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 $\boldsymbol{p}_i$ be the location on the surface corresponding to pixel i on each image, for $i = 1, \ldots, N$. From (8), the intensity at pixel i of image j can be approximated by

$$Y_j(\boldsymbol{p}_i, \boldsymbol{n}_i) \approx \rho(\boldsymbol{p}_i) \sum_{m=1}^{M} x_{m,j} \cdot S_m(\boldsymbol{n}_i), \tag{10}$$

for $1 \leq i \leq N$, $1 \leq j \leq J$. We can put all equations in (10) into a single matrix form by defining matrices $\boldsymbol{Y} \in \mathbb{R}^{N \times J}$; $\boldsymbol{\Phi} \in \mathbb{D}^N$ (the set of all size-N diagonal matrices); $\boldsymbol{X} \in \mathbb{R}^{M \times J}$; and $\boldsymbol{S} \in \mathbb{R}^{N \times M}$ as

$$\boldsymbol{Y}_{ij} = Y_j(\boldsymbol{p}_i), \quad \boldsymbol{\Phi}_{ii} = \rho(\boldsymbol{p}_i), \quad \boldsymbol{X}_{mj} = x_{m,j},$$

and

$$\boldsymbol{S}_{im} = S_m(\boldsymbol{n}_i), \tag{11}$$

for $1 \leq i \leq N, 1 \leq j \leq J, 1 \leq m \leq M$, and $\boldsymbol{M}_{ij}$ denotes the entry of matrix $\boldsymbol{M}$ at row i and column j. With these notations, (10) becomes

$$\boldsymbol{Y} \approx \boldsymbol{\Phi S X}, \tag{12}$$

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

B. The Matrix Factorization

For analytical purpose, we first assume the noiseless case when the approximation in (12) is replaced with exact equation

$$\boldsymbol{Y} = \boldsymbol{\Phi S X}, \tag{13}$$

In light of (13), the inverse rendering problem becomes a matrix factorization in which the image matrix $\boldsymbol{Y} \in \mathbb{R}^{N \times J}$ and SH matrix $\boldsymbol{S} \in \mathbb{R}^{N \times M}$ are known (can be computed from given images and 3D model of the object); the albedo matrix $\boldsymbol{\Phi} \in \mathbb{D}^N$ and lighting matrix $\boldsymbol{X} \in \mathbb{R}^{M \times J}$ are to be recovered. Once the albedo and lighting matrices are reconstructed, say $\hat{\boldsymbol{\Phi}}$ and $\hat{\boldsymbol{X}}$, the forward rendering becomes an obvious matrix multiplication

$$\boldsymbol{y}_{\text{new}} = \hat{\boldsymbol{\Phi}} \boldsymbol{S} \boldsymbol{x}_{\text{new}},$$

where $\boldsymbol{y}_{\text{new}}$ corresponds to a novel image under a virtual lighting generated by a vector $\boldsymbol{x}_{\text{new}}$. The rest of the paper is dedicated to the problem of solving $(\boldsymbol{\Phi}, \boldsymbol{X})$ from (13) given $(\boldsymbol{Y}, \boldsymbol{S})$, which will be referred to as *the matrix factorization*.

We first note this matrix factorization must always have a solution $(\boldsymbol{\Phi}_0, \boldsymbol{X}_0) \in \mathbb{D}^N \times \mathbb{R}^{M \times J}$, where $\mathbb{D}^N$ denotes the set of all $N \times N$ diagonal matrices, as the image matrix $\boldsymbol{Y}$ is collected according to the above lighting model. The recovery of $(\boldsymbol{\Phi}, \boldsymbol{X})$ is possible only if $(\boldsymbol{\Phi}_0, \boldsymbol{X}_0)$ is the unique solution of the matrix factorization, within a scaling. Therefore, we need to assume further that: (1) $\boldsymbol{\Phi}_0 \in \mathbb{D}^N_*$, or all diagonal entries of $\boldsymbol{\Phi}_0$ are nonzero, and (2) $\boldsymbol{X}_0 \in \mathbb{R}^{M \times J}_{\text{full}}$, or $\boldsymbol{X}_0$ has full row rank. The need for these assumptions will be justified later in the proof of uniqueness. Here, we just want to make a few remarks on them. The first assumption that all albedos are nonzero, does not restrict ourselves because the zero-intensity pixels corresponding to zero albedos can be masked out of the equations. Likewise, the second assumption that $\boldsymbol{X}_0$ has full row rank is reasonable and often made in classical algorithms for array signal processing such as MUSIC (MUltiple SIgnal Classification) [24] and ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques) [25]. It essentially means that the images are to be taken under diversified lighting conditions. This assumption also implicitly requires that the number of images J must be greater than the dimension M of the approximation subspace.

V. Noiseless Case: Uniqueness and Exact Solution

We start with the noiseless case when (13) holds, ignoring both the approximation error and measurement noise. The first question one should ask when dealing with an inverse problem is when the recovery is unique. This section provides a necessary and sufficient condition on the SH matrix $\boldsymbol{S}$ such that the matrix factorization (13) is unique (up to a scaling factor), and an SVD-based method for recovery.

A. Uniqueness

Before stating the condition for uniqueness, we introduce some new notions of matrix full rank (FR) which will be useful later. The results of this subsection might be also of independent interests. Consider a tall matrix $\boldsymbol{S} \in \mathbb{R}^{N \times M}$

with $N > M$. Let $\mathcal{N}$ denote the set $\{1, 2, \ldots, N\}$. For any subset $\mathcal{J}$ of $\mathcal{N}$, let $\mathcal{J}^c$ be the complement of $\mathcal{J}$ in $\mathcal{N}$, and let $\boldsymbol{S}_{\mathcal{J}}$ be the submatrix constituted from the rows of $\boldsymbol{S}$ indexed by $\mathcal{J}$. We say $\mathcal{J}$ is a nontrivial subset of $\mathcal{N}$ if $\mathcal{J} \neq \mathcal{N}$ and $\mathcal{J} \neq \emptyset$.

Definition 1 ([26], [27]): The Kruskal rank of $\boldsymbol{S}$, denoted by $\text{krank}(\boldsymbol{S})$ *is the maximal number k such that any k rows of $\boldsymbol{S}$ are linearly independent.*[3] *$\boldsymbol{S}$ is said to have* Kruskal full rank (KFR) *if* $\text{krank}(\boldsymbol{S}) = M$.

Definition 2: $\boldsymbol{S}$ is said to have strong full rank (SFR) *if it has no zero rows and* $\text{rank}(\boldsymbol{S}_{\mathcal{N}\setminus\{i\}}) = M$ *for all* $i \in \mathcal{N}$.

Definition 3: $\boldsymbol{S}$ is said to have nonseparable full rank (NSFR) *if it has no zero rows, and* $\text{rank}(\boldsymbol{S}) = M$, *and there does not exist a nontrivial subset $\mathcal{J}$ of $\mathcal{N}$ such that* $\text{rank}(\boldsymbol{S}_{\mathcal{J}}) + \text{rank}(\boldsymbol{S}_{\mathcal{J}^c}) = M$.

In words, a matrix has NSFR if it has full column rank and its rows cannot be separated into two groups whose ranks add to the rank of the matrix. That justifies the term "nonseparable full rank." The following result gives another description of NSFR.

Proposition 1: $\boldsymbol{S}$ has NSFR if and only if $\text{rank}(\boldsymbol{S}_{\mathcal{J}}) + \text{rank}(\boldsymbol{S}_{\circ\mathcal{J}}) \geq M + 1$, *for every nontrivial subset $\mathcal{J}$ of $\mathcal{N}$.*

Proof: Suppose that $\boldsymbol{S}$ has NSFR, then for every nontrivial subset $\mathcal{J}$ of $\mathcal{N}$ we have

$$\text{rank}(\boldsymbol{S}_{\mathcal{J}}) + \text{rank}(\boldsymbol{S}_{\mathcal{J}^c}) \neq M.$$

On the other hand

$$\text{rank}(\boldsymbol{S}_{\mathcal{J}}) + \text{rank}(\boldsymbol{S}_{\mathcal{J}^c}) \geq \text{rank}(\boldsymbol{S}) = M.$$

Hence

$$\text{rank}(\boldsymbol{S}_{\mathcal{J}}) + \text{rank}(\boldsymbol{S}_{\mathcal{J}^c}) \geq M + 1, \quad \forall\, \emptyset \neq \mathcal{J} \subsetneq \mathcal{N} \quad (14)$$

completing the "only if" part. For the "if" part, suppose (14) holds, then it is obvious that $\text{rank}(\boldsymbol{S}_{\mathcal{J}}) + \text{rank}(\boldsymbol{S}_{\mathcal{J}^c}) \neq M$, for every nontrivial subset $\mathcal{J}$ of $\mathcal{N}$. We only need to show that $\boldsymbol{S}$ has no zero rows and $\text{rank}(\boldsymbol{S}) = M$. Indeed, for every $i \in \mathcal{N}$ we have

$$\text{rank}(\boldsymbol{S}_{\{i\}}) + \text{rank}(\boldsymbol{S}_{\mathcal{N}\setminus\{i\}}) \geq M + 1. \quad (15)$$

Since $\text{rank}(\boldsymbol{S}_{\{i\}}) \leq 1$ and $\text{rank}(\boldsymbol{S}_{\mathcal{N}\setminus\{i\}}) \leq \text{rank}(\boldsymbol{S}) \leq M$, (15) implies that $\text{rank}(\boldsymbol{S}_{\{i\}}) = 1$, and $\text{rank}(\boldsymbol{S}_{\mathcal{N}\setminus\{i\}}) = M$. From $\text{rank}(\boldsymbol{S}_{\{i\}}) = 1, \forall i \in \mathcal{N}$, we can deduce that $\boldsymbol{S}$ has no zero rows. Meanwhile, $\text{rank}(\boldsymbol{S}_{\mathcal{N}\setminus\{i\}}) = M$ implies that $\text{rank}(\boldsymbol{S}) = M$, completing the "if" part. □

The definition of NSFR is not very intuitive. The following fact gives us a sense of how strong NSFR is, in comparison to KFR and SFR, whose definitions are more intuitive.

Proposition 2: The following implications hold: KFR $\Rightarrow$ NSFR $\Rightarrow$ SFR $\Rightarrow$ FR. The reverse implications are generally not true. However, for $N = M + 1$, KFR, NSFR, and SFR are equivalent.

Proof: See Appendix A. □

Fig. 2 illustrates the four concepts of full rank in the relations described in Proposition 2. We would like to emphasize

[3]Kruskal rank is typically defined for fat matrices, requiring linear independence of any k columns. Here, we define it for tall matrices to conform to the context of the matrix factorization problem stated in Section IV.

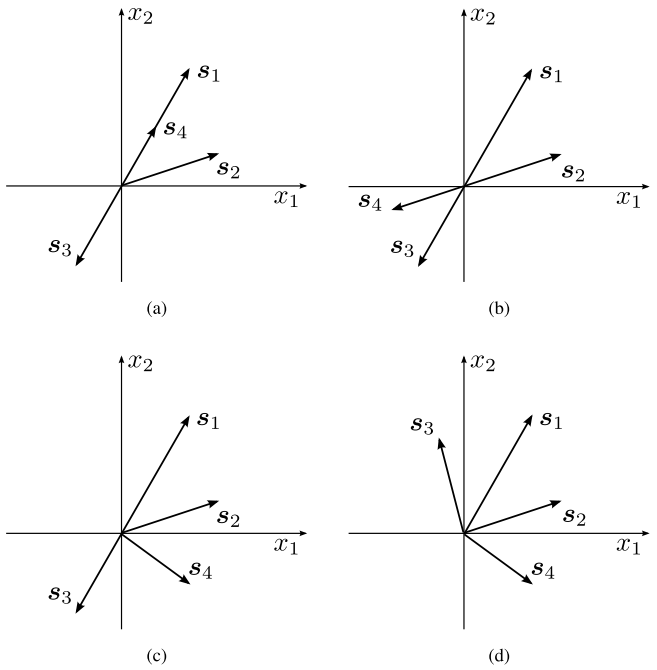


Fig. 2. Illustration of different notions of full rank for $M = 2$, $N = 4$ and $\boldsymbol{S} = [\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3, \boldsymbol{s}_4]^T$. (a) Full rank. (b) Strong full rank. (c) Nonseparable full rank. (d) Kruskal full rank.

that the number of rows of $\boldsymbol{S}$ in our matrix factorization problem is typically large, resulting in costly computations. The following result will be useful later when we want to reduce the size of the problem.

