NON-CONVEX TV DENOISING CORRUPTED BY IMPULSE NOISE

Yoon Mo Jung
Department of Mathematics, Sungkyunkwan University
Suwon 16419, Korea

Taeuk Jeong
Department of Computational Science and Engineering, Yonsei University
Seoul 03722, Korea

Sangwoon Yun
Department of Mathematics Education, Sungkyunkwan University
Seoul 03063, Korea

(Communicated by Sung Ha Kang)

Abstract. We propose a non-convex type total variation model for impulse noise removal by incorporating TV and the quasi-norm $\ell_q$, $0 < q < 1$. Since the proposed model is non-convex and non-smooth, an iteratively reweighted algorithm is adapted and combined with a linearized ADMM. The convergence of the proposed algorithm is established and numerical results are given to illustrate the validity and efficiency of the proposed model.

1. Introduction. We consider the image denoising problem degraded by impulse noise, which is caused by faulty pixels in camera or defective memory positions in hardware, for example [1, 4]. Two types of impulse noise are widely considered; salt-and-pepper noise and random-valued noise. Salt-and-pepper noise takes only the maximum and the minimum values in its range at noisy pixels. In the case of random-valued noise, any random value can be taken [4].

To remove impulse noise, median type filters are widely used as a robust estimator, since noise statistics are characterized by probability densities having heavier tails than that of Gaussian [1]. To overcome the limitation of the median filter, diverse remedies and successors have been proposed such as the adaptive median filter [9] and the multistate median filter [6]. We refer the interested reader to the paper [4] and the references therein.

On the other hand, denoising models using the $\ell_1$ norm as a data-fidelity term have been proposed to deal with outliers and impulse noise [13, 14]. Actually the median and the $\ell_1$ norm are closely related, since the median is a minimizer of the mean absolute error (MAE) and thus the sample median is the maximum likelihood (ML) estimator in the presence of Laplacian noise. To preserve edges and maintain...
sharp discontinuities, total variation (TV) is applied for the regularizer in addition to the ℓ₁ norm as fidelity term [3, 13].

However, TV often causes staircase artifacts, favoring piecewise constant solutions even in smooth regions. To alleviate the drawback in smooth regions and to promote piecewise smooth solutions, higher-order TV is applied for Gaussian noise [5, 13]. Since a heavy-tailed distribution is often observed in natural image gradients, a hyper-Laplacian distribution $p(x) \propto e^{-k|\nabla x|^{q}}$ with $0.5 \leq \alpha \leq 0.8$ is introduced [11]. Combining the idea of higher-order TV and hyper-Laplacian, we apply a non-convex type TV regularizer [8, 15], which is composed of the ℓₚ norm (TVₚ) and TV, to remove impulse noise.

Since the proposed model is non-convex and non-smooth, we adapt an iteratively reweighted algorithm, which is widely applied to the ℓₚ norm minimization arising in compressed sensing [2, 7]. We establish the convergence property of the proposed algorithm by showing that the values of the objective function are nonincreasing in iterates; see Theorem 3.1. Numerical experiments show the efficiency and stability of the proposed model in impulse noise reduction.

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

2. NCTVℓ₁ denoising model for impulse noise. Under the presence of impulse noise, an observed image $b \in \mathbb{R}^{M \times N}$ can be expressed as

$$(1) \quad b = u + \eta,$$

where $u$ is the underlying clean image and $\eta$ denotes impulse noise.

To reduce the noise effect and preserve sharp discontinuities at the same time, the following total variation regularized ℓ₁ norm (TVℓ₁) model is considered in [13]:

$$(2) \quad \min_{u \in U} c \|\nabla u\| + \|u - b\|_1,$$

where $c > 0$, $U = [0, C]^n$ with $0 < C < \infty$, and $\|\nabla u\| = \sum_{i,j} \sqrt{((u_x)_{i,j})^2 + ((u_y)_{i,j})^2}$.

As noticed in Introduction, to alleviate the drawback in smooth regions and cope with a heavy-tailed distribution in image gradients, we propose the nonconvex TV regularized ℓ₁ norm (NCTVℓ₁) model:

$$(3) \quad \min_{u \in U} c_1 \|\nabla u\|_{q_1}^{q_1} + c_2 \|\nabla^2 u\|_{q_2}^{q_2} + \|u - b\|_1,$$

where $0 < q_1, q_2 < 1$ and $c_1, c_2 \geq 0$. Here, we denote $\|\nabla u\|_{q_1} = \sum_{i,j} \|((u_x)_{i,j})^{|q_1|} = \sum_{i,j} \sqrt{((u_x)_{i,j})^2 + ((u_y)_{i,j})^2}^{q_1/2}$ and $\|\nabla^2 u\|_{q_2} = \sum_{i,j} \sqrt{((u_{xx})_{i,j})^2 + ((u_{yy})_{i,j})^2 + ((u_{xy})_{i,j})^2}^{q_2/2}$. Due to non-convexity and non-smoothness, it is a challenging minimization problem. Thus we apply the iteratively reweighted ℓ₁ algorithm, given in the next section.

---

1 Although the ℓₚ “norm” is not a formal norm for $q < 1$, it is widely referred as norm. We follow this convention.
3. Iteratively reweighted $\ell_1$ algorithm for NCTV$\ell_1$. In this section, we describe the iteratively reweighted $\ell_1$ (IR$\ell_1$) algorithm for the NCTV$\ell_1$ denoising model \cite{[6]}. For simplicity of notation, we let $b \in \mathbb{R}_n^+$ and $u \in \mathbb{R}_n^+$ be vectorized versions of the two dimensional observed image and the reconstructed image of the $n = M \times N$ size, respectively.

