WEIGHTED ANISOTROPIC – ISOTROPIC TOTAL VARIATION FOR POISSON DENOISING

Kevin Bui[1], Yifei Lou[2], Fredrick Park[3], *and* Jack Xin[1]

[1]Department of Mathematics; University of California, Irvine; Irvine, CA 92697, United States
[2]Department of Mathematical Sciences; University of Texas, Dallas; Richardson, TX 75080, United States
[3]Department of Mathematics & Computer Science; Whittier College; Whittier, CA 90602, United States

ABSTRACT

Poisson noise commonly occurs in images captured by photon-limited imaging systems such as in astronomy and medicine. As the distribution of Poisson noise depends on the pixel intensity value, noise levels vary from pixels to pixels. Hence, denoising a Poisson-corrupted image while preserving important details can be challenging. In this paper, we propose a Poisson denoising model by incorporating the weighted anisotropic–isotropic total variation (AITV) as a regularization. We then develop an alternating direction method of multipliers with a combination of a proximal operator for an efficient implementation. Lastly, numerical experiments demonstrate that our algorithm outperforms other Poisson denoising methods in terms of image quality and computational efficiency.

Index Terms— Poisson noise, total variation, nonconvex optimization, ADMM, proximal operator

1. INTRODUCTION

In various applications such as astronomy [1] and medicine [2], photon-counting devices are utilized to capture images. However, these images are susceptible to Poisson noise, in which any measured intensity at each pixel is a realization of a Poisson random variable with mean and variance equal to its true value. Recall that a Poisson random variable with mean and variance has a probability distribution function $\mathbb{P}_\mu(n) = \frac{e^{-\mu}\mu^n}{n!}$, $n \geq 0$. Let g be a clean image of size $M \times N$. If g is corrupted by Poisson noise, then its corresponding noisy measurement f can be formulated as

$$f_{i,j} \sim \text{Poisson}(g_{i,j}),\ 1 \leq i \leq M, 1 \leq j \leq N.$$

One general approach for Poisson denoising is by the maximum a posteriori (MAP)

$$\prod_{i,j} \mathbb{P}(u_{i,j}|f_{i,j}) = \prod_{i,j} \frac{e^{-u_{i,j}} u_{i,j}^{f_{i,j}}}{(f_{i,j})!} \frac{\mathbb{P}(u_{i,j})}{\mathbb{P}(f_{i,j})},$$

with respect to an estimated image u from a noisy input f. Taking the negative logarithm yields the following optimization problem to minimize:

$$\sum_{i,j} u_{i,j} - f_{i,j} \log u_{i,j} - \log \mathbb{P}(u_{i,j}).$$

The term $-\log \mathbb{P}(u_{i,j})$ can be regarded as an image prior. A classic choice is the total variation (TV) [3]. Le *et al.* [4] derived a TV-regularized model for Poisson denoising:

$$\min_u \lambda\langle u - f \log u, \mathbb{1}\rangle + \|\nabla u\|_{2,1}, \tag{1}$$

where $\mathbb{1}$ denotes the all-one vector and $\|\nabla u\|_{2,1}$ is the isotropic form of TV. TV has been a popular regularizer for denoising, but it only performs well on piecewise-smooth images. For natural images, TV tends to introduce staircasing artifacts, lose image contrasts, and smear out fine details [5]. Several variants of TV have been proposed including nonlocal TV [6] and total generalized variation [7]. Unfortunately, these TV variants are numerically complicated to compute. As an alternative, fractional-order total variation (FOTV) has a convenient, compact discrete form and was demonstrated to perform well in Poisson denoising [8].

More complex models such as nonlocal methods [9, 10, 11, 12, 13, 14] and convolutional neural networks (CNNs) [15, 16, 17] have been developed to perform Poisson denoising. However, nonlocal methods are computationally expensive due to the needs to compare image patches in groups in terms of their similarity and to enforce a low-rank structure. As highly nonconvex models, CNNs require an adequately large training set, hyperparameter tuning (e.g., batch size, learning rate, number of layers, etc.), and heavy computational resources such as GPUs. Since CNNs demonstrate the promising potential of nonconvex modeling, one direction is to develop a nonconvex variant of (1) to avoid most of their computational limitations.

