

ROBUST SURFACE RECONSTRUCTION FROM GRADIENTS VIA ADAPTIVE DICTIONARY REGULARIZATION

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

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

ABSTRACT

This paper introduces a novel approach to robust surface reconstruction from photometric stereo normal vector maps that is particularly well-suited for reconstructing surfaces from noisy gradients. Specifically, we propose an adaptive dictionary learning based approach that attempts to simultaneously integrate the gradient fields while sparsely representing the spatial patches of the reconstructed surface in an adaptive dictionary domain. We show that our formulation learns the underlying structure of the surface, effectively acting as an adaptive regularizer that enforces a smoothness constraint on the reconstructed surface. Our method is general and may be coupled with many existing approaches in the literature to improve the integrity of the reconstructed surfaces. We demonstrate the performance of our method on synthetic data as well as real photometric stereo data and evaluate its robustness to noise.

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

1. INTRODUCTION

Imaging techniques such as photometric stereo [1] allow one to efficiently estimate the normal vector map of an object. The primary goal of such methods is to ultimately derive a three-dimensional representation of the object, which requires some flavor of numerical integration of the gradient fields defined by the normal vector map. Robust photometric stereo—the problem of accurately determining the normal map of a non-ideal surface or from noisy data—has attracted considerable attention in recent years [2, 3, 4]. In this work, we seek to develop a robust approach to the problem of reconstructing surfaces from gradient fields that can accurately estimate a 3D representation of an object in the presence of noise.

The problem of reconstructing a surface from estimates of its photometric stereo gradient fields has been investigated since the late 1980s. The seminal works of Simchony *et al.* [5] and Frankot and Chellappa [6] seek to solve the problem through a least squares approach, utilizing efficient discrete

Fourier transform (DFT) or discrete cosine transform (DCT) based solvers. Harker and O’Leary [7] propose a modified “global” least squares problem and extend this method to incorporate regularization [8], solving a Sylvester equation to obtain the solution. Recently, Quéau and Durou [9] introduced a weighted-least squares formulation as well as formulations minimizing total-variation and incorporating the ℓ_1 norm to promote sparsity. Further attempts at applying a regularization term while integrating the gradients have also been proposed at the expense of computation time [10, 11]. Additional approaches include line-integral based methods [12, 13] and reconstructions based on the calculus of variations [14, 15, 16]. A range of other methods have also been proposed with mixed results [17, 18, 19, 20, 21, 22].

Our work builds on these previous works, specifically those that utilize a least squares-type formulation to relate the underlying surface and its gradient fields. In particular, we propose a novel adaptive dictionary learning framework that enables the robust estimation of surfaces from noisy gradients. Dictionary learning [23, 24, 25] has, in recent years, been successfully applied to many imaging applications, e.g., [26, 27, 28]. In dictionary learning models, one typically seeks to learn sparse representations of local regions (patches) of the data. These models often induce a type of smoothness constraint on the underlying data that, in the case of surface reconstruction, we show leads to robust reconstructions with desirable noise rejection properties. Our framework is general and can be easily combined with any existing method that utilizes a least squares-type objective to estimate the underlying surface.

2. SURFACE RECONSTRUCTION FROM GRADIENT FIELDS

Let $n(x, y) \in \mathbb{R}^3$ denote the normal vector of a differentiable surface $z(x, y)$ at position (x, y) , and let $n_1(x, y)$, $n_2(x, y)$, and $n_3(x, y)$ denote the x , y , and z components of this vector, respectively. Under this ideal model, one can relate the x and y derivatives of the surface z to its normal vectors via the relation

$$\frac{\partial z(x, y)}{\partial x} = -p(x, y), \quad \frac{\partial z(x, y)}{\partial y} = q(x, y), \quad (1)$$

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.

