ROBUST PHOTOMETRIC STEREO USING LEARNED IMAGE AND GRADIENT DICTIONARIES

Andrew J. Wagenmaker, Brian E. Moore, and Raj Rao Nadakuditi

Department of EECS, University of Michigan, Ann Arbor, MI, USA

ABSTRACT

Photometric stereo is a method for estimating the normal vectors of an object from images of the object under varying lighting conditions. Motivated by several recent works that extend photometric stereo to more general objects and lighting conditions, we study a new robust approach to photometric stereo that utilizes dictionary learning. Specifically, we propose and analyze two approaches to adaptive dictionary regularization for the photometric stereo problem. First, we propose an image preprocessing step that utilizes an adaptive dictionary learning model to remove noise and other non-idealities from the image dataset before estimating the normal vectors. We also propose an alternative model where we directly apply the adaptive dictionary regularization to the normal vectors themselves during estimation. We study the practical performance of both methods through extensive simulations, which demonstrate the state-of-the-art performance of both methods in the presence of noise.

Index Terms— Dictionary learning, photometric stereo, sparse representations.

1. INTRODUCTION

Photometric stereo [1] is a method for estimating the normal vectors of an object from images of the object under varying lighting conditions. Since its inception, a significant amount of work has been done extending photometric stereo to more general conditions. This body of work has been divided into two primary areas. Uncalibrated photometric stereo seeks to solve the photometric stereo problem when the lighting directions are unknown [2–5], while robust photometric stereo algorithms attempt to estimate the normal vectors of an object when the surface violates the assumptions of the underlying model (usually, the Lambertian reflectance model). In this work, we are primarily concerned with the latter problem.

The Lambertian reflectance model states that the observed intensity of a point on a surface is linearly proportional to the direction the surface is illuminated and the object's normal vectors [1]. While this assumption holds in some cases, shadows, specularities, and other non-idealities can cause this model to break down. A variety of techniques have been developed to compensate for these corruptions. Several works seek to model non-Lambertian effects as outliers and employ a framework to estimate these outliers and discard them from the data, leaving only Lambertian data behind [6–10]. Of particular interest in this category are the more recent works by Wu *et al*. [11] and Ikehata *et al*. [12]. These works model the outliers as a sparse matrix, casting the problem as a matrix completion problem, and use robust PCA or sparse regression, respectively, to solve for some Lambertian representation of the data. Other works have proposed more complicated reflectance models to account for non-Lambertian effects, eliminating the need to discard non-ideal data [13–21]. The current state-of-the-art in this category are the works by Ikehata *et al*. [22] and Shi *et al*. [23].

In this work, we propose two new approaches for photometric stereo that are robust to noisy data. Our methods utilize a dictionary learning model [24, 25] to handle non-idealities and impose some adaptive structure on the data. Our approach is motivated in part by the recent success of dictionary learning in other imaging domains [26, 27]. We demonstrate the viability and performance of our methods on several datasets with varying degrees of non-ideality—including the recently proposed DiLiGenT dataset [28]—comparing them to the performance of state-of-the-art methods. In particular, we investigate the ability of our methods to handle general, non-sparse errors and noise.

The rest of this paper is organized as follows. Section 2 introduces the photometric stereo problem. Section 3 presents our dictionary learning based formulations and details their implementation. Finally, in Section 4, we demonstrate the performance of our methods on a variety of datasets.

2. PHOTOMETRIC STEREO PROBLEM

The Lambertian reflectance model states that, given an image of a *Lambertian* object, the light intensity observed at point (x, y) on the surface satisfies

$$I(x, y) = \rho(x, y)\ell^T n(x, y), \tag{1}$$

where $I(x, y)$ is the image intensity, $\ell \in \mathbb{R}^3$ is the direction of the light source incident on the surface, $n(x, y)$ is the nor-

This work was supported in part by the following grants: ONR grant N00014-15-1-2141, DARPA Young Faculty Award D14AP00086, and ARO MURI grants W911NF-11-1-0391 and 2015-05174-05.

mal vector of the surface, and $\rho(x,y)$ is the surface albedo—a measure of the reflectivity of the surface.

