Low Rank Poisson Denoising (LRPD): A Low Rank Approach using split Bregman Algorithm for Poisson Noise Removal from Images

Prashanth Kumar G.[1] and Rajiv Ranjan Sahay[2]
[1,2]Computational Vision Laboratory,
[1]Advanced Technology Development Centre, [2]Department of Electrical Engineering,
[1,2]Indian Institute of Technology Kharagpur, India
[1]prashanth02.b@gmail.com, [2]rajivsahay@gmail.com

Abstract

Occurrence of Poisson noise in captured observations is inevitable in various real imaging applications ranging from medical imaging to night vision imaging. Restoration of fine details of an image is difficult when it is corrupted by Poisson noise. Recently, low rank approaches outperformed several state-of-the-art techniques for image denoising, deblurring, image completion, super-resolution, etc. The ability of low rank techniques to preserve fine details, even though the image is corrupted by severe noise, motivated us to develop an optimization framework wherein, we propose to use a low rank prior for Poisson noise removal. In the proposed low rank Poisson denoising (LRPD) algorithm, we resort to split Bregman technique to solve an appropriate objective function. We incorporate the forward-backward splitting scheme to minimize the first subproblem and the weighted nuclear norm minimization (WNNM) for the second subproblem of split Bregman algorithm to arrive at the final solution. We conduct several experiments on both simulated and real-world Poisson noisy data and show the superiority of the proposed method over other state-of-the-art Poisson denoising techniques.

1. Introduction

Poisson denoising is one of the active research topics in the field of computer vision. Poisson noise typically occurs in real-world applications such as night vision, biomedical imaging, microscopic imaging, astronomy imaging, etc. wherein images are obtained by counting the number of photons that hit the sensor. When an image is captured in low-intensity illumination or with short exposure times, the main source of noise can be well modeled as Poissonian [1], [2]. The strength of Poisson noise rely on the pixel intensity, hence is strongly signal dependent and not additive. Therefore, the traditional denoising algorithms [3], [4], [5], [6], [7], [8], [9], [10], [11], developed to tackle additive white Gaussian noise (AWGN) are not suitable.

In image restoration, there is a rich literature of Poisson denoising algorithms and they are broadly classified into two categories: 1) variance stabilization transformation (VST) based and 2) non VST methods. In VST method, the image is preprocessed using a nonlinear VST such as Anscombe [12], [13] or Haar-Fisz transform [14], [15], [16] to alleviate the signal dependency of Poisson noise. The transformed image noise statistics are approximately treated as additive white Gaussian with constant variance and the underlying clean image is estimated using any of the well-known Gaussian denoising algorithms available in literature. The underlying clean image is finally estimated by applying the inverse VST [17], [18] to the denoised version of the transformed image. The main drawback of VST approach is that the performance deteriorates (i.e., the SNR decreases) if the photon-counting rate is less due to low intensity.

Providing an alternative to the classical VST techniques [12], [19], the authors of [20] proposed a wavelet domain filter, which can be interpreted as a data-adaptive Wiener filter using wavelet basis. The seminal work by [21], introduced a Haar domain threshold for multi-scale Poisson image denoising [22] and a noise-free coefficient was estimated under the Bayesian setting. One of the key advantages of Bayesian methods are that they allow incorporation of prior knowledge into the estimation procedure. In particular, the framework of [1] and [23] involves a decomposition of the Poisson process likelihood function across scales, allowing computationally efficient intensity estimation by means of a scale-recursive scheme. Further improvements in this field include Bayesian Skellam (noise distribution of Haar wavelet coefficients) mean estimators [24] and unbiased estimate of risk [25], [26] aimed at the recovery of noise-free wavelet coefficients from Poisson noisy images. The authors in [27] proposed a hybrid minimum risk shrinkage operator for denoising the multi-scale Poisson image that effectively produced a denoised image with minimum attainable ℓ_2 error.

To alleviate the limitations of the VST approach, many authors have proposed non-VST Poisson denoising algorithms [28], [29], [30], [31], [32], [33], [34], [35], [36] which directly rely on the statistics of the Poisson noise. In [28], the authors have proposed a Poisson denoising algorithm based on the idea of non-local means (NLM) and used stochastic distances to measure the similarities between the non-local patches instead of Euclidean distance. The method proposed in [29] uses a dictionary learning strategy with a sparse coding algorithm by directly relying on the Poisson statistical model. Relying directly on Poisson noise properties, a denoising technique was proposed in [31] which combines sparse patch-based representations and the elements of dictionary learning. It employs both an adaptation of principal component analysis (PCA) [37] and sparsity-regularized convex optimization algorithms [32] for photon-limited images in non-local framework. The non-local PCA (NLPCA) and the non-local sparse PCA (NL-SPCA) are the two versions of the work proposed in [31]. Particularly, the NL-SPCA have yielded better performance in image recovery by integrating ℓ_1 regularization term to the objective function which is minimized. The authors in [34] developed an efficient Poisson denoising model which uses the newly-developed trainable nonlinear reaction diffusion (TNRD) model [38] which has outperformed all state-of-the-art Poisson denoising methods in terms of computational efficiency and recovery quality. Even though the performance of [34] is superior compared to existing techniques, it is limited to the natural image denoising. However, Poisson noise often arises also in real-world applications such as biomedical imaging, fluorescence microscopy, astronomy, etc. training the diffusion process for these specific class of images is tedious due to non-availability of large amount of annotated data and the difference in the pixel intensity values.

In recent years, patch-based techniques [3], [39] have shown significant improvement in performance for image restoration problems by exploiting the self-similarity property of the image. If we collect all similar patches together in an image and construct a matrix by stacking all those similar patches as the column vectors and it will be of low rank [40]. Low rank minimization methods [9], [41], [42], [43], [11] have achieved great success in low-level vision problems. The work proposed in [41] uses non-local low rank regularization approach to recover the image in a compressive sensing framework. For a low rank matrix, the nuclear norm is defined as the sum of its singular values and the minimization of nuclear norm (NNM) is presented in [42] for video denoising problem. Since all singular values are equally penalized in NNM, the major edge and texture information will be degraded due to shrinkage of large singular values. The limitation of NNM was overcome by penalizing each singular value by different weights depending on their magnitudes, leading to a technique called weighted nuclear norm minimization (WNNM) [11], [43] which has outperformed all state-of-the-art denoising techniques. The WNNM proposed in [11], [43] penalize the smaller singular values by larger weights and the significant singular values by smaller weights and hence preserve large-scale sharp edges and small-scale fine image details more effectively. The authors in [44] proposed an algorithm which combines low rank and TV priors to recover the deblurred image in the presence of salt-and-pepper noise and additive Gaussian noise. In [45], the authors have used weighted nuclear norm (WNN) as a prior to estimate the structure of a 3D object by incorporating new shrinkage operator to penalize the singular values. The ability of low rank (WNN) prior to preserve large-scale sharp edges and small-scale fine details more effectively, motivated us to develop the proposed algorithm for Poisson noise removal.