where we have defined $p(x, y) := n_1(x, y)/n_3(x, y)$ and $q(x, y) := n_2(x, y)/n_3(x, y)$. In practice, the estimated (e.g., by photometric stereo) normal vectors of a surface and its gradient fields will not exactly satisfy (1), so one must instead find a function $z(x, y)$ with derivatives *close* to $p(x, y)$ and $q(x, y)$ in an appropriate sense, often by minimizing a variational problem of the form

$$\int \int_{\Omega} \left(\frac{\partial z(x, y)}{\partial x} - p(x, y) \right)^2 + \left(\frac{\partial z(x, y)}{\partial y} - q(x, y) \right)^2 dx dy. \quad (2)$$

When our data is instead sampled on a discrete grid, we will not have access to a continuous normal map $n(x, y)$ but will instead have a matrix $N \in \mathbb{R}^{m \times n \times 3}$ containing the normal vectors of the object on the grid. Following (1), we can compute matrices $P \in \mathbb{R}^{m \times n}$ and $Q \in \mathbb{R}^{m \times n}$ containing the measured gradients, and our goal then becomes to estimate the matrix $Z \in \mathbb{R}^{m \times n}$ containing the values of the surface $z(x, y)$ sampled on the grid. The discrete analogue of (2) is commonly expressed [5, 6, 7] as a standard least squares problem of the form

$$z^* = \arg \min_z \|Az - v\|_2^2, \quad (3)$$

where $z = \text{vec}(Z) \in \mathbb{R}^{mn}$ is the vectorized surface, A is a numerical differentiation operator, and the vector v is an appropriate function of the measured gradients, P and Q . Solving this problem yields a representation of our surface that is optimal in the least squares sense.

Note that the specific forms of A and v can vary. One possible formulation is

$$A = \begin{bmatrix} D_n \otimes I_m \\ I_n \otimes D_m \end{bmatrix}, \quad v = \begin{bmatrix} \text{vec}(P) \\ \text{vec}(Q) \end{bmatrix}, \quad (4)$$

where D_n is the discrete first differences matrix, and \otimes denotes the Kronecker product. However, multiple models are possible, each yielding reconstructed surfaces with different properties. Importantly, the dictionary learning based approach that we introduce in Section 3 can be coupled with any least squares model of the form (3), so our proposed approach is quite flexible.

3. ADAPTIVE DICTIONARY LEARNING REGULARIZATION

Given normal vectors corrupted by noise or other non-idealities, solving (3) directly generally produces a rough, bumpy surface, even when the underlying true surface is smooth. Thus, in this work, we propose an adaptive dictionary regularizer that can be combined with the least squares model (3) to more accurately estimate the underlying surface. In particular, we propose to solve the dictionary learning

problem

$$\arg \min_{z, B, D} \frac{1}{2} \|Az - v\|_2^2 + \lambda \left(\sum_{j=1}^c \|P_j z - D b_j\|_2^2 + \mu^2 \|B\|_0 \right) \quad (5)$$

s.t. $\|d_i\|_2 = 1, \|b_j\|_\infty \leq a, \forall i, j.$

Here, P_j is a patch extraction operator that extracts a vectorized $c_x \times c_y$ spatial patch from z , $D \in \mathbb{R}^{c_x c_y \times K}$ is a dictionary matrix whose columns d_i are vectorized $c_x \times c_y$ atoms, and $B \in \mathbb{R}^{K \times c}$ is a matrix of sparse codes, where column b_j of B defines the (usually sparse) linear combination of atoms used to represent the patch $P_j z$ of z . Also, $\|\cdot\|_0$ is the familiar ℓ_0 “norm” and $\lambda, \mu > 0$ are regularization parameters.

We include the constraints $\|b_j\|_\infty \leq a$, where a is typically very large, to prevent pathologies that could theoretically arise during minimization since (5) is non-coercive with respect to B , but the constraint is inactive in practice [29]. In addition, we constrain the columns of D to be unit-norm to avoid scaling ambiguity between D and B [30]. Note that (5) is a non-convex problem.

By solving (5), we are attempting to estimate our surface z by numerically integrating its gradient fields through a least squares functional while simultaneously enforcing that local patches of the reconstructed surface should have sparse representations with respect to the dictionary D . As D itself is learned, our proposed algorithm can automatically adapt to the underlying properties of the surface and its gradients. Intuitively, since the same dictionary of atoms is used to (sparsely) represent all patches, the atoms can learn universal features of the surface. This effectively equips (5) with global information to estimate each patch, which (as we will show) can yield robust reconstructions when the gradients are noisy/corrupted.