Proposition 3: $\boldsymbol{S}$ has NSFR if there exists a subset $\mathcal{J}$ of $\mathcal{N}$ such that $\boldsymbol{S}_{\mathcal{J}}$ has NSFR.

Proof: See Appendix B □

With the NSFR carefully defined, we are ready to state the necessary and sufficient condition for the uniqueness of the matrix factorization up to a scaling factor. This is certainly the best we can hope for because if $(\boldsymbol{\Phi}_0, \boldsymbol{X}_0)$ is a solution then so is $\left(\alpha\boldsymbol{\Phi}_0, \frac{1}{\alpha}\boldsymbol{X}_0\right)$, for any scalar $\alpha \neq 0$. In the remainder of the paper, we frequently use the $\odot$ symbol to denote point-wise operators between matrices. Namely,

$$\begin{aligned} \boldsymbol{A} \odot \boldsymbol{B} &\triangleq \{a_{ij} \cdot b_{ij}\}_{i,j}, \\ \boldsymbol{A} \odot /\boldsymbol{B} &\triangleq \{a_{ij}/b_{ij}\}_{i,j}, \\ \boldsymbol{A}\odot^{\alpha} &\triangleq \{a_{ij}^{\alpha}\}_{i,j}. \end{aligned}$$

Theorem 1: Given a pair of matrices $(\boldsymbol{Y}, \boldsymbol{S})$, the matrix factorization (13) *has no solutions $(\boldsymbol{\Phi}, \boldsymbol{X}) \in \mathbb{D}^N \times \mathbb{R}^{M\times J}$ other than $\left(\alpha\boldsymbol{\Phi}_0, \frac{1}{\alpha}\boldsymbol{X}_0\right)$, for some scalar $\alpha \neq 0$, if and only if $\boldsymbol{S}$ has NSFR.*

Proof: For brevity, we only show here the proof for the sufficiency part; the proof for the necessity part is in Appendix E.

Suppose $\boldsymbol{S}$ has NSFR and there exist $(\tilde{\boldsymbol{\Phi}}, \tilde{\boldsymbol{X}}) \in \mathbb{D}^N \times \mathbb{R}^{M\times J}$ such that

$$\boldsymbol{Y} = \boldsymbol{\Phi}_0\boldsymbol{S}\boldsymbol{X}_0 = \tilde{\boldsymbol{\Phi}}\boldsymbol{S}\tilde{\boldsymbol{X}}. \quad (16)$$

We will show that $(\tilde{\boldsymbol{\Phi}}, \tilde{\boldsymbol{X}}) = (\alpha\boldsymbol{\Phi}_0, \frac{1}{\alpha}\boldsymbol{X}_0)$ for some $\alpha \neq 0$. Assume conversely that $\tilde{\boldsymbol{\Phi}}$ is not a scaled version of $\boldsymbol{\Phi}_0$. We first note that, by definition, $\boldsymbol{S}$ has full column rank,

i.e., $\text{rank}(\boldsymbol{S}) = M$. Let $\boldsymbol{\Phi}_0 = \text{diag}(\boldsymbol{\varphi})$, and $\tilde{\boldsymbol{\Phi}} = \text{diag}(\tilde{\boldsymbol{\varphi}})$, where $\boldsymbol{\varphi} \in \mathbb{R}_*^N$ and $\tilde{\boldsymbol{\varphi}} \in \mathbb{R}^N$. Since $\boldsymbol{S}$ has no zero rows and $\boldsymbol{X}_0$ has full row rank, each row of the product matrix $\boldsymbol{S}\boldsymbol{X}_0$ must be different from zero vector. Furthermore, since $\varphi_i \neq 0, \forall i \in \mathcal{N}$, each row of $\boldsymbol{Y}$ must also be different from zero vector. From this and (16), we can deduce that $\tilde{\varphi}_i \neq 0, \forall i \in \mathcal{N}$.

It is then valid to put $\boldsymbol{\psi} = \boldsymbol{\varphi} \odot /\tilde{\boldsymbol{\varphi}}$, and $\boldsymbol{\Psi} = \text{diag}(\boldsymbol{\psi})$. From (16) we get

$$\boldsymbol{\Psi}\boldsymbol{S}\boldsymbol{X}_0 = \boldsymbol{S}\tilde{\boldsymbol{X}}. \tag{17}$$

Since $\boldsymbol{X}_0$ has full row rank, (17) is equivalent to

$$\boldsymbol{\Psi}\boldsymbol{S} = \boldsymbol{S}\tilde{\boldsymbol{X}}\boldsymbol{X}_0^{\dagger}, \tag{18}$$

where $\boldsymbol{X}_0^{\dagger} = \boldsymbol{X}_0^T(\boldsymbol{X}_0\boldsymbol{X}_0^T)^{-1}$ is the pseudo-inverse of $\boldsymbol{X}_0$. Consider the following matrix

$$\hat{\boldsymbol{S}} = [\boldsymbol{S}|\boldsymbol{\Psi}\boldsymbol{S}] = \begin{bmatrix} \boldsymbol{s}_1^T & \psi_1\boldsymbol{s}_1^T \\ \boldsymbol{s}_2^T & \psi_2\boldsymbol{s}_2^T \\ \cdots & \cdots \\ \boldsymbol{s}_N^T & \psi_N\boldsymbol{s}_N^T \end{bmatrix}, \tag{19}$$

where $\boldsymbol{s}_i^T \in \mathbb{R}^M$ be the i-th row vector of $\boldsymbol{S}$ for $i \in \mathcal{N}$. Before proceeding we need the following lemmas whose proofs can be found in Appendices C and D.

Lemma 1: $\text{rank}(\hat{\boldsymbol{S}}_{\mathcal{I}}) = \text{rank}(\boldsymbol{S}_{\mathcal{I}})$, *for all* $\mathcal{I} \subset \mathcal{N}$. *Especially,* $\text{rank}(\hat{\boldsymbol{S}}) = \text{rank}(\boldsymbol{S}) = M$.

Lemma 2: For any $\mathcal{I} \subset \mathcal{N}$, *if* $\mathcal{B}$ *is a subset of* $\mathcal{I}$ *such that* $\{\boldsymbol{s}_i\}_{i\in\mathcal{B}}$ *is a basis for* $\mathcal{R}(\boldsymbol{S}_{\mathcal{I}}^T)$ *then* $\{[\boldsymbol{s}_i^T|\psi_i\boldsymbol{s}_i^T]^T\}_{i\in\mathcal{B}}$ *is a basis for* $\mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{I}}^T)$.

Let us now continue with the proof of Thm. 1. The hypothesis that $\tilde{\boldsymbol{\Phi}}$ is not a scaled version of $\boldsymbol{\Phi}_0$ implies that elements of the set $\{\psi_i\}_{i=1}^N$ are not all equal. Thus, there must exist a nontrivial subset $\mathcal{J}$ of $\mathcal{N}$ such that $\psi_j = \theta, \forall j \in \mathcal{J}$ and $\psi_j \neq \theta, \forall j \in \mathcal{J}^c$, for some fixed $\theta \neq 0$.

Now that $\boldsymbol{S}$ has NSFR, from Proposition 1 we must have

$$\text{rank}(\boldsymbol{S}_{\mathcal{J}}) + \text{rank}(\boldsymbol{S}_{\mathcal{J}^c}) \geq M + 1. \tag{20}$$

Combining with Lemma 1 yields

$$\text{rank}(\hat{\boldsymbol{S}}_{\mathcal{J}}) + \text{rank}(\hat{\boldsymbol{S}}_{\mathcal{J}^c}) \geq M + 1. \tag{21}$$

Expressing this in terms of dimensions gives

$$\begin{aligned} M + 1 &\leq \dim\left(\mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{J}}^T)\right) + \dim\left(\mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{J}^c}^T)\right) \\ &= \dim\left(\mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{J}}^T) + \mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{J}^c}^T)\right) + \dim\left(\mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{J}}^T) \cap \mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{J}^c}^T)\right) \\ &= \dim\left(\mathcal{R}(\hat{\boldsymbol{S}}^T)\right) + \dim\left(\mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{J}}^T) \cap \mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{J}^c}^T)\right) \\ &= M + \dim\left(\mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{J}}^T) \cap \mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{J}^c}^T)\right). \end{aligned}$$

Therefore

$$\dim\left(\mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{J}}^T) \cap \mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{J}^c}^T)\right) \geq 1.$$

There thus exits a nonzero vector $\hat{\boldsymbol{s}}$ in $\mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{J}}^T) \cap \mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{J}^c}^T)$. Let $\mathcal{K} \subset \mathcal{J}$ and $\mathcal{L} \subset \mathcal{J}^c$ such that $\{\boldsymbol{s}_k\}_{k\in\mathcal{K}}$ is a basis of $\mathcal{R}(\boldsymbol{S}_{\mathcal{J}}^T)$ and $\{\boldsymbol{s}_\ell\}_{\ell\in\mathcal{L}}$ is a basis of $\mathcal{R}(\boldsymbol{S}_{\mathcal{J}^c}^T)$. It follows from Lemma 2 that $\{[\boldsymbol{s}_k^T|\theta\boldsymbol{s}_k^T]^T\}_{k\in\mathcal{K}}$ is a basis of $\mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{J}}^T)$ and $\{[\boldsymbol{s}_\ell^T|\psi_\ell\boldsymbol{s}_\ell^T]^T\}_{\ell\in\mathcal{L}}$ is a basis of $\mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{J}^c}^T)$. Then $\hat{\boldsymbol{s}}$ can be represented in two different ways as

$$\hat{\boldsymbol{s}} = \sum_{k\in\mathcal{K}} \alpha_k[\boldsymbol{s}_k^T|\theta\boldsymbol{s}_k^T]^T = \sum_{\ell\in\mathcal{L}} \beta_\ell[\boldsymbol{s}_\ell^T|\psi_\ell\boldsymbol{s}_\ell^T]^T. \tag{22}$$

Hence

$$\sum_{k\in\mathcal{K}} \alpha_k\boldsymbol{s}_k = \sum_{\ell\in\mathcal{L}} \beta_\ell\boldsymbol{s}_\ell = \sum_{\ell\in\mathcal{L}} \frac{\beta_\ell\psi_\ell}{\theta}\boldsymbol{s}_\ell. \tag{23}$$

By the linear independence of $\{\boldsymbol{s}_\ell\}_{\ell\in\mathcal{L}}$ we have

$$\beta_\ell = \beta_\ell\psi_\ell/\theta, \quad \forall\ell \in \mathcal{L}. \tag{24}$$

Since $\hat{\boldsymbol{s}} \neq \boldsymbol{0}$, it follows from (22) that there exists a $\ell^* \in \mathcal{L}$ such that $\beta_{\ell^*} \neq 0$. This together with (24) yields $\psi_{\ell^*} = \theta$, contradicting to the above hypothesis that $\psi_j \neq \theta$ for all $j \in \mathcal{J}^c$. Therefore $\{\psi_i\}_{i=1}^N$ must be all equal, i.e. $\tilde{\boldsymbol{\Phi}} = \alpha\boldsymbol{\Phi}$ for some $\alpha \neq 0$. From (16), because $\boldsymbol{S}$ has full column rank, it must be that

$$\tilde{\boldsymbol{X}} = \boldsymbol{S}^{\dagger}\tilde{\boldsymbol{\Phi}}^{-1}\boldsymbol{Y} = \frac{1}{\alpha}\boldsymbol{S}^{\dagger}\boldsymbol{\Phi}_0^{-1}\boldsymbol{Y} = \frac{1}{\alpha}\boldsymbol{X}_0, \tag{25}$$

where $\boldsymbol{S}^{\dagger} = (\boldsymbol{S}^T\boldsymbol{S})^{-1}\boldsymbol{S}^T$ is the pseudo-inverse of $\boldsymbol{S}$. This completes the proof of the sufficiency part. □

B. Exact Solution

This subsection is to solve the matrix factorization (13) given that it has unique solution, i.e. $\boldsymbol{S}$ has NSFR by Thm. 1. Note that since $\boldsymbol{S}$ has full column rank, once $\boldsymbol{\Phi}$ is known, $\boldsymbol{X}$ can be uniquely recovered by

$$\boldsymbol{X} = \boldsymbol{S}^{\dagger}(\boldsymbol{\Phi}^{-1}\boldsymbol{Y}). \tag{26}$$

Therefore, we can focus on finding $\boldsymbol{\Phi}$ which has a diagonal structure. We propose an SVD-based algorithm to exactly recover $\boldsymbol{\Phi}$. The algorithm is based on the following theorem, our second main result in the noiseless case.