The key challenge in Poisson intensity estimation problem is that the mean and the variance of the observed count are same and also the Poisson noise is completely signal dependent. Hence, building of an optimization framework is not straight forward. In the proposed work, we formulate an appropriate objective function with low rank prior constraint. We resort to split Bregman technique [46], [47] to solve the proposed objective function in terms of easily solvable subproblems. Since direct gradient descent method can't be applied due to numerical instability and hence, we use proximal gradient descent method [48], [49] by incorporating a forward-backward splitting scheme [50] to solve the first subproblem of the split Bregman algorithm. Solution of the second subproblem is obtained using weighted nuclear norm minimization (WNNM) [11], [43].

Organization of rest of the paper is as follows. The detailed description of the proposed LRPD algorithm is presented in section 2. Experimental results of the proposed approach are compiled in section 3. Finally, concluding remarks of the proposed method are provided in section 4.

2. The proposed LRPD algorithm

In this section we describe the steps involved to develop the mathematical model of the proposed Poisson denoising algorithm in variational approach with low rank regularization constraints using split Bregman technique [46].

2.1. Mathematical modeling of Poisson denoising

Let $\mathbf{x} \in \mathbb{R}^{N^2\times 1}$ be the original clean image and $\mathbf{y} \in \mathbb{Z}_+^{N^2\times 1}$ denotes a noise image degraded by Poisson noise (both $\mathbf{x}$ and $\mathbf{y}$ are lexicographically arranged column vectors). Our aim is to recover the clean image $\mathbf{x}$ from Poisson noisy observation $\mathbf{y}$. Each observed pixel value y_i in $\mathbf{y}$ given pixel x_i in $\mathbf{x}$ is assumed to be Poisson distributed independent random variable with mean and variance equal to x_i. The likelihood of Poisson distribution is given by

$$P(\mathbf{y}|\mathbf{x}) = \begin{cases} \prod_{i=1}^{N^2} \frac{x_i^{y_i} \exp(-x_i)}{y_i!}, & \text{if } x_i > 0 \\ \delta_0(y_i), & \text{if } x_i = 0 \end{cases} \tag{1}$$

where x_i and y_i are the i^{th} components in $\mathbf{x}$ and $\mathbf{y}$ respectively and δ_0 is the Kronecker delta function. We know that the occurrence of Poisson noise is proportional to the pixel intensity x_i and is completely signal dependent. Therefore, the noise level in image $\mathbf{x}$ is usually defined as the peak value in $\mathbf{x}$. This is evident, since the effect of the Poisson noise decreases (i.e., SNR increases) as the intensity value x_i increases and vice versa. The negative Poisson log-likelihood function is given by

$$\begin{aligned} F(\mathbf{x}) = -\log\big(P(\mathbf{y}|\mathbf{x})\big) &= -\log\left(\prod_{i=1}^{N^2} \frac{x_i^{y_i} \exp(-x_i)}{y_i!}\right) \\ &= \sum_{i=1}^{N^2} \Big\{x_i - y_i \log x_i + \log(y_i!)\Big\} \end{aligned} \tag{2}$$

Since $\log(y_i!)$ is a constant and therefore, it can be neglected. Minimizing the negative Poisson log-likelihood $F(\mathbf{x}) = -\log\big(P(\mathbf{y}|\mathbf{x})\big)$ leads to the following data-fidelity term

$$F(\mathbf{x}) = \sum_{i=1}^{N^2} \{x_i - y_i \log x_i\} = \langle \mathbf{x} - \mathbf{y} \log \mathbf{x}, \mathbf{1} \rangle \tag{3}$$

and is also known as the so called Csiszár I-divergence model [51] and this data-fidelity term has been widely used in previous Poisson image denoising algorithms [35], [36], [28], [34].

However, the straightforward direct gradient descent $\big(\frac{\partial F(\mathbf{x})}{\partial \mathbf{x}} = 1 - \frac{\mathbf{y}}{\mathbf{x}}\big)$ is not applicable [34] in practice for this problem, because of two reasons. Firstly, there exists numerical instability when $\mathbf{x}$ leads very close to zero and secondly, the output image $\mathbf{x}$ after one iteration may become negative and the non-positive values of $\mathbf{x}$ will breach the constraint of the data-fidelity term $F(\mathbf{x}) = \langle \mathbf{x} - \mathbf{y} \log \mathbf{x}, 1 \rangle$.

As a consequence, we resort to the proximal gradient descent technique [48], [49] which avoids the gradient formula $\big(\frac{\partial F(\mathbf{x})}{\partial \mathbf{x}} = 1 - \frac{\mathbf{y}}{\mathbf{x}}\big)$ and can solve numerical instability problem and prevents negative values in the output.

2.2. The proximal gradient algorithm

The proximal gradient descent method [48], [49] is applicable to optimization problems which are composed of a smooth, differentiable function $G(\mathbf{x})$ and possibly a non-smooth function $F(\mathbf{x})$:

$$\Phi(\mathbf{x}) = \arg\min_{\mathbf{x}} F(\mathbf{x}) + G(\mathbf{x}) \tag{4}$$

where, $\Phi(\mathbf{x})$ is the objective function to be minimized. The solution of Eq. (4) can be obtained using forward-backward splitting scheme [50]. The basic update rule to solve Eq. (4) is given by

$$\mathbf{x}^{t+1} = \big(I - \tau\partial F\big)^{-1}\big(\mathbf{x}^t - \tau\nabla G(\mathbf{x}^t)\big) \tag{5}$$

where I denotes identity matrix, ∂F is sub-differential of non-smooth function $F(\mathbf{x})$, $\big(\mathbf{x}^t - \tau\nabla G(\mathbf{x}^t)\big)$ is the forward gradient descent step, τ represents the step size of the forward gradient step and the term $\big(I - \tau\partial F\big)^{-1}$ denotes the proximal mapping and is also called as backward step [49]. Let us assume, the forward gradient step as

$$\tilde{\mathbf{x}} = \mathbf{x}^{\mathbf{t}} - \tau\nabla G(\mathbf{x}^{\mathbf{t}}) \tag{6}$$

The proximal mapping $\big(I - \tau\partial F\big)^{-1}(\tilde{\mathbf{x}})$ with respect to F is given by the following optimization problem [34], [49]

$$\big(I - \tau\partial F\big)^{-1}(\tilde{\mathbf{x}}) = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{x} - \tilde{\mathbf{x}}\|_2^2 + \tau F(\mathbf{x}) \tag{7}$$

substituting $F(\mathbf{x})$ from Eq. (3) in Eq. (7), we get

$$\big(I - \tau\partial F\big)^{-1}(\tilde{\mathbf{x}}) = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{x} - \tilde{\mathbf{x}}\|_2^2 + \tau\big(\mathbf{x} - \mathbf{y}\log\mathbf{x}\big) \tag{8}$$

which leads to a point-wise solution for the i^{th} element and is given by

$$\big(I - \tau\partial F\big)^{-1}(\tilde{x}_i) = \arg\min_{x_i} \left\{\frac{1}{2}\big(x_i - \tilde{x}_i\big)^2 + \tau\big(x_i - y_i\log x_i\big)\right\} \tag{9}$$

The solution of Eq. (9) can be obtained by setting the gradient with respect to x_i to zero,

$$\begin{gathered} \frac{\partial}{\partial x_i}\left[\frac{1}{2}\big(x_i - \tilde{x}_i\big)^2 + \tau\big(x_i - y_i\log x_i\big)\right] = 0 \\ x_i - \tilde{x}_i + \tau - \frac{\tau y_i}{x_i} = 0 \end{gathered} \tag{10}$$