To overcome the non-convexity of NCTV$\ell_1$ \cite{[6]}, the following convex approximation of the proposed model is solved at each iteration of IR$\ell_1$:

$$\min_{u \in U} c_1 \sum_{i=1}^n \nu_1((\nabla u^{(k)})_i)\|\nabla u_i\| + c_2 \sum_{i=1}^n \nu_2((\nabla^2 u^{(k)})_i)\|\nabla^2 u_i\| + \|u - b\|_1,$$

where $\nu_1((\nabla u^{(k)})_i) = \frac{1}{\|((\nabla u^{(k)})_i)\| + \eta^{(k)}}$, and $\nu_2((\nabla^2 u^{(k)})_i) = \frac{1}{\|((\nabla^2 u^{(k)})_i)\| + \eta^{(k)}}$, with $\eta^{(k)} \geq 0$. The formal procedure of IR$\ell_1$ is given in Algorithm 1

**Algorithm 1 IR$\ell_1$**

- Update $u^{(k+1)}$ and $\eta^{(k+1)}$ from $u^{(k)}$ and $\eta^{(k)}$:
  - 1.: $u^{(k+1)} = \arg\min_{u \in U} c_1 \sum_{i=1}^n \nu_1((\nabla u^{(k)})_i)\|\nabla u_i\|$
    $+ c_2 \sum_{i=1}^n \nu_2((\nabla^2 u^{(k)})_i)\|\nabla^2 u_i\| + \|u - b\|_1$, if
  - 2.: $\eta^{(k+1)} = \nu \eta^{(k)}$ with $\nu \in (0, 1)$.

Theorem 3.1 establishes the convergence property of IR$\ell_1$. The proof is based on that the objective value decreases as the iteration of IR$\ell_1$ proceeds. In what follows, $\hat{u}$ is called a stationary point of the problem (3).

**Theorem 3.1.** Let $\{u^{(k)}\}$ be a sequence generated by IR$\ell_1$ and $u^*$ an accumulation point of the sequence. Assume that $((\nabla u^*)_i) \neq 0$ and $((\nabla^2 u^*)_i) \neq 0$ for all i. Then $u^*$ is a stationary point of the problem (3).

**Proof.** To clear up the proof, we define the function $f$ as follows:

$$f(u, t, s, \eta) := \sum_{i=1}^n \left(c_1 q_1 \left(\frac{\|((\nabla u)_i) + \eta\|_{t_1}^{1-q_1} + 1 - q_1 t_1^{q_1}}{q_1} \right) + c_2 q_2 \left(\frac{\|((\nabla^2 u)_i) + \eta\|_{s_2}^{1-q_2} + 1 - q_2 s_2^{q_2}}{q_2} \right)\right),$$

and consider the following equivalent problem to (4):

$$\min_{u \in U} f(u, t, s, \eta) + \|u - b\|_1,$$

with $t_i = \|((\nabla u^{(k)})_i) + \eta^{(k)}\|_{t_1}^{1-q_1}$, $s_i = \|((\nabla^2 u^{(k)})_i) + \eta^{(k)}\|_{s_2}^{1-q_2}$ for all i and $\eta = \eta^{(k)}$. By letting $t_i^{(k)} = \|((\nabla u^{(k)})_i) + \eta^{(k)}\|_{t_1}^{1-q_1}$, $s_i^{(k)} = \|((\nabla^2 u^{(k)})_i) + \eta^{(k)}\|_{s_2}^{1-q_2}$ for all i, we observe that

$$\min_{u \in U} f(u^{(k+1)}, t^{(k)}, s^{(k)}, \eta^{(k)}) + \|u^{(k+1)} - b\|_1 \leq f(u^{(k)}, t^{(k)}, s^{(k)}, \eta^{(k)}) + \|u^{(k)} - b\|_1,$$

with $t_i^{(k)} = \|((\nabla u^{(k)})_i) + \eta^{(k)}\|_{t_1}^{1-q_1}$, $s_i^{(k)} = \|((\nabla^2 u^{(k)})_i) + \eta^{(k)}\|_{s_2}^{1-q_2}$ for all i, and $\|u^{(k+1)} - b\|_1 \leq \min_{u \in U} f(u^{(k)}, t^{(k)}, s^{(k)}, \eta^{(k)}) + \|u^{(k)} - b\|_1$.\]
Next, we consider the minimization problem

\[
\min_{t, s \in (0, \infty)^n} f(u, t, s, \eta) + \|u - b\|_1,
\]

with \(u = u^{(k+1)}\) and \(\eta = \eta^{(k)}\). Then the solution is easy:

\[
t_i = \|(\nabla u^{(k+1)})_i\| + \eta^{(k)}
\]

and

\[
s_i = \|(\nabla^2 u^{(k+1)})_i\| + \eta^{(k)}, \quad \forall i.
\]