Studies have demonstrated that nonconvex regularizers can preserve edges better than convex models [18]. For example, $\text{TV}^p(0 < p < 1)$ preserves edges better than the convex TV model in Gaussian denoising [19]. Another nonconvex TV variant is called weighted anisotropic–isotropic TV (AITV) [20] that recovers sharper images than TV and

TV^p for images corrupted by Gaussian noise [20]. Despite its recent success, AITV has not yet been applied in Poisson denoising and is slow in computations. In this paper, we propose a variational Poisson denoising model with the AITV regularization and improve the efficiency by taking advantage of a proximal operator [21] over its original implementation that uses the difference-of-convex algorithm (DCA).

2. THE PROPOSED APPROACH

2.1. Model Formulation

Two popular numerical approximations of TV are the isotropic TV and the anisotropic TV. Specifically, we represent an image as an $M \times N$ matrix and define an Euclidean space $X := \mathbb{R}^{M \times N}$ with the standard inner product $\langle \cdot, \cdot \rangle_X$ and the Euclidean norm $\| \cdot \|_2$. We will omit the subscript X and use $\langle \cdot, \cdot \rangle$ for the sake of brevity.

To discretize the image gradient, we define another Euclidean space $Y := X \times X$. The discrete gradient operator $\nabla : X \to Y$ is given by $(\nabla u)_{i,j} = ((\nabla_x u)_{i,j}, (\nabla_y u)_{i,j})$, where ∇_x, ∇_y are the horizontal and vertical difference operators. For the space Y, we define the inner product by

$$\langle p, q \rangle = \langle p_1, q_1 \rangle_X + \langle p_2, q_2 \rangle_X,$$

for $p = (p_1, p_2), q = (q_1, q_2) \in Y$. We also define the following norms on Y:

$$\|p\|_1 = \sum_{i=1}^{M} \sum_{j=1}^{N} |(p_1)_{i,j}| + |(p_2)_{i,j}|,$$

$$\|p\|_2 = \sqrt{\sum_{i=1}^{M} \sum_{j=1}^{N} |(p_1)_{i,j}|^2 + |(p_2)_{i,j}|^2},$$

$$\|p\|_{2,1} = \sum_{i=1}^{M} \sum_{j=1}^{N} \sqrt{(p_1)_{i,j}^2 + (p_2)_{i,j}^2}.$$

By our definitions, the isotropic TV and anisotropic TV can be formulated as $\|\nabla u\|_{2,1}$ and $\|\nabla u\|_1$, respectively. Unfortunately both of these fail in recovering oblique edges [22]. To mitigate this artifact, Lou *et al.* [20] proposed the AITV regularizer $\|\nabla u\|_1 - \alpha \|\nabla u\|_{2,1}$, where the parameter $\alpha \in [0, 1]$ controls the sparsity of the gradient at each pixel. Replacing the isotropic TV in (1) with AITV for a pre-defined $\alpha \in [0, 1]$, we arrive at the proposed model:

$$\min_u \lambda \langle u - f \log u, \mathbb{1} \rangle + \|\nabla u\|_1 - \alpha \|\nabla u\|_{2,1}. \quad (2)$$

2.2. Numerical Algorithm

We develop an alternating direction method of multipliers (ADMM) [23] to solve for (2). By introducing two auxiliary variables $v \in X$ and $w = (w_x, w_y) \in Y$, we have the following constrained optimization problem:

$$\min_{u,v,w} \quad \lambda \langle v - f \log v, \mathbb{1} \rangle + \|w\|_1 - \alpha \|w\|_{2,1}$$
$$\text{s.t.} \quad u = v \quad \text{and} \quad \nabla u = w. \quad (3)$$