Theorem 2: $\boldsymbol{\Phi} = \text{diag}(\boldsymbol{\varphi})$ *is a solution to* (13) *if and only if* $\boldsymbol{z} \stackrel{\Delta}{=} \boldsymbol{\varphi}\odot^{-1}$ *is a nontrivial solution to*

$$\left((\boldsymbol{I} - \boldsymbol{S}\boldsymbol{S}^{\dagger}) \odot (\boldsymbol{Y}\boldsymbol{Y}^T)\right)\boldsymbol{z} = \boldsymbol{0}. \tag{27}$$

Proof: Let $\boldsymbol{z} \stackrel{\Delta}{=} \boldsymbol{\varphi}\odot^{-1}$, then it is clear that $\boldsymbol{\varphi}$ solves $\boldsymbol{Y} = \text{diag}(\boldsymbol{\varphi})\boldsymbol{S}\boldsymbol{X}$ if and only if $\boldsymbol{z}$ solves

$$\text{diag}(\boldsymbol{z})\boldsymbol{Y} = \boldsymbol{S}\boldsymbol{X}. \tag{28}$$

It is equivalent to that each column of the LHS is in the range space of $\boldsymbol{S}$, i.e.

$$(\text{diag}(\boldsymbol{z})\boldsymbol{Y})_j \in \mathcal{R}(\boldsymbol{S}), \quad \forall j = 1, \ldots, J. \tag{29}$$

Note that a vector lies in some subspace if and only if its orthogonal projection onto the subspace is equal to itself. Thus, (29) can be written equivalently as

$$\boldsymbol{P}_{\mathcal{R}(\boldsymbol{S})}(\text{diag}(\boldsymbol{z})\boldsymbol{Y})_j = (\text{diag}(\boldsymbol{z})\boldsymbol{Y})_j, \quad \forall j = 1, \ldots, J.$$

Replacing the projection matrix $\boldsymbol{P}_{\mathcal{R}(\boldsymbol{S})}$ by $\boldsymbol{S}\boldsymbol{S}^{\dagger}$, we arrive at

$$(\boldsymbol{I} - \boldsymbol{S}\boldsymbol{S}^{\dagger})(\text{diag}(\boldsymbol{z})\boldsymbol{Y})_j = \boldsymbol{0}, \quad \forall j = 1, \ldots, J. \tag{30}$$

Algorithm 1 SVD-Based Factorization

Inputs: $\boldsymbol{Y}$, $\boldsymbol{S}$
Outputs: Estimates $\hat{\boldsymbol{\Phi}}, \hat{\boldsymbol{X}}$ of $\boldsymbol{\Phi}, \boldsymbol{X}$

1. Compute $\boldsymbol{M} = (\boldsymbol{I} - \boldsymbol{S}\boldsymbol{S}^{\dagger}) \odot (\boldsymbol{Y}\boldsymbol{Y}^{T})$.
2. Compute the eigenvalue decomposition of $\boldsymbol{M}$.
3. Let $\boldsymbol{z}_{\min}$ be the eigenvector associated with the smallest eigenvalue of $\boldsymbol{M}$.
4. Return $\hat{\boldsymbol{\Phi}} = \operatorname{diag}(\boldsymbol{z}_{\min} \odot^{-1})$; $\hat{\boldsymbol{X}} = \boldsymbol{S}^{\dagger} \operatorname{diag}(\boldsymbol{z}_{\min}) \boldsymbol{Y}$.

Let $\boldsymbol{Q} \stackrel{\Delta}{=} \boldsymbol{I} - \boldsymbol{S}\boldsymbol{S}^{\dagger}$. For each $j = 1, \ldots, J$, let $\boldsymbol{D}_j \stackrel{\Delta}{=} \operatorname{diag}((\boldsymbol{Y})_j)$, then we can rewrite (30) as

$$\boldsymbol{Q}\boldsymbol{D}_j \boldsymbol{z} = \boldsymbol{0}, \quad \forall j = 1, \ldots, J. \tag{31}$$

Define matrix $\boldsymbol{A}$ by

$$\boldsymbol{A} \stackrel{\Delta}{=} \begin{bmatrix} \boldsymbol{Q}\boldsymbol{D}_1 \\ \vdots \\ \boldsymbol{Q}\boldsymbol{D}_J \end{bmatrix}. \tag{32}$$

With this notation, we can compact all J equations in (31) into a single one as

$$\boldsymbol{A}\boldsymbol{z} = \boldsymbol{0}, \tag{33}$$

which is in turn equivalent to

$$\boldsymbol{A}^T \boldsymbol{A}\boldsymbol{z} = \boldsymbol{0}. \tag{34}$$

We can complete the proof at this point by using the following lemma whose proof is given in Appendix F.

Lemma 3: For matrix $\boldsymbol{A}$ *defined in* (32)*, we have*

$$\boldsymbol{A}^T \boldsymbol{A} = (\boldsymbol{I} - \boldsymbol{S}\boldsymbol{S}^{\dagger}) \odot (\boldsymbol{Y}\boldsymbol{Y}^T).$$

□

Thm. 2 naturally gives rise to Alg. 1, in which we solve (27) for $\boldsymbol{z}$ by picking the eigenvector associated with the smallest eigenvalue of the positive definite matrix $(\boldsymbol{I} - \boldsymbol{S}\boldsymbol{S}^{\dagger}) \odot (\boldsymbol{Y}\boldsymbol{Y}^T)$. Once $\boldsymbol{\Phi}$ is found, $\boldsymbol{X}$ can be computed using (26).

The following corollary is very useful in checking whether a matrix $\boldsymbol{S}$ has NSFR. It can be deduced directly from Thm. 1 and Thm. 2 by setting $\boldsymbol{\Phi}_0 = \boldsymbol{I}_N$, $\boldsymbol{X}_0 = \boldsymbol{I}_M$, and $\boldsymbol{Y} = \boldsymbol{S}$.

Corollary 1: A matrix $\boldsymbol{S} \in \mathbb{R}^{N \times M}$ *has NSFR if and only if*

$$\operatorname{rank}\Big((\boldsymbol{I} - \boldsymbol{S}\boldsymbol{S}^{\dagger}) \odot (\boldsymbol{S}\boldsymbol{S}^T)\Big) = N - 1.$$

We conclude this subsection by giving a few comments on this result. Without identifying the NSFR with the uniqueness of the corresponding matrix factorization, we can hardly see the connection between Definition 3 and Corollary 1. When fixing M and growing N, checking if an $N \times M$ matrix has NSFR 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.

VI. Noisy Case: Algorithms

It is important to note that the equation in (13) can never be the case due to the approximation and measurement errors. Taking into account these noises, (13) should be remodeled as

$$\boldsymbol{Y} = \boldsymbol{\Phi}(\boldsymbol{S}\boldsymbol{X} + \boldsymbol{W}_1) + \boldsymbol{W}_2, \tag{35}$$

where $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ are respectively the approximation and measurement noises. Here, we assume both of them are deterministic unknown quantities. It is natural to pose the following least squares problem

$$\min \left\| \begin{bmatrix} \boldsymbol{W}_1 \\ \boldsymbol{W}_2 \end{bmatrix} \right\|_{\mathrm{F}} \text{ s.t. } \boldsymbol{Y} = \boldsymbol{\Phi}(\boldsymbol{S}\boldsymbol{X} + \boldsymbol{W}_1) + \boldsymbol{W}_2, \quad \text{(P0)}$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm. Solving (P0) in the presence of both $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ is generally hard. In the sequel, we only tackle the two special cases of this noisy problem when either $\boldsymbol{W}_1$ or $\boldsymbol{W}_2$ is negligible.

A. No Measurement Error

If there is no measurement error, i.e. $\boldsymbol{W}_2 = \boldsymbol{0}$, (35) can be rewritten as

$$\boldsymbol{\Phi}^{-1}\boldsymbol{Y} = \boldsymbol{S}\boldsymbol{X} + \boldsymbol{W}_1.$$

Therefore, (P0) becomes

$$\min_{\substack{\boldsymbol{\Phi} \in \mathbb{D}_*^N, \|\boldsymbol{\Phi}^{-1}\|_F = 1 \\ \boldsymbol{X} \in \mathbb{R}^{M \times J}}} \|\boldsymbol{\Phi}^{-1}\boldsymbol{Y} - \boldsymbol{S}\boldsymbol{X}\|_F, \quad \text{(P1)}$$

where the constraint $\|\boldsymbol{\Phi}^{-1}\|_F = 1$ is added to resolve the scaling ambiguity. The solution to this problem turns out to be identical to the exact solution in the noiseless case.

Theorem 3: The solution to the least squares problem (P1) is given by

$$\boldsymbol{\Phi}_{LS} = \operatorname{diag}(\boldsymbol{z}_{\min} \odot^{-1}), \tag{36}$$

and

$$\boldsymbol{X}_{LS} = \boldsymbol{S}^{\dagger} \operatorname{diag}(\boldsymbol{z}_{\min}) \boldsymbol{Y}, \tag{37}$$

where $\boldsymbol{z}_{\min}$ *is the eigenvector associated with the smallest eigenvalue of* $(\boldsymbol{I} - \boldsymbol{S}\boldsymbol{S}^{\dagger}) \odot (\boldsymbol{Y}\boldsymbol{Y}^T)$.

Proof: Fixing $\boldsymbol{\Phi} \in \mathbb{D}_*^N$ such that $\|\boldsymbol{\Phi}^{-1}\|_F = 1$, the least squares solution for $\boldsymbol{X}$ is given by

$$\begin{aligned} \boldsymbol{X}_{\mathrm{LS}}(\boldsymbol{\Phi}) &= \arg \min_{\boldsymbol{X} \in \mathbb{R}^{M \times J}} \|\boldsymbol{\Phi}^{-1}\boldsymbol{Y} - \boldsymbol{S}\boldsymbol{X}\|_F \\ &= \boldsymbol{S}^{\dagger}\boldsymbol{\Phi}^{-1}\boldsymbol{Y}. \end{aligned} \tag{38}$$

Substituting into (P1) and using the matrix notations in Sec. V we can solve for the least squares solution of $\boldsymbol{\Phi}$ as

$$\begin{aligned} \boldsymbol{\Phi}_{\mathrm{LS}} &= \arg \min_{\boldsymbol{\Phi} \in \mathbb{D}_*^N, \|\boldsymbol{\Phi}^{-1}\|_F = 1} \|\boldsymbol{\Phi}^{-1}\boldsymbol{Y} - \boldsymbol{S}\boldsymbol{X}_{LS}(\boldsymbol{\Phi})\|_F \\ &= \arg \min_{\boldsymbol{\Phi} \in \mathbb{D}_*^N, \|\boldsymbol{\Phi}^{-1}\|_F = 1} \|\boldsymbol{\Phi}^{-1}\boldsymbol{Y} - \boldsymbol{S}\boldsymbol{S}^{\dagger}\boldsymbol{\Phi}^{-1}\boldsymbol{Y}\|_F \\ &= \arg \min_{\boldsymbol{\Phi} \in \mathbb{D}_*^N, \|\boldsymbol{\Phi}^{-1}\|_F = 1} \|\boldsymbol{Q}\boldsymbol{\Phi}^{-1}\boldsymbol{Y}\|_F \\ &= \arg \min_{\boldsymbol{\Phi} \in \mathbb{D}_*^N, \|\boldsymbol{\Phi}^{-1}\|_F = 1} \|\operatorname{vec}(\boldsymbol{Q}\boldsymbol{\Phi}^{-1}\boldsymbol{Y})\|_2 \end{aligned} \tag{39}$$