Letting \(t^{(k+1)}_i = \|(\nabla u^{(k+1)})_i\| + \eta^{(k)}\) and \(s^{(k+1)}_i = \|(\nabla^2 u^{(k+1)})_i\| + \eta^{(k)}\) for all \(i\), we have

\[
f(u^{(k+1)}, t^{(k+1)}, s^{(k+1)}, \eta^{(k)}) + \|u^{(k+1)} - b\|_1 \leq f(u^{(k+1)}, t^{(k)}, s^{(k)}, \eta^{(k)}) + \|u^{(k+1)} - b\|_1
\]

and

\[
f(u^{(k+1)}, t^{(k+1)}, s^{(k+1)}, \eta^{(k)}) + \|u^{(k+1)} - b\|_1
\]

\[
= \sum_{i=1}^{n} c_1 q_1 \frac{\|(\nabla u^{(k+1)})_i\| + \eta^{(k)}}{t^{(k+1)}_i} + \frac{1 - a_2}{q_2} (t^{(k+1)}_i q_2)
\]

\[
+ \sum_{i=1}^{n} c_2 q_2 \frac{\|(\nabla^2 u^{(k+1)})_i\| + \eta^{(k)}}{s^{(k+1)}_i} + \frac{1 - a_2}{q_2} (s^{(k+1)}_i q_2) + \|u^{(k+1)} - b\|_1
\]

\[
= \sum_{i=1}^{n} \left( c_1 \left( \|(\nabla u^{(k+1)})_i\| + \eta^{(k)} \right) q_1 + c_2 \left( \|(\nabla^2 u^{(k+1)})_i\| + \eta^{(k)} \right) q_2 \right) + \|u^{(k+1)} - b\|_1.
\]

If we choose \(\eta^{(k+1)} = \nu \eta^{(k)}\) with \(\nu < 1\), then, for all \(i\),

\[
(\|(\nabla u^{(k+1)})_i\| + \eta^{(k+1)} q_1) \leq (\|(\nabla u^{(k+1)})_i\| + \eta^{(k)} q_1),
\]

\[
(\|(\nabla^2 u^{(k+1)})_i\| + \eta^{(k+1)} q_2) \leq (\|(\nabla^2 u^{(k+1)})_i\| + \eta^{(k)} q_2).
\]

The above inequalities and the equation (10) imply that

\[
f(u^{(k+1)}, t^{(k+1)}, s^{(k+1)}, \eta^{(k+1)}) + \|u^{(k+1)} - b\|_1
\]

\[
\leq f(u^{(k+1)}, t^{(k+1)}, s^{(k+1)}, \eta^{(k)}) + \|u^{(k+1)} - b\|_1.
\]

The inequalities (10), (8), and (10) imply that the sequence \(\{f(u^{(k)}, t^{(k)}, s^{(k)}, \eta^{(k)}) + \|u^{(k)} - b\|_1\}\) is non-increasing. By a similar procedure as in (10), we get

\[
f(u^{(k)}, t^{(k)}, s^{(k)}, \eta^{(k)}) + \|u^{(k)} - b\|_1
\]

\[
= \sum_{i=1}^{n} \left( c_1 \left( \|(\nabla u^{(k)}_i\| + \eta^{(k)} \right) q_1 + c_2 \left( \|(\nabla^2 u^{(k)}_i\| + \eta^{(k)} \right) q_2 \right) + \|u^{(k)} - b\|_1.
\]

Since \(u^*\) is an accumulation point of \(\{u^{(k)}\}\), there exists a subsequence \(K\) such that \(\{u^{(k)}\}_K \to u^*\). And we let \(t^*_i = \|(\nabla u^*)_i\|\) and \(s^*_i = \|(\nabla^2 u^*)_i\|\) for all \(i\). From the continuity of \(f\) and \(\|\cdot\|_1\) with \(\{u^{(k)}\}_K \to u^*\) and \(\eta^{(k)} \to 0\),

\[
\{f(u^{(k)}, t^{(k)}, s^{(k)}, \eta^{(k)}) + \|u^{(k)} - b\|_1\}_K \to c_1 \|\nabla u^*\|_V + c_2 \|\nabla^2 u^*\|_W + \|u^* - b\|_1.
\]

This convergence property with the monotonicity of \(f(u^{(k)}, t^{(k)}, s^{(k)}, \eta^{(k)}) + \|u^{(k)} - b\|_1\) implies that the function value \(f(u^{(k)}, t^{(k)}, s^{(k)}, \eta^{(k)}) + \|u^{(k)} - b\|_1\) converges to \(c_1 \|\nabla u^*\|_V + c_2 \|\nabla^2 u^*\|_W + \|u^* - b\|_1\). Also, from the inequality (10), we further have

\[
f(u^{(k+1)}, t^{(k+1)}, s^{(k+1)}, \eta^{(k+1)}) + \|u^{(k+1)} - b\|_1 \to c_1 \|\nabla u^*\|_V + c_2 \|\nabla^2 u^*\|_W + \|u^* - b\|_1.
\]
In addition, by the equivalence between (11) and (13) with $t_i = \|\nabla (u^{(k)})_i\| + \eta^{(k)}$, $s_i = \|\nabla^2 (u^{(k)})_i\| + \eta^{(k)}$ for all $i$ and $\eta = \eta^{(k)}$

$$f(u, t^{(k)}, s^{(k)}, \eta^{(k)}) + \|u - b\|_1 \geq f(u^{(k+1)}, t^{(k)}, s^{(k)}, \eta^{(k)}) + \|u^{(k+1)} - b\|_1, \quad \forall u \in U.$$ 