If we fix the position of a camera facing our surface and vary the position of the light source over d unique locations, we can write d equations of the form (1), which can be stacked to form the system of equations

$$[I_1(x,y)\ \dots\ I_d(x,y)]^T = [\ell_1\ \dots\ \ell_d]^T \rho(x,y) n(x,y). \quad (2)$$

Assuming each of the d images has size $m_1 \times m_2$, Equation (2) can then be solved $m_1 m_2$ times to obtain the normal vectors of the object at each point on its surface. These equations can also be combined into a single matrix equation. Indeed, let us define the observation matrix

$$Y := [\mathbf{vec}(I_1)\ \dots\ \mathbf{vec}(I_d)] \in \mathbb{R}^{m_1 m_2 \times d}, \quad (3)$$

where $\mathbf{vec}(I_j) := [I_j(1,1)\ \dots\ I_j(m_1,m_2)]^T$. Assuming our light source is at infinity and there is no variation in illumination from point to point on our object, one can succinctly express (2) as

$$Y = NL, \quad (4)$$

where $N = [\rho(1,1)n(1,1)\ \dots\ \rho(m_1,m_2)n(m_1,m_2)]^T \in \mathbb{R}^{m_1 m_2 \times 3}$ and $L = [\ell_1\ \dots\ \ell_d] \in \mathbb{R}^{3 \times d}$. For simplicity, we assume that $\|\ell_k\|_2 = 1$, and, without loss of generality, we assume that $n(x,y)$ are *unit* normals.

Given $d \geq 3$ images and their corresponding light directions, one can solve (4) exactly to obtain the normal vector matrix N, from which one can compute the full 3D representation of the underlying surface [29].

In theory, (4) should hold exactly for a Lambertian surface, but, in practice, due to noise and other non-idealities, one only expects that $Y \approx NL$. In the latter case, one can instead collect $d > 3$ measurements and solve the overdetermined least squares problem

$$\min_N \|Y - NL\|_F^2, \quad (5)$$

which has the convenient closed-form solution $\hat{N} = YL^\dagger$, where † denotes the Moore-Penrose pseudoinverse.

3. DICTIONARY LEARNING MODELS

In this section, we propose two adaptive dictionary learning methods for estimating the normal vectors of a surface from (possibly) noisy images, Y. Intuitively, these models seek to learn a locally sparse representation of the data with respect to a collection of learned basis "atoms" that capture the underlying local structure of the data.

3.1. Preprocessing Images through Dictionary Learning (DLPI)

Our first approach applies dictionary learning to the data in a preprocessing step before estimating the normal vectors. This formulation represents the input image data Y as locally sparse in an adaptive dictionary domain—thereby removing non-idealities that are not well-represented by the dictionary. Specifically, we propose to solve the problem

$$\min_{v,B,D} \frac{1}{2}\|y - v\|_2^2 + \lambda\left(\textstyle\sum_{j=1}^{c} \|P_j v - Db_j\|_2^2 + \mu^2 \|B\|_0\right)$$
$$\text{s.t. } \|d_i\|_2 = 1,\ \|b_j\|_\infty \leq a,\ \forall i,j. \quad (6)$$

Here, $y = \mathbf{vec}(Y)$ and P_j is a matrix that extracts a (vectorized) 3D patch of dimensions $c_x \times c_y \times c_z$ from v, where c_x and c_y are the dimensions of the patches extracted from each image and c_z is the number of distinct images whose patches are combined to form the 3D patch. $D \in \mathbb{R}^{c_x c_y c_z \times K}$ is a dictionary matrix whose columns d_i are the (learned) dictionary atoms, and $B \in \mathbb{R}^{K \times c}$ is a sparse coding matrix whose columns b_j define (usually sparse) linear combinations of dictionary atoms used to represent each patch. Also, $\|\cdot\|_0$ is the familiar ℓ_0 "norm", and $\lambda, \mu > 0$ are parameters.