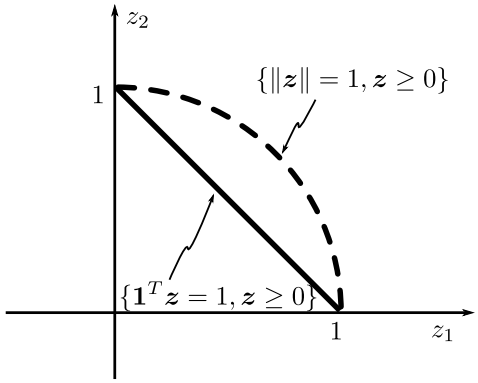


Fig. 3. Illustrating the convexification of the feasible set of interest in $\mathbb{R}^2$. The original nonconvex feasible set, depicted by the dashed arc, is replaced with the chord connecting the two endpoints of the arc.

$$\begin{aligned} &= \arg \min_{\boldsymbol{\Phi}=\text{diag}(z\odot^{-1}),\|z\|_2=1} \left\| \begin{bmatrix} \boldsymbol{Q}\boldsymbol{D}_1 \\ \vdots \\ \boldsymbol{Q}\boldsymbol{D}_J \end{bmatrix} z \right\|_2 \\ &= \arg \min_{\boldsymbol{\Phi}=\text{diag}(z\odot^{-1}),\|z\|_2=1} \|\boldsymbol{A}\boldsymbol{z}\|_2 \\ &= \arg \min_{\boldsymbol{\Phi}=\text{diag}(z\odot^{-1}),\|z\|_2=1} \|\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{z}\|_2 \\ &= \text{diag}\left(z_{\min}\odot^{-1}\right), \end{aligned} \tag{40}$$

where the operator vec$(\cdot)$ in (39) stacks columns of a matrix into a big column vector; (40) follows from Lemma 3 and the definition of $z_{\min}$. Substituting (40) into (38) yields

$$\boldsymbol{X}_{\text{LS}} = \boldsymbol{S}^{\dagger}\,\text{diag}(z_{\min})\boldsymbol{Y},$$

completing the proof. □

With the aid of Thm. 3, Alg. 1 finds not only the exact solution in the noiseless case but also the least squares solution of problem (P1). We want to recall, however, that the diagonal entries of $\boldsymbol{\Phi}$ are the albedos of the object, and so must be all positive. It follows that the vector $\boldsymbol{z} = \boldsymbol{\varphi}\odot^{-1}$ must be pointwise positive as well. Therefore the positivity constraint on $\boldsymbol{z}$ should also be incorporated into the recovery. In order to do so, we first note that, finding the eigenvector corresponding to the smallest eigenvalue of $\boldsymbol{M}$ is nothing but solving the optimization problem

$$\min \|\boldsymbol{M}\boldsymbol{z}\|_2 \quad \text{s.t.} \quad \|\boldsymbol{z}\|_2 = 1. \tag{41}$$

Now we can adjust (41) by adding the positivity constraint as

$$\min \|\boldsymbol{M}\boldsymbol{z}\|_2 \quad \text{s.t.} \quad \|\boldsymbol{z}\|_2 = 1 \text{ and } \boldsymbol{z} \geq 0. \tag{42}$$

Solving the optimization problem (42) is hard due to the nonconvexity of the feasible set (See Fig. 3). The easiest way to convexify this problem is to replace the arc $\{\|\boldsymbol{z}\|_2 = 1, \boldsymbol{z} \geq 0\}$ on the unit sphere with its chord $\{\mathbf{1}^T\boldsymbol{z} = 1, \boldsymbol{z} \geq 0\}$. This means the problem is modified as

$$\min \|\boldsymbol{M}\boldsymbol{z}\|_2 \quad \text{s.t.} \quad \mathbf{1}^T\boldsymbol{z} = 1 \text{ and } \boldsymbol{z} \geq 0. \tag{43}$$

Solving (43) is now easy using some well-developed convex programming method. The above modification of Alg. 1 is summarized in Alg. 2 under the name optimization-based factorization.

Algorithm 2 Optimization-Based Factorization

Inputs: $\boldsymbol{Y}, \boldsymbol{S}$
Outputs: Estimates $\hat{\boldsymbol{\Phi}}, \hat{\boldsymbol{X}}$ of $\boldsymbol{\Phi}, \boldsymbol{X}$

1. Compute $\boldsymbol{M} = (\boldsymbol{I} - \boldsymbol{S}\boldsymbol{S}^{\dagger}) \odot (\boldsymbol{Y}\boldsymbol{Y}^T)$.
2. Use a convex programming method to find $\boldsymbol{z}_*$ that minimizes $\|\boldsymbol{M}\boldsymbol{z}\|_2$ s.t. $\mathbf{1}^T\boldsymbol{z} = 1$, $\boldsymbol{z} \geq 0$.
3. Return $\hat{\boldsymbol{\Phi}} = \text{diag}(\boldsymbol{z}_*\odot^{-1})$; $\hat{\boldsymbol{X}} = \boldsymbol{S}^{\dagger}\,\text{diag}(\boldsymbol{z}_*)\boldsymbol{Y}$.

B. No Approximation Error

If there is no approximation error, i.e. $\boldsymbol{W}_1 = \boldsymbol{0}$, (35) simplifies to

$$\boldsymbol{Y} = \boldsymbol{\Phi}\boldsymbol{S}\boldsymbol{X} + \boldsymbol{W}_2,$$

and so (P0) becomes

$$\min_{\boldsymbol{\Phi}\in\mathbb{D}_*^N, \boldsymbol{X}\in\mathbb{R}^{M\times J}} \|\boldsymbol{Y} - \boldsymbol{\Phi}\boldsymbol{S}\boldsymbol{X}\|_F. \tag{P2}$$

Following the same strategy for solving (P1), we first fix $\boldsymbol{\Phi}$ and solve for $\boldsymbol{X}$ as

$$\begin{aligned} \boldsymbol{X}_{\text{LS}}(\boldsymbol{\Phi}) &= (\boldsymbol{\Phi}\boldsymbol{S})^{\dagger}\boldsymbol{Y} = ((\boldsymbol{\Phi}\boldsymbol{S})^T(\boldsymbol{\Phi}\boldsymbol{S}))^{-1}(\boldsymbol{\Phi}\boldsymbol{S})^T\boldsymbol{Y} \\ &= (\boldsymbol{S}^T\boldsymbol{\Phi}^2\boldsymbol{S})^{-1}\boldsymbol{S}^T\boldsymbol{\Phi}\boldsymbol{Y}. \end{aligned} \tag{44}$$

Substituting into (P2) we get

$$\boldsymbol{\Phi}_{\text{LS}} = \min_{\boldsymbol{\Phi}\in\mathbb{D}_*^N} \|\boldsymbol{Y} - \boldsymbol{\Phi}\boldsymbol{S}(\boldsymbol{S}^T\boldsymbol{\Phi}^2\boldsymbol{S})^{-1}\boldsymbol{S}^T\boldsymbol{\Phi}\boldsymbol{Y}\|_F. \tag{45}$$

This nonlinear least squares problem seems to be intractable. Instead of solving (45), we can heuristically solve for $\boldsymbol{\Phi}$ and $\boldsymbol{X}$ using alternating least squares. In particular, (44) solves for $\boldsymbol{X}$ in terms of $\boldsymbol{\Phi}$. Conversely, if $\boldsymbol{X}$ is fixed, then $\boldsymbol{\Phi}_{\text{LS}}(\boldsymbol{X}) = \text{diag}(\hat{\boldsymbol{\varphi}})$ can be solved element-wise as

$$\hat{\varphi}_i = \arg\min_{\varphi_i\in\mathbb{R}} \|\boldsymbol{y}_i^T - \varphi_i\boldsymbol{s}_i^T\boldsymbol{X}\|_2 = \frac{\boldsymbol{s}_i^T\boldsymbol{X}\boldsymbol{y}_i}{\boldsymbol{s}_i^T\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{s}_i}, \tag{46}$$

for $i = 1, \ldots, N$, where $\boldsymbol{s}_i^T$ and $\boldsymbol{y}_i^T$ are i-th rows of $\boldsymbol{S}$ and $\boldsymbol{Y}$ respectively. If we insist in the positivity of albedos, it is clear that (46) should be modified as

$$\hat{\varphi}_i = \arg\min_{\varphi_i\geq\varepsilon} \|\boldsymbol{y}_i^T - \varphi_i\boldsymbol{s}_i^T\boldsymbol{X}\|_2 = \left[\frac{\boldsymbol{s}_i^T\boldsymbol{X}\boldsymbol{y}_i}{\boldsymbol{s}_i^T\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{s}_i}\right]_+, \tag{47}$$

where ε is an arbitrarily small positive number, and $[\cdot]_+$ is a short notation for $\max(\cdot, \varepsilon)$. Putting (44) and (47) together, we arrive at the so called alternating factorization, described in Alg. 3.

VII. Computational Issues

Implementation of Algs. 1 and 2 is costly due to the large size of the SVD/optimization problem. The size N of vector $\boldsymbol{z}$ is equal to the number of pixels in an image which is typically large. Nonetheless, we can reduce the size of the problem significantly by observing that

$$\boldsymbol{Y}_{\mathcal{J}} = \boldsymbol{\Phi}_{\mathcal{J}}\boldsymbol{S}_{\mathcal{J}}\boldsymbol{X}, \quad \forall\mathcal{J} \subset \mathcal{N}, \tag{48}$$

Algorithm 3 Alternating Factorization

Inputs: $\boldsymbol{Y}$, $\boldsymbol{S}$
Outputs: Estimates $\hat{\boldsymbol{\Phi}}, \hat{\boldsymbol{X}}$ of $\boldsymbol{\Phi}, \boldsymbol{X}$.

1. Initialize $\hat{\boldsymbol{\Phi}} = \boldsymbol{I}_N$.
2. Set $\hat{\boldsymbol{X}} = (\boldsymbol{S}^T \hat{\boldsymbol{\Phi}}^2 \boldsymbol{S})^{-1} \boldsymbol{S}^T \hat{\boldsymbol{\Phi}} \boldsymbol{Y}$.
3. Set $\hat{\varphi}_i = \left[\frac{\boldsymbol{s}_i^T \hat{\boldsymbol{X}} \boldsymbol{y}_i}{\boldsymbol{s}_i^T \hat{\boldsymbol{X}} \hat{\boldsymbol{X}}^T \boldsymbol{s}_i} \right]_+, \quad i = 1, \ldots, N.$
4. Repeat steps 2 and 3 until improvement in mean squared error is negligible.

where $\boldsymbol{\Phi}_{\mathcal{J}} = \mathrm{diag}(\{\varphi_i\}_{i \in \mathcal{J}})$. Now $\boldsymbol{\Phi}_{\mathcal{J}}$ and $\boldsymbol{X}$ can be found by solving a similar SVD/optimization problem of size $|\mathcal{J}| \ll N$. Once $\boldsymbol{X}$ is found, the full matrix $\boldsymbol{\Phi}$ can be recovered element-wise as in Alg. 3. If the equation is exact, the subset $\mathcal{J}$ can be chosen arbitrarily as long as $\boldsymbol{S}_{\mathcal{J}}$ has NSFR. Recall from Proposition 3 that $\boldsymbol{S}_{\mathcal{J}}$ being NSFR also implies that $\boldsymbol{S}$ is NSFR. However, if it is noisy, a careful selection of $\mathcal{J}$ for stable recovery has to be done. The following result characterizes the deviation of a singular vector associated with the smallest singular value of a matrix from the actual solution of the corresponding homogeneous equation under some perturbation. This will serve as a guidance for choosing a "good" subset $\mathcal{J}$.

Proposition 4: Suppose $\boldsymbol{A}$ is a matrix with condition number $\kappa(\boldsymbol{A})$ and $\delta\boldsymbol{A}$ is a perturbation such that $\boldsymbol{A} + \delta\boldsymbol{A}$ has a nontrivial null space. Let $\boldsymbol{z}$ and $\boldsymbol{z} + \delta\boldsymbol{z}$ be unit norm singular vectors associated with the smallest singular values of $\boldsymbol{A}$ and $\boldsymbol{A} + \delta\boldsymbol{A}$, respectively. Then we have the following bound

$$\|\delta \boldsymbol{z}\|_2 \lessapprox 1 + 2\kappa^2(\boldsymbol{A}) \frac{\|\delta \boldsymbol{A}\|_2}{\|\boldsymbol{A}\|_2}, \tag{49}$$

where the approximate inequality $\lessapprox$ assumes that $(\delta\boldsymbol{A})^T \delta\boldsymbol{A} \approx \boldsymbol{0}$.