Then its augmented Lagrangian is written as

$$\mathcal{L}_\beta(u, v, w, y, z) = \lambda \langle v - f \log v, \mathbb{1} \rangle + \|w\|_1 - \alpha \|w\|_{2,1}$$
$$+ \langle y, u - v \rangle + \frac{\beta}{2} \|u - v\|_2^2 + \langle z, \nabla u - w \rangle + \frac{\beta}{2} \|\nabla u - w\|_2^2,$$

where $y \in X$, $z = (z_x, z_y) \in Y$ are the Lagrange multipliers and $\beta > 0$ is a penalty parameter. As a result, ADMM iterates as follows:

$$u_{k+1} = \arg\min_u \mathcal{L}_\beta(u, v_k, w_k, y_k, z_k) \quad (4a)$$

$$v_{k+1} = \arg\min_v \mathcal{L}_\beta(u_{k+1}, v, w_k, y_k, z_k) \quad (4b)$$

$$w_{k+1} = \arg\min_w \mathcal{L}_\beta(u_{k+1}, v_{k+1}, w, y_k, z_k) \quad (4c)$$

$$y_{k+1} = y_k + \beta_k(u_{k+1} - v_{k+1}) \quad (4d)$$

$$z_{k+1} = z_k + \beta_k(\nabla u_{k+1} - w_{k+1}) \quad (4e)$$

$$\beta_{k+1} = \sigma \beta_k, \quad (4f)$$

where $\sigma > 1$. The last step (4f) is inspired from [24] to accelerate the numerical convergence of ADMM. If σ is too large, the algorithm might stop too early, yielding an unsatisfactory solution. Hence, σ needs to be chosen carefully.

We derive closed-form solutions for the subproblems (4a)-(4c). The first-order optimality condition for (4a) is

$$\beta_k(I - \Delta)u_{k+1} = \beta_k v_k - y_k - \nabla^\top(z_k - \beta_k w_k), \quad (5)$$

where $\Delta = -\nabla^\top \nabla$ is the Laplacian operator. By assuming periodic boundary condition for u, (5) can be solved efficiently by the 2D discrete Fourier transform $\mathcal{F}$ [25], thus leading to an update of u_{k+1} to be

$$u_{k+1} = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(\beta_k v_k - y_k) - \mathcal{F}(\nabla)^* \circ \mathcal{F}(z_k - \beta_k w_k)}{\beta_k \mathcal{F}(I - \Delta)}\right), \quad (6)$$

where $\mathcal{F}^{-1}$ is the inverse Fourier transform, the superscript $*$ denotes complex conjugate, $\circ$ denotes the componentwise product, and the division is componentwise as well. By taking derivative of (4b) with respect to v and setting it to zero, we get the closed-form solution for

$$v_{k+1} = \frac{r_k + \sqrt{r_k^2 + 4\lambda \beta_k f}}{2\beta_k}, \quad (7)$$

where $r_k = \beta_k u_{k+1} + y_k - \lambda \mathbb{1}$ and all the operations (square root, square, and division) are componentwise. Lastly, the w-subproblem (4c) can be decomposed independently at each

pixel (i, j), i.e.,

$$(w_{i,j})_{k+1} = \arg\min_{w_{i,j}} \|w_{i,j}\|_1 - \alpha\|w_{i,j}\|_2$$
$$+ \frac{\beta_k}{2} \left\| w_{i,j} - \left((\nabla u_{k+1})_{i,j} + \frac{(z_k)_{i,j}}{\beta_k} \right) \right\|_2^2. \quad (8)$$

The optimization problem for each component of w is a special case of the proximal operator for $\ell_1 - \alpha\ell_2$, defined by

$$\text{prox}(x, \alpha, \beta) = \arg\min_y \|y\|_1 - \alpha\|y\|_2 + \frac{\|x-y\|_2^2}{2\beta}. \quad (9)$$

With the help of the proximal operator (9), we obtain a closed-form solution to update every $w_{i,j}$ by

$$(w_{i,j})_{k+1} = \text{prox}\left((\nabla u_{k+1})_{i,j} + \frac{(z_k)_{i,j}}{\beta_k}, \alpha, \frac{1}{\beta_k} \right).$$

As derived in [21], the proximal operator for $\ell_1 - \alpha\ell_2$ has a closed-form solution formulated in Lemma 1.