We impose the constraint $\|b_j\|_\infty \leq a$, where a is typically very large, since (6) is non-coercive with respect to B, but the constraint is typically inactive in practice [30]. Without loss of generality, we impose a unit-norm constraint on the dictionary atoms d_i to avoid scaling ambiguity between D and B [31]. We allow the possibility that patches from $c_z > 1$ input images can be combined into a 3D patch to allow the dictionary atoms to learn correlated features between images, but one can set $c_z = 1$ to work with 2D per-image patches.

Once we have solved (6), we reshape v (back) into an $m_1 m_2 \times d$ matrix whose columns are vectorized (now preprocessed) images, and then we estimate the associated normal vectors using the standard least squares model (5). Henceforth, we refer to this approach as the Dictionary Learning with Preprocessed Imgaes (DLPI) method.

3.2. Normal Vector Computation through Dictionary Learning (DLNV)

We next propose modifying (5) by applying an adaptive dictionary regularization term to the normal vectors, N, under the Lambertian model (5). Specifically, we propose to solve the problem

$$\min_{n,B,D} \frac{1}{2}\|y - An\|_2^2 + \lambda\left(\textstyle\sum_{j=1}^{w} \|P_j n - Db_j\|_2^2 + \mu^2 \|B\|_0\right)$$
$$\text{s.t. } \|d_i\|_2 = 1,\ \|b_j\|_\infty \leq a,\ \forall i,j. \quad (7)$$

Here $y = \mathbf{vec}(Y)$, $A = L^T \otimes I$—where $\otimes$ denotes the Kronecker product and I is the $m_1 m_2 \times m_1 m_2$ identity matrix—and $n = \mathbf{vec}(N)$. Also, P_j denotes a patch extraction matrix that extracts (vectorized) patches of dimensions $w_x \times w_y \times w_z$ from n. All other terms are defined analogously to the corresponding terms in (6) with appropriate dimensions.

The dictionary learning terms in (7) encourage the estimated normal vectors to be well-represented by sparse linear combinations of a few (learned) dictionary atoms. Intuitively, this acts as an adaptive regularization that yields normal vectors that are more robust to noise and other non-idealities in the data. Henceforth, we refer to this approach as the Dictionary Learning on Normal Vectors (DLNV) method.

3.3. Algorithms for DLPI and DLNV

We propose solving (6) and (7), respectively, via block coordinate descent-type algorithms where we alternate between updating n and v, respectively, with (D, B) fixed and then updating (D, B) with n or v held fixed. We omit the (D, B) updates here due to space considerations, but the precise update expressions can be found in [30, 32].

3.3.1. v update

Solving (6) for v with D and B fixed yields the problem

$$\min_{v} \; \frac{1}{2} \|y - v\|_2^2 + \lambda \textstyle\sum_{j=1}^{c} \|P_j v - Db_j\|_2^2. \tag{8}$$

Equation (8) is a simple least squares problem whose solution v satisfies the normal equation

$$\left(I + 2\lambda \textstyle\sum_{j=1}^{c} P_j^T P_j\right) v = y + 2\lambda \textstyle\sum_{j=1}^{c} P_j^T D b_j, \tag{9}$$

where I denotes the identity matrix. The matrix pre-multiplying v in (9) is diagonal, so its inverse can be cheaply computed, allowing us to efficiently update v.

