J. Korean Math. Soc. **55** (2018), No. 3, pp. 719–734
https://doi.org/10.4134/JKMS.j170432
pISSN: 0304-9914 / eISSN: 2234-3008

ITERATIVE REWEIGHTED ALGORITHM FOR NON-CONVEX POISSONIAN IMAGE RESTORATION MODEL

Taeuk Jeong, Yoon Mo Jung, and Sangwoon Yun

Abstract. An image restoration problem with Poisson noise arises in many applications of medical imaging, astronomy, and microscopy. To overcome ill-posedness, Total Variation (TV) model is commonly used owing to edge preserving property. Since staircase artifacts are observed in restored smooth regions, higher-order TV regularization is introduced. However, sharpness of edges in the image is also attenuated. To compromise benefits of TV and higher-order TV, the weighted sum of the non-convex TV and non-convex higher order TV is used as a regularizer in the proposed variational model. The proposed model is non-convex and non-smooth, and so it is very challenging to solve the model. We propose an iterative reweighted algorithm with the proximal linearized alternating direction method of multipliers to solve the proposed model and study convergence properties of the algorithm.

1. Introduction

Images degraded by blurring and further corrupted by Poisson noise appear in various applications including medical [22], biological [20], and astronomical [3] imaging, for example; the recovery of those images is an important task and an active research area [1]. Poisson data occur in imaging processes where images are obtained by counting particles, such as photons hitting a detector. We refer to the survey paper [1] for a snap shot view of the literature and the comprehensive reference therein.

More specifically, we assume that observations follow a Poisson distribution whose mean is linearly related to the underlying image. Let $b \in \Re_+^n$ and $u \in \Re_+^n$

Received June 28, 2017; Accepted September 6, 2017.
2010 *Mathematics Subject Classification.* Primary 90C26, 49M37.
Key words and phrases. iterative reweighted algorithm, proximal linearized alternating direction method of multipliers, Poisson image restoration, non-convex.
The second author was financially supported by the National Research Foundation of Korea (NRF) NRF-2016R1A5A1008055 and NRF-2016R1D1A1B03931337.
The third author was financially supported by the National Research Foundation of Korea (NRF) NRF-2014R1A1A2056038, NRF-2016R1A5A1008055 and NRF-2016R1D1A1B 03934371.

be vectorized versions of the two dimensional observed image and the image to be estimated of the size $M \times N$ ($n = MN$), respectively. Now the image b can be represented as

$$b \sim \text{Poisson}(Au), \tag{1}$$

where $A \in \Re^{n\times n}$ is a linear operator, the identity operator for pure denoising or a convolution with a blur kernel for deblurring. It is a typical inverse problem and thus ill-posed in most cases of interest.

To bypass such difficulties, a *maximum a posteriori* (MAP) estimate is widely used, which consists of a fitting term and a regularization term. From the following likelihood function of the Poisson noise model (1)

$$p(b|Au) = \prod_{i=1}^{n} \frac{((Au)_i)^{b_i}}{b_i!} \exp(-(Au)_i),$$

where x_i denotes the ith component of x, the negative log-likelihood

$$\langle \mathbf{1},\, Au\rangle - \langle b,\, \log(Au)\rangle \tag{2}$$

can be derived [9]. By adopting the Total Variation (TV) as a regularizer, the following convex minimization model

$$\min_{u\in U} \langle \mathbf{1},\, Au\rangle - \langle b,\, \log(Au)\rangle + \lambda TV(u), \tag{3}$$

where $\lambda > 0$, $TV(u) = \sum_{i=1}^{n} \sqrt{((\nabla u)_{i,1})^2 + ((\nabla u)_{i,2})^2}$, and $U = [\upsilon, C]^n$ with $0 < \upsilon \le C < \infty$, is successfully applied for Poisson image restoration problems [9,21]; the TV regularizer provides relatively well-preserved and sharp edges in the restoration. Here, υ is a positive constant to ensure taking the logarithm of u and C is a maximum value of u, for example, 255 if 8 bit unsigned integer is used, or 1 if rescaled.

However, the TV term often introduces so called *staircase artifacts*, an intrinsic property of TV favoring piecewise constant solutions even in smooth regions. To remove the drawback of the TV term in smooth regions, higher-order TV is applied for Gaussian noise, which may promote piecewise smooth solutions [5, 14]. The major challenge in the model using only higher-order derivatives is to maintain sharp edges in the reconstruction. To compromise benefits of both lower- and higher-order TV's, hybrid regularizers that use both lower- and higher-order derivatives have been proposed aiming to preserve discontinuities between smooth regions and keep smoothness within each region [12, 15].

In contrast to those convex models with Gaussian noise, Krishnan and Fergus [11] proposed a non-convex TV model based on the statistical observation that natural image gradients have a heavy-tailed distribution. This distribution is modeled by a hyper-Laplacian, $p(x) \propto e^{-k|x|^\alpha}$ with $0.5 \le \alpha \le 0.8$. Chartrand [6] also shows that a non-convex regularizer $\|\nabla u\|^\alpha$ with $0 < \alpha < 1$ preserves shapes better than the convex one.

Combining the idea of hybrid models and non-convex models, Oh et al. [18] proposed a variational model with the non-convex hybrid TV regularizer with Gaussian noise:

$$\min_u (1-c)\|\nabla u\|^{\alpha_1} + c\|\nabla^2 u\|^{\alpha_2} + \frac{\mu}{2}\|u-b\|^2, \tag{4}$$

where $0 \le c \le 1$ and $0 < \alpha_1, \alpha_2 < 1$. The superiority of the non-convex type TV regularizer is a good balance between preserving edges and alleviating staircase artifacts and is demonstrated by providing experimental results [13, 18].

By adopting the model in [18], we propose the non-convex weighted total variation (NWTV) model for Poisson image restoration problems:

$$\min_{u\in U}\ c_1\|\nabla u\|^{\alpha_1} + c_2\|\nabla^2 u\|^{\alpha_2} + \langle \mathbf{1},\, Au\rangle - \langle b,\, \log(Au)\rangle, \tag{5}$$

where $c_1,\ c_2 \ge 0$. In words, we use a weighted sum of the non-convex first order TV and the non-convex second order TV as a regularizer which could preserve image details such as edges well while reducing staircase artifacts.