Lemma 1 ([21]). *Given $x \in \mathbb{R}^n$, $\beta > 0$, and $\alpha \in [0, 1]$, the optimal solution to (9) is given by one of the following cases:*

1. *When $\|x\|_\infty > \beta$, we have $x^* = (\|\xi\|_2 + \alpha\beta)\frac{\xi}{\|\xi\|_2}$, where $\xi = \text{sign}(x) \circ \max(|x| - \beta, 0)$.*

2. *When $(1-\alpha)\beta < \|x\|_\infty \le \beta$, then x^* is a 1-sparse vector such that one chooses $i \in \arg\max_j(|x_j|)$ to define $x_i^* = (|x_i| + (\alpha - 1)\beta)\text{sign}(x_i)$ and set the remaining elements to 0.*

3. *When $\|x\|_\infty \le (1-\alpha)\beta$, then $x^* = 0$.*

The overall ADMM framework to solve (2) is described in Algorithm 1. By emulating the proof of [24, Theorem 2], we have $\|u_{k+1} - u_k\|_2 \to 0$, which corresponds to the stopping criterion in Algorithm 1. Algorithm 1 is expected to converge within a reasonable number of iterations. Although global convergence was proven for nonconvex ADMM [26], it may not be guaranteed for our algorithm since the gradient operator does not satisfy the necessary surjectivity condition.

Algorithm 1 ADMM for (2)

Require: Noisy image f, fidelity parameter λ, penalty parameter β_0, penalty multiplier $\sigma > 1$.
1: Initialize u_0, w_0, z_0.
2: Set $k = 0$.
3: **while** $\frac{\|u_k - u_{k-1}\|_2}{\|u_k\|_2} > \epsilon$ **do**
4:     Compute u_{k+1} by (6).
5:     Compute v_{k+1} by (7).
6:     Compute w_{k+1} by (8).
7:     $y_{k+1} = y_k + \beta_k(u_{k+1} - v_{k+1})$.
8:     $z_{k+1} = z_k + \beta_k(\nabla u_{k+1} - w_{k+1})$.
9:     $\beta_{k+1} = \sigma\beta_k$.
10:     $k := k + 1$.
11: **end while**
12: **return** Denoised image $u^* = u_{k+1}$.

(a) River   (b) Butterfly   (c) Starfish

(d) Penguin   (e) Boat
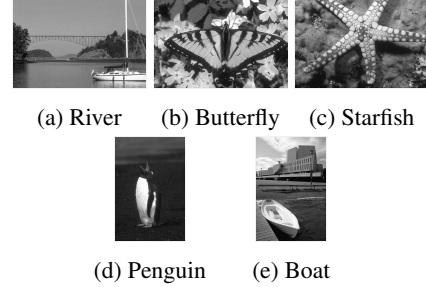
Fig. 1: Original natural images for testing. The image sizes are either 321×481 (top) or 481×321 (bottom).

3. NUMERICAL RESULTS

We evaluate the AITV-regularized Poisson denoising model (2) on five grayscale images selected from the Berkeley Segmentation Dataset [27]. The original images are shown in Figure 1. We compare our proposed AITV model with the classical TV [4], non-local PCA (NL-PCA) [10], and a recent Poisson denoising method by FOTV [8]. We use the MATLAB codes provided by the authors of NL-PCA and FOTV. Note that FOTV (including TV as its special case) is solved by ADMM, which is different from Algorithm 1 in that its penalty parameter β is fixed and it solves a nonlinear equation per iteration. Quantitatively, we evaluate the performance of image denoising by peak-signal-to-noise ratio (PSNR) and structural similiarity index (SSIM). The experiments are performed in MATLAB R2021b on a Dell laptop with a 1.80 GHz Intel Core i7-8565U processor and 16.0 GB of RAM. The code is available at `https://github.com/kbui1993/Official_AITV_Poisson_Denoising`.