3.3.2. n update

On the other hand, solving (7) for n with D and B fixed yields the problem

$$\min_{n} \; \frac{1}{2} \|y - An\|_2^2 + \lambda \textstyle\sum_{j=1}^{w} \|P_j n - Db_j\|_2^2. \tag{10}$$

Note that while (10) is also a least squares problem, its normal equation cannot be easily inverted as in (8) due to the presence of the A matrix. We therefore adopt a proximal gradient scheme [33]. The cost function in (10) can be written in the form $f(n) + g(n)$ where $f(n) = \frac{1}{2} \|y - An\|_2^2$ and $g(n) = \lambda \sum_{j=1}^{w} \|P_j n - Db_j\|_2^2$. The proximal updates thus become

$$n^{k+1} = \mathbf{prox}_{\tau g}(n^k - \tau \nabla f(n^k)), \tag{11}$$

where

$$\mathbf{prox}_{\tau g}(z) := \arg\min_{x} \; \frac{1}{2} \|z - x\|_2^2 + \tau g(x). \tag{12}$$

Define $\tilde{n}^k := n^k - \tau \nabla f(n^k)$. Then (11) and (12) imply that n^{k+1} satisfies the normal equation

$$\left(I + 2\tau\lambda \textstyle\sum_{j=1}^{w} P_j^T P_j\right) n^{k+1} = \tilde{n}^k + 2\tau\lambda \textstyle\sum_{j=1}^{w} P_j^T D b_j. \tag{13}$$

As in (9), the matrix multiplying n^{k+1} in (13) is diagonal and can be efficiently inverted, yielding n^{k+1}. Note that proximal gradient is one of a wealth of available iterative schemes for minimzing the (quadratic) objective (10).

4. RESULTS

We now empirically demonstrate the performance of our proposed methods on several real-world datasets. To obtain quantitative results, we rely primarily on the DiLiGenT dataset [28]. This dataset contains images of a variety of surfaces and provides the true normal vectors of each object, allowing us to evaluate the performance of each method against a ground truth. We quantify error by measuring the mean angular difference between true normal vectors and estimated normal vectors.

For each experiment, we compare the results of our method to Wu *et al.*'s robust PCA (RPCA) method [11], Ikehata *et al.*'s sparse regression (SR) method [12], and Ikehata *et al.*'s constrained bivariate regression (CBR) method [22]. We also compare against the simple least squares (LS) method (5). With the exception of LS, each method contains one or more tunable parameters that dictate their performance. For each method, we sweep the parameters across a wide range of values, including any values recommended by the authors in this sweep. The reported results are the errors produced by the optimal parameter values.

To evaluate the ability of our method to robustly reject non-idealities, we add Poisson noise to the images in the original datasets. In each case, we run the experiment over multiple noise realizations and average the results.

4.1. Varying Noise Levels

We first simulate the addition of Poisson noise to the images, varying the signal-to-noise-ratio (SNR). Figure 1 illustrates the results of these simulations on a 20-image subset of the DiLiGenT Cat dataset.

As Figure 1 shows, in the low SNR (high noise) regime, both dictionary learning approaches significantly outperform existing approaches and are able to produce much cleaner normal vectors. The performance of all methods becomes comparable in the high SNR (low noise) regime, although our proposed dictionary learning based approaches are less sensitive to changes in the noise strength.

4.2. Varying Number of Images

Table 1 illustrates the accuracy of the estimated normal vectors of each algorithm as a function of the number of images used in the reconstruction. We ran this experiment on the DiLiGenT Bear dataset, sweeping from 5 to 96 images and adding Poisson noise with 10 dB SNR.

Table 1 shows that our proposed DLNV and DLPI algorithms significantly outperform existing methods on small datasets, achieving nearly 15 degree improvements in mean angular error. These results imply that, while dictionary learning based approaches generally perform well in low SNR (high noise) regimes, they are particularly robust to noise on small datasets compared to existing methods.

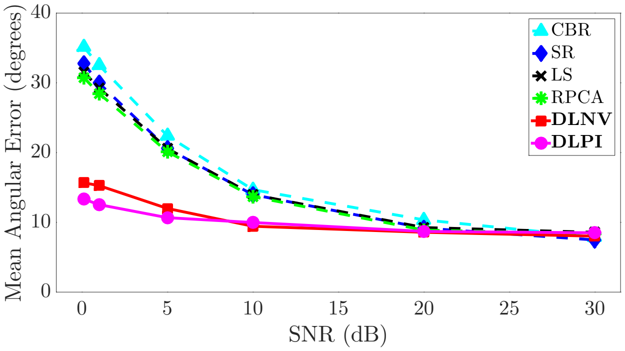


Fig. 1: Sweeping SNR on the DiLiGenT Cat [28] dataset with 20 images.