Taking limits on both sides of this inequality with $K$ and (11) imply that

$$f(u, t^*, s^*, 0) + \|u - b\|_1 \geq c_1 \|\nabla u^*\|_q^2 + c_2 \|\nabla^2 u^*\|_{q^2} + \|u^* - b\|_1 \quad \forall u \in U.$$ 

By using $t^*_i = \|\nabla (u^*)_i\|$ and $s^*_i = \|\nabla^2 (u^*)_i\|$ for all $i$, we have

$$f(u^*, t^*, s^*, 0) + \|u^* - b\|_1 = c_1 \|\nabla u^*\|_q^2 + c_2 \|\nabla^2 u^*\|_{q^2} + \|u^* - b\|_1.$$ 

Thus $u^* \in \arg \min_{u \in U} f(u, t^*, s^*, 0) + \|u - b\|_1$, and so $u^*$ is a stationary point of $\min_{u \in U} f(u, t^*, s^*, 0) + \|u - b\|_1$, that is,

$$0 \in c_1 q_1 (t^*_i)^{q_1-1} \partial \|\nabla (u^*)_i\| + c_2 q_2 (s^*_i)^{q_2-1} \partial \|\nabla^2 (u^*)_i\| + \partial \|u^* - b\|_1 + (N_U(u^*))_i.$$ 

The relations $t^*_i = \|\nabla (u^*)_i\|$ and $s^*_i = \|\nabla^2 (u^*)_i\|$ for all $i$ imply that

$$0 \in c_1 q_1 \|\nabla^2 u^*_i\|^{q_1-1} \partial \|\nabla (u^*)_i\| + c_2 q_2 \|\nabla^2 u^*_i\|^{q_2-1} \partial \|\nabla^2 (u^*)_i\| + \partial \|u^* - b\|_1 + (N_U(u^*))_i,$$ 

for all $i$.

Therefore, $u^*$ is a stationary point of the problem (33).

Although the problem (11) is convex, we further need an efficient algorithm, because of the nonlinearity and non-smoothness of $\|\nabla (u)_i\|$, $\|\nabla^2 (u)_i\|$, and $\|u - b\|_1$.

For this purpose, we apply linearized ADMM (LADMM); we introduce new variables $v = \nabla u$, $w = \nabla^2 u$, $z = u - b$ and then reformulate the subproblem (11) as a linearly constrained minimization problem:

$$\min_{u \in U, v, w, z} c_1 \sum_{i=1}^n \nu_1 (\|\nabla (u^{(k)})_i\|_q) \|v_i\| + c_2 \sum_{i=1}^n \nu_2 (\|\nabla^2 (u^{(k)})_i\|_q) \|w_i\| + \|z\|_1$$

subject to $v = \nabla u$, $w = \nabla^2 u$, $z = u - b$.

To describe the algorithmic framework of LADMM, we introduce the corresponding augmented Lagrangian function and linearized augmented Lagrangian function:

$$\mathcal{L}(v, w, z, u, \lambda, \tau, \xi; u^{(k)}) = c_1 \sum_{i=1}^n \nu_1 (\|\nabla (u^{(k)})_i\|_q) \|v_i\| - \langle \lambda, v - \nabla u \rangle + \frac{\beta}{2} \|v - \nabla u\|^2$$

$$+ c_2 \sum_{i=1}^n \nu_2 (\|\nabla^2 (u^{(k)})_i\|_q) \|w_i\| - \langle \tau, w - \nabla^2 u \rangle + \frac{\gamma}{2} \|w - \nabla^2 u\|^2$$

$$+ \|z\|_1 \times (\xi, z - (u - b)) + \frac{\delta}{2} \|z - (u - b)\|^2.$$ (13)

$$\mathcal{L}(v, w, z, u, \lambda, \tau, \xi; u^{(k)})$$

$$= c_1 \sum_{i=1}^n \nu_1 (\|\nabla (u^{(k)})_i\|_q) \|v_i\| + c_2 \sum_{i=1}^n \nu_2 (\|\nabla^2 (u^{(k)})_i\|_q) \|w_i\| + \|z\|_1$$

$$+ \langle \Psi(u), u - \hat{u} \rangle + \frac{\rho}{2} \|u - \hat{u}\|^2,$$ (14)

where $\Psi(u) = \nabla^T (\lambda + \beta (\nabla u - v)) + (\nabla^2)^T (\tau + \gamma (\nabla^2 \hat{u} - w)) + \xi + \delta (\hat{u} - b - z)$ and $\rho$ is a constant satisfying $\rho > \beta \nabla^T \nabla + \gamma (\nabla^2)^T \nabla^2 + \delta I$. 

Inverse Problems and Imaging Volume 11, No. 4 (2017), 689–702
Finally, we describe LADMM using the augmented Lagrangian function (13) and the linearized augmented Lagrangian function (14) to solve the linearly constrained reformulation (12).