Hence, the following quadratic equation is obtained

$$x_i^{\,2} + \big(\tau - \tilde{x}_i\big)x_i - \tau y_i = 0 \tag{11}$$

Let $a = 1$, $b = (\tau - \tilde{x}_i)$ and $c = -\tau y_i$, the solution of Eq. (11) has two real roots and we choose the positive one due to the constraint of $x_i > 0$. Therefore, the point-wise solution of Eq. (9) is

$$\hat{\mathbf{x}}_i = \big(I - \tau\partial F\big)^{-1}(\tilde{\mathbf{x}}_i) = \frac{\tilde{\mathbf{x}}_i - \tau + \sqrt{(\tau - \tilde{\mathbf{x}}_i)^2 + 4\tau\mathbf{y}_i}}{2} \tag{12}$$

where $\hat{\mathbf{x}}_i$ is the real positive root of Eq. (11). Note that, $\hat{\mathbf{x}}_i$ is always positive, if $\tau > 0$ and $\mathbf{y}_i > 0$, as

$$\sqrt{(\tau - \tilde{\mathbf{x}}_i)^2 + 4\tau\mathbf{y}_i} > |\tau - \tilde{\mathbf{x}}_i| \geq 0 \tag{13}$$

Hence, this update rule can assure $\hat{\mathbf{x}}_i > 0$ in each iteration if we ensure $\tau > 0$ in each update step. The proof of Eq. (13) and assurance of $\hat{\mathbf{x}}_i > 0$ is given in Appendix 1.

Finally, from Eqs. (5), (6) and Eq. (12), we can summarize the mathematical formulation to remove Poisson noise using proximal gradient descent method with an objective function composed of a smooth function and a convex (possibly non-smooth) term as follows

$$\mathbf{x}^{t+1} = \frac{\tilde{\mathbf{x}}^{t+1} - \tau + \sqrt{(\tau - \tilde{\mathbf{x}}^{t+1})^2 + 4\tau\mathbf{y}}}{2} \tag{14}$$

and

$$\tilde{\mathbf{x}}^{t+1} = \mathbf{x}^t - \tau\nabla G(\mathbf{x}^t) \tag{15}$$

where $\mathbf{x}$ is estimate of the latent image, $\mathbf{y}$ is the observed image corrupted by Poisson noise, t is iteration number and $G(\mathbf{x})$ can be treated as smooth function or a regularization constraint. In Eq. (15), parameter τ plays a role of step-size in the proximal gradient algorithm, which should be carefully selected in order to obtain faithful results.

2.3. The proposed LRPD algorithm using split Bregman framework

The proposed objective function for Poisson denoising using low rank prior i.e., weighted nuclear norm (WNN) prior as a regularizer is given by

$$\Phi(\mathbf{x}) = \arg\min_{\mathbf{x}} F(\mathbf{x}) + F_2(\mathbf{x}) \tag{16}$$

where $F(\mathbf{x}) = \mu(\mathbf{x} - \mathbf{y}\log\mathbf{x})$ be the data fidelity term derived from the negative log-likelihood of the Poisson distribution and $F_2(\mathbf{x})$ represents WNN prior, then Eq. (16) can be rewritten as

$$\Phi(\mathbf{x}) = \arg\min_{\mathbf{x}} \left\{ \mu(\mathbf{x} - \mathbf{y}\log\mathbf{x}) + \sum_{p\in\aleph} \|\mathbf{R}_p\mathbf{x}\|_{\mathbf{w},*} \right\} \tag{17}$$

where $\|\cdot\|_{\mathbf{w},*}$ denotes weighted nuclear norm, μ is a constant, $\aleph$ denotes the set of indices of all reference patches of $\mathbf{x}$. At first, the operator $\mathbf{R}_p$ collects all the patches similar to the reference patch located at p^{th} position in $\mathbf{x}$ and stacks all those patches as column vectors to construct a matrix $\mathbf{R}_p\mathbf{x}$ which is assumed to be of low rank.

Solving the model in Eq. (17) is not straightforward due to the date fidelity term and hence, we introduce an auxiliary variable $\mathbf{z}$ to the prior term in Eq. (17) and the corresponding constrained objective function is thus given by

$$\Phi(\mathbf{x}) = \arg\min_{\mathbf{x}} \left\{ \mu(\mathbf{x} - \mathbf{y}\log\mathbf{x}) + \sum_{p\in\aleph} \|\mathbf{R}_p\mathbf{z}\|_{\mathbf{w},*} \right\} \quad \text{s.t.} \quad \mathbf{z} = \mathbf{x} \tag{18}$$

To minimize Eq. (18), we use split Bregman algorithm [46] to split it into an easily solvable subproblems. To convert Eq. (18) into an unconstrained one, a quadratic penalty term is added and the corresponding unconstrained objective function is thus given by

$$(\mathbf{x}^{l+1}, \mathbf{z}^{l+1}) = \arg\min_{\mathbf{x},\mathbf{z}} \left\{ \mu(\mathbf{x}^l - \mathbf{y}\log\mathbf{x}^l) + \sum_{p\in\aleph} \|\mathbf{R}_p\mathbf{z}^l\|_{\mathbf{w},*} + \frac{\lambda}{2}\|\mathbf{z}^l - \mathbf{x}^l - \mathbf{b}^l\|_2^2 \right\} \tag{19}$$

and

$$\mathbf{b}^{l+1} = \mathbf{b}^l + (\mathbf{x}^{l+1} - \mathbf{z}^{l+1}) \tag{20}$$

where l represents the number of Bregman iterations and λ is a constant used to control the quality of the output. Splitting Eq. (19) into two separate subproblems, we get

$$\mathbf{x}^{l+1} = \arg\min_{\mathbf{x}} \left\{ \mu(\mathbf{x}^l - \mathbf{y}\log\mathbf{x}^l) + \frac{\lambda}{2}\|\mathbf{z}^l - \mathbf{x}^l - \mathbf{b}^l\|_2^2 \right\} \tag{21}$$

$$\mathbf{z}^{l+1} = \arg\min_{\mathbf{z}} \sum_{p\in\aleph} \|\mathbf{R}_p\mathbf{z}^l\|_{\mathbf{w},*} + \frac{\lambda}{2}\|\mathbf{z}^1 - \mathbf{x}^{l+1} - \mathbf{b}^l\|_2^2 \tag{22}$$

and

$$\mathbf{b}^{l+1} = \mathbf{b}^l + (\mathbf{x}^{l+1} - \mathbf{z}^{l+1}) \tag{23}$$

Eq. (21) to Eq. (23) represents the variational approach of the proposed LRPD algorithm developed to remove Poisson noise from an image using split Bregman framework.