Since both terms are non-convex and non-smooth, we solve the proposed model by adapting an iterative reweighted algorithm (IRA) [4, 8, 17] with the proximal linearized alternating direction method of multipliers (PLADM) [7, 10]. We also analyze the convergence properties of IRA for the proposed model by using the framework of a block nonlinear Gauss-Seidel method. Since we assume only differentiability and/or strong convexity of the fitting term, the convergence analysis includes the iterative reweighted algorithm solving the non-convex hybrid variational models proposed in [13, 18]. We note that the convergence properties of the methods in [13, 18] have not been studied. Numerical experiments demonstrate the efficiency and stability of the proposed NWTV model for Poisson image restoration problems.

In our notation, for any $x \in \Re^n$, x_j denotes the jth component of x, and $\|x\|_p = \left(\sum_{j=1}^n |x_j|^p\right)^{1/p}$ for $1 \le p < \infty$. For simplicity, we write $\|x\| = \|x\|_2$. The identity matrix is denoted by I and the vector of zero entries is denoted by 0. $\mathbf{1}$ is the vector of all ones. Unless otherwise specified, $\{x^k\}$ denotes the sequence $x^0, x^1, \ldots$.

2. Iterative reweighted algorithm

In this section, we describe an iterative reweighted algorithm for solving the proposed non-convex weighted total variation model (5) and establish its convergence.

The proposed model (5) is non-convex and non-smooth, and so gradient based algorithms are not suitable for applying to solve the model (5). Hence we propose an iterative reweighted algorithm (IRA) for solving (5), i.e., we use the following convex model at each iteration:

$$u^{k+1} = \arg\min_{u\in U}\ f(u) + c_1\nu_1(\nabla u^k)\|\nabla u\| + c_2\nu_2(\nabla^2 u^k)\|\nabla^2 u\|, \tag{6}$$

where

$$f(u) = \langle \mathbf{1},\, Au\rangle - \langle b,\, \log(Au)\rangle$$

and, with $\eta^k \geq 0$,

$$\nu_1(\nabla u^k) = \frac{\alpha_1}{(\|\nabla u^k\| + \eta^k)^{1-\alpha_1}}, \quad \text{and} \quad \nu_2(\nabla^2 u^k) = \frac{\alpha_2}{(\|\nabla^2 u^k\| + \eta^k)^{1-\alpha_2}}.$$

Here, ∇u is given by $(\nabla u)_{i,j} = ((D_x^+ u)_{i,j}, (D_y^+ u)_{i,j})$, where $D_x^+ u$ and $D_y^+ u$ denote the first order forward difference operator with the reflexive boundary condition:

$$(D_x^+ u)_{i,j} = \begin{cases} u_{i,j+1} - u_{i,j} & \text{if } 1 \leq j < N \\ 0 & \text{if } j = N, \end{cases}$$

$$(D_y^+ u)_{i,j} = \begin{cases} u_{i+1,j} - u_{i,j} & \text{if } 1 \leq i < N \\ 0 & \text{if } i = N. \end{cases}$$

Similarly, with the reflexive boundary condition, the first order backward difference operators $D_x^- u$ and $D_y^- u$ are denoted by

$$(D_x^- u)_{i,j} = \begin{cases} 0 & \text{if } j = 1 \\ u_{i,j} - u_{i,j-1} & \text{if } 1 < j \leq N, \end{cases}$$

$$(D_y^- u)_{i,j} = \begin{cases} 0 & \text{if } i = 1 \\ u_{i,j} - u_{i-1,j} & \text{if } 1 < i \leq N. \end{cases}$$

Based on the above operators, we derive the second order difference operator as follows:

$$(\nabla^2 u)_{i,j} = \begin{pmatrix} D_x^-(D_x^+ u)_{i,j} & D_x^+(D_y^+ u)_{i,j} \\ D_y^-(D_x^- u)_{i,j} & D_y^-(D_y^+ u)_{i,j} \end{pmatrix}.$$

They can be considered as approximations of the gradient and the Hessian matrix using finite differences.

We have two versions of IRA depending on the choice of η^k. In the first version of IRA, we reduce η^k gradually to zero and we describe it in Algorithm 1. The second version of IRA fixes η^k for all k and is later given in Algorithm 3.

Algorithm 1 IRA

Update u^{k+1} and η^{k+1} from u^k and η^k:

1. $u^{k+1} = \arg\min_{u \in U} f(u) + c_1 \nu_1(\nabla u^k)\|\nabla u\| + c_2 \nu_2(\nabla^2 u^k)\|\nabla^2 u\|$.
2. $\eta^{k+1} = \nu \eta^k$ with $\nu \in (0, 1)$.

In the first step, we apply the proximal linearized alternating direction method of multipliers to solve the subproblem (6). We describe those methods in Section 3.

In the second step, the parameter $\nu < 1$. Thus $\eta^k \to 0$ and so IRA is the exact algorithm for solving the NWTV model (5). But, in this case, $\nu_1(\nabla \tilde{u})$ and $\nu_2(\nabla^2 \tilde{u})$ in (6) may not be defined when $\tilde{u}$ is an accumulation point of

the sequence $\{u^k\}$ generated by IRA. Hence we need some assumptions for establishing the convergence of IRA.

Next, we give a proof for the convergence of IRA. We define the following function

$$
\begin{aligned}
F(u,t,\eta) := f(u) &+ c_1\alpha_1\left(\frac{\|\nabla u\|+\eta}{t_1^{1-\alpha_1}} + \frac{1-\alpha_1}{\alpha_1}t_1^{\alpha_1}\right) \\
&+ c_2\alpha_2\left(\frac{\|\nabla^2 u\|+\eta}{t_2^{1-\alpha_2}} + \frac{1-\alpha_2}{\alpha_2}t_2^{\alpha_2}\right)
\end{aligned} \tag{7}
$$

and consider the following minimization problem

$$
\min_{u\in U, t\in(0,\infty)^2, \eta\geq 0} F(u,t,\eta). \tag{8}
$$

If we apply the three-block nonlinear Gauss-Seidel (ThBGS) method to the problem (8), we alternatively minimize the problem with respect to u, with respect to t, and then with respect to η but inexactly. The algorithmic framework for ThBGS is given in Algorithm 2.