\[
\begin{align*}
    v^{(\ell+1)} &\leftarrow \arg\min_v L(v, w^{(\ell)}, z^{(\ell)}, u^{(\ell)}, \lambda^{(\ell)}, \tau^{(\ell)}, \xi^{(\ell)}; u^{(k)}) \\
    w^{(\ell+1)} &\leftarrow \arg\min_w L(v^{(\ell+1)}, w, z^{(\ell)}, u^{(\ell)}, \lambda^{(\ell)}, \tau^{(\ell)}, \xi^{(\ell)}; u^{(k)}) \\
    z^{(\ell+1)} &\leftarrow \min_z L(v^{(\ell+1)}, w^{(\ell+1)}, z, u^{(\ell)}, \lambda^{(\ell)}, \tau^{(\ell)}, \xi^{(\ell)}; u^{(k)}) \\
    u^{(\ell+1)} &\leftarrow \arg\min_u L(v^{(\ell+1)}, w^{(\ell+1)}, z^{(\ell+1)}, u, \lambda^{(\ell)}, \tau^{(\ell)}, \xi^{(\ell)}; u^{(\ell)}, u^{(k)}) \\
    \lambda^{(\ell+1)} &\leftarrow \lambda^{(\ell)} - \beta(u^{(\ell+1)} - \nabla u^{(\ell+1)}) \\
    \tau^{(\ell+1)} &\leftarrow \tau^{(\ell)} - \gamma(w^{(\ell+1)} - \nabla^2 u^{(\ell+1)}) \\
    \xi^{(\ell+1)} &\leftarrow \xi^{(\ell)} - \delta(z^{(\ell+1)} - u^{(\ell+1)} + b)
\end{align*}
\]

The first four minimizations are easy to solve, since we have the following closed form solutions.

\[
\begin{align*}
    v_i^{(\ell+1)} &= \text{shrink} \left( \nabla u_i^{(\ell)} + \frac{\lambda_i^{(\ell)}}{\beta}, \frac{c_1 \nu_1((\nabla u^{(k)})_i)}{\beta} \right) \\
    &= \max \left\{ \left\| \nabla u_i^{(\ell)} + \frac{\lambda_i^{(\ell)}}{\beta} \right\| - \frac{c_1 \nu_1((\nabla u^{(k)})_i)}{\beta}, 0 \right\} \left\| \nabla u_i^{(\ell)} + \frac{\lambda_i^{(\ell)}}{\beta} \right\| \\
    (i = 1, \ldots, n) \\
    w_i^{(\ell+1)} &= \text{shrink} \left( \nabla^2 u_i^{(\ell)} + \frac{\tau_i^{(\ell)}}{\gamma}, \frac{c_2 \nu_2((\nabla^2 u^{(k)})_i)}{\gamma} \right) \\
    &= \max \left\{ \left\| \nabla^2 u_i^{(\ell)} + \frac{\tau_i^{(\ell)}}{\gamma} \right\| - \frac{c_2 \nu_2((\nabla^2 u^{(k)})_i)}{\gamma}, 0 \right\} \left\| \nabla^2 u_i^{(\ell)} + \frac{\tau_i^{(\ell)}}{\gamma} \right\| \\
    (i = 1, \ldots, n) \\
    z_i^{(\ell+1)} &= \min \left( u_i^{(\ell)} - b_i + \frac{\xi_i^{(\ell)}}{\delta}, \frac{1}{\delta} \right) \\
    &= \max \left\{ \left\| u_i^{(\ell)} - b_i + \frac{\xi_i^{(\ell)}}{\delta} \right\| - \frac{1}{\delta}, 0 \right\} \left\| u_i^{(\ell)} - b_i + \frac{\xi_i^{(\ell)}}{\delta} \right\| \\
    (i = 1, \ldots, n) \\
    u_i^{(\ell+1)} &= \text{median} \left\{ 0, \left( u_i^{(\ell)} - \frac{1}{\rho} \Psi(u^{(\ell)}) \right) \right\} C \quad (i = 1, \ldots, n),
\end{align*}
\]

where the shrinkage operator is adopted from [17] along with the convention \(0 \cdot (0/0) = 0\). The convergence properties of LADMM for linearizing the quadratic penalty term can be found in [11].

4. Algorithm for TV\(f_1\). To show the soundness of the proposed model, we compare it with the TV\(f_1\) model. However, because of the non-smoothness property of \(\|\nabla u\|\) and \(\|u - b\|_1\) in [2], we apply LADMM, similar to the subproblem (12). By using new variables \(v = \nabla u\) and \(z = u - b\), the following linearly constrained
reformulation of the problem (2) is considered.

\[
\min_{u \in U, v, z} \|z\|_1 + c\|v\|
\]

subject to \(v = \nabla u, \ z = u - b\).

Then the corresponding augmented Lagrangian function and linearized augmented Lagrangian function are as follows:

\[
\mathcal{L}(v, z, u, \lambda, \xi) = c\|v\| - \langle \lambda, v - \nabla u \rangle + \frac{\beta}{2}\|v - \nabla u\|^2
\]

\[
+ \|z\|_1 - \langle \xi, z - (u - b) \rangle + \frac{\delta}{2}\|z - (u - b)\|^2.
\]

\[
\mathcal{L}(v, z, u, \lambda, \xi; \bar{u}) = c\|v\| + \|z\|_1 + \langle \Psi(\bar{u}), u - \bar{u} \rangle + \frac{\rho}{2}\|u - \bar{u}\|^2,
\]

where \(\Psi(\bar{u}) = \nabla^T(\lambda + \beta(\nabla \bar{u} - v)) + \xi + \delta(\bar{u} - b - z)\) and \(\rho\) is a constant satisfying \(\rho I \geq \beta \nabla^T \nabla + \delta I\).