In Eq. (21), due to the data-fidelity term, direct gradient descent method can not be used to solve it. Therefore, the solution can be obtained by exploiting proximal gradient method as presented in section 2.2. Let the terms $F(\mathbf{x}) = \mu(\mathbf{x}^{\mathbf{l}} - \mathbf{y}\log\mathbf{x}^l)$ and $G(\mathbf{x}) = \frac{\lambda}{2}\|\mathbf{z}^l - \mathbf{x}^l - \mathbf{b}^l\|_2^2$, comparing to Eqs. (4), (5), (6) and Eq. (14), the solution of Eq. (21) can be obtained using proximal gradient descent method and is given by

$$\mathbf{x}^{l+1} = \frac{\tilde{\mathbf{x}}^{l+1} - \tau\mu + \sqrt{(\tau\mu - \tilde{\mathbf{x}}^{l+1})^2 + 4\tau\mu\mathbf{y}}}{2} \tag{24}$$

where $\tilde{\mathbf{x}} = \mathbf{x} - \tau\nabla G(\mathbf{x})$ is forward gradient step. From Eq. (15) we have, $\nabla G(\mathbf{x}) = \frac{\partial}{\partial\mathbf{x}}\big(\frac{\lambda}{2}\|\mathbf{z}^l - \mathbf{x}^l - \mathbf{b}^l\|_2^2\big)$ and $\tilde{\mathbf{x}}^{l+1}$ will be updated as

$$\begin{aligned} \tilde{\mathbf{x}}^{l+1} &= \mathbf{x}^l - \tau\frac{\partial}{\partial\mathbf{x}}\left(\frac{\lambda}{2}\|\mathbf{z}^l - \mathbf{x}^l - \mathbf{b}^l\|_2^2\right) \\ &= \mathbf{x}^l - \tau\lambda\big(\mathbf{x}^l - (\mathbf{z}^l - \mathbf{b}^l)\big) \end{aligned} \tag{25}$$

Next, the optimal solution of Eq. (22) is obtained using weighted nuclear norm minimization(WNNM) [11], [43].

The input image $\mathbf{z}$ is divided into several patches, and for each p^{th} patch, the operator $\mathbf{R}_p$ collects all the patches similar to the p^{th} reference patch and stacks all collected patches together as column vectors to construct a matrix $\mathbf{R}_p\mathbf{z}$ corresponding to the p^{th} reference patch which is assumed to be low rank. The clean patch corresponding to the p^{th} reference patch is estimated from the constructed low rank matrix $\mathbf{R}_p\mathbf{z}$ using WNNM [11] and finally the whole image $\mathbf{z}$ is reconstructed by aggregating all the estimated clean patches. Therefore, to estimate the optimal solution of $\mathbf{z}$, we rewrite Eq. (22) in WNNM framework and the corresponding modified minimization problem of Eq. (22) is given by

$$\mathbf{R}_p\mathbf{z}^{l+1} = \arg\min_{\mathbf{R}_p\mathbf{z}} \Big\{ \|\mathbf{R}_p\mathbf{z}^l\|_{\mathbf{w},*} + \frac{\lambda}{2}\|\mathbf{R}_p\mathbf{z}^1 - (\mathbf{R}_p\mathbf{x}^{l+1} + \mathbf{R}_p\mathbf{b}^l)\|_F^2 \Big\} \tag{26}$$

where, $\mathbf{R}_p\mathbf{x}^{l+1}$ and $\mathbf{R}_p\mathbf{b}^l$ are assumed to be low rank matrices corresponding to the p^{th} reference patch in $\mathbf{x}^{l+1}$ and $\mathbf{b}^l$, respectively.

According to the Theorem 1 in [11], given the singular value decomposition (SVD) of matrix $(\mathbf{R}_p\mathbf{x}^{l+1} + \mathbf{R}_p\mathbf{b}^l)$,

$$\big(\mathbf{R}_p\mathbf{x}^{l+1} + \mathbf{R}_p\mathbf{b}^l\big) = \mathbf{U}_p\mathbf{\Sigma}_p\mathbf{V}_p^T \tag{27}$$

where $\mathbf{\Sigma}_p = \begin{pmatrix} \mathrm{diag}(\sigma_{xb,1}, ., \sigma_{xb,n}) \\ 0 \end{pmatrix}$ and $(\sigma_{xb,1}, ., \sigma_{xb,n})$ are singular values of the low rank matrix $\big(\mathbf{R}_p\mathbf{x}^{l+1} + \mathbf{R}_p\mathbf{b}^l\big)$ then the global optimum of the above problem is

$$\mathbf{R}_p\mathbf{z}^{l+1} = \mathbf{U}_p \begin{pmatrix} diag(d_1, d_2, ..., d_n) \\ 0 \end{pmatrix} \mathbf{V}_p^T \tag{28}$$

where $(d_1, d_2, ..., d_n)$ is the solution of the following convex optimization problem:

$$\min_{d_1, d_2, \ldots, d_n} \sum_{i=1}^{n} (\sigma_{xb,i} - d_i)^2 + \frac{2w_i}{\lambda} d_i \quad \text{s.t.} \quad d_1 \geq d_2 \geq .... \geq d_n \geq 0. \tag{29}$$

The closed-form solution of Eq. (29) according to the Remark 1 in [11] is given by

$$d_i = \begin{cases} 0, & \text{if } c_2 < 0 \\ \frac{c_1 + \sqrt{c_2}}{2}, & \text{if } c_2 \geq 0 \end{cases} \tag{30}$$

where $c_1 = \sigma_{xb,i} - \epsilon$, $c_2 = (\sigma_{xb,i} + \epsilon)^2 - \frac{8C}{\lambda}$, for $(i = 1, 2, .., n)$, $\epsilon > 0$ is a very small value, C is a constant set as $\sqrt{2N_p}$ and N_p represents number of similar patches considered to construct the low rank matrix.

After finding the solutions $\mathbf{x}^{l+1}$ and $\mathbf{z}^{l+1}$, we update the variable $\mathbf{b}^{l+1}$ according to Eq. (23). The above steps are together repeated for several Bregman iterations to obtain the final estimate of the latent image from the Poisson noisy image.