Proof: Let $\sigma_1(\boldsymbol{A})$ and $\sigma_n(\boldsymbol{A})$ be the largest and smallest singular values of $\boldsymbol{A}$ respectively. $\boldsymbol{z}$ being the singular vector associated with the smallest singular value of $\boldsymbol{A}$ means

$$\boldsymbol{A}^T \boldsymbol{A} \boldsymbol{z} = \sigma_n^2(\boldsymbol{A}) \boldsymbol{z}. \tag{50}$$

Since $\boldsymbol{A} + \delta\boldsymbol{A}$ has nontrivial null space, its smallest singular value must be 0, and so

$$(\boldsymbol{A} + \delta\boldsymbol{A})(\boldsymbol{z} + \delta\boldsymbol{z}) = \boldsymbol{0},$$

which is also equivalent to

$$(\boldsymbol{A} + \delta\boldsymbol{A})^T (\boldsymbol{A} + \delta\boldsymbol{A})(\boldsymbol{z} + \delta\boldsymbol{z}) = \boldsymbol{0}. \tag{51}$$

Ignoring the second order term $(\delta\boldsymbol{A})^T \delta\boldsymbol{A}$, we can rearrange terms of (51) and take the norm of both sides to get

$$\begin{aligned}
\|\boldsymbol{A}^T \boldsymbol{A} \delta\boldsymbol{z}\|_2 &= \|\boldsymbol{A}^T \boldsymbol{A} \boldsymbol{z} + \left(\boldsymbol{A}^T \delta\boldsymbol{A} + (\delta\boldsymbol{A})^T \boldsymbol{A} \right) (\boldsymbol{z} + \delta\boldsymbol{z})\|_2 \\
&= \|\sigma_n^2(\boldsymbol{A}) \boldsymbol{z} + \left(\boldsymbol{A}^T \delta\boldsymbol{A} + (\delta\boldsymbol{A})^T \boldsymbol{A} \right) (\boldsymbol{z} + \delta\boldsymbol{z})\|_2 \\
&\leq \sigma_n^2(\boldsymbol{A}) \|\boldsymbol{z}\|_2 + \| \left(\boldsymbol{A}^T \delta\boldsymbol{A} + (\delta\boldsymbol{A})^T \boldsymbol{A} \right) (\boldsymbol{z} + \delta\boldsymbol{z})\|_2 \\
&\leq \sigma_n^2(\boldsymbol{A}) + \|\boldsymbol{A}^T \delta\boldsymbol{A} + (\delta\boldsymbol{A})^T \boldsymbol{A}\|_2 \cdot \|\boldsymbol{z} + \delta\boldsymbol{z}\|_2 \\
&\leq \sigma_n^2(\boldsymbol{A}) + \|\boldsymbol{A}^T \delta\boldsymbol{A}\|_2 + \|(\delta\boldsymbol{A})^T \boldsymbol{A}\|_2 \\
&\leq \sigma_n^2(\boldsymbol{A}) + 2\|\boldsymbol{A}\|_2 \|\delta\boldsymbol{A}\|_2 \\
&= \sigma_n^2(\boldsymbol{A}) + 2\sigma_1^2(\boldsymbol{A}) \frac{\|\delta\boldsymbol{A}\|_2}{\|\boldsymbol{A}\|_2}.
\end{aligned}$$

On the other hand

$$\|\boldsymbol{A}^T \boldsymbol{A} \delta\boldsymbol{z}\|_2 \geq \sigma_n^2(\boldsymbol{A}) \|\delta\boldsymbol{z}\|_2. \tag{52}$$

Hence

$$\|\delta\boldsymbol{z}\|_2 \leq 1 + 2\frac{\sigma_1^2(\boldsymbol{A})}{\sigma_n^2(\boldsymbol{A})} \cdot \frac{\|\delta\boldsymbol{A}\|_2}{\|\boldsymbol{A}\|_2} = 1 + 2\kappa^2(\boldsymbol{A}) \frac{\|\delta\boldsymbol{A}\|_2}{\|\boldsymbol{A}\|_2}. \tag{53}$$

□

Algorithm 4 RANSAC-Based Factorization

Inputs: $\boldsymbol{Y}, \boldsymbol{S}, \bar{N}, L$
Outputs: Estimates $\hat{\boldsymbol{\Phi}}, \hat{\boldsymbol{X}}$ of $\boldsymbol{\Phi}, \boldsymbol{X}$

1. Randomly generate L subsets of length $\bar{N}$ of $\mathcal{N}$.
2. For each subset $\mathcal{J}$, compute $\boldsymbol{M}_{(\mathcal{J})} = (\boldsymbol{I} - \boldsymbol{S}_{\mathcal{J}} \boldsymbol{S}_{\mathcal{J}}^{\dagger}) \odot (\boldsymbol{Y}_{\mathcal{J}} \boldsymbol{Y}_{\mathcal{J}}^T)$.
3. Choose $\mathcal{J}$ with largest $\rho(\mathcal{J})$ as defined in (54).
4. Use Alg. 2 to solve for $\hat{\boldsymbol{\Phi}}_{\mathcal{J}}, \hat{\boldsymbol{X}}$ from $\boldsymbol{Y}_{\mathcal{J}}, \boldsymbol{S}_{\mathcal{J}}$.
5. Set $\hat{\varphi}_i = \left[\frac{\boldsymbol{s}_i^T \hat{\boldsymbol{X}} \boldsymbol{y}_i}{\boldsymbol{s}_i^T \hat{\boldsymbol{X}} \hat{\boldsymbol{X}}^T \boldsymbol{s}_i} \right]_+, \quad i = 1, \ldots, N.$
6. Return $\hat{\boldsymbol{X}}$ in step 4 and $\hat{\boldsymbol{\Phi}} = \mathrm{diag}(\hat{\varphi}_1, \ldots, \hat{\varphi}_N)$.

Back to the issue of choosing a "good" subset $\mathcal{J}$, for a particular $\mathcal{J}$, we want to solve for the eigenvector associated with the smallest eigenvalue of the matrix $\boldsymbol{M}_{(\mathcal{J})} = (\boldsymbol{I} - \boldsymbol{S}_{\mathcal{J}} \boldsymbol{S}_{\mathcal{J}}^{\dagger}) \odot (\boldsymbol{Y}_{\mathcal{J}} \boldsymbol{Y}_{\mathcal{J}}^T)$. On the one hand, Proposition 4 suggests that in order to have a stable solution, $\mathcal{J}$ should be chosen so that the condition number of $\boldsymbol{M}_{(\mathcal{J})}$ is small. On the other hand, we also want the nullity of $\boldsymbol{M}_{(\mathcal{J})}$ to be numerically close to 1 in some sense so that the eigenvector associated with the smallest eigenvalue is close to the real solution. One heuristic way to measure this closeness is to compute the drop from the second smallest eigenvalue $\lambda_{N-1}(\boldsymbol{M}_{\mathcal{J}})$ to the smallest eigenvalue $\lambda_N(\boldsymbol{M}_{\mathcal{J}})$ relatively to the largest eigenvalue $\lambda_1(\boldsymbol{M}_{\mathcal{J}})$ of $\boldsymbol{M}_{(\mathcal{J})}$. Thus, we heuristically combined these two observations to define a single quantity associated with a subset $\mathcal{J}$ to measure the numerical rank-$(N-1)$ of $\boldsymbol{M}_{(\mathcal{J})}$ as

$$\begin{aligned}
\rho(\mathcal{J}) &= \frac{1}{\kappa(\boldsymbol{M}_{\mathcal{J}})} \cdot \frac{\lambda_{N-1}(\boldsymbol{M}_{\mathcal{J}}) - \lambda_N(\boldsymbol{M}_{\mathcal{J}})}{\lambda_1(\boldsymbol{M}_{\mathcal{J}})} \\
&= \frac{\lambda_N(\boldsymbol{M}_{\mathcal{J}}) \lambda_{N-1}(\boldsymbol{M}_{\mathcal{J}}) - \lambda_N^2(\boldsymbol{M}_{\mathcal{J}})}{\lambda_1^2(\boldsymbol{M}_{\mathcal{J}})}.
\end{aligned} \tag{54}$$

We will search for $\mathcal{J}$ with largest $\rho(\mathcal{J})$ over all subsets of $\mathcal{N}$. However, this leads to an NP combinatorial optimization problem. We propose instead an algorithm to choose the best $\mathcal{J}$ in terms of $\rho(\mathcal{J})$ over just a relatively small number of random subsets of $\mathcal{N}$. Alg. 4 is similar, to some extent, to the RANSAC (RANdom SAmple Consensus) [8], a commonly used algorithm in computer vision for model fitting. We will therefore call it RANSAC-based factorization.

Fig. 4. Dataset 1: 12 images of a real object under different directional light sources.

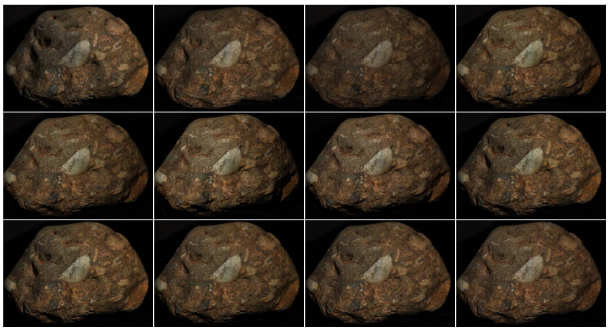


Fig. 5. Dataset 2: 12 images of a real object under different directional light sources.

VIII. Experimental Results

A. Dealing With Color Images

For inverse rendering from color images, we need to solve (13) for the three color channels simultaneously to eliminate the scaling ambiguity of the albedos across channels. Specifically, let $\boldsymbol{Y}_{\mathrm{R}}, \boldsymbol{Y}_{\mathrm{G}}, \boldsymbol{Y}_{\mathrm{B}}$ be the image matrices for red, green, and blue channels respectively. Also, let $\boldsymbol{\varphi}_{\mathrm{R}}, \boldsymbol{\varphi}_{\mathrm{G}}, \boldsymbol{\varphi}_{\mathrm{B}}$ be the albedo vectors associated with red, green, and blue channels respectively. Then the actual matrix factorization we want implement is given by the concatenating equation

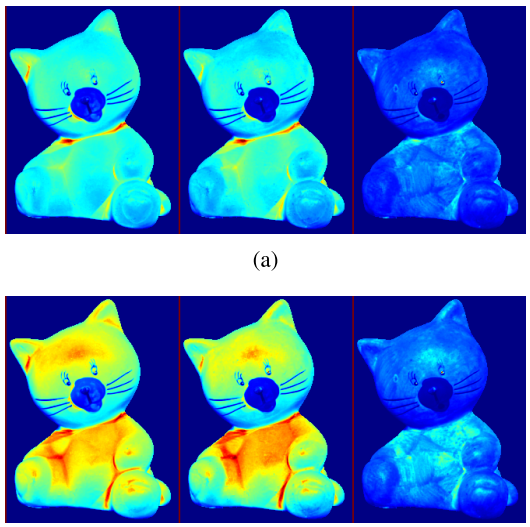
$$\begin{bmatrix} \boldsymbol{Y}_{\mathrm{R}} \\ \boldsymbol{Y}_{\mathrm{G}} \\ \boldsymbol{Y}_{\mathrm{B}} \end{bmatrix} = \mathrm{diag}\left(\begin{bmatrix} \boldsymbol{\varphi}_{\mathrm{R}} \\ \boldsymbol{\varphi}_{\mathrm{G}} \\ \boldsymbol{\varphi}_{\mathrm{B}} \end{bmatrix} \right) \cdot \begin{bmatrix} \boldsymbol{S} \\ \boldsymbol{S} \\ \boldsymbol{S} \end{bmatrix} \boldsymbol{X}. \tag{55}$$

Forming this equation is important for implementations, but the equation itself is mathematically the same as (13) and does not change any of the above theoretical results.