The algorithmic framework to solve (19) is comparable to (13).

\[
\begin{align*}
&v^{(\ell+1)} \leftarrow \arg \min_v \mathcal{L}(v, z^{(\ell)}, u^{(\ell)}, \lambda^{(\ell)}, \xi^{(\ell)}) \\
&z^{(\ell+1)} \leftarrow \arg \min_z \mathcal{L}(v^{(\ell+1)}, z, u^{(\ell)}, \lambda^{(\ell)}, \xi^{(\ell)}) \\
&u^{(\ell+1)} \leftarrow \arg \min_{u \in U} \mathcal{L}(v^{(\ell+1)}, z^{(\ell+1)}, u, \lambda^{(\ell)}, \xi^{(\ell)}; u^{(\ell)}) \\
&\lambda^{(\ell+1)} \leftarrow \lambda^{(\ell)} - \beta(v^{(\ell+1)} - \nabla u^{(\ell+1)}) \\
&\xi^{(\ell+1)} \leftarrow \xi^{(\ell)} - \delta(z^{(\ell+1)} - u^{(\ell+1)} + b)
\end{align*}
\]

The first three steps have the following closed form solutions.

\[
v_i^{(\ell+1)} = \text{shrink} \left( \nabla u_i^{(\ell)} + \frac{\lambda_i^{(\ell)}}{\beta}, \frac{c}{\beta} \right)
\]

\[
= \max \left\{ \left\| \nabla u_i^{(\ell)} + \frac{\lambda_i^{(\ell)}}{\beta} \right\| - \frac{c}{\beta} ; 0 \right\} \frac{\nabla u_i^{(\ell)} + \frac{\lambda_i^{(\ell)}}{\beta}}{\left\| \nabla u_i^{(\ell)} + \frac{\lambda_i^{(\ell)}}{\beta} \right\|} \quad (i = 1, \ldots, n)
\]

\[
z_i^{(\ell+1)} = \text{shrink} \left( u_i^{(\ell)} - b + \frac{\xi_i^{(\ell)}}{\delta}, \frac{1}{\delta} \right)
\]

\[
= \max \left\{ \left\| u_i^{(\ell)} - b + \frac{\xi_i^{(\ell)}}{\delta} \right\| - \frac{1}{\delta} ; 0 \right\} \frac{u_i^{(\ell)} - b + \frac{\xi_i^{(\ell)}}{\delta}}{\left\| u_i^{(\ell)} - b + \frac{\xi_i^{(\ell)}}{\delta} \right\|} \quad (i = 1, \ldots, n)
\]

\[
u_i^{(\ell+1)} = \text{median} \left\{ 0, \left( u_i^{(\ell)} - \frac{1}{\rho} \Psi(u_i^{(\ell)}) \right) \right\}_i \quad (i = 1, \ldots, n).
\]

5. **Numerical experiments.** In this section, we report numerical results on the proposed impulse noise removal model. The performance of NCTV\(\ell_1\) (3) using IR\(\ell_1\) is compared to those of TV\(\ell_1\) (2), and the standard median filter. To show the efficiency of the model, we further compare it with the penalized model which is recently proposed in (12) and called AMM. This model uses the anisotropic TV and is solved by alternating minimization method. When the penalty parameters go to infinity, this model approaches (2) with replacing TV by the anisotropic TV.
For experiments, salt-and-pepper noise and random valued impulse noise (RVIN) are generated by the `imnoise` function in Matlab. All algorithms are implemented in Matlab and experiments are performed on a laptop with Intel Core i7 3.4 GHz PC with 32 GB memory, running Windows 7.

We use the relative difference for the stopping criterion of the proposed NCTV\(\ell_1\)
\[
||u^{(k+1)} - u^{(k)}|| \leq C_t||u^{(k)}||,
\]
where \(C_t > 0\) is a stopping tolerance and set to \(C_t = 5 \times 10^{-3}\). For the inner loop, we choose the same criterion with \(C_t = 1 \times 10^{-3}\) or we stop after 20 iterations.

For quality evaluation, SNR and PSNR are employed, defined by
\[
SNR = 10 \log_{10} \frac{P_s}{P_n},
\]
\[
PSNR = 10 \log_{10} \frac{255^2}{P_n},
\]
where \(P_s = \frac{1}{n}||u||^2\) and \(P_n = \frac{1}{n}||u - \hat{u}||^2\) with the original and reconstructed images \(u\) and \(\hat{u}\), respectively.

To see the benefit of the NCTV\(\ell_1\) model clearly, the proposed and two comparison models, the median filter and TV\(\ell_1\) are applied to an one-dimensional piecewise smooth image under salt-and-pepper noise with various levels. The results are given in Figures 1 and 2. The parameters \(q_1 = q_2 = 0.9, c_1 = c_2 = 2.5, \beta = 1.5, \gamma = 0.01, \delta = 2.5, \text{ and } \rho = 8.0\) are chosen for NCTV\(\ell_1\), and \(c = 2.5, \beta = 0.9, \delta = 2.5, \text{ and } \rho = 8.0\) are for TV\(\ell_1\) regardless of the noise levels. The NCTV\(\ell_1\) model produces robust reconstruction in various noise levels, while two existing models suffer from the undesirable non-removed spikes and the number of those spikes increases as the
noise level increases. The proposed model outperforms two models over $8 - 10\text{dB}$ for noise level 30% and about 10dB for 50%, in terms of SNR.