Algorithm 2 ThBGS

Update u^{k+1}, t^{k+1}, η^{k+1} from u^k, t^k, η^k:

1. $u^{k+1} = \arg\min_{u\in U} F(u, t^k, \eta^k)$.
2. $t^{k+1} = \arg\min_{t\in(0,\infty)^2} F(u^{k+1}, t, \eta^k)$.
3. $\eta^{k+1} = \nu\eta^k$ with $\nu \in (0,1)$.

By taking the derivative of F with respect to t, we have

$$
\nabla_t F(u,t,\eta) = \begin{pmatrix} c_1\alpha_1\left((\alpha_1-1)(\|\nabla u\|+\eta)t_1^{\alpha_1-2} + (1-\alpha_1)t_1^{\alpha_1-1}\right) \\ c_2\alpha_2\left((\alpha_2-1)(\|\nabla^2 u\|+\eta)t_2^{\alpha_2-2} + (1-\alpha_2)t_2^{\alpha_2-1}\right) \end{pmatrix}.
$$

With $u = u^{k+1}$ and $\eta = \eta^k$, the solution of $\nabla_t F(u,t,\eta) = 0$ is

$$
t_1^{k+1} = \|\nabla u^{k+1}\| + \eta^k \text{ and } t_2^{k+1} = \|\nabla^2 u^{k+1}\| + \eta^k. \tag{9}
$$

Thus, t has a closed form solution (9) at the second step. By ignoring the constant terms, the minimization problem in the first step can be expressed as follows:

$$
\min_{u\in U} f(u) + c_1\nu_1(\nabla u^k)\|\nabla u\| + c_2\nu_2(\nabla^2 u^k)\|\nabla^2 u\|,
$$

in which the objective function is the same one in the problem (6). Hence u^{k+1} in ThBGS is exactly same as u^{k+1} in IRA. Therefore the iterative reweighted algorithm for solving the proposed model (5) is equivalent to the three-block Gauss-Seidel method for solving the problem (8).

The next theorem shows that every accumulation point of the sequence generated by IRA is a stationary point of the problem (5).

In what follows, $\hat{u}$ is a stationary point of $f(u)+c_1\|\nabla u\|^{\alpha_1}+c_2\|\nabla^2 u\|^{\alpha_2}+\iota_U$, if

$$0 \in \nabla f(\hat{u}) + c_1\alpha_1\|\nabla\hat{u}\|^{\alpha_1-1}\partial(\|\nabla\hat{u}\|) + c_2\alpha_2\|\nabla^2\hat{u}\|^{\alpha_2-1}\partial(\|\nabla^2\hat{u}\|) + N_U(\hat{u}),$$

under conditions that $\nabla\hat{u} \neq 0$ and $\nabla^2\hat{u} \neq 0$. Here, ι_U denotes the indicator function of U, i.e., $\iota_U(u) = 0$ if $u \in U$ and $\iota_U(u) = \infty$ if $u \notin U$. ∂ is the set of subgradients and N_U is the normal cone of U [19].

Theorem 2.1. *Let $\{u^k\}$ be the sequence generated by IRA. Suppose that u^* is an accumulation point of the sequence and assume that $\nabla u^* \neq 0$ and $\nabla^2 u^* \neq 0$. Then u^* is a stationary point of the problem* (5).

Proof. Let $F(u,t,\eta)$ be defined in (7). By the first step in Algorithm 2, we can observe that

$$F(u^{k+1},t^k,\eta^k) \leq F(u^k,t^k,\eta^k). \tag{10}$$

Also, by the second step in Algorithm 2, we can observe that

$$F(u^{k+1},t^{k+1},\eta^k) \leq F(u^{k+1},t^k,\eta^k). \tag{11}$$

Since $t_1^{k+1} = \|\nabla u^{k+1}\| + \eta^k$ and $t_2^{k+1} = \|\nabla^2 u^{k+1}\| + \eta^k$, we have

$$\begin{aligned} F(u^{k+1},t^{k+1},\eta^k) &= f(u^{k+1}) + c_1\alpha_1\left(\frac{\|\nabla u^{k+1}\|+\eta^k}{(t_1^{k+1})^{1-\alpha_1}} + \frac{1-\alpha_1}{\alpha_1}(t_1^{k+1})^{\alpha_1}\right) \\ &\quad + c_2\alpha_2\left(\frac{\|\nabla^2 u^{k+1}\|+\eta^k}{(t_2^{k+1})^{1-\alpha_2}} + \frac{1-\alpha_2}{\alpha_2}(t_2^{k+1})^{\alpha_2}\right) \\ &= f(u^{k+1}) + c_1(\|\nabla u^{k+1}\|+\eta^k)^{\alpha_1} + c_2(\|\nabla^2 u^{k+1}\|+\eta^k)^{\alpha_2}. \end{aligned} \tag{12}$$

If $\eta^{k+1} = \nu\eta^k$ with $\nu < 1$, then we obtain

$$\begin{aligned} (\|\nabla u^{k+1}\|+\eta^{k+1})^{\alpha_1} &\leq (\|\nabla u^{k+1}\|+\eta^k)^{\alpha_1}, \\ (\|\nabla^2 u^{k+1}\|+\eta^{k+1})^{\alpha_2} &\leq (\|\nabla^2 u^{k+1}\|+\eta^k)^{\alpha_2}. \end{aligned}$$

These inequalities together with (12) imply

$$F(u^{k+1},t^{k+1},\eta^{k+1}) \leq F(u^{k+1},t^{k+1},\eta^k). \tag{13}$$

The inequalities (10), (11), and (13) yield that the sequence $\{F(u^k,t^k,\eta^k)\}$ is non-increasing. Since u^* is an accumulation point of $\{u^k\}$, there exists a subsequence K such that $\{u^k\}_K \to u^*$. And we let $t_1^* = \|\nabla u^*\|$ and $t_2^* = \|\nabla^2 u^*\|$.