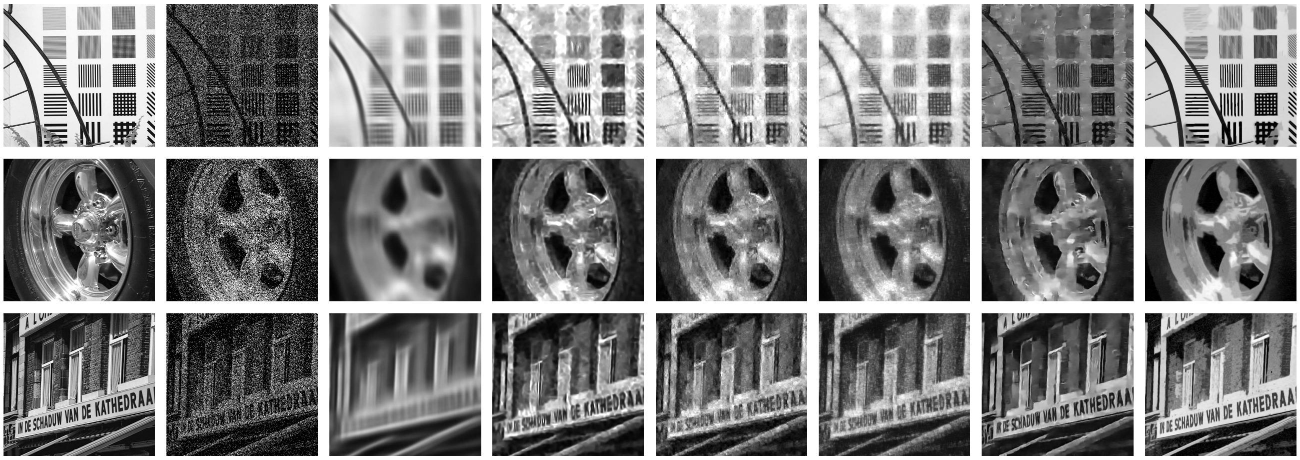
3. Experimental results

In this Section, we demonstrate the superiority of the proposed LRPD algorithm over other state-of-the-art Poisson denoising algorithms. The performance of our method is compared with several existing Poisson denoising techniques, namely, non-local sparse PCA (NLSPCA) [31], combined generalized Anscombe transform and BM3D (GAT+BM3D) [18], Poisson unbiased risk estimate with linear expansion threshold (PURE-LET) [25], hybrid Skellam minimum risk shrinkage operator (HMRSO) [27] and trained reaction diffusion models for Poisson denoising (TRDPD) [34]. We evaluate the proposed LRPD method both quantitatively and qualitatively on simulated and the real-world Poisson noisy images. We conduct synthetic experiments on the few images taken from the McGill calibrated color image database [52] and a few images from the Berkeley segmentation dataset (BSD) [53]. In real-world experiments, we test our algorithm on the real sensor data provided by the authors of [27] captured using FUJIFILM X-PRO 1 camera and the image captured using the laboratory microscope Nikon ECLIPSE LV150N.

In both simulated and real-world experiments, the same Poisson noisy image is assumed as initial estimate of the denoised output. The parameters such as step size τ, constants μ and λ are different for all experiments and we tune them empirically to achieve better performance. We run fixed number of iterations to estimate $\mathbf{x}^{l+1}$ in the first subproblem, $\mathbf{v}^{l+1}$ is estimated using weighted nuclear norm minimization (WNNM) proposed in [11] followed by updating the variable $\mathbf{b}^{l+1}$ according to Eq. (23). All these steps are together repeated for several Bregman iterations to obtain final estimate of the denoised image. We choose values of the parameters for WNNM, such as reference patch of size 7×7, size of the search window to find the similar patches is 30×30, number of non-local similar patches 80. The best results obtained after optimally tuning the values of parameters τ, μ and λ are presented. To evaluate performance of the proposed method with state-of-the-art techniques quantitatively, we use structural similarity index metric (SSIM) [54] and peak signal-to-noise ratio (PSNR).

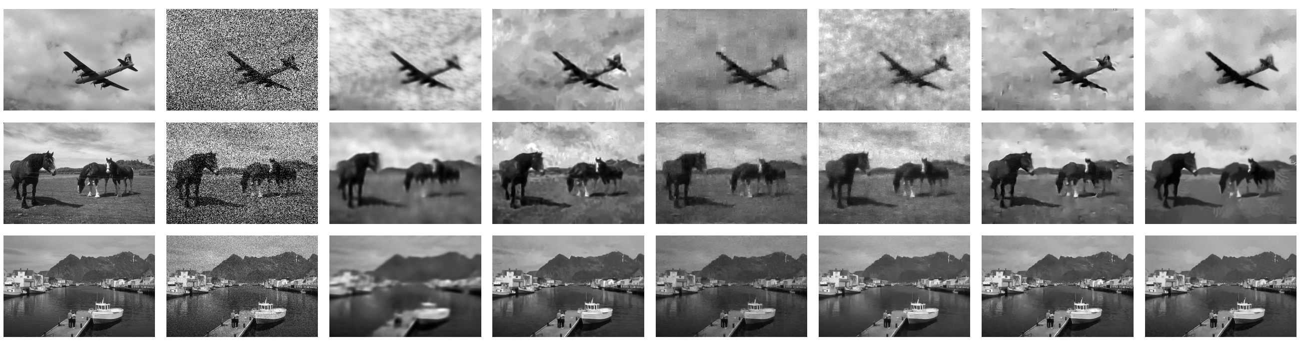
3.1. Synthetic experiments

To conduct synthetic experiments, we consider few images from the McGill calibrated color image database [52] and few images from the Berkeley segmentation dataset [53]. Initially, Poisson noisy images are generated synthetically for different peak intensity values followed by denoising them using the proposed LRPD algorithm. In the first simulated experiments, we consider zoomed portion of the three images from the McGill calibrated color image database [52] with average peak pixel intensity of 0.5 and the corresponding results are shown in Fig. 1. The last column of Fig. 1 shows the superiority of the proposed LRPD



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 1. Results of zoomed portion of the images taken from McGill calibrated color image database [52] with average peak pixel intensity of 0.5. (a) Clean images. (b) Noisy images. (c)-(h) Results obtained using state-of-the-art methods in [31], [18], [55], [27], [34] and the proposed LRPD algorithm, respectively.



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 2. Results of the images taken from the BSD dataset [53]. First, second and third rows show the results corresponding to the peak pixel values = 1, 2 and 40, respectively. (a) Clean images. (b) Noisy images. (c)-(h) Results obtained using state-of-the-art methods in [31], [18], [55], [27], [34] and the proposed LRPD algorithm, respectively.

algorithm. The sharp edges of the lines and small boxes are efficiently estimated for the first image, fine structure of the wheel is recovered for the second image and in the third image, the brick structure on the wall, image details and text on the board are more effectively reconstructed (zoom and see) using the proposed method as compared to other state-of-the-art Poisson denoising techniques.

In the next synthetic experiments, we consider few images taken from the BSD dataset [53] and conduct experiments for different peak pixel values = 1, 2, 4, 20, and 40. The results obtained for peak pixel values=1, 2 and 40 are shown in Fig. 2. First and second columns of Fig. 2 show the original and the corresponding Poisson noisy images, the results obtained using state-of-the-art Poisson denoising algorithms and our method are shown in Fig. 2 (c)-(h), respectively. The superiority of the proposed LRPD algorithm is visually discern from the last column of Fig. 2. The quantitative evaluation in terms of PSNR (in dB) and SSIM for the images taken from the McGill database [52] and BSD dataset [53] are summarized in Table 1 and 2, respectively.

In all our synthetic experiments, the results of NLSPCA [31] looks smoothed estimate under both high and low degree of Poisson noise level. Even though the GAT+BM3D [18] method preserves image contrast and looks good at first glance, careful observation reveals the presence of blocking artifacts and the wrong texture in the denoised image. Under low peak intensity, the blocking artifact and wrong textures are more dominant which leads to degraded denoised image. The PURE-LET [55] and the HMRSO [27] methods have yielded better noise suppression under both low Poisson noise level (i.e., larger peak value) and high noise level (i.e., smaller peak values). However, artifact were introduced in the denoised image and also the fine image details were completely degraded. Although, the TRDPD [34] method outperformed in some synthetic experiments, it failed to recover the fine details and the denoised images suffer from smoothing artifacts. The main drawback of the TRDPD [34] method is that it needs separate training of diffusion model for each different peak value. In contrast with all the existing techniques, our approach outperforms state-of-the-art Poisson denoising methods, especially with respect to recovery of sharp edges and fine image details even though the images are corrupted by severe Poisson noise.