![Figure 3. Reconstructed subimages of the “lori” image (256x256) under the noise level 20%. The first row is salt-and-pepper noise, and the second row is RVIN. PSNRs: (a) 12.00dB, (b) 29.55dB, (c) 29.63dB, (d) 31.24dB, (e) 13.35dB, (f) 29.31dB, (g) 29.59dB, (h) 30.72dB.](image)

Those models are applied to the four test images “lori”, “boat”, “cameraman”, and “lighthouse” under salt-and-pepper noise and RVIN with various levels, given in Figures 3 - 6. The model parameters $q_1$, $q_2$, $c_1$, and $c_2$ are chosen by the best average PSNR on the chosen 3 training images (“babara”, “boat”, and “goldhill” of the size 512x512). The algorithm parameters $\beta$, $\gamma$, $\rho$, and $\delta$ are heuristically chosen and fixed for all experiments. The parameters are set to $q_1 = q_2 = 0.9$, $c_1 = 0.3$, $c_2 = 0.4$, $\beta = 0.9$, $\gamma = 0.01$, $\delta = 2.5$, and $\rho = 8.0$ for NCTV1, regardless of noise types and levels. For TV1, the model parameter $c$ is differently set to noise levels ($c = 0.6$ for 20%, $c = 0.7$ for 30%, $c = 0.8$ for 40%) for better performances and the others are fixed to $\beta = 0.9$, $\delta = 2.5$, and $\rho = 8.0$. For the parameters of AMM, we refer to [12].

In Figure 3, TV1 yields non-smooth edges around hat and face line in the recovered image, while NCTV1 preserves smooth edges. For AMM, some unevenness is observed in the plain region. Figure 4 shows the reconstructed subimages of the “boat” image under the noise level 30%. The edges of poles on the boat and border lines of the boat body are non-smooth and rough in the reconstructions by TV1. In Figure 5, TV1 and AMM provide jagged edges at the camera body and tripod regions. In contrast, the proposed NCTV1 shows well-preserved and smooth edges and consistent visual results in various noise levels. Figure 6 illustrates the staircase artifacts in the house and bottom regions of the recovered image by TV1. The sky region is unevenly recovered in the AMM case.
Tables [1][3] show the performance comparison among three methods. They include PSNR under various noise types and levels. In addition, we report the computational time, the number of iterations, and, if applicable, the total number of inner iterations. Overall, NCTV\(\ell_1\) improves PSNR by 3–5dB comparing TV\(\ell_1\). Considering AMM, PSNR is 2–3dB higher. However, due to the complexity introduced by \(\ell_q\) norms and the Hessian term, NCTV\(\ell_1\) is computationally slower than the TV\(\ell_1\) model.

6. Conclusions. In this paper, we propose a non-convex variational model for impulse noise removal. The new model uses a non-convex regularizer which is composed of \(\ell_q\) norm and TV. To bypass non-convexity and non-smoothness, the iteratively reweighted \(\ell_1\) algorithm is applied. We show the convergence property of the proposed algorithm and demonstrate the effectiveness of the proposed model via numerical simulations.
Figure 5. Reconstructed subimages of the “cameraman” image (512 × 512) under the noise level 40%. The first row is salt-and-pepper noise, and the second row is RVIN. PSNRs: (a) 9.06dB, (b) 28.44dB, (c) 28.96dB, (d) 32.65dB, (e) 10.43dB, (f) 28.68dB, (g) 28.62dB, (h) 32.09dB.

Figure 6. Reconstructed subimages of the “lighthouse” image (512 × 512) under the noise level 40%. The first row is salt-and-pepper noise, and the second row is RVIN. PSNR: (a) 9.48dB, (b) 24.62dB, (c) 24.91dB, (d) 25.54dB, (e) 11.07dB, (f) 24.78dB, (g) 24.95dB, (h) 25.47dB.
Table 1. The performance comparison under the noise level 20%.

| image size | Noise Type | PSNR | Method | Iter (Total inners) | time | PSNR |
|------------|------------|------|--------|---------------------|------|------|
| lori 256x256 | salt-and-pepper (20%) | 12.00 | TV$\ell_1$ | 57 | 0.21 | 29.55 |
| | | | AMM | 51(255) | 0.60 | 29.63 |
| | | | NCTV$\ell_1$ | 5(100) | 0.88 | 31.24 |
| | RVIN (20%) | 13.35 | TV$\ell_1$ | 56 | 0.21 | 29.31 |
| | | | AMM | 50(250) | 0.59 | 29.59 |
| | | | NCTV$\ell_1$ | 5(100) | 0.86 | 30.72 |
| boat 512x512 | salt-and-pepper (20%) | 12.31 | TV$\ell_1$ | 49 | 1.27 | 29.88 |
| | | | AMM | 44(220) | 3.75 | 30.31 |
| | | | NCTV$\ell_1$ | 4(74) | 4.80 | 33.74 |
| | RVIN (20%) | 13.81 | TV$\ell_1$ | 48 | 1.17 | 30.39 |
| | | | AMM | 43(215) | 3.54 | 30.42 |
| | | | NCTV$\ell_1$ | 4(76) | 4.89 | 33.65 |
| cameraman 512x512 | salt-and-pepper (20%) | 10.60 | TV$\ell_1$ | 46 | 1.14 | 32.09 |
| | | | AMM | 47(235) | 3.90 | 33.05 |
| | | | NCTV$\ell_1$ | 4(70) | 4.61 | 36.92 |
| | RVIN (20%) | 13.43 | TV$\ell_1$ | 54 | 1.18 | 32.05 |
| | | | AMM | 45(225) | 3.76 | 32.65 |
| | | | NCTV$\ell_1$ | 4(73) | 4.83 | 37.99 |
| lighthouse 512x512 | salt-and-pepper (20%) | 9.48 | TV$\ell_1$ | 55 | 1.44 | 27.43 |
| | | | AMM | 52(260) | 4.52 | 27.00 |
| | | | NCTV$\ell_1$ | 5(100) | 6.63 | 27.85 |
| | RVIN (20%) | 11.07 | TV$\ell_1$ | 54 | 1.34 | 27.52 |
| | | | AMM | 50(250) | 4.17 | 26.96 |
| | | | NCTV$\ell_1$ | 5(100) | 6.56 | 27.80 |