By the definition of t^k and (12), we can verify that $F(u^k,t^k,\eta^k) = f(u^k) + c_1(\|\nabla u^k\|+\eta^k)^{\alpha_1} + c_2(\|\nabla^2 u^k\|+\eta^k)^{\alpha_2}$. From the continuity of f, $\|\cdot\|^{\alpha_1}$, and $\|\cdot\|^{\alpha_2}$, with $\{u^k\}_K \to u^*$ and $\eta^k \to 0$, we have $\{F(u^k,t^k,\eta^k)\}_K \to f(u^*) + c_1\|\nabla u^*\|^{\alpha_1} + c_2\|\nabla^2 u^*\|^{\alpha_2}$. This together with the monotonicity of $\{F(u^k,t^k,\eta^k)\}$ implies that $F(u^k,t^k,\eta^k) \to f(u^*) + c_1\|\nabla u^*\|^{\alpha_1} + c_2\|\nabla^2 u^*\|^{\alpha_2}$. Using this relation and (10), we further have

$$F(u^{k+1},t^k,\eta^k) \to f(u^*) + c_1\|\nabla u^*\|^{\alpha_1} + c_2\|\nabla^2 u^*\|^{\alpha_2}. \tag{14}$$

In addition, by the definition of u^{k+1}, we know that $F(u, t^k, \eta^k) \geq F(u^{k+1}, t^k, \eta^k)$ for all $u \in U$. Taking limits on both sides of this inequality with K and (14) imply that

$$F(u, t^*, 0) \geq f(u^*) + c_1\|\nabla u^*\|^{\alpha_1} + c_2\|\nabla^2 u^*\|^{\alpha_2} \quad \forall\, u \in U.$$

Since $t_1^* = \|\nabla u^*\|$ and $t_2^* = \|\nabla^2 u^*\|$,

$$F(u^*, t^*, 0) = f(u^*) + c_1\|\nabla u^*\|^{\alpha_1} + c_2\|\nabla^2 u^*\|^{\alpha_2},$$

and thus $u^* \in \arg\min_{u\in U} F(u, t^*, 0)$.

Hence u^* is a stationary point of $\min_{u\in U} F(u, t^*, 0)$, that is,

$$0 \in \nabla f(u^*) + c_1\alpha_1(t_1^*)^{\alpha_1-1}\partial\|\nabla u^*\| + c_2\alpha_2(t_2^*)^{\alpha_2-1}\partial\|\nabla^2 u^*\| + N_U(u^*).$$

The identities $t_1^* = \|\nabla u^*\|$ and $t_2^* = \|\nabla^2 u^*\|$ imply that

$$0 \in \nabla f(u^*)+c_1\alpha_1\|\nabla u^*\|^{\alpha_1-1}\partial(\|\nabla u^*\|)+c_2\alpha_2\|\nabla^2 u^*\|^{\alpha_2-1}\partial(\|\nabla^2 u^*\|)+N_U(u^*).$$

Therefore, u^* is a stationary point of the problem (5). □

Remark 2.2. In general, $\nabla u^* \neq 0$ and $\nabla^2 u^* \neq 0$ in natural scene images, and so, neither constant nor affine image may be considered. This shows the soundness of the assumption.

If $A = I$, the denoising case, a stronger result holds; the next theorem guarantees the convergence of the entire sequence.

Theorem 2.3. *Let $\{u^k\}$ be the sequence generated by IRA with $A = I$. If $\{u^k\}$ converges to u^* with $\nabla u^* \neq 0$ and $\nabla^2 u^* \neq 0$, then u^* is a stationary point of the problem* (5).

Proof. $f(u)$ is twice differentiable on the set U and $\nabla_u^2 f(u) = \mathrm{Diag}(b \circ u^{-2})$ for all $u \in U$. Since $\nabla_u^2 f(u) \succeq \zeta I$ with $\zeta := \frac{\min_j b_j}{C^2}$ for all $u \in U$, the function f is strongly convex [16], i.e.,

$$f(u) \geq f(v) + \langle \nabla f(v),\, u - v\rangle + \frac{\zeta}{2}\|u - v\|^2. \tag{15}$$

By Fermat's rule [19, Theorem 10.1],

$$u^{k+1} \in \arg\min_u\, \langle \nabla f(u^{k+1}),\, u\rangle + P(u),$$

where $P(u) := c_1\alpha_1 \left(\frac{\|\nabla u\|}{(t^k)_1^{1-\alpha_1}}\right) + c_2\alpha_2 \left(\frac{\|\nabla^2 u\|}{(t^k)_2^{1-\alpha_2}}\right)$. This implies that

$$\langle \nabla f(u^{k+1}),\, u^{k+1}\rangle + P(u^{k+1}) \leq \langle \nabla f(u^{k+1}),\, u^k\rangle + P(u^k). \tag{16}$$

From the strong convexity of f (15), we have

$$\begin{aligned} F(u^{k+1}, t^k, \eta^k) - F(u^k, t^k, \eta^k) &= f(u^{k+1}) - f(u^k) + P(u^{k+1}) - P(u^k) \\ &\leq \langle \nabla f(u^{k+1}),\, u^{k+1} - u^k\rangle - \frac{\zeta}{2}\|u^{k+1} - u^k\|^2 \\ &\quad + P(u^{k+1}) - P(u^k) \end{aligned}$$

$$\leq -\frac{\zeta}{2}\|u^{k+1}-u^k\|^2, \tag{17}$$

where the second inequality uses (16).

The above inequality (17) together with the inequalities (10), (11), (13), the boundedness of $\{\eta^k\}$, and the compactness of the set U implies that

$$\frac{\zeta}{2}\|u^{k+1}-u^k\|^2 \leq F(u^k,t^k,\eta^k) - F(u^{k+1},t^k,\eta^k) \to 0.$$

Since U is compact, there exists a $\tilde{u}$ such that $u^k \to \tilde{u}$. By similar arguments as the proof of Theorem 2.1 with $u^* = \tilde{u}$, we conclude that $\tilde{u}$ is a stationary point of the problem (18) with $A = I$. □

In the second step of IRA, if the parameter $\nu = 1$, i.e., η^k is fixed for all k, it becomes the inexact algorithm for solving the NWTV model (5). In other words, it solves the following approximated problem

$$\min_{u\in U} f(u) + c_1(\|\nabla u\| + \bar{\eta})^{\alpha_1} + c_2(\|\nabla^2 u\| + \bar{\eta})^{\alpha_2}, \tag{18}$$

where $\bar{\eta} = \eta^0 > 0$. The algorithm for solving the above problem (18) is given in Algorithm 3.