	NLSPCA [31]	**GAT+BM3D** [18]	**PURE-LET** [55]	**HMRSO** [27]	**TRDPD** [34]	**Ours**
Image 1	11.93 / 0.418	13.13 / 0.423	13.35 / 0.412	13.53 / 0.452	11.67 / 0.383	**15.49 / 0.692**
Image 2	18.71 / 0.465	20.83 / 0.552	20.73 / 0.529	20.63 / 0.519	20.55 / 0.548	**20.87 / 0.553**
Image 3	14.28 / 0.368	15.81 / 0.451	15.071 / 0.441	15.88 / 0.461	17.319 / **0.576**	**17.39** / 0.572

Table 1. Quantitative evaluation of state-of-the-art methods and our method for the images taken from the McGill database [52].

	NLSPCA [31]	**GAT+BM3D** [18]	**PURE-LET** [55]	**HMRSO** [27]	**TRDPD** [34]	**Ours**
Peak = 1	24.76 / 0.858	24.01 / 0.881	24.05 / 0.659	24.17 / 0.582	26.13 / 0.900	**26.52 / 0.914**
Peak = 2	22.73 / 0.517	23.10 / 0.502	22.77 / 0.464	23.30 / 0.420	23.74 / 0.558	**23.88 / 0.562**
Peak = 4	21.09 / 0.634	23.29 / 0.655	22.31 / 0.567	23.26 / 0.501	**23.83 / 0.714**	23.74 / 0.683
Peak = 20	21.93 / 0.677	26.76 / 0.847	25.13 / 0.833	26.62 / 0.778	27.01 / 0.856	**27.15 / 0.862**
Peak = 40	22.23 / 0.598	28.67 / 0.799	27.40 / 0.704	28.87 / 0.760	29.06 / 0.809	**29.12 / 0.813**

Table 2. Quantitative evaluation of state-of-the-art methods and our method for the images taken from the BSD dataset [53].

3.2. Real-world data experiments

In real-world data experiments, we use the real sensor Poisson noisy image provided by the authors of [27]. The image of the “aquarium” was captured by the authors of [27] using FUJIFILM X-PRO 1 camera with an average Poisson count value (peak pixel value) of 18.0411 and is as shown in Fig. 3.

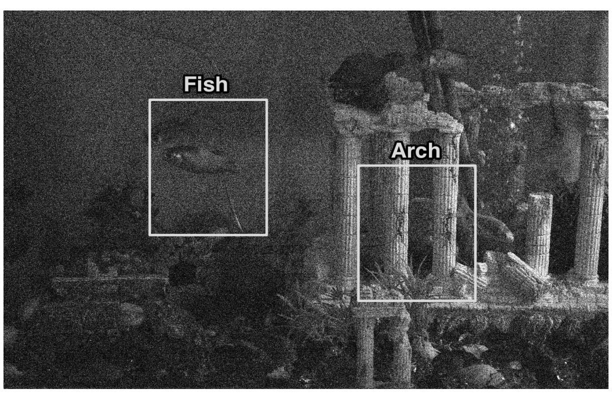


Figure 3. Real sensor image data captured by the authors of [27] using FUJIFILM X-PRO 1 camera. Two different regions are taken for testing and are marked as ‘Fish’ and ‘Arch’, respectively.

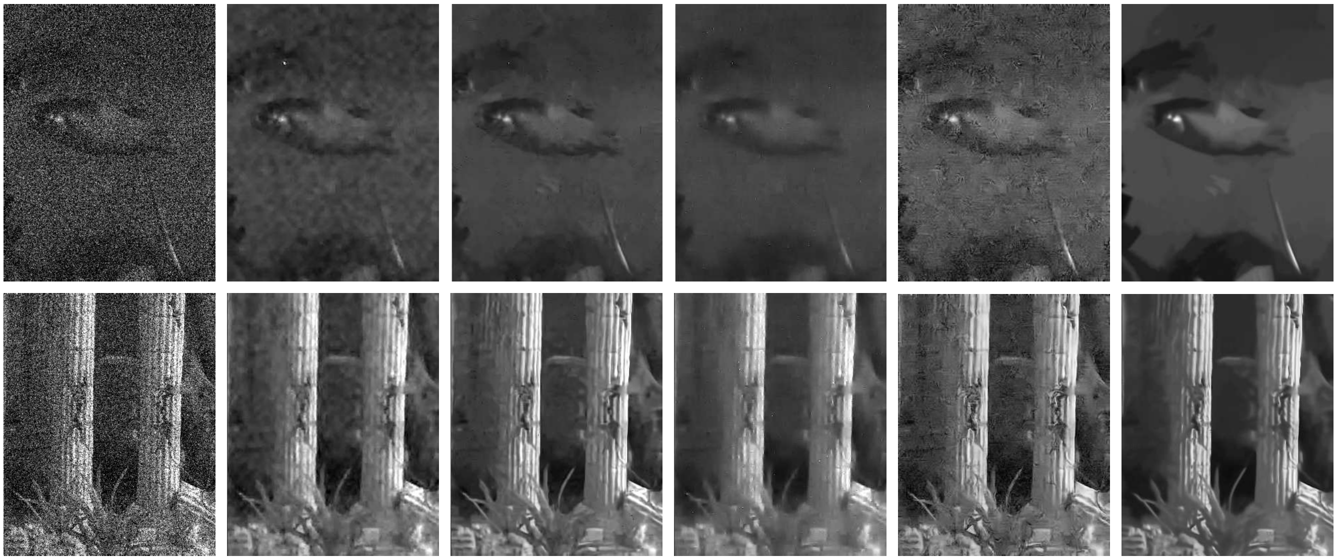
To check the effectiveness of the proposed LRPD algorithm, we consider two different regions from the real sensor noisy image marked as ‘Fish’ and ‘Arch’ regions as shown in Fig. 3. The zoomed portions of the noisy image and the corresponding results obtained using the state-of-the-art methods and our method are shown in Fig. 4. First and second rows of Fig. 4 show the results corresponding to ‘Fish’ and ‘Arch’ regions, respectively.

In the next real-world experiment, we consider the Poisson noisy image of a text-strip captured using NIKON ECLIPSE LV150N microscope using $2.5\times$ objective under low illumination condition with an average Poisson count value (peak pixel value) of 15.237. Fig. 5(a) show the captured microscopic noisy image, Figs (b)-(f) shows the results obtained using state-of-the-art methods and the proposed LRPD method. respectively.

In Fig. 4 and Fig. 5, at first glance, GAT+BM3D results appears to be visually superior, but the closer examination reveals that it introduced false textures (we can clearly observe by zooming) into the background which is strictly not acceptable in image recovery. Also the edge details, objects corners are unnaturally smoothed to give it a waxy appearance as observed by the authors of [27]. Even though the PURE-LET method recovered the edge details, it also introduced false texture and artifacts in the estimated results. The HMRSO method produced better noise suppression but failed to recover the sharp edge details. Even though the TRDPD method outperformed in few synthetic experiments, it failed to recover the image details in the real sensor data and also introduced smoothed artifacts in the estimated output. Since in real sensor data, the peak pixel values are arbitrary and therefore, the TRDPD method fails to denoise the real sensor image efficiently and it needs training of separate diffusion model for each peak value which is practically a tedious process. In comparison with all the existing Poisson denoising techniques, the proposed LRPD method outperforms in real-world experiments also in terms of preserving the sharp edges and the fine image details more efficiently.