B. Simulations

Simulations are performed on two real datasets,[4] each of which has $J = 12$ images of an object as shown in Figs. 4 and 5, using MATLAB with the convex programming `cvx` provided by [28]. The images were taken under different directional light sources. The light directions of the 12 sources can be calibrated and will be used as a ground truth for the inverse rendering.

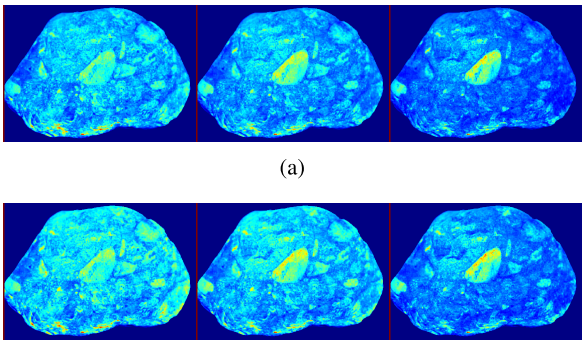
[4]The data and code for photometric stereo are available at: http://pages.cs.wisc.edu/~csverma/CS766_09/Stereo/stereo.html.



(a)

(b)

Fig. 6. Dataset 1: Reconstruction of albedo maps in red, green and blue channels from left to right using either (a) alternating or (b) RANSAC-based factorization.



(a)

(b)

Fig. 7. Dataset 2: Reconstruction of albedo maps in red, green and blue channels from left to right using either (a) alternating or (b) RANSAC-based factorization.

The normal map of the object was estimated from the images and light directions using photometric stereo. We want to note that the light directions were only used to estimate the geometry. In our inverse rendering, the lighting, however, is assumed to be unknown and arbitrary. Of course, in practice we are aiming at estimating the geometry directly using some depth-based 3D modeling framework. The spherical harmonic matrix $\boldsymbol{S}$ was then computed from the surface normals of the object using (9) and (11), with $M = 9$. The size of each image is 340×512; excluding zero-intensity pixels results in $N = 35983$ for dataset 1, and $N = 72561$ for dataset 2. We can easily confirm that the spherical harmonic matrix $\boldsymbol{S}$ indeed has NSFR for both datasets by using Proposition 3 and Corollary 1. In particular, we randomly chose an index subset $\mathcal{J}$ and check if the rank of the matrix $(\boldsymbol{I} - \boldsymbol{S}_{\mathcal{J}}\boldsymbol{S}_{\mathcal{J}}^{\dagger}) \odot (\boldsymbol{S}_{\mathcal{J}}\boldsymbol{S}_{\mathcal{J}}^{T})$ is equal to $|\mathcal{J}| - 1$.

The albedo and lighting matrices $\hat{\boldsymbol{\Phi}}$ and $\hat{\boldsymbol{X}}$ were estimated using either Alg. 3 or 4. The reconstructed albedo maps for 3 color channels using both methods are visualized in Figs. 6 and 7 using the `jet` color map. To quantify the

TABLE I

DATASET 1: ANGLES BETWEEN THE RECONSTRUCTIONS OF LIGHT SOURCE DIRECTIONS AND THEIR GROUND TRUTHS USING DIFFERENT METHODS: ALTERNATING AND RANSAC-BASED FACTORIZATIONS. THE MEAN ERRORS OVER ALL 12 LIGHT SOURCES ARE 6.8 DEGREES FOR ALTERNATING AND 2.7 DEGREES FOR RANSAC-BASED

source	1	2	3	4	5	6	7	8	9	10	11	12
Alternating	9.6	4.9	1.2	8.6	8.1	8.9	7.9	5.6	7.5	7.5	3.0	8.5
RANSAC	6.3	6.3	1.6	3.6	1.7	3.4	2.1	1.6	0.4	1.0	2.0	2.1

TABLE II

DATASET 2: ANGLES BETWEEN THE RECONSTRUCTIONS OF LIGHT SOURCE DIRECTIONS AND THEIR GROUND TRUTHS USING DIFFERENT METHODS: ALTERNATING AND RANSAC-BASED FACTORIZATIONS. THE MEAN ERRORS OVER ALL 12 LIGHT SOURCES ARE 9.9 DEGREES FOR ALTERNATING AND 1.8 DEGREES FOR RANSAC-BASED

source	1	2	3	4	5	6	7	8	9	10	11	12
Alternating	18.0	7.4	12.0	9.5	7.7	9.8	10.5	11.2	9.1	8.7	8.4	7.0
RANSAC	4.5	0.9	4.3	0.8	1.7	0.8	0.8	1.5	1.1	0.2	1.6	3.1

(a) Original (b) 29.3686 dB (c) 29.8326 dB

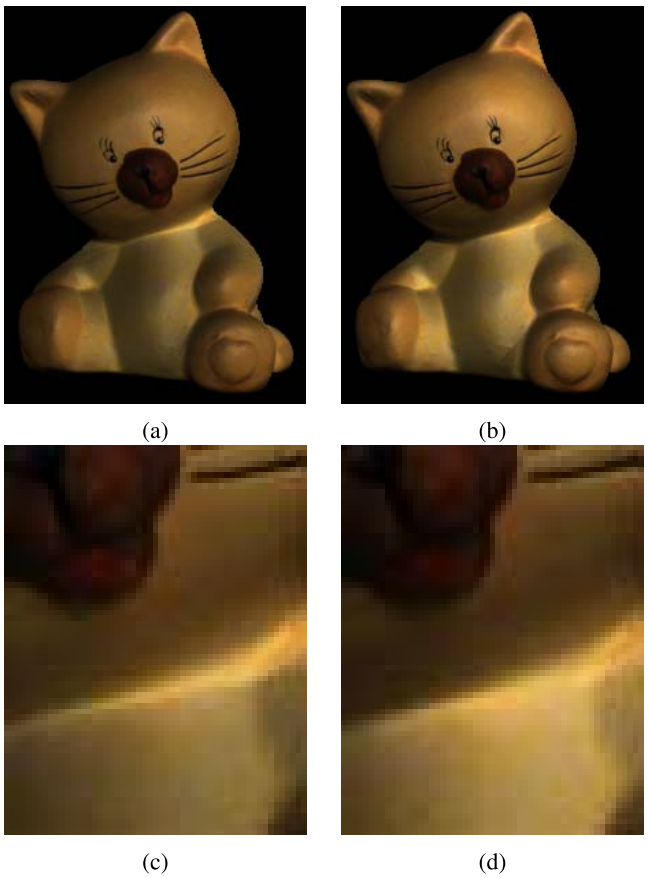
Fig. 8. Dataset 1: Reconstruction of one of the original images (a) using either alternating (b) or RANSAC-based factorization (c) with corresponding SNRs. The average SNRs over all given twelve images are 24.9786 dB for alternating and 24.8493 dB for RANSAC-based.

(a) Original (b) 28.6698 dB (c) 29.4742 dB

Fig. 9. Dataset 2: Reconstruction of one of the original images (a) using either alternating (b) or RANSAC-based factorization (c) with corresponding SNRs. The average SNRs over all given twelve images are 27.1829 dB for alternating and 27.4583 dB for RANSAC.

accuracy of our model, the twelve original images in $\boldsymbol{Y}$ were compared in SNR (Signal-to-Noise Ratio) to the reconstructed images by forming $\hat{\boldsymbol{Y}} = \hat{\boldsymbol{\Phi}}\boldsymbol{S}\hat{\boldsymbol{X}}$. The results are shown in Figs. 8 and 9 for one of the images in each dataset.

Although both methods can reconstruct the original images reasonably well, the reconstruction error itself does not quantify the accuracy of estimating albedos and lighting individually. Therefore we did another step to estimate the light directions for which we have ground truths. In particular, from (9), the x, y, z components of the light directions can be extracted from rows $2, 3, 4$ of the lighting matrix $\boldsymbol{X}$, respectively. The angle between each recovered light direction and its ground truth was then computed for all 12 sources and shown in Tabs. I and II. It can be seen that the RANSAC-based

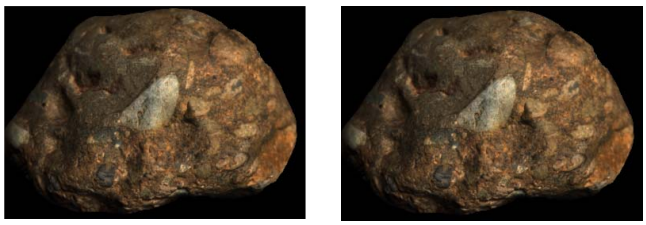


(a) (b) (c) (d)

Fig. 10. Dataset 1: Random relighting using (a) alternating and (b) RANSAC-based factorization. The magnifications of an interconnection region are shown respectively below in (c) and (d).

is much more accurate than the alternating factorization, with a mean error of 2.7 degrees versus 6.8 degrees for dataset 1, and 1.8 degrees versus 9.9 degrees for dataset 2.

For the purpose of relighting, new images were rendered by randomly generating a novel lighting matrix $\boldsymbol{X}_{\text{new}}$ and computing $\boldsymbol{Y}_{\text{new}} = \hat{\boldsymbol{\Phi}}\boldsymbol{S}\boldsymbol{X}_{\text{new}}$. The relit images using alternating and RANSAC-based factorizations are demonstrated in Figs. 10 and 11. It can be subjectively observed from Fig. 10 that relighting using the alternating



(a) (b)

Fig. 11. Dataset 2: Random relighting using (a) alternating and (b) RANSAC-based factorization.

factorization behaves poorly at the interconnection regions of the object where our convex assumption is violated. It is because the measurement noise model that the algorithm is based on does not account for approximation error resulting from the nonconvexity of the object. The difference between the two relighting methods is less visible in Fig. 11 because the 'rock' object does not have abrupt transitions like the neck of the 'cat' object. We remark that the implementations of the alternating algorithm were faster. In contrast, relighting using RANSAC-based factorization, at the cost of increasing the computations, produces very realistic looking images that are well-behaved at the interconnection regions.

IX. Concluding Remarks

We have studied the inverse rendering problem of a Lambertian convex object with distant light sources. Under these assumptions, all possible images of the object live close to a low-dimensional linear subspace spanned by the first few spherical harmonics. The inverse rendering thus becomes a factorization of the image matrix into a product of a diagonal albedo and a lighting matrix with a known SH matrix in the middle. This special matrix factorization is unique if and only if the SH matrix associated with the object has nonseparable full rank, a stronger notion of full rank. In the noiseless case, the exact solution (up to some scale) can be found via an SVD-based algorithm. When an approximation error is present, the SVD-based algorithm also yields the least squares solution to the factorization. An optimization-based algorithm is introduced to enforce the positivity of albedos. Because the number of pixels in each images is typically large, a heuristic RANSAC-based algorithm is proposed to reduce the size of the optimization problem. When a measurement noise is present in the model, an alternating algorithm is introduced instead for the factorization.

Appendix A
Proof of Proposition 2

A. KFR⇒NSFR

Suppose $\mathcal{N}_1$ and $\mathcal{N}_2$ are nonempty disjoint subsets such that $\mathcal{N}_1 \cup \mathcal{N}_2 = \mathcal{N}$. We will show that $\text{rank}(\boldsymbol{S}_{\mathcal{N}_1}) + \text{rank}(\boldsymbol{S}_{\mathcal{N}_2}) \geq M+1$. Indeed, since $\boldsymbol{S}$ has full Kruskal rank, any k rows of $\boldsymbol{S}$ are linearly independent for all $k \leq M$. Thus, $\text{rank}(\boldsymbol{S}_{\mathcal{N}_1}) = \min\{M, |\mathcal{N}_1|\}$, and $\text{rank}(\boldsymbol{S}_{\mathcal{N}_2}) = \min\{M, |\mathcal{N}_2|\}$. It follows that

$$\begin{aligned}\text{rank}(\boldsymbol{S}_{\mathcal{N}_1}) + \text{rank}(\boldsymbol{S}_{\mathcal{N}_2}) &= \min\{M, |\mathcal{N}_1|\} + \min\{M, |\mathcal{N}_2|\} \\ &\geq \min\{2M, M + |\mathcal{N}_1|, M \\ &\qquad + |\mathcal{N}_2|, |\mathcal{N}_1| + |\mathcal{N}_2|\} \\ &\overset{(a)}{\geq} \min\{M+1, N\} = M+1,\end{aligned}$$

where (a) follows from $|\mathcal{N}_1| + |\mathcal{N}_2| = |\mathcal{N}| = N$.