Algorithm 3 IRA-approx

Update u^{k+1} from u^k with $\eta^k = \bar{\eta}$ for all k:

1. $u^{k+1} = \arg\min_{u\in U} f(u) + c_1\nu_1(\nabla u^k)\|\nabla u\| + c_2\nu_2(\nabla^2 u^k)\|\nabla^2 u\|$.

The following theorem establishes the convergence for IRA-approx. In this case, the assumptions $\nabla u^* \neq 0$ and $\nabla^2 u^* \neq 0$ are not necessary due to the fixed $\eta^k = \bar{\eta}$.

Theorem 2.4. *Let $\{u^k\}$ be the sequence generated by IRA-approx. Suppose that u^* is an accumulation point of the sequence. Then u^* is a stationary point of the problem* (18).

Proof. By proceeding as the proof of Theorem 2.1 with $\eta^k = \bar{\eta}$ for all k, we have that

$$F(u^{k+1},t^{k+1},\bar{\eta}) \leq F(u^{k+1},t^k,\bar{\eta}) \leq F(u^k,t^k,\bar{\eta}) \quad \forall k, \tag{19}$$

i.e., $\{F(u^k,t^k,\bar{\eta})\}$ is non-increasing.

From the definition of t^k, we have

$$F(u^k,t^k,\bar{\eta}) = f(u^k) + c_1(\|\nabla u^k\| + \bar{\eta})^{\alpha_1} + c_2(\|\nabla^2 u^k\| + \bar{\eta})^{\alpha_2}.$$

One can show that there exists a subsequence K such that $\{u^k\}_K \to u^*$ and $\{F(u^k,t^k,\bar{\eta})\}_K \to f(u^*) + c_1(\|\nabla u^*\| + \bar{\eta})^{\alpha_1} + c_2(\|\nabla^2 u^*\| + \bar{\eta})^{\alpha_2}$, similar to the proof of Theorem 2.1.

Since the sequence $\{F(u^k, t^k, \bar{\eta})\}$ is non-increasing, $F(u^k, t^k, \bar{\eta}) \to f(u^*) + c_1(\|\nabla u^*\| + \bar{\eta})^{\alpha_1} + c_2(\|\nabla^2 u^*\| + \bar{\eta})^{\alpha_2}$. By using (19), we obtain

$$F(u^{k+1}, t^k, \bar{\eta}) \to f(u^*) + c_1(\|\nabla u^*\| + \bar{\eta})^{\alpha_1} + c_2(\|\nabla^2 u^*\| + \bar{\eta})^{\alpha_2}. \tag{20}$$

By further proceeding as the proof of Theorem 2.1 with $\eta^k = \bar{\eta}$ for all k, u^* is a stationary point of the problem (18). □

Similar to Theorem 2.3, the convergence of the entire sequence is also guaranteed for denoising.

Theorem 2.5. *Let $\{u^k\}$ be the sequence generated by IRA-approx with $A = I$. Then the sequence $\{u^k\}$ converges to a stationary point of the problem* (5).

Proof. By proceeding as the proof of Theorems 2.3 with $\eta^k = \bar{\eta}$ and 2.4, we conclude that the sequence $\{u^k\}$ converges to a stationary point of the problem (18) with $A = I$. □

Remark 2.6. The analysis of Theorems 2.1 and 2.4 can still hold when the function f is replaced by any differentiable function such as $\frac{\mu}{2}\|Au - b\|^2$. Also, the analysis of Theorems 2.3 and 2.5 are valid when the function f is replaced by any differentiable strongly convex function, for example, $\frac{\mu}{2}\|u - b\|^2$. Thus, those theorems also guarantee the convergence of the algorithms in [13, 18], which have not been studied yet.

3. Proximal linearized alternating direction method of multipliers

In this section, we briefly describe the proximal linearized alternating direction method of multipliers (PLADM) to solve the subproblem (6). The convergence properties of this method can be found in [7, 10].

Because of the nonlinearlity and non-smoothness properties of $\|\nabla u\|$ and $\|\nabla^2 u\|$, we introduce new variables $v = \nabla u$, $w = \nabla^2 u$, and $z = Au$ and then consider the following reformulation of the subproblem (6):

$$\begin{aligned} &\min_{u \in U, v, w, z} c_1\nu_1(\nabla u^k)\|v\| + c_2\nu_2(\nabla^2 u^k)\|w\| + \langle \mathbf{1},\, z\rangle - \langle b,\, \log(z)\rangle \\ &\text{subject to } v = \nabla u,\ w = \nabla^2 u,\ z = Au. \end{aligned} \tag{21}$$

To apply PLADM to solve the problem (21), we introduce the augmented Lagrangian function $\mathcal{L}$ and the linearized augmented Lagrangian function $\mathcal{L}_u$ with respect to the variable u:

$$\begin{aligned} \mathcal{L}(v, w, u, z, \lambda, \tau, \xi; u^k) = {} & c_1\nu_1(\nabla u^k)\|v\| - \langle \lambda,\, v - \nabla u\rangle + \frac{\beta}{2}\|v - \nabla u\|^2 \\ & + \ c_2\nu_2(\nabla^2 u^k)\|w\| - \langle \tau,\, w - \nabla^2 u\rangle + \frac{\gamma}{2}\|w - \nabla^2 u\|^2 \\ & + \langle \mathbf{1},\, z\rangle - \langle b,\, \log(z)\rangle - \langle \xi,\, z - Au\rangle + \frac{\rho}{2}\|z - Au\|^2, \end{aligned}$$

where β, γ, and ρ are positive penalty parameters controlling the proximity of v, w, and z to ∇u, $\nabla^2 u$ and Au, respectively, and

$$\begin{aligned}
&\mathcal{L}_u(v,w,u,z,\lambda,\tau,\xi;u^\ell,u^k)\\
&= c_1\nu_1(\nabla u^k)\|v\| - \langle\lambda,\, v\rangle + \ c_2\nu_2(\nabla^2 u^k)\|w\| - \langle\tau,\, w\rangle + \langle\mathbf{1},\, z\rangle\\
&\quad - \langle b,\, \log(z)\rangle - \langle\xi,\, z\rangle + \langle\Psi^\ell + A^T(\xi+\rho(Au^\ell - z),\, u-u^\ell\rangle + \frac{\delta}{2}\|u-u^\ell\|^2,
\end{aligned}$$

where $\Psi^\ell = \nabla^T(\lambda+\beta(\nabla u^\ell - v)) + (\nabla^2)^T(\tau+\gamma(\nabla^2 u^\ell - w))$ and δ is a penalty parameter for a simple quadratic approximation $\mathcal{L}_u$ of $\mathcal{L}$ with respect to u at u^ℓ.