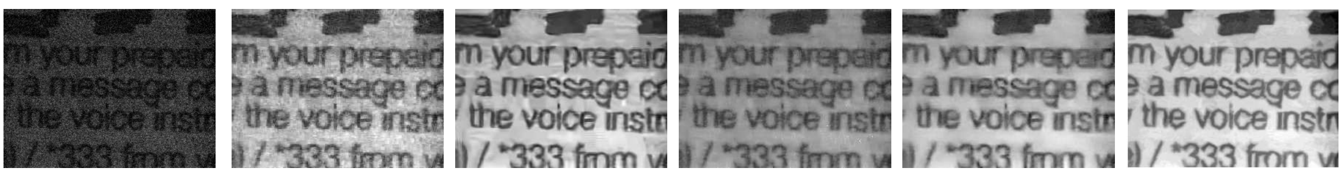
4. Conclusions

In the proposed algorithm, we developed a variational approach to remove Poisson noise from the images using low rank prior. We formulated the appropriate objective function to overcome the drawback of variance stabilization transform (VST) in the context of Poisson denoising by exploiting proximal gradient descent method. To optimize the proposed objective function split Bregman technique is used. The ability of WNN prior is exploited in the proposed work to recover sharp discontinuities and fine image details more effectively even in the presence of severe Poisson noise. Since the LRPD algorithm is developed in an optimization framework, it is more advantageous as it can also be used to denoise Poisson noisy images which occurs in various real applications such as astronomical imaging, biomedical and microscopic imaging, night vision, etc. Both qualitative and quantitative evaluations on various synthetic and real-world images concludes that the proposed algorithm outperforms several existing state-of-the-art Poisson denoising techniques.



(a) Noisy image (b) PURE-LET (c) GAT+BM3D (d) HMRSO (e) TRDPD (f) Ours

Figure 4. Results of real-world Poisson noisy image captured by the authors of [27] using a FUJIFILM X-PRO 1 camera. First and second rows show the results corresponding to 'Fish' and 'Arch' regions, respectively. (a) noisy images. (b)-(f) Results obtained using state-of-the-art methods in [55], [18], [27], [34] and the proposed LRPD algorithm, respectively.



(a) Real noisy image (b) PURE-LET (c) GAT+BM3D (d) HMRSO (e) TRDPD (f) Ours

Figure 5. Results of real-world Poisson noisy image captured using NIKON LV150N microscope. (a) noisy image. (b)-(f) Results obtained using the methods in [55], [18], [27], [34] and the proposed LRPD algorithm, respectively.

Appendix - 1

The proof of condition $\hat{\mathbf{x}} > 0$, and Eq. 13 is given here in Appendix 1. Let the point-wise solution is given by

$$\hat{\mathbf{x}}_i = \frac{\tilde{\mathbf{x}}_i - \tau + \sqrt{(\tau - \tilde{\mathbf{x}}_i)^2 + 4\tau\mathbf{y}_i}}{2}$$

If $\tau > 0$ and $\mathbf{y}_i > 0$, the term $(\tau - \tilde{\mathbf{x}}_i)^2 + 4\tau\mathbf{y}_i$ is positive and hence,

$$(\tau - \tilde{\mathbf{x}}_i)^2 + 4\tau\mathbf{y}_i > (\tau - \tilde{\mathbf{x}}_i)^2 \geq 0$$

take square root on both side

$$\sqrt{(\tau - \tilde{\mathbf{x}}_i)^2 + 4\tau\mathbf{y}_i} > |\tau - \tilde{\mathbf{x}}_i| \geq 0$$

since, $\sqrt{(\tau - \tilde{\mathbf{x}}_i)^2 + 4\tau\mathbf{y}_i} > |\tau - \tilde{\mathbf{x}}_i|$, we have

$$\sqrt{(\tau - \tilde{\mathbf{x}}_i)^2 + 4\tau\mathbf{y}_i} - |\tau - \tilde{\mathbf{x}}_i| > 0$$

therefore,

$$(\tau - \tilde{\mathbf{x}}_i) + \sqrt{(\tau - \tilde{\mathbf{x}}_i)^2 + 4\tau\mathbf{y}_i} > 0$$

$$\text{and, } -(\tau - \tilde{\mathbf{x}}_i) + \sqrt{(\tau - \tilde{\mathbf{x}}_i)^2 + 4\tau\mathbf{y}_i} > 0$$

this is always true even though the value of τ is very large as compared to $\tilde{\mathbf{x}}_i$. Now for any value of $\tau > 0$, even if $(\tau > \tilde{\mathbf{x}}_i)$ or $(\tau < \tilde{\mathbf{x}}_i)$, the value of $\hat{\mathbf{x}}_i$ would be always

$$\hat{\mathbf{x}}_i = \frac{\pm(\tau - \tilde{\mathbf{x}}_i) + \sqrt{(\tau - \tilde{\mathbf{x}}_i)^2 + 4\tau\mathbf{y}_i}}{2} > 0$$

Therefore, the value of $\hat{\mathbf{x}}_i$ is always positive for any value of $\tau > 0$ and $\mathbf{y}_i > 0$ which assure $\hat{\mathbf{x}}_i > 0$ in each iteration if we ensure $\tau > 0$.

Acknowledgments: We would like to thank the authors of [27] for providing the real-world Poisson noisy images and [18], [31], [34], [55] for sharing their source-codes. We would like to thank Yunjin Chen for useful discussions which helped in our work.

References

[1] K. E. Timmermann and R. D. Nowak, "Multiscale modeling and estimation of Poisson processes with application to photon-limited imaging," *IEEE Transactions on Information Theory*, vol. 45, no. 3, pp. 846–862, 1999.

[2] X. Jin and K. Hirakawa, "Approximations to camera sensor noise," SPIE, vol. 8655, pp. 86 550H–1–86 550H–7, 2013.

[3] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2005, pp. 60–65.

[4] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[5] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2419–2434, 2009.

[6] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," in *Proceedings of IEEE International Conference on Computer Vision*, 2011, pp. 479–486.

[7] W. Dong, L. Zhang, and G. Shi, "Centralized sparse representation for image restoration," in *Proceedings of IEEE International Conference on Computer Vision*, 2011, pp. 1259–1266.

[8] S. Wang, L. Zhang, and Y. Liang, "Nonlocal spectral prior model for low-level vision," in *Proceedings of Asian Conference on Computer Vision*, 2012, pp. 231–244.

[9] W. Dong, G. Shi, and X. Li, "Nonlocal image restoration with bilateral variance estimation: a low-rank approach," *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 700–711, 2013.

[10] J. Xu, L. Zhang, W. Zuo, D. Zhang, and X. Feng, "Patch group based nonlocal self-similarity prior learning for image denoising," in *Proceedings of IEEE International Conference on Computer Vision*, 2015, pp. 244–252.

[11] 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.

[12] F. J. Anscombe, "The transformation of Poisson, binomial and negative-binomial data," *Biometrika*, vol. 35, no. 3/4, pp. 246–254, 1948.