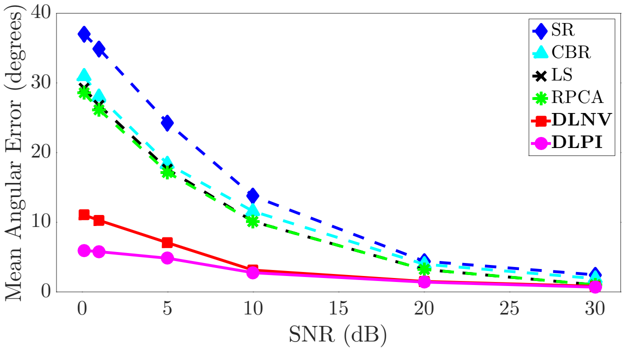
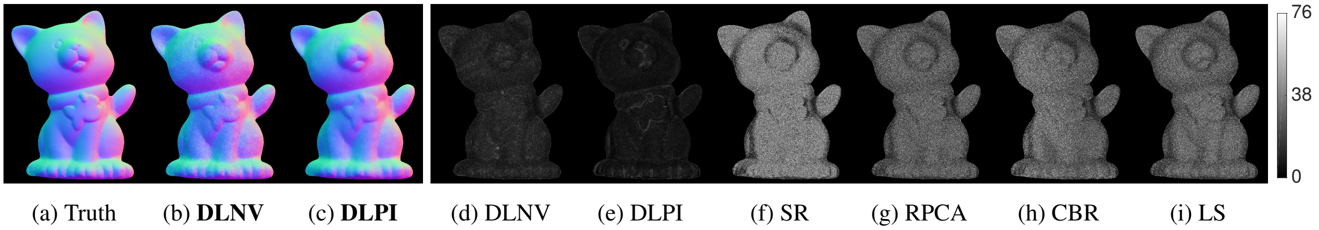


Fig. 2: Sweeping SNR on the Hippo dataset [34] with 20 images.



(a) Truth (b) **DLNV** (c) **DLPI** (d) DLNV (e) DLPI (f) SR (g) RPCA (h) CBR (i) LS

Fig. 3: Normal vector plots and error maps computed from the Cat dataset [34] with 20 images and SNR = 1 dB. Error maps plot angular error (in degrees) in the normal vectors at each point on surface.

No. of Images	**DLPI**	**DLNV**	SR	RPCA	CBR	LS
5	**20.91**	21.43	34.29	33.12	31.10	34.25
15	**9.73**	10.20	14.89	14.44	14.86	14.90
25	**9.23**	9.52	12.12	11.78	12.76	12.14
35	**9.13**	9.18	11.23	10.93	11.99	11.24
45	8.96	**8.90**	10.62	10.34	11.71	10.64
55	8.86	**8.78**	10.24	9.97	11.51	10.25
65	8.89	**8.77**	10.00	9.75	11.29	10.02
75	8.79	**8.68**	9.73	9.50	11.21	9.750
85	8.72	**8.62**	9.57	9.34	11.20	9.58
96	8.69	**8.61**	9.43	9.22	11.05	9.45

Table 1: Mean angular error sweeping number of images on the DiLiGenT Bear dataset [28] with SNR = 10 dB.

4.3. Analysis of non-DiLiGenT datasets

In addition to the DiLiGenT dataset, we also consider the dataset from[1] [34]. This dataset contains images of several real objects without ground truth normal vectors. To obtain an estimate of the ground truth normal vectors, we assume the objects in this dataset follow a truly Lambertian model and compute normal vectors from the uncorrupted dataset using the simple least squares model (5). While the Lambertian assumption may not hold exactly, the objects are matte in appearance – the primary characteristic of Lambertian surfaces – so these vectors are a reasonable approximation of the true normal vectors. This approach allows us to isolate the robustness of each method to noise when our data otherwise perfectly follow the modeling assumptions. Figure 2 depicts the results of these experiments.