B. NSFR⇒SFR

For $i \in \mathcal{N}$, since $\boldsymbol{S}$ has NSFR, it must be that

$$\text{rank}(\boldsymbol{S}_{\{i\}}) + \text{rank}(\boldsymbol{S}_{\mathcal{N}\setminus\{i\}}) \geq M+1.$$

However, since $\boldsymbol{S}$ contains no zero rows, it is clear that $\text{rank}(\boldsymbol{S}_{\{i\}}) = 1$. Hence, $\text{rank}(\boldsymbol{S}_{\mathcal{N}\setminus\{i\}}) \geq M$. On the other hand, $\text{rank}(\boldsymbol{S}_{\mathcal{N}\setminus\{i\}}) \leq \text{rank}(\boldsymbol{S}) = M$. Therefore, $\text{rank}(\boldsymbol{S}_{\mathcal{N}\setminus\{i\}}) = M$, or $\boldsymbol{S}$ has strong full rank.

C. SFR⇒FR

If $\boldsymbol{S}$ has strong full rank, then for $i = 1$ we have $\text{rank}(\boldsymbol{S}) \geq \text{rank}(\boldsymbol{S}_{\mathcal{N}\setminus\{1\}}) = M$. Since $\boldsymbol{S} \in \mathbb{R}^{N\times M}$, we also have $\text{rank}(\boldsymbol{S}) \leq M$. Thus, $\text{rank}(\boldsymbol{S}) = M$, or $\boldsymbol{S}$ has full rank.

Now consider the reverse implications. If $N = M+1$, it is obvious that the definitions of strong full rank and Kruskal full rank coincide. Thus, SFR⇒KFR⇒NSFR. Finally, for $N \geq M+2$ we can show that the reverse implications are not true by constructing simple counterexamples. We leave that to the readers due to the space constraint. For a visual proof, see Figure 2.

Appendix B
Proof of Proposition 3

Let $\mathcal{N}_1$ and $\mathcal{N}_2$ be nonempty disjoint subsets of $\mathcal{N}$ such that $\mathcal{N} = \mathcal{N}_1 \cup \mathcal{N}_2$. We need to show that $\text{rank}(\boldsymbol{S}_{\mathcal{N}_1}) + \text{rank}(\boldsymbol{S}_{\mathcal{N}_2}) \geq M+1$. Indeed, consider the following cases:

If $\mathcal{J} \subset \mathcal{N}_1$ then

$$\begin{aligned}\text{rank}(\boldsymbol{S}_{\mathcal{N}_1}) + \text{rank}(\boldsymbol{S}_{\mathcal{N}_2}) &\geq \text{rank}(\boldsymbol{S}_{\mathcal{J}}) + \text{rank}(\boldsymbol{S}_{\mathcal{N}_2}) \\ &= M + \text{rank}(\boldsymbol{S}_{\mathcal{N}_2}) \geq M+1.\end{aligned}$$

If $\mathcal{J} \subset \mathcal{N}_2$, similarly we also have $\text{rank}(\boldsymbol{S}_{\mathcal{N}_1}) + \text{rank}(\boldsymbol{S}_{\mathcal{N}_2}) \geq M+1$.

If $\mathcal{J}$ is not a subset of either $\mathcal{N}_1$ or $\mathcal{N}_2$, put $\mathcal{J}_1 = \mathcal{J} \cap \mathcal{N}_1$, and $\mathcal{J}_1 = \mathcal{J} \cap \mathcal{N}_1$. It is evident that $\mathcal{J}_1, \mathcal{J}_2$ are nonempty disjoint subsets of $\mathcal{J}$ such that $\mathcal{J} = \mathcal{J}_1 \cup \mathcal{J}_2$. Since $\boldsymbol{S}_{\mathcal{J}}$ has nonseparable full rank, it must be that

$$\text{rank}(\boldsymbol{S}_{\mathcal{N}_1}) + \text{rank}(\boldsymbol{S}_{\mathcal{N}_2}) \geq \text{rank}(\boldsymbol{S}_{\mathcal{J}_1}) + \text{rank}(\boldsymbol{S}_{\mathcal{J}_2}) \geq M+1.$$

Hence, $\boldsymbol{S}$ has NSFR, completing the proof.

Appendix C
Proof of Lemma 1

For all $\mathcal{I} \subset \mathcal{N}$, we can write

$$\hat{\boldsymbol{S}}_{\mathcal{I}} = [\boldsymbol{S}_{\mathcal{I}} | (\boldsymbol{\Psi}\boldsymbol{S})_{\mathcal{I}}] \overset{(a)}{=} [\boldsymbol{S}_{\mathcal{I}} | \boldsymbol{S}_{\mathcal{I}} \tilde{\boldsymbol{X}} \boldsymbol{X}^{\dagger}],$$

where (a) follows from (18). Note that every column of $\boldsymbol{S}_{\mathcal{I}} \tilde{\boldsymbol{X}} \boldsymbol{X}^{\dagger}$ is a linear combination of columns of $\boldsymbol{S}_{\mathcal{I}}$. This implies that $\mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{I}}) = \mathcal{R}(\boldsymbol{S}_{\mathcal{I}})$. Hence

$$\text{rank}(\hat{\boldsymbol{S}}_{\mathcal{I}}) = \text{rank}(\boldsymbol{S}_{\mathcal{I}}), \quad \forall \mathcal{I} \subset \mathcal{N}.$$

APPENDIX D
PROOF OF LEMMA 2

We first show that $\{[\boldsymbol{s}_i^T | \psi_i \boldsymbol{s}_i^T]^T\}_{i\in\mathcal{B}}$ are linearly independent. Suppose $\{\alpha_i\}_{i\in\mathcal{B}}$ satisfies

$$\sum_{i\in\mathcal{B}} \alpha_i [\boldsymbol{s}_i^T | \psi_i \boldsymbol{s}_i^T]^T = \boldsymbol{0},$$

which implies $\sum_{i\in\mathcal{B}} \alpha_i \boldsymbol{s}_i = \sum_{i\in\mathcal{B}} \alpha_i \psi_i \boldsymbol{s}_i = \boldsymbol{0}$. By the independence of $\{\boldsymbol{s}_i\}_{i\in\mathcal{B}}$, it must be that $\alpha_i = 0, \forall i \in \mathcal{B}$, or $\{[\boldsymbol{s}_i^T | \psi_i \boldsymbol{s}_i^T]^T\}_{i\in\mathcal{B}}$ are linearly independent.

Next, it follows from Lemma 1 that

$$\dim\left(\mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{I}}^T)\right) = \dim\left(\mathcal{R}(\boldsymbol{S}_{\mathcal{I}}^T)\right) = |\mathcal{B}|.$$

It means that $\{[\boldsymbol{s}_i^T | \psi_i^T \boldsymbol{s}_i^T]^T\}_{i\in\mathcal{B}}$ is a set of linearly independent vectors in the space $\mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{I}}^T)$ whose cardinality is equal to the dimension of the space. Thus, it is a basis for $\mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{I}}^T)$.

APPENDIX E
PROOF OF THE NECESSITY PART OF THEOREM 1

Suppose that $\boldsymbol{S}$ does not have NSFR. We will construct a solution $(\tilde{\boldsymbol{\Phi}}, \tilde{\boldsymbol{S}})$ to (13) such that either $\tilde{\boldsymbol{\Phi}}$ is not a scaled version of $\boldsymbol{\Phi}_0$ or $\tilde{\boldsymbol{X}}$ is not a scaled version of $\boldsymbol{X}_0$. Consider the following cases.

Case 1: $\boldsymbol{S}$ has a zero row, say the first row. It follows that the first row of $\boldsymbol{Y}$ is also zero. We can therefore simply choose $\tilde{\boldsymbol{X}} = \boldsymbol{X}_0$ and $\tilde{\boldsymbol{\Phi}} = \text{diag}(2, 1, \ldots, 1) \odot \boldsymbol{\Phi}_0$.

Case 2: $\boldsymbol{S}$ does not have full column rank. It means that columns of $\boldsymbol{S}$ are linearly dependent. There thus exists a matrix $\tilde{\boldsymbol{X}}$ that is not a scaled version of $\boldsymbol{X}_0$ such that $\boldsymbol{S}\tilde{\boldsymbol{X}} = \boldsymbol{S}\boldsymbol{X}_0$. We then choose $(\boldsymbol{\Phi}_0, \tilde{\boldsymbol{X}})$ as a solution.

Case 3: $\boldsymbol{S}$ has full column rank with no zero rows. It follows from the definition of NSFR that there exists a nontrivial subset $\mathcal{J}$ of $\mathcal{N}$ such that $\text{rank}(\boldsymbol{S}_{\mathcal{J}}) + \text{rank}(\boldsymbol{S}_{\mathcal{J}^c}) = M$. Let $\boldsymbol{\Psi} = \text{diag}(\boldsymbol{\psi})$, where $\psi_j = 2, \forall j \in \mathcal{J}$ and $\psi_j = 1, \forall j \in \mathcal{J}^c$. Consider the matrix $\hat{\boldsymbol{S}} = [\boldsymbol{S} | \boldsymbol{\Psi}\boldsymbol{S}]$.

We first show that $\text{rank}(\hat{\boldsymbol{S}}_{\mathcal{J}}) = \text{rank}(\boldsymbol{S}_{\mathcal{J}})$. Let $\{\boldsymbol{s}_k\}_{k\in\mathcal{K}}$ be a basis for $\mathcal{R}(\boldsymbol{S}_{\mathcal{J}}^T)$, we can show that $\{[\boldsymbol{s}_k^T | 2\boldsymbol{s}_k^T]^T\}_{k\in\mathcal{K}}$ is a basis for $\mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{J}}^T)$. Indeed, similarly to the proof of Lemma 2, those vectors are independent. Moreover, any row of $\hat{\boldsymbol{S}}_{\mathcal{J}}$ can be written as

$$[\boldsymbol{s}_j^T | 2\boldsymbol{s}_j^T] = \sum_{k\in\mathcal{K}} \alpha_k [\boldsymbol{s}_k^T | 2\boldsymbol{s}_k^T], \quad \forall j \in \mathcal{J}, \tag{56}$$

where $\boldsymbol{s}_j = \sum_{k\in\mathcal{K}} \alpha_k \boldsymbol{s}_k$ is the linear representation of $\boldsymbol{s}_j$ in terms of $\{\boldsymbol{s}_k\}_{k\in\mathcal{K}}$. This implies that $\{[\boldsymbol{s}_k^T | 2\boldsymbol{s}_k^T]^T\}_{k\in\mathcal{K}}$ is a basis for $\mathcal{R}(\hat{\boldsymbol{S}}_{\mathcal{J}}^T)$, and so $\text{rank}(\hat{\boldsymbol{S}}_{\mathcal{J}}) = \text{rank}(\boldsymbol{S}_{\mathcal{J}}) = |\mathcal{K}|$. Similarly, $\text{rank}(\hat{\boldsymbol{S}}_{\mathcal{J}^c}) = \text{rank}(\boldsymbol{S}_{\mathcal{J}^c})$. Therefore

$$\begin{aligned}\text{rank}(\hat{\boldsymbol{S}}) &\le \text{rank}(\hat{\boldsymbol{S}}_{\mathcal{J}}) + \text{rank}(\hat{\boldsymbol{S}}_{\mathcal{J}^c}) \\ &= \text{rank}(\boldsymbol{S}_{\mathcal{J}}) + \text{rank}(\boldsymbol{S}_{\mathcal{J}^c}) = M.\end{aligned} \tag{57}$$

On the other hand, if $\{\boldsymbol{s}_i\}_{i\in\mathcal{B}}$ is a basis for $\mathcal{R}(\boldsymbol{S}^T)$, then similarly to the proof of Lemma 2, $\{[\boldsymbol{s}_i | \psi_i \boldsymbol{s}_i]\}_{i\in\mathcal{B}}$ are linearly independent. It means that the dimension of $\mathcal{R}(\hat{\boldsymbol{S}}^T)$ is at least $|\mathcal{B}|$, or

$$\text{rank}(\hat{\boldsymbol{S}}) \ge \text{rank}(\boldsymbol{S}) = M. \tag{58}$$

Combining (57) and (58) yields $\text{rank}(\hat{\boldsymbol{S}}) = \text{rank}(\boldsymbol{S}) = M$. It follows that each column of the matrix $\boldsymbol{\Psi}\boldsymbol{S}$ must be a linear combination of columns of $\boldsymbol{S}$. In other words, there exists a matrix $\boldsymbol{\Gamma}$ such that

$$\boldsymbol{\Psi}\boldsymbol{S} = \boldsymbol{S}\boldsymbol{\Gamma}. \tag{59}$$

Choose $\tilde{\boldsymbol{\Phi}} = \text{diag}(\tilde{\boldsymbol{\varphi}})$, where $\tilde{\varphi}_j = \varphi_j/2, \forall j \in \mathcal{J}$ and $\tilde{\varphi}_j = \varphi_j, \forall j \in \mathcal{J}^c$. Left-multiplying by $\tilde{\boldsymbol{\Phi}}$ and right-multiplying by $\boldsymbol{X}_0$ to both sides of (59) we get

$$\boldsymbol{\Phi}_0 \boldsymbol{S} \boldsymbol{X}_0 = \tilde{\boldsymbol{\Phi}} \boldsymbol{S} \tilde{\boldsymbol{X}}, \tag{60}$$

where $\tilde{\boldsymbol{X}} = \boldsymbol{\Gamma}\boldsymbol{X}_0$. It shows that $(\tilde{\boldsymbol{\Phi}}, \tilde{\boldsymbol{X}})$ is a solution to (13). However, by the construction, $\tilde{\boldsymbol{\Phi}}$ is not a scaled version of $\boldsymbol{\Phi}$.