We have the following framework of PLADM to solve the reformulation (21):

$$\text{(22)}\qquad \begin{cases}
v^{\ell+1} \leftarrow \arg\min\limits_{v} \mathcal{L}(v, w^\ell, u^\ell, z^\ell, \lambda^\ell, \tau^\ell, \xi^\ell; u^k)\\
w^{\ell+1} \leftarrow \arg\min\limits_{w} \mathcal{L}(v^{\ell+1}, w, u^\ell, z^\ell, \lambda^\ell, \tau^\ell, \xi^\ell; u^k)\\
z^{\ell+1} \leftarrow \arg\min\limits_{z} \mathcal{L}(v^{\ell+1}, w^{\ell+1}, u^\ell, z, \lambda^\ell, \tau^\ell, \xi^\ell; u^k)\\
u^{\ell+1} \leftarrow \arg\min\limits_{u\in U} \mathcal{L}_u(v^{\ell+1}, w^{\ell+1}, u, z^{\ell+1}, \lambda^\ell, \tau^\ell, \xi^\ell; u^\ell, u^k)\\
\lambda^{\ell+1} \leftarrow \lambda^\ell - \beta(v^{\ell+1} - \nabla u^{\ell+1})\\
\tau^{\ell+1} \leftarrow \tau^\ell - \gamma(w^{\ell+1} - \nabla^2 u^{\ell+1})\\
\xi^{\ell+1} \leftarrow \xi^\ell - \rho(z^{\ell+1} - Au^{\ell+1}).
\end{cases}$$

The first four steps have the closed form solutions. In what follows, $\mathcal{P}_U(u)$ denotes the projection of u onto the set U and the shrinkage operator in the first two steps is adopted from [23]:

$$\begin{aligned}
v_i^{\ell+1} &= \text{shrink}\left(\nabla u_i^\ell + \frac{\lambda_i^\ell}{\beta}, \frac{c_1\nu_1(\nabla u^k)}{\beta}\right)\\
&= \max\left\{\left\|\nabla u_i^\ell + \frac{\lambda_i^\ell}{\beta}\right\| - \frac{c_1\nu_1(\nabla u^k)}{\beta}, 0\right\}\frac{\nabla u_i^\ell + \frac{\lambda_i^\ell}{\beta}}{\left\|\nabla u_i^\ell + \frac{\lambda_i^\ell}{\beta}\right\|} \quad (i=1,\dots,n),\\
w_i^{\ell+1} &= \text{shrink}\left(\nabla^2 u_i^\ell + \frac{\tau_i^\ell}{\gamma}, \frac{c_2\nu_2(\nabla^2 u^k)}{\gamma}\right)\\
&= \max\left\{\left\|\nabla^2 u_i^\ell + \frac{\tau_i^\ell}{\gamma}\right\| - \frac{c_2\nu_2(\nabla^2 u^k)}{\gamma}, 0\right\}\frac{\nabla^2 u_i^\ell + \frac{\tau_i^\ell}{\gamma}}{\left\|\nabla^2 u_i^\ell + \frac{\tau_i^\ell}{\gamma}\right\|} \quad (i=1,\dots,n),\\
z^{\ell+1} &= \frac{q^\ell + \sqrt{(q^\ell)^2 + 4\rho b}}{2\rho},\\
u^{\ell+1} &= \mathcal{P}_U\left(u^\ell - \frac{r^\ell}{\delta}\right),
\end{aligned}$$

where $q^\ell = \rho Au^\ell + \xi^\ell - \mathbf{1}$ and $r^\ell = \nabla^T(\lambda^\ell + \beta(\nabla u^\ell - v^{\ell+1})) + (\nabla^2)^T(\tau^\ell + \gamma(\nabla^2 u^\ell - w^{\ell+1})) + A^T(\xi^\ell + \rho(Au^\ell - z^{\ell+1}))$. We note that the alternating direction method

(a) boat (b) cameraman (c) goldhill (d) lena (e) peppers

FIGURE 1. Dataset for numerical experiments: Five standard test images of the size 512×512.

of multipliers (ADMM) [2, 24] can also be applied to solve the reformulation (21) and its framework is different from PLADM at the fourth step in which it requires an inversion. PLADM has an advantage over ADMM when the cost for an inversion is expensive.

4. Numerical experiments

In this section, we report numerical results on image restoration problems with Poisson noise, including denoising and deblurring. We assume that the dynamic range of all image data is $[0, 1]$, by normalization if necessary.

We compare the proposed NWTV model (5) with IRA to the total variation regularized variational (TVRV) model (3) with PIDSplit+ [21] and the convex case of the model (5) with $\alpha_1 = \alpha_2 = 1$ (L1L1). L1L1 is solved by ADMM.

Five well-known standard test images (boat, cameraman, goldhill, lena, and peppers) of size 512×512 in Figure 1 are used for our numerical experiments and all algorithms are implemented in Matlab. We stop all the algorithms by the following stopping condition:

$$||u^{k+1} - u^k|| \leq C_t||u^k||, \tag{23}$$

where $C_t > 0$ is the stopping tolerance and set to $C_t = 5 \times 10^{-3}$ for the proposed methods. Similarly, the inner solver PLADM (22) is stopped by the condition (23) with $C_t = 10^{-3}$. The following peak signal-to-noise ratio (PSNR) between the original and reconstructed image is calculated and used for quality comparison:

$$PSNR = 10 \log_{10} \left(\frac{255^2 n}{\|\bar{u} - u^*\|^2} \right),$$

where n is the size of the image, $\bar{u}$ is the original image, and u^* is the recovered image.