3.1. Dictionary Learning on Surfaces (DLS) Algorithm

We propose to solve (5) via a block coordinate descent-type algorithm where we alternate between updating z with (D, B) fixed and updating (D, B) with z fixed. Henceforward, we refer to this algorithm as the Dictionary Learning on Surfaces (DLS) method.

3.1.1. (D, B) updates

Let P be the matrix with columns $P_j z$. With z fixed, the minimization of (5) with respect to (D, B) is

$$\min_{B, D} \|P - DB\|_F^2 + \mu^2 \|B\|_0 \quad (6)$$

s.t. $\|d_i\|_2 = 1, \|b_j\|_\infty \leq a, \forall i, j.$

We solve (6) via a block coordinate descent method where we iteratively minimize the cost with respect to the i th row of B and the i th atom, d_i , of D for every $1 \leq i \leq K$ with all other variables held fixed. A full derivation of this step can be found in [29, 31], which we omit here due to space considerations.

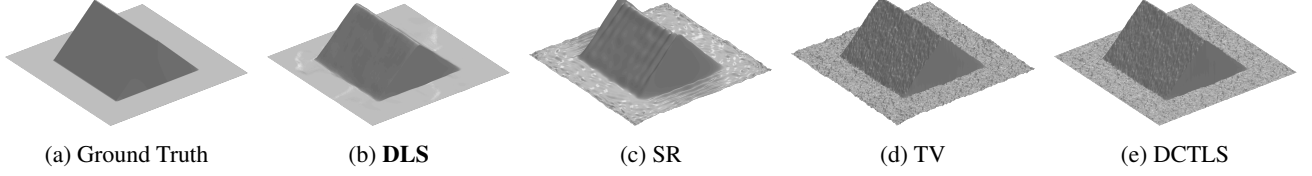


Fig. 1: Surface reconstructions of the Tent dataset with SNR = 20 dB.

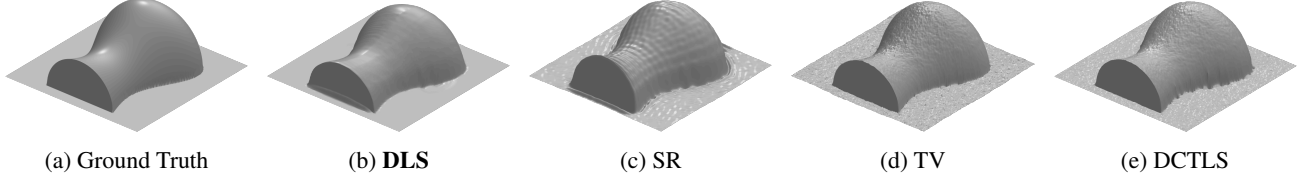


Fig. 2: Surface reconstructions of the Vase dataset with SNR = 30 dB.

3.1.2. z update

With D and B fixed, our problem becomes

$$\min_z \frac{1}{2} \|Az - v\|_2^2 + \lambda \sum_{j=1}^c \|P_j z - Db_j\|_2^2. \quad (7)$$

The cost in (7) can be written in the form $f(z) + g(z)$, where

$$f(z) = \frac{1}{2} \|Az - v\|_2^2, \quad g(z) = \lambda \sum_{j=1}^c \|P_j z - Db_j\|_2^2. \quad (8)$$

We utilize a proximal gradient [32] strategy to solve (7), iteratively updating z according to

$$z^{k+1} = \text{prox}_{\tau g}(z^k - \tau \nabla f(z^k)), \quad (9)$$

where

$$\text{prox}_{\tau g}(y) := \arg \min_x \frac{1}{2} \|y - x\|_2^2 + \tau g(x). \quad (10)$$

Defining $\tilde{z}^k := z^k - \tau \nabla f(z^k)$, we see that (9) and (10) imply that z^{k+1} satisfies the normal equation

$$\left(I + 2\tau \lambda \sum_{j=1}^c P_j^T P_j \right) z^{k+1} = \tilde{z}^k + 2\tau \lambda \sum_{j=1}^c P_j^T Db_j. \quad (11)$$

The matrix on the left hand side of (11) is diagonal, so its inverse can be cheaply computed to solve for z^{k+1} . Note that (7) is a simple least squares problem and, as such, could be minimized by other methods (e.g., conjugate gradients).