APPENDIX F
PROOF OF LEMMA 3

Let $\boldsymbol{Q} = \boldsymbol{I} - \boldsymbol{S}\boldsymbol{S}^\dagger$. It is easy to check that $\boldsymbol{Q}$ is a projection matrix, i.e., $\boldsymbol{Q}^2 = \boldsymbol{Q} = \boldsymbol{Q}^T$. Since $\boldsymbol{D}_i$ is a diagonal matrix for all j, we have

$$(\boldsymbol{Q}\boldsymbol{D}_j)^T = \boldsymbol{D}_j^T \boldsymbol{Q}^T = \boldsymbol{D}_j \boldsymbol{Q}, \quad j = 1, \ldots, J.$$

Hence

$$\boldsymbol{A}^T = [\boldsymbol{D}_1 \boldsymbol{Q} | \boldsymbol{D}_2 \boldsymbol{Q} | \cdots | \boldsymbol{D}_K \boldsymbol{Q}].$$

It follows that

$$\boldsymbol{A}^T\boldsymbol{A} = \sum_{j=1}^{J} \boldsymbol{D}_j \boldsymbol{Q}^2 \boldsymbol{D}_j = \sum_{j=1}^{J} \boldsymbol{D}_j \boldsymbol{Q} \boldsymbol{D}_j. \tag{61}$$

Thus for any $i, k \in \{1, 2 \ldots, N\}$, we get

$$\begin{aligned}(\boldsymbol{A}^T\boldsymbol{A})_{ik} &= \sum_{j=1}^{J} \sum_{m,n=1}^{N} (\boldsymbol{D}_j)_{im} (\boldsymbol{Q})_{mn} (\boldsymbol{D}_j)_{nk} \\ &= \sum_{j=1}^{J} (\boldsymbol{D}_j)_{ii} (\boldsymbol{Q})_{ik} (\boldsymbol{D}_j)_{kk} \qquad (62) \\ &= (\boldsymbol{Q})_{ik} \sum_{j=1}^{J} (\boldsymbol{Y})_{ij} (\boldsymbol{Y})_{kj} \qquad (63) \\ &= (\boldsymbol{Q})_{ik} \sum_{j=1}^{J} (\boldsymbol{Y})_{ij} (\boldsymbol{Y}^T)_{jk} \\ &= (\boldsymbol{Q})_{ik} (\boldsymbol{Y}\boldsymbol{Y}^T)_{ik},\end{aligned}$$

where (62) follows from the diagonality of $\boldsymbol{D}_i$s and (63) follows from the definition of $\boldsymbol{D}_i$s. This implies that $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{Q} \odot (\boldsymbol{Y}\boldsymbol{Y}^T)$. Substituting $\boldsymbol{Q}$ by $\boldsymbol{I} - \boldsymbol{S}\boldsymbol{S}^\dagger$ completes the proof.

ACKNOWLEDGMENTS

The authors would like to thank S. Liu for valuable suggestions and for reviewing early drafts of this paper.

REFERENCES

[1] R. Ramamoorthi, "A signal-processing framework for forward and inverse rendering," Ph.D. dissertation, Dept. Comput. Sci., Stanford Univ., Stanford, CA, USA, Aug. 2002.

[2] J. Lehtinen, "A framework for precomputed and captured light transport," *ACM Trans. Graph.*, vol. 26, no. 4, pp. 13:1–13:22, Oct. 2007.

[3] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar, "Acquiring the reflectance field of a human face," in *Proc. 27th Annu. Conf. Comput. Graph. Interact. Techn. (SIGGRAPH)*, New Orleans, LA, USA, Jul. 2000, pp. 145–156.

[4] S. Izadi *et al.*, "KinectFusion: Real-time 3D reconstruction and interaction using a moving depth camera," in *Proc. 24th Annu. ACM Symp. User Inter. Softw. Technol. (UIST)*, 2011, pp. 559–568.

[5] R. Ramamoorthi and P. Hanrahan, "A signal-processing framework for inverse rendering," in *Proc. 28th Annu. Conf. Comput. Graph. Interact. Techn. (SIGGRAPH)*, Los Angeles, CA, USA, Aug. 2001, pp. 117–128.

[6] R. Basri and D. W. Jacobs, "Lambertian reflectance and linear subspaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 2, pp. 218–233, Feb. 2003.

[7] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

[8] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981.

[9] R. Ramamoorthi and P. Hanrahan, "An efficient representation for irradiance environment maps," in *Proc. 28th Annu. Conf. Comput. Graph. Interact. Techn. (SIGGRAPH)*, Los Angeles, CA, USA, Aug. 2001, pp. 497–500.

[10] R. Ramamoorthi and P. Hanrahan, "A signal-processing framework for reflection," *ACM Trans. Graph.*, vol. 23, no. 4, pp. 1004–1042, Oct. 2004.

[11] R. Basri, D. Jacobs, and I. Kemelmacher, "Photometric stereo with general, unknown lighting," *Int. J. Comput. Vis.*, vol. 72, no. 3, pp. 239–257, May 2007.

[12] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: A factorization method," *Int. J. Comput. Vis.*, vol. 9, no. 2, pp. 137–154, Nov. 1992.

[13] P. N. Belhumeur and D. J. Kriegman, "What is the set of images of an object under all possible illumination conditions?" *Int. J. Comput. Vis.*, vol. 28, no. 3, pp. 245–260, Jul. 1998.

[14] K. Kanatani, *Geometric Computation for Machine Vision*. New York, NY, USA: Oxford Univ. Press, 1993.

[15] A. D. Bue, J. Xavier, L. Agapito, and M. Paladini, "Bilinear factorization via augmented lagrange multipliers," in *Proc. 11th Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2010, pp. 283–296.

[16] M. Nießner, M. Zollhöfer, S. Izadi, and M. Stamminger, "Real-time 3D reconstruction at scale using voxel hashing," *ACM Trans. Graph.*, vol. 32, no. 6, pp. 169:1–169:11, Nov. 2013.

[17] M. Zollhöfer *et al.*, "Real-time non-rigid reconstruction using an RGB-D camera," *ACM Trans. Graph.*, vol. 33, no. 4, pp. 156:1–156:12, Jul. 2014.

[18] C. Wu, M. Zollhöfer, M. Nießner, M. Stamminger, S. Izadi, and C. Theobalt, "Real-time shading-based refinement for consumer depth cameras," in *Proc. SIGGRAPH Asia*, Dec. 2014.

[19] Y. Sato, M. D. Wheeler, and K. Ikeuchi, "Object shape and reflectance modeling from observation," in *Proc. 24th Annu. Conf. Comput. Graph. Interact. Techn. (SIGGRAPH)*, Los Angeles, CA, USA, 1997, pp. 379–387.

[20] Y. Yu, P. Debevec, J. Malik, and T. Hawkins, "Inverse global illumination: Recovering reflectance models of real scenes from photographs," in *Proc. 26th Annu. Conf. Comput. Graph. Interact. Techn. (SIGGRAPH)*, Los Angeles, CA, USA, 1999, pp. 215–224.

[21] S. R. Marschner, S. H. Westin, E. P. F. Lafortune, K. E. Torrance, and D. P. Greenberg, "Image-based BRDF measurement including human skin," in *Proc. 10th Eurograph. Conf. Rendering (EGWR)*, Granada, Spain, Jun. 1999, pp. 131–144.

[22] W. Matusik, H. Pfister, M. Brand, and L. McMillan, "Efficient isotropic BRDF measurement," in *Proc. 14th Eurograph. Workshop Rendering (EGRW)*, Leuven, Belgium, 2003, pp. 241–247.

[23] D. B. Goldman, B. Curless, A. Hertzmann, and S. M. Seitz, "Shape and spatially-varying BRDFs from photometric stereo," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 6, pp. 1060–1071, Jun. 2010.

[24] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[25] R. Roy and T. Kailath, "ESPRIT—Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.

[26] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra Appl.*, vol. 18, no. 2, pp. 95–138, 1977.

[27] K. Lee, Y. Bresler, and M. Junge, "Subspace methods for joint sparse recovery," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3613–3641, Jun. 2012.

[28] M. Grant and S. Boyd. (Aug. 2012). *CVX: Matlab Software for Disciplined Convex Programming*. [Online]. Available: http://cvxr.com/cvx/

Ha Q. Nguyen was born in Haiphong, Vietnam, in 1983. He received the B.S. degree in mathematics from the Hanoi National University of Education, Hanoi, Vietnam, in 2005, the M.S. degree in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2009, and the Ph.D. degree in electrical and computer engineering from the University of Illinois at Urbana-Champaign, Urbana, IL, USA, in 2014.

He was a Lecturer of Electrical Engineering with International University, Vietnam National University, Ho Chi Minh City, Vietnam, from 2009 to 2011. He is currently a Post-Doctoral Research Associate with the Biomedical Imaging Group, Swiss Federal Institute of Technology Lausanne, Lausanne, Switzerland. His research interests include image processing, computational imaging, data compression, and sampling theory.

Dr. Nguyen was a fellow of the Vietnam Education Foundation, Cohort 2007. He was a recipient of the Best Student Paper Award at the IEEE International Conference on Acoustics, Speech and Signal Processing in 2014 for his paper (with P. A. Chou and Y. Chen) on compression of human body sequences using graph wavelet filter banks.

Minh N. Do (M'01–SM'07–F'14) was born in Vietnam in 1974. He received the B.Eng. degree in computer engineering from the University of Canberra, Canberra, ACT, Australia, in 1997, and the D.Sc. degree in communication systems from the Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, Switzerland, in 2001.

He has been with the faculty of the University of Illinois at Urbana-Champaign (UIUC), Urbana, IL, USA, since 2002, where he is currently a Professor with the Department of Electrical and Computer Engineering, and holds joint appointments with the Coordinated Science Laboratory, the Beckman Institute for Advanced Science and Technology, and the Department of Bioengineering. His research interests include image and multidimensional signal processing, wavelets and multiscale geometric analysis, computational imaging, augmented reality, and visual information representation.

Prof. Do was a recipient of the Silver Medal from the 32nd International Mathematical Olympiad in 1991, the University Medal from the University of Canberra in 1997, the Doctorate Award from EPFL in 2001, the CAREER Award from the National Science Foundation in 2003, and the Young Author Best Paper Award from IEEE in 2008. He was named a Beckman Fellow at the Center for Advanced Study, UIUC, in 2006, and received a Xerox Award for Faculty Research from the College of Engineering, UIUC, in 2007. He was a member of the IEEE Signal Processing Theory and Methods Technical Committee, and Image, Video, and Multidimensional Signal Processing Technical Committee, and an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING. He is the Co-Founder and Chief Scientist of Personify Inc., a spinoff from UIUC to commercialize depth-based visual communication.