Table 1: PSNR/SSIM of the four denoising methods under three peak values of the original images in Figure 1. **Bold** indicates the best value. The last column is the average over the five testing images.

| Peak	Method	river	butterfly	starfish	penguin	boat	Avg.
	Noisy	22.03/0.36	22.61/0.53	22.73/0.56	24.75/0.35	22.46/0.53	22.92/0.47
	NL-PCA	29.35/0.71	28.15/**0.83**	28.52/**0.83**	33.41/0.88	28.23/0.79	29.53/0.81
80	TV	30.39/0.84	28.40/**0.83**	28.31/0.81	33.40/0.87	28.40/0.80	29.78/0.83
	FOTV	30.20/0.84	28.50/0.82	**28.63**/0.82	33.25/0.87	28.23/0.78	29.76/0.83
	AITV	**30.70/0.85**	28.46/**0.83**	28.33/0.81	**33.89/0.90**	**28.69/0.81**	**30.01/0.84**
	Noisy	20.38/0.30	20.54/0.46	21.08/0.48	23.08/0.28	20.81/0.47	21.18/0.40
	NL-PCA	28.42/0.75	**27.46/0.80**	**27.88/0.80**	32.81/0.85	27.54/0.75	28.82/0.79
55	TV	29.06/0.83	27.26/0.80	27.32/0.77	32.51/0.87	27.29/0.76	28.69/0.80
	FOTV	28.93/0.76	27.40/0.78	27.62/0.79	32.03/0.80	27.17/0.76	28.63/0.78
	AITV	**29.67/0.83**	27.43/0.80	27.36/0.78	**33.26/0.88**	**27.67/0.77**	**29.08/0.81**
	Noisy	17.74/0.23	17.86/0.36	18.41/0.36	20.42/0.18	18.21/0.37	18.53/0.30
	NL-PCA	27.57/0.68	**25.87/0.72**	**26.29/0.73**	31.96/0.85	26.16/**0.72**	**27.57**/0.74
30	TV	27.14/0.61	25.46/0.70	25.71/0.70	29.68/0.71	25.62/0.68	26.72/0.68
	FOTV	27.11/0.68	25.58/0.71	25.94/0.72	28.69/0.61	25.51/0.68	26.57/0.68
	AITV	**28.18/0.79**	25.71/0.75	25.79/0.72	**31.99/0.87**	**26.17**/0.72	**27.57/0.77**

The parameters for each method are carefully tuned for the best PSNR. As the TV, FOTV, and AITV models are solved by ADMM, their parameters are nearly the same. We fix the penalty parameter $\beta = 10^{-3}$ and find the optimal fidelity parameter λ from $\{3, 5, 8, 10, 12, 15, 20\}$. The frac-

1022

(a) clean   (b) noisy   (c) NL-PCA   (d) TV   (e) FOTV   (f) AITV
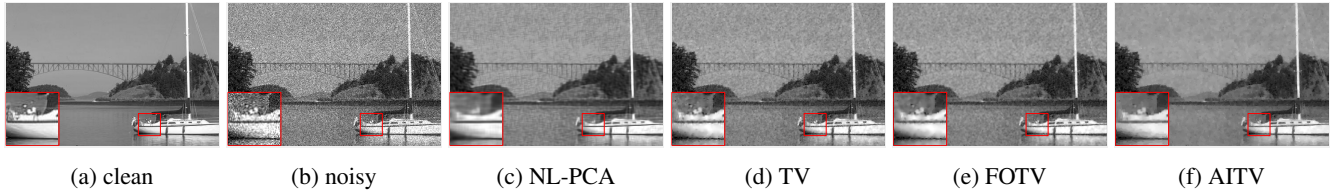
Fig. 2: Poisson denoising results for Figure 1a with peak value 30.

Method	Avg. Time (s)
NL-PCA	20.05
TV	16.59
FOTV	17.47
AITV	1.89

Table 2: Average computational time in seconds.
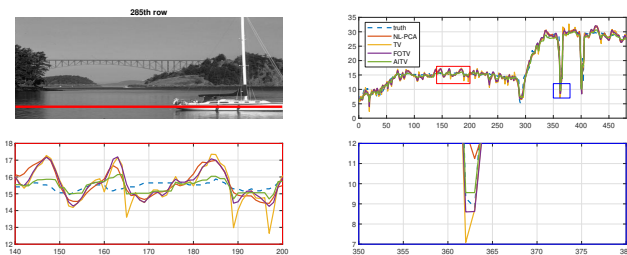
Fig. 3: Reconstructed line profiles from Figure 2 over the ground truth line profile for the 285th row of Figure 1a.