4. RESULTS

In this section, we numerically evaluate our proposed DLS method on several datasets. In each case we compare our method to the spectral regularization method (SR) [8], the

isotropic total variation (TV) method [9], and DCT based least squares (DCTLS) [5]. For methods that include tunable parameters, we sweep over a wide range of values, reporting the best results obtained. Our proposed DLS method can incorporate any least squares based solver by simply defining A and v in (5) accordingly. For all results given here, we use the least squares cost found in [5]. To evaluate the robustness of each algorithm, we add Gaussian noise to the data.

For our proposed DLS method, we use dictionary atoms of size 8×8 pixels and a square 64×64 dictionary D , initialized with a DCT matrix. We extracted patches from z using a spatial stride of two pixels in each direction, allowing adjacent patches to overlap. Finally, we initialized z as the vectorized surface produced by solving the stand-alone least squares problem in [5], and initialized $B = 0$.

4.1. Synthetic Surface Reconstructions

To quantitatively evaluate our method, we first considered two synthetic datasets, which we call “Tent” and “Vase”, for which we have analytic expressions for $z = f(x, y)$. Given $f(x, y)$, we can differentiate to obtain the gradients, $\partial f(x, y)/\partial x$ and $\partial f(x, y)/\partial y$, and sample on a discrete grid. After reconstructing the surface from these gradients subject to additive noise, we evaluate the integrity of the reconstructions against the ground truth, $f(x, y)$, via the SSIM metric [34]. In these experiments, we add noise directly to the gradient fields to achieve a prescribed signal-to-noise ratio (SNR).

Figures 1 and 2 show the reconstructed surfaces produced by each algorithm. As these images illustrate, the proposed DLS method produces significantly smoother surfaces from noisy data compared to the existing methods. Intuitively, the locally sparse model imposed by the dictionary regularization denoises the surfaces, while the adaptive nature of the dictionary allows DLS to represent and reconstruct both sharp edges and smooth regions on a data-dependent basis.

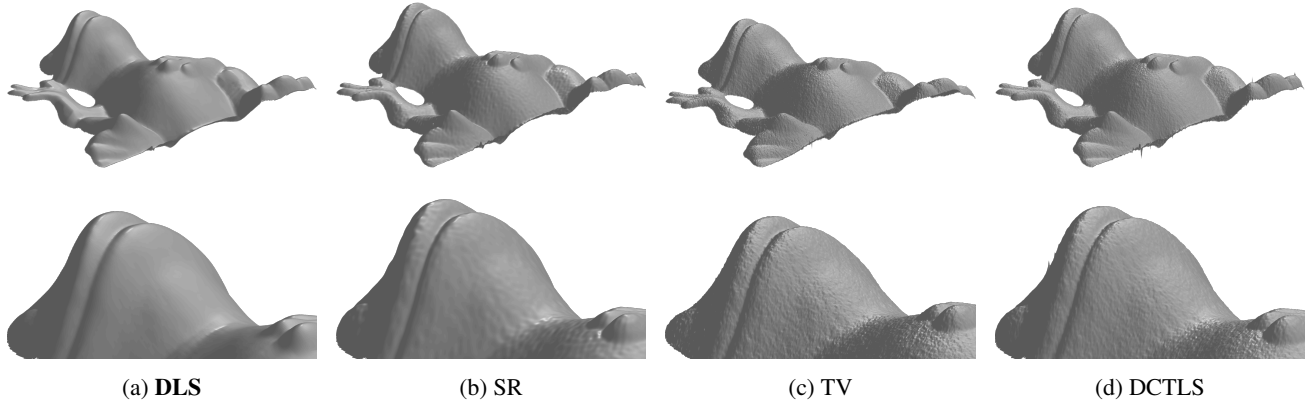


Fig. 3: Surface reconstructions of the photometric stereo Frog dataset [33] with SNR = 17 dB.