From Figure 2, we see that, for high SNR cases where corruptions are minimal, all methods converge to (nearly) zero mean angular error, as expected, since most methods are based on a Lambertian model. However, in the low SNR (high noise) regime, we see that, as in our previous results, both proposed dictionary learning based methods are significantly more robust to noise and produce much more accurate reconstructions. Figure 3 illustrates the normal vectors obtained by the dictionary learning based approaches and error maps of all methods on the Cat dataset with an SNR of 1dB. Intuitively, the proposed adaptive dictionary learning methods are able to learn local features of data that effectively denoise the images (DLPI) or normal vectors (DLNV).

5. CONCLUSION

In this work, we investigated two dictionary learning based methods for robust photometric stereo. Each method seeks to represent some form of our data—either the original images or the estimated normal vectors—as sparse with respect to an adaptive (learned) dictionary. We showed that both approaches are significantly more robust to noise than existing methods. The results presented here indicate that DLPI usually outperforms DLNV, but which method performs best in general may depend on underlying properties of the data. We leave this nuanced investigation for future work.

[1] The data can be found at `http://vision.seas.harvard.edu/qsfs/Data.html`

6. REFERENCES

[1] R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Optical Engineering*, vol. 19, no. 1, pp. 191139–191139, 1980.

[2] H. Hayakawa, "Photometric stereo under a light source with arbitrary motion," *JOSA A*, vol. 11, no. 11, pp. 3079–3089, 1994.

[3] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille, "The bas-relief ambiguity," *International Journal of Computer Vision*, vol. 35, no. 1, pp. 33–44, Nov. 1999.

[4] A. L. Yuille, D. Snow, R. Epstein, and P. N. Belhumeur, "Determining generative models of objects under varying illumination: Shape and albedo from multiple images using SVD and integrability," *International Journal of Computer Vision*, vol. 35, no. 3, pp. 203–222, 1999.

[5] A. S. Georghiades, "Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo," in *ICCV*, 2003, vol. 2, pp. 816–823.

[6] S. Barsky and M. Petrou, "The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows," *IEEE PAMI*, vol. 25, no. 10, pp. 1239–1252, Oct. 2003.

[7] M. Chandraker, S. Agarwal, and D. Kriegman, "Shadowcuts: Photometric stereo with shadows," in *CVPR*, June 2007, pp. 1–8.

[8] F. Verbiest and L. Van Gool, "Photometric stereo with coherent outlier handling and confidence estimation," in *CVPR*, June 2008, pp. 1–8.

[9] C. Yu, Y. Seo, and S. W. Lee, "Photometric stereo from maximum feasible lambertian reflections," in *ECCV*, 2010, pp. 115–126.

[10] T-P Wu and C-K Tang, "Photometric stereo via expectation maximization," *IEEE PAMI*, vol. 32, no. 3, pp. 546–560, Mar. 2010.

[11] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma, "Robust photometric stereo via low-rank matrix completion and recovery," *ACCV*, pp. 703–717, 2011.

[12] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa, "Robust photometric stereo using sparse regression," in *CVPR*, 2012, pp. 318–325.

[13] M. Oren and S. K. Nayar, "Generalization of the lambertian model and implications for machine vision," *International Journal of Computer Vision*, vol. 14, no. 3, pp. 227–251, 1995.

[14] A. Hertzmann and S. M. Seitz, "Example-based photometric stereo: Shape reconstruction with general, varying BRDFs," *IEEE PAMI*, vol. 27, no. 8, pp. 1254–1264, Aug. 2005.

[15] N. G. Alldrin and D. J. Kriegman, "Toward reconstructing surfaces with arbitrary isotropic reflectance: A stratified photometric stereo approach," in *ICCV*, 2007, pp. 1–8.