4.1. Image denoising

To test the performances of the proposed NWTV model (5) for Poisson noise removal, the model parameters α_1, α_2, c_1, and c_2 are chosen by the best average PSNR on the chosen 5 test images and additional 4 standard test images. The

parameters α_1 and α_2 are simply selected among four pairs $(0.7, 0.7)$, $(0.7, 0.9)$, $(0.9, 0.7)$, and $(0.9, 0.9)$, based on the paper [18]. The algorithm parameters β, γ, ρ, and δ are heuristically chosen as 0.4, 0.2, 1.0, and 0.8, respectively and fixed for all experiments. The parameter value λ of TVRV model and all parameters of L1L1 are tuned to provide the high performance on the images. Parameter are as follows: For TVRV, $\lambda = 0.10$. For L1L1, $c_1 = 0.05$, $c_2 = 0.05$. For NWTV, $\alpha_1 = 0.7$, $\alpha_2 = 0.7$, $c_1 = 0.10$, $c_2 = 0.20$.

Table 1 reports the CPU time (in seconds) and the PSNR values for TVRV, L1L1, and NWTV. Comparing to TVRV, the proposed NWTV model shows 0.49dB and 1.07dB improvements in terms of average PSNR values. Comparing to L1L1, the NWTV model shows better performance but improvements are less than the case of TVRV. The recovered images are given in Figure 2.

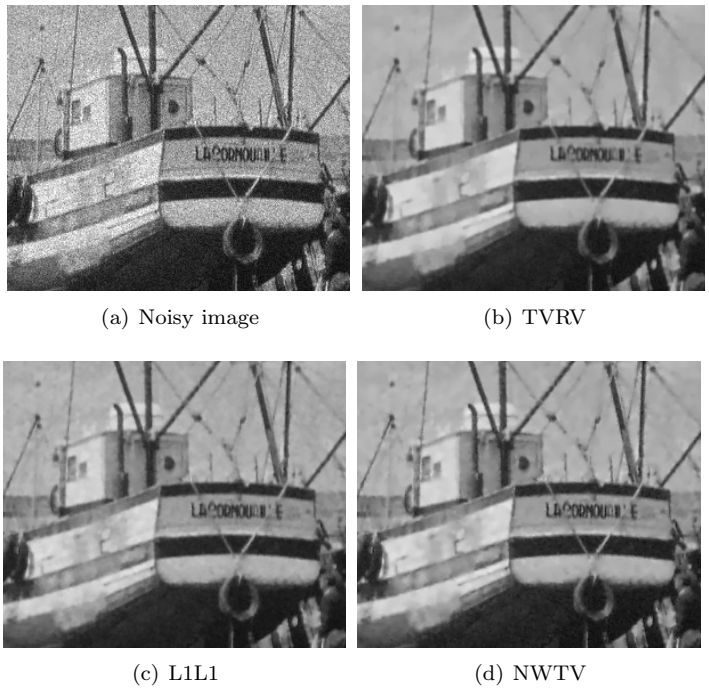
TABLE 1. Comparison for denoising

Denoising	Noisy	TVRV		L1L1		NWTV	
image	PSNR	time	PSNR	time	PSNR	time	PSNR
boat	22.74	0.53	29.89	0.46	30.08	1.85	30.55
cameraman	23.34	0.53	32.96	0.45	32.99	1.85	33.36
goldhill	23.67	0.53	29.83	0.46	30.02	1.85	30.47
lena	23.11	0.52	32.19	0.46	32.08	1.84	32.68
peppers	23.05	0.52	31.15	0.45	31.50	1.87	32.22

4.2. Image deblurring

To perform restoring images degraded by blurring and Poisson noise, we consider two cases; 7×7 Gaussian blur with standard deviation 1.5 and 7×7 uniform blur. We again compare the performance of NWTV to TVRV and L1L1. The parameters are chosen under the same criteria as for the denoising case. The parameters β, γ, ρ, and δ are set to 0.5, 0.05, 2.0, and 0.9, except PLADM for Gaussian blur. In that case, γ is 0.2. The parameters for Gaussian blur are as follows: For TVRV, $\lambda = 0.17$. For L1L1, $c_1 = 0.08$, $c_2 = 0.08$. For NWTV, $\alpha_1 = 0.7$, $\alpha_2 = 0.7$, $c_1 = 0.10$, $c_2 = 0.20$. The parameters for uniform blur are $\lambda = 0.20$ for TVRV, $c_1 = 0.10$, $c_2 = 0.10$ for L1L1, and, $\alpha_1 = 0.7$, $\alpha_2 = 0.7$, $c_1 = 0.20$, $c_2 = 0.40$ for NWTV.

Similar to Table 1, Tables 2 and 3 report the results of deblurring under Gaussian blur and uniform blur, respectively. Comparing to TVRV and L1L1, the proposed NWTV model shows similar improvements to denoising cases, in terms of average PSNR values. For example, the PSNR value of NWTV for the "goldhill" image is 0.36dB higher than that of TVRV under the uniform blur. The restored images of uniform blur case are given in Figure 3.



(a) Noisy image (b) TVRV

(c) L1L1 (d) NWTV

FIGURE 2. Denoising on the subimage of "boat"

TABLE 2. Comparison for deblurring under Gaussian blur

Denoising	Noisy	TVRV		L1L1		NWTV-PLADM	
image	PSNR	time	PSNR	time	PSNR	time	PSNR
boat	17.06	0.67	24.42	0.55	24.56	3.70	24.60
cameraman	17.69	0.90	25.84	0.63	25.97	3.74	25.93
goldhill	18.01	0.63	25.27	0.63	25.30	3.72	25.54
lena	17.67	0.68	26.65	0.63	26.66	3.83	26.76
peppers	17.63	0.68	26.47	0.54	26.55	3.80	26.71

5. Conclusion

In this paper, we propose a non-convex weighted total variation model for image restoration problems with Poisson noise and propose an iterative reweighted algorithm with the alternating direction method of multipliers or the linearized alternating direction method of multipliers. We also analyze the convergence properties of the iterative reweighted algorithm for the proposed non-convex and nonsmooth model by using the framework of the coordinate