Table 2. The performance comparison under the noise level 30%.

| image size | Noise Type | PSNR | Method | Iter (Total inners) | time | PSNR |
|------------|------------|------|--------|---------------------|------|------|
| lori 256x256 | salt-and-pepper (30%) | 10.23 | TV$\ell_1$ | 65 | 0.25 | 27.51 |
| | | | AMM | 58(290) | 0.69 | 28.04 |
| | | | NCTV$\ell_1$ | 6(120) | 1.05 | 29.57 |
| | RVIN (30%) | 11.57 | TV$\ell_1$ | 65 | 0.24 | 28.10 |
| | | | AMM | 57(285) | 0.69 | 28.35 |
| | | | NCTV$\ell_1$ | 6(120) | 1.07 | 29.49 |
| boat 512x512 | salt-and-pepper (30%) | 10.55 | TV$\ell_1$ | 58 | 1.48 | 28.49 |
| | | | AMM | 54(270) | 4.87 | 28.84 |
| | | | NCTV$\ell_1$ | 5(93) | 6.21 | 31.44 |
| | RVIN (30%) | 12.04 | TV$\ell_1$ | 58 | 1.42 | 27.71 |
| | | | AMM | 50(250) | 4.21 | 28.87 |
| | | | NCTV$\ell_1$ | 5(93) | 5.93 | 31.15 |
| cameraman 512x512 | salt-and-pepper (30%) | 10.29 | TV$\ell_1$ | 55 | 1.34 | 30.41 |
| | | | AMM | 57(285) | 4.83 | 31.00 |
| | | | NCTV$\ell_1$ | 5(91) | 5.92 | 35.57 |
| | RVIN (30%) | 11.7 | TV$\ell_1$ | 54 | 1.33 | 30.65 |
| | | | AMM | 52(260) | 4.39 | 30.72 |
| | | | NCTV$\ell_1$ | 5(92) | 5.94 | 35.22 |
| lighthouse 512x512 | salt-and-pepper (30%) | 10.75 | TV$\ell_1$ | 64 | 1.65 | 25.98 |
| | | | AMM | 61(305) | 5.33 | 25.98 |
| | | | NCTV$\ell_1$ | 6(120) | 7.98 | 26.64 |
| | RVIN (30%) | 12.31 | TV$\ell_1$ | 64 | 1.6 | 25.96 |
| | | | AMM | 57(285) | 4.75 | 25.92 |
| | | | NCTV$\ell_1$ | 6(120) | 7.87 | 26.53 |
Table 3. The performance comparison under the noise level 40%.

| Image size | Noise Type       | PSNR | Method | Iter (Total inners) | time | PSNR  |
|------------|------------------|------|--------|---------------------|------|-------|
| lori       | salt-and-pepper  | 8.99 | TV$_{\ell_1}$ | 80 | 0.30 | 26.8 |
| 256x256    | AMM              | 76(380) | 0.92 | 27.23 |
|           | NCTV$_{\ell_1}$ | 7(140) | 1.21 | 28.67 |
| boat       | salt-and-pepper  | 9.33 | TV$_{\ell_1}$ | 70 | 1.75 | 27.22 |
| 512x512    | AMM              | 68(340) | 6.01 | 27.54 |
|           | NCTV$_{\ell_1}$ | 5(100) | 6.45 | 29.31 |
| cameraman  | salt-and-pepper  | 9.06 | TV$_{\ell_1}$ | 67 | 1.61 | 28.44 |
| 512x512    | AMM              | 69(345) | 5.72 | 28.96 |
|           | NCTV$_{\ell_1}$ | 5(100) | 6.48 | 32.65 |
| lighthouse | salt-and-pepper  | 9.48 | TV$_{\ell_1}$ | 78 | 2.02 | 24.62 |
| 512x512    | AMM              | 79(395) | 6.79 | 24.91 |
|           | NCTV$_{\ell_1}$ | 7(140) | 9.27 | 25.54 |

REFERENCES

[1] A. C. Bovik, *Handbook of Image and Video Processing*, Academic Press, Inc., Orlando, FL, USA, 2005.

[2] E. J. Candès, M. B. Wakin and S. P. Boyd, Enhancing sparsity by reweighted $\ell_1$ minimization, *J. Fourier Anal. Appl.*, 14 (2008), 877–905.

[3] A. Chambolle and T. Pock, A first-order primal-dual algorithm for convex problems with applications to imaging, *J. Math. Imaging Vision*, 40 (2011), 120–145.

[4] R. Chan, C.-W. Ho and M. Nikolova, Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization, *IEEE Transactions on Image Processing*, 14 (2005), 1479–1485.

[5] T. Chan, A. Marquina and P. Mulet, High-order total variation-based image restoration *SIAM Journal on Scientific Computing*, 22 (2000