tional order for FOTV is optimized among $\{1.2, 1.4, 1.6, 1.8\}$. The AITV parameter α is optimized among $\{0.1i\}_{i=1}^{5}$. We fix $\sigma = 1.75$ in the proposed Algorithm 1. The stopping conditions are up to 300 iterations with a relative error stopping criterion $\frac{\|u_k - u_{k-1}\|_2}{\|u_k\|_2} < 10^{-5}$. For NL-PCA, we tune the patch size and the number of clusters, which are selected from $\{3, 5, 7, 9\}$ and $\{15, 20, 25\}$, respectively.

Since Poisson noise depends on the pixel intensity, we control the noise level by changing the peak value of an image. In particular, before adding Poisson noise to an image, we rescale its peak value to 80, 55, and 30 with a lower peak corresponding to a noisier image. Table 1 records the PSNR and SSIM metrics of the denoised images by various competing methods, showing that the proposed AITV model achieves the best results in most cases. The average results are also shown in the last column, showcasing that AITV is consistently the best under the three peak values.

We examine Figure 1a. The image has peak value 30 before adding Poisson noise. The denoised results are presented in Figure 2. By focusing on the sky, the denoised image by AITV looks less noisy compared to the other methods. From the enlarged window with red boundary, the result appears sharper for AITV than the other methods. Figure 3 compares a line profile from a denoised image with its original, showing that AITV is the least noisy.

Table 2 reports the average computational time over all

the testing scenarios (any combination of five images and three peak levels), demonstrating that the proposed Algorithm 1 is nearly one order of magnitude (2 seconds) faster than the other methods (about 20 seconds). Overall, the proposed AITV model solved by our designed ADMM algorithm is the most efficient in computation time and most effective in PSNR and SSIM.

4. CONCLUSION

In this work, we formulated a variational Poisson model with AITV regularization. To solve the model, we designed an efficient ADMM incorporating the $\ell_1 - \alpha\ell_2$ proximal operator that numerically converges within seconds. Our experiments demonstrated the efficiency of the proposed approach over several representative Poisson denoising methods in terms of quantitative measures, visual appearance, and computational time. For future directions, we will extend the proposed model to color images and design a globally convergent algorithm. Extending to deep learning, we will unfold the proposed ADMM algorithm in a similar fashion as [28].

5. REFERENCES

[1] H. Lantéri and C. Theys, "Restoration of astrophysical images—the case of Poisson data with additive Gaussian noise," *EURASIP Journal on Advances in Signal Processing*, vol. 2005, no. 15, pp. 1–14, 2005.

[2] Y. Vardi, L. A. Shepp, and L. Kaufman, "A statistical model for positron emission tomography," *Journal of the American Statistical Association*, vol. 80, no. 389, pp. 8–20, 1985.

[3] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.

[4] T. Le, R. Chartrand, and T. J. Asaki, "A variational approach to reconstructing images corrupted by Poisson noise," *Journal of Mathematical Imaging and Vision*, vol. 27, no. 3, pp. 257–263, 2007.

[5] M. Lysaker, A. Lundervold, and X.-C. Tai, "Noise removal using fourth-order partial differential equation

with applications to medical magnetic resonance images in space and time," *IEEE Transactions on Image Processing*, vol. 12, no. 12, pp. 1579–1590, 2003.

[6] S. Holla Kayyar and P. Jidesh, "Non-local total variation regularization approach for image restoration under a Poisson degradation," *Journal of Modern Optics*, vol. 65, no. 19, pp. 2231–2242, 2018.

[7] X.-d. Wang, X.-c. Feng, W.-w. Wang, and W.-j. Zhang, "Iterative reweighted total generalized variation based Poisson noise removal model," *Applied Mathematics and Computation*, vol. 223, pp. 264–277, 2013.

[8] M. R. Chowdhury, J. Zhang, J. Qin, and Y. Lou, "Poisson image denoising based on fractional-order total variation," *Inverse Problems & Imaging*, vol. 14, no. 1, 2020.