The surfaces obtained by SR, TV, and DCTLS are much more sensitive to the noisy gradients. Indeed, while they retain the general shape of the surface, they exhibit significantly more corruption. In particular, the spectral regularization method seems to introduce a systematic “rippling” into the reconstructions.

SNR (dB)	DLS	SR	TV	DCTLS
1	0.969	0.944	0.918	0.924
5	0.971	0.950	0.938	0.944
10	0.976	0.956	0.957	0.962
20	0.988	0.969	0.979	0.983
30	0.995	0.978	0.989	0.992
40	0.997	0.985	0.994	0.996
50	0.998	0.988	0.996	0.998
60	0.999	0.989	0.997	0.998

Table 1: Quality of Tent reconstructions in SSIM as a function of SNR.

SNR (dB)	DLS	SR	TV	DCTLS
1	0.958	0.930	0.889	0.894
5	0.966	0.934	0.911	0.915
10	0.971	0.942	0.933	0.936
20	0.977	0.961	0.965	0.966
30	0.982	0.975	0.981	0.981
40	0.990	0.982	0.990	0.989
50	0.993	0.984	0.993	0.992
60	0.995	0.985	0.995	0.993

Table 2: Quality of Vase reconstructions in SSIM as a function of SNR.

Tables 1 and 2 numerically corroborate the qualitative results from Figures 1 and 2. In the low SNR regime, DLS significantly outperforms the other approaches. As SNR increases, the gap decreases. When the data is essentially noiseless, DLS, TV, and DCTLS are all able to reconstruct the surfaces with comparably negligible error.

4.2. Photometric Stereo

We now return to the problem of reconstructing a 3D representation of an object from normal vectors obtained through photometric stereo. We consider a dataset containing 10 images, each taken under a unique, known lighting direction, and corrupt the images with Gaussian noise at a prescribed SNR. We compute the normal vectors from the noisy images using the standard least squares approach [2]. Given the normal vectors, we then compute gradient fields as discussed in Section 2, and from those we generate a 3D reconstruction of the object using each method. Figure 3 illustrates the results of this procedure on the Frog dataset,¹ which contains real images of a frog statue [33].

The reconstructions in Figure 3 showcase the ability of the proposed DLS approach to produce a smoother surface from the noisy gradients compared to the existing methods. The denoising capability of DLS may prove valuable when running photometric stereo on real-world data, where noise and other non-idealities are inevitable.

5. CONCLUSION

In this work, we explored the use of adaptive dictionary learning based regularization for the estimation of surfaces from their gradient fields. We showed that our proposed dictionary learning approach is able to effectively reject the addition of noise to gradient fields/images and produce more accurate and smooth representations of the underlying surfaces compared to existing methods. Our dictionary learning framework is very general and would be straightforward to combine with many existing algorithms. In future work, we plan to investigate these combinations and perform a more thorough study of the influence of the various parameters of our dictionary learning model on the reconstructed surfaces.