TABLE 3. Comparison for deblurring under the uniform blur

Denoising image	Noisy PSNR	TVRV time	TVRV PSNR	L1L1 time	L1L1 PSNR	NWTV time	NWTV PSNR
boat	16.80	0.68	23.50	0.54	23.60	4.15	23.75
cameraman	17.40	0.90	24.51	0.63	24.59	4.25	24.62
goldhill	17.80	0.69	24.55	0.63	24.54	4.29	24.91
lena	17.51	0.67	25.71	0.63	25.73	4.27	25.96
peppers	17.45	0.68	25.50	0.63	25.63	4.17	25.77

(a) Noisy image

(b) TVRV

(c) L1L1

(d) NWTV

FIGURE 3. Restoration of uniform blur

descent method. The convergence analysis guarantees that there exists a convergent subsequence to a stationary point. We further show that the whole sequence converges to a stationary point when the fidelity term is strongly convex, for example, denoising problems. The numerical simulations show that

the NWTV model with IRA performs better than the TVRV model with PID-Split+ and the convex model (5) with $\alpha_1 = \alpha_2 = 1$ in terms of the PSNR value.

References

[1] M. Bertero, P. Boccacci, G. Desidera, and G. Vicidomini, *Image deblurring with Poisson data: from cells to galaxies*, Inverse Problems **25** (2009), no. 12, 123006, 26 pp.

[2] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, *Distributed optimization and statistical learning via the alternating direction method of multipliers*, Foundations and Trends in Machine Learning **3** (2011), 1–122.

[3] E. Bratsolis and M. Sigelle, *A spatial regularization method preserving local photometry for Richardson-Lucy restoration*, Astron. Astrophys. **375** (2001), 1120–1128.

[4] E. J. Candé, M. B. Wakin, and S. P. Boyd, *Enhancing sparsity by reweighted l_1 minimization*, J. Fourier Anal. Appl. **14** (2008), no. 5-6, 877–905.

[5] T. Chan, A. Marquina, and P. Mulet, *High-order total variation-based image restoration*, SIAM J. Sci. Comput. **22** (2000), no. 2, 503–516.

[6] R. Chartrandn, *Nonconvex regularization for shape preservation*, in Proc. IEEE Intl. Conf. Image Processing (ICIP) **1** (2007), 293–296.

[7] D.-Q. Chen and L.-Z. Cheng, *Fast linearized alternating direction minimization algorithm with adaptive parameter selection for multiplicative noise removal*, J. Comput. Appl. Math. **257** (2014), 29–45.

[8] I. Daubechies, R. Devore, M. Fornasier, and C. S. Güntuürk, *Iteratively reweighted least squares minimization for sparse recovery*, Comm. Pure Appl. Math. **63** (2010), no. 1, 1–38.

[9] M. A. T. Figueiredo and J. M. Bioucas-Dias, *Restoration of Poissonian images using alternating direction optimization*, IEEE Trans. Image Process. **19** (2010), no. 12, 3133–3145.

[10] T. Jeong, H. Woo, and S. Yun, *Frame-based Poisson image restoration using a proximal linearized alternating direction method*, Inverse Problems **29** (2013), no. 7, 075007, 20 pp.

[11] D. Krishnan, and R. Fergus, *Fast image deconvolution using hyper-Laplacian priors*, in Proc. Adv. Neural Inf. Process. Syst. **22** (2009), 1–9.

[12] F. Li, C. Shen, J. Fan, and C. Shen, *Image restoration combining a total variational filter and a fourth-order filter*, J. Vis. Commun. Image Represent. **18** (2007), 322–330.

[13] R. W. Liu and T. Xu, *A robust alternating direction method for constrained hybrid variational deblurring model*, Preprint, 2013.

[14] M. Lysaker, A. Lundervold, and X.-C. Tai, *Noise removal using fourth-order partial differential equation with applaication to medical magnetic resonance images in space and time*, IEEE Trans. Image Process. **12** (2003), 1579–1590.

[15] M. Lysaker and X.-C. Tai, *Iterative image restoration combining total variation minimization and a second-order functional*, Int. J. Comput. Vis. **66** (2006), 5–18.

[16] Y. Nesterov, *Introductory Lectures on Convex Optimization*, Applied Optimization, **87**, Kluwer Academic Publishers, Boston, MA, 2004.

[17] P. Ochs, A. Dosovitskiy, T. Brox, and T. Pock, *On iteratively reweighted algorithms for nonsmooth nonconvex optimization in computer vision*, SIAM J. Imaging Sci. **8** (2015), no. 1, 331–372.

[18] S. Oh, H. Woo, S. Yun, and M. Kang, *Non-convex hybrid total variation for image denoising*, J. Vis. Comm. Image R. **24** (2013), 332–344.

[19] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis*, Grundlehren der Mathematischen Wissenschaften, **317**, Springer-Verlag, Berlin, 1998.

[20] P. Sarder and A. Nehorai, *Deconvolution methods for 3-d fluorescence microscopy images*, IEEE Signal Process. Mag. **23** (2006), 32–45.
[21] S. Setzer, G. Steidl, and T. Teuber, *Deblurring Poissonian images by split Bregman techniques*, J. Vis. Commun. Image R. **21** (2010), 193–199.
[22] L. A. Shepp and Y. Vardi, *Maximum likelihood reconstruction for emission tomography*, IEEE Trans. Med. Image **1** (1982), 113–122.
[23] Y. Wang, J. Yang, W. Yin, and Y. Zhang, *A new alternating minimization algorithm for total variation image reconstruction*, SIAM J. Imaging Sci. **1** (2008), no. 3, 248–272.
[24] C. Wu and X.-C. Tai, *Augmented Lagrangian method, dual methods, and split Bregman iteration for ROF, vectorial TV, and high order models*, SIAM J. Imaging Sci. **3** (2010), no. 3, 300–339.

Taeuk Jeong
Department of Computational Science and Engineering
Yonsei University
Seoul 03722, Korea
Email address: iamlogin@yonsei.ac.kr

Yoon Mo Jung
Department of Mathematics
Sungkyunkwan University
Suwon 16419, Korea
Email address: yoonmojung@skku.edu

Sangwoon Yun
Department of Mathematics Education
Sungkyunkwan University
Seoul 03063, Korea
Email address: yswmathedu@skku.edu