[9] P. G. Kumar and R. Ranjan Sahay, "Low rank poisson denoising (LRPD): A low rank approach using split bregman algorithm for poisson noise removal from images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, June 2019.

[10] J. Salmon, Z. Harmany, C.-A. Deledalle, and R. Willett, "Poisson noise reduction with non-local pca," *Journal of Mathematical Imaging and Vision*, vol. 48, no. 2, pp. 279–294, 2014.

[11] Z. Zha, X. Yuan, B. Wen, J. Zhou, J. Zhang, and C. Zhu, "From rank estimation to rank approximation: Rank residual constraint for image restoration," *IEEE Transactions on Image Processing*, vol. 29, pp. 3254–3269, 2019.

[12] M. Zhao, Y.-W. Wen, M. Ng, and H. Li, "A nonlocal low rank model for poisson noise removal," *Inverse Problems & Imaging*, vol. 15, no. 3, pp. 519, 2021.

[13] Z. Zha, B. Wen, X. Yuan, J. Zhou, and C. Zhu, "Simultaneous nonlocal low-rank and deep priors for poisson denoising," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Pro- cess.*, 2022, pp. 2320–2324.

[14] Z. Zha, X. Yuan, J. Zhou, C. Zhu, and B. Wen, "Image restoration via simultaneous nonlocal self-similarity priors," *IEEE Transactions on Image Processing*, vol. 29, pp. 8561–8576, 2020.

[15] W. Feng, P. Qiao, and Y. Chen, "Fast and accurate Poisson denoising with trainable nonlinear diffusion," *IEEE Transactions on Cybernetics*, vol. 48, no. 6, pp. 1708–1719, 2017.

[16] T. Remez, O. Litany, R. Giryes, and A. M. Bronstein, "Class-aware fully convolutional Gaussian and Poisson denoising," *IEEE Transactions on Image Processing*, vol. 27, no. 11, pp. 5707–5722, 2018.

[17] H. Ren, M. El-Khamy, and J. Lee, "DN-Resnet: Efficient deep residual network for image denoising," in *Proc. Asian Conf. Comput. Vis.* Springer, 2018, pp. 215–230.

[18] C. Zeng and C. Wu, "On the edge recovery property of noncovex nonsmooth regularization in image restoration," *SIAM Journal on Numerical Analysis*, vol. 56, no. 2, pp. 1168–1182, 2018.

[19] X. Chen, M. K. Ng, and C. Zhang, "Non-Lipschitz ℓ_p-regularization and box constrained model for image restoration," *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4709–4721, 2012.

[20] Y. Lou, T. Zeng, S. Osher, and J. Xin, "A weighted difference of anisotropic and isotropic total variation model for image processing," *SIAM Journal on Imaging Sciences*, vol. 8, no. 3, pp. 1798–1823, 2015.

[21] Y. Lou and M. Yan, "Fast l1–l2 minimization via a proximal operator," *Journal of Scientific Computing*, vol. 74, no. 2, pp. 767–785, 2018.

[22] L. Condat, "Discrete total variation: New definition and minimization," *SIAM Journal on Imaging Sciences*, vol. 10, no. 3, pp. 1258–1290, 2017.

[23] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, et al., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[24] S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, and L. Zhang, "Weighted nuclear norm minimization and its applications to low level vision," *International Journal of Computer Vision*, vol. 121, no. 2, pp. 183–208, 2017.

[25] Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," *SIAM Journal on Imaging Sciences*, vol. 1, no. 3, pp. 248–272, 2008.

[26] Y. Wang, W. Yin, and J. Zeng, "Global convergence of ADMM in nonconvex nonsmooth optimization," *Journal of Scientific Computing*, vol. 78, no. 1, pp. 29–63, 2019.

[27] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. Int. Conf. Comput. Vis.*, 2001, vol. 2, pp. 416–423 vol.2.

[28] C. Zheng, D. Shi, and W. Shi, "Adaptive unfolding total variation network for low-light image enhancement," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2021, pp. 4439–4448.