[13] D. L. Donoho, "Nonlinear wavelet methods for recovery of signals, densities, and spectra from indirect and noisy data," in *Proceedings of Symposia in Applied Mathematics*. Citeseer, 1993.

[14] M. Fisz, "The limiting distribution of a function of two independent random variables and its statistical application," *Colloquium Mathematicum*, vol. 3, pp. 138–146, 1955.

[15] P. Fryzlewicz and G. P. Nason, "A Haar-Fisz algorithm for Poisson intensity estimation," *Journal of Computational and Graphical Statistics*, vol. 13, no. 3, pp. 621–638, 2004.

[16] P. Fryzlewicz, V. Delouille, and G. P. Nason, "Goes-8 x-ray sensor variance stabilization using the multiscale data-driven Haar–Fisz transform," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 56, no. 1, pp. 99–116, 2007.

[17] B. Zhang, J. M. Fadili, and J.-L. Starck, "Wavelets, ridgelets, and curvelets for Poisson noise removal," *IEEE Transactions on Image Processing*, vol. 17, no. 7, pp. 1093–1108, 2008.

[18] M. Makitalo and A. Foi, "Optimal inversion of the Anscombe transformation in low-count Poisson image denoising," *IEEE Transactions on Image Processing*, vol. 20, no. 1, pp. 99–109, 2011.

[19] L. Azzari and A. Foi, "Variance stabilization for noisy+ estimate combination in iterative Poisson denoising," *IEEE Signal Processing Letters*, vol. 23, no. 8, pp. 1086–1090, 2016.

[20] R. D. Nowak and R. G. Baraniuk, "Wavelet-domain filtering for photon imaging systems," *IEEE Transactions on Image Processing*, vol. 8, no. 5, pp. 666–678, 1999.

[21] R. Nowak and E. D. Kolaczyk, "A Bayesian multiscale framework for Poisson inverse problems," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1999, pp. 1741–1744.

[22] R. D. Nowak and E. D. Kolaczyk, "A statistical multiscale framework for Poisson inverse problems," *IEEE Transactions on Information Theory*, vol. 46, no. 5, pp. 1811–1825, 2000.

[23] E. D. Kolaczyk, "Bayesian multiscale models for Poisson processes," *Journal of the American Statistical Association*, vol. 94, no. 447, pp. 920–933, 1999.

[24] K. Hirakawa and P. J. Wolfe, "Skellam shrinkage: Wavelet-based intensity estimation for inhomogeneous Poisson data," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 1080–1093, 2012.

[25] F. Luisier, T. Blu, and M. Unser, "Image denoising in mixed Poisson–Gaussian noise," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 696–708, 2011.

[26] Y. Le Montagner, E. D. Angelini, and J.-C. Olivo-Marin, "An unbiased risk estimator for image denoising in the presence of mixed Poisson–Gaussian noise," *IEEE Transactions on Image Processing*, vol. 23, no. 3, pp. 1255–1268, 2014.

[27] W. Cheng and K. Hirakawa, "Minimum risk wavelet shrinkage operator for Poisson image denoising," *IEEE Transactions on Image Processing*, vol. 24, no. 5, pp. 1660–1671, 2015.

[28] A. A. Bindilatti and N. D. Mascarenhas, "A nonlocal Poisson denoising algorithm based on stochastic distances," *IEEE Signal Processing Letters*, vol. 20, no. 11, pp. 1010–1013, 2013.

[29] R. Giryes and M. Elad, "Sparsity-based Poisson denoising with dictionary learning," *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 5057–5069, 2014.

[30] X. Zhang, Y. Lu, and T. Chan, "A novel sparsity reconstruction method from Poisson data for 3d bioluminescence tomography," *Journal of Scientific Computing*, vol. 50, no. 3, pp. 519–535, 2012.

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

[32] Z. T. Harmany, R. F. Marcia, and R. M. Willett, "This is spiral-tap: Sparse Poisson intensity reconstruction algorithmstheory and practice," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 1084–1096, 2012.

[33] A. Rond, R. Giryes, and M. Elad, "Poisson inverse problems by the plug-and-play scheme," *Journal of Visual Communication and Image Representation*, vol. 41, pp. 96–108, 2016.

[34] 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, 2018.

[35] 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.

[36] M. A. Figueiredo and J. M. Bioucas-Dias, "Restoration of poissonian images using alternating direction optimization," *IEEE Transactions on Image Processing*, vol. 19, no. 12, pp. 3133–3145, 2010.

[37] M. Collins, S. Dasgupta, and R. E. Schapire, "A generalization of principal components analysis to the exponential family," in *Proceedings of Advances in Neural Information Processing Systems*, 2002, pp. 617–624.

[38] Y. Chen and T. Pock, "Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1256–1272, 2017.

[39] M. Elad, M. A. Figueiredo, and Y. Ma, "On the role of sparse and redundant representations in image processing," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 972–982, 2010.

[40] M. Fazel, "Matrix rank minimization with applications," Ph.D. dissertation, PhD thesis, Stanford University, 2002.

[41] W. Dong, G. Shi, X. Li, Y. Ma, and F. Huang, "Compressive sensing via nonlocal low-rank regularization," *IEEE Transactions on Image Processing*, vol. 23, no. 8, pp. 3618–3632, 2014.

[42] H. Ji, C. Liu, Z. Shen, and Y. Xu, "Robust video denoising using low rank matrix completion," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2010, pp. 1791–1798.

[43] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 2862–2869.

[44] L. Ma, L. Xu, and T. Zeng, "Low rank prior and total variation regularization for image deblurring," *Journal of Scientific Computing*, vol. 70, no. 3, pp. 1336–1357, 2017.

[45] G. P. Kumar and R. R. Sahay, "Accurate structure recovery via weighted nuclear norm: A low rank approach to shape-from-focus," in *Proceedings of IEEE International Conference on Computer Vision*, 2017, pp. 563–574.

[46] T. Goldstein and S. Osher, "The split Bregman method for ℓ_1-regularized problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.

[47] J.-F. Cai, S. Osher, and Z. Shen, "Split Bregman methods and frame based image restoration," *Multiscale Modeling & Simulation*, vol. 8, no. 2, pp. 337–369, 2009.

[48] Y. Nesterov, *Introductory lectures on convex optimization: A basic course*. Springer Science & Business Media, 2013, vol. 87.

[49] N. Parikh, S. Boyd *et al.*, "Proximal algorithms," *Foundations and Trends in Optimization*, vol. 1, no. 3, pp. 127–239, 2014.

[50] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Modeling & Simulation*, vol. 4, no. 4, pp. 1168–1200, 2005.

[51] I. Csiszar, "Why least squares and maximum entropy? an axiomatic approach to inference for linear inverse problems," *The Annals of Statistics*, pp. 2032–2066, 1991.

[52] A. Olmos and F. A. Kingdom, "A biologically inspired algorithm for the recovery of shading and reflectance images," *Perception*, vol. 33, no. 12, pp. 1463–1473, 2004.

[53] 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 *Proceedings of IEEE International Conference on Computer Vision*, 2001, pp. 416–423.

[54] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[55] F. Luisier, C. Vonesch, T. Blu, and M. Unser, "Fast interscale wavelet denoising of Poisson-corrupted images," *Signal Processing*, vol. 90, no. 2, pp. 415–427, 2010.