[16] H-S Chung and J. Jia, "Efficient photometric stereo on glossy surfaces with wide specular lobes," in *CVPR*, June 2008, pp. 1–8.

[17] N. Alldrin, T. Zickler, and D. Kriegman, "Photometric stereo with non-parametric and spatially-varying reflectance," in *CVPR*, June 2008, pp. 1–8.

[18] D. B. Goldman, B. Curless, A. Hertzmann, and S. M. Seitz, "Shape and spatially-varying BRDFs from photometric stereo," *IEEE PAMI*, vol. 32, no. 6, pp. 1060–1071, 2010.

[19] T. Higo, Y. Matsushita, and K. Ikeuchi, "Consensus photometric stereo," in *CVPR*, June 2010, pp. 1157–1164.

[20] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi, "Elevation angle from reflectance monotonicity: Photometric stereo for general isotropic reflectances," *ECCV*, pp. 455–468, 2012.

[21] M. Chandraker, J. Bai, and R. Ramamoorthi, "On differential photometric reconstruction for unknown, isotropic BRDFs," *IEEE PAMI*, vol. 35, no. 12, pp. 2941–2955, 2013.

[22] S. Ikehata and K. Aizawa, "Photometric stereo using constrained bivariate regression for general isotropic surfaces," in *CVPR*, 2014, pp. 2179–2186.

[23] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi, "Bi-polynomial modeling of low-frequency reflectances," *IEEE PAMI*, vol. 36, no. 6, pp. 1078–1091, June 2014.

[24] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. on Image Proc.*, vol. 15, no. 12, pp. 3736–3745, 2006.

[25] M. Aharon, M. Elad, and A. Bruckstein, "k-svd: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. on Signal Proc.*, vol. 54, no. 11, pp. 4311–4322, 2006.

[26] S. Ravishankar and Y. Bresler, "MR image reconstruction from highly undersampled k-space data by dictionary learning," *IEEE Trans. on Med. Imag.*, vol. 30, no. 5, pp. 1028–1041, 2011.

[27] S. Ravishankar, B. E. Moore, R. R. Nadakuditi, and J. A. Fessler, "LASSI: A low-rank and adaptive sparse signal model for highly accelerated dynamic imaging," in *IVMSP Workshop*, 2016, pp. 1–5.

[28] B. Shi, Z. Wu, Z. Mo, D. Duan, S-K Yeung, and P. Tan, "A benchmark dataset and evaluation for non-lambertian and uncalibrated photometric stereo," in *CVPR*, 2016.

[29] T. Simchony, R. Chellappa, and M. Shao, "Direct analytical methods for solving poisson equations in computer vision problems," *IEEE PAMI*, vol. 12, no. 5, pp. 435–446, May 1990.

[30] S. Ravishankar, R. R. Nadakuditi, and J. A. Fessler, "Efficient sum of outer products dictionary learning (SOUP-DIL) - the ℓ_0 method," *arXiv preprint arXiv:1511.08842*, 2015.

[31] R. Gribonval and K. Schnass, "Dictionary identification–sparse matrix-factorization via l_1 -minimization," *IEEE Trans. on Inform. Theory*, vol. 56, no. 7, pp. 3523–3539, 2010.

[32] S. Ravishankar, B. E. Moore, R. R. Nadakuditi, and J. A. Fessler, "Efficient learning of dictionaries with low-rank atoms," in *Proc. IEEE Global Conference on Signal and Information Processing*, 2016.

[33] N. Parikh and S. Boyd, "Proximal algorithms," *Found. Trends Optim.*, vol. 1, no. 3, pp. 127–239, Jan. 2014.

[34] Y. Xiong, A. Chakrabarti, R. Basri, S. J. Gortler, D. W. Jacobs, and T. E. Zickler, "From shading to local shape," *IEEE PAMI*, vol. 37, no. 1, pp. 67–79, 2015.