¹This dataset 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] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma, "Robust photometric stereo via low-rank matrix completion and recovery," in *ACCV*, 2010, pp. 703–717.
- [3] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa, "Robust photometric stereo using sparse regression," in *CVPR*, 2012, pp. 318–325.
- [4] S. Ikehata and K. Aizawa, "Photometric stereo using constrained bivariate regression for general isotropic surfaces," in *CVPR*, June 2014, pp. 2187–2194.
- [5] 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.
- [6] R. T. Frankot and R. Chellappa, "A method for enforcing integrability in shape from shading algorithms," *IEEE PAMI*, vol. 10, no. 4, pp. 439–451, 1988.
- [7] M. Harker and P. O'Leary, "Least squares surface reconstruction from measured gradient fields," in *CVPR*, 2008, pp. 1–7.
- [8] M. Harker and P. O'Leary, "Regularized reconstruction of a surface from its measured gradient field," *Journal of Mathematical Imaging and Vision*, vol. 51, no. 1, pp. 46–70, 2015.
- [9] Y. Quéau and J-D Durou, "Edge-preserving integration of a normal field: Weighted least-squares, TV and l^1 approaches," in *International Conference on Scale Space and Variational Methods in Computer Vision*, 2015, pp. 576–588.
- [10] A. Agrawal, R. Raskar, and R. Chellappa, "What is the range of surface reconstructions from a gradient field?," in *ECCV*, 2006, pp. 578–591.
- [11] H-S Ng, T-P Wu, and C-K Tang, "Surface-from-gradients without discrete integrability enforcement: A gaussian kernel approach," *IEEE PAMI*, vol. 32, no. 11, pp. 2085–2099, 2010.
- [12] Z. Wu and L. Li, "A line-integration based method for depth recovery from surface normals," *Computer Vision, Graphics, and Image Processing*, vol. 43, no. 1, pp. 53–66, 1988.
- [13] A. Robles-Kelly and E. R. Hancock, "A graph-spectral method for surface height recovery," *Pattern Recognition*, vol. 38, no. 8, pp. 1167–1186, 2005.
- [14] B. K. P. Horn and M. J. Brooks, "The variational approach to shape from shading," *Computer Vision, Graphics, and Image Processing*, vol. 33, no. 2, pp. 174–208, 1986.
- [15] J. Balzer and T. Mörwald, "Isogeometric finite-elements methods and variational reconstruction tasks in vision—a perfect match," in *CVPR*, 2012, pp. 1624–1631.
- [16] J-D Durou and F. Courteille, "Integration of a normal field without boundary condition," in *PACV*, 2007, pp. 8–p.
- [17] I. Horowitz and N. Kiryati, "Depth from gradient fields and control points: Bias correction in photometric stereo," *Image and Vision Computing*, vol. 22, no. 9, pp. 681–694, 2004.
- [18] K. M. Lee and C-C J. Kuo, "Surface reconstruction from photometric stereo images," *JOSA A*, vol. 10, no. 5, pp. 855–868, 1993.
- [19] B. Karaçalı and W. Snyder, "Reconstructing discontinuous surfaces from a given gradient field using partial integrability," *Computer Vision and Image Understanding*, vol. 92, no. 1, pp. 78–111, 2003.
- [20] B. Karaçalı and W. Snyder, "Noise reduction in surface reconstruction from a given gradient field," *International Journal of Computer Vision*, vol. 60, no. 1, pp. 25–44, 2004.
- [21] P. Kovesi, "Shapelets correlated with surface normals produce surfaces," in *ICCV*, 2005, vol. 2, pp. 994–1001.
- [22] J. Balzer, "A gauss-newton method for the integration of spatial normal fields in shape space," *Journal of Mathematical Imaging and Vision*, vol. 44, no. 1, pp. 65–79, 2012.
- [23] 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.
- [24] 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.
- [25] K. Kreutz-Delgado, J. F. Murray, B. D. Rao, K. Engan, T-W Lee, and T. J. Sejnowski, "Dictionary learning algorithms for sparse representation," *Neural Computation*, vol. 15, no. 2, pp. 349–396, 2003.
- [26] S. Ravishanker 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. Ravishanker, 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] S. Ravishanker, B. Moore, R. R. Nadakuditi, and J. Fessler, "Low-rank and adaptive sparse signal (LASSI) models for highly accelerated dynamic imaging," *IEEE Trans. on Med. Imag.*, 2017.
- [29] S. Ravishanker, 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.
- [30] 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.
- [31] S. Ravishanker, 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, pp. 222–226.
- [32] N. Parikh and S. Boyd, "Proximal algorithms," *Found. Trends Optim.*, vol. 1, no. 3, pp. 127–239, Jan. 2014.
- [33] 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.
- [34] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. on Image Proc.*, vol. 13, no. 4, pp. 600–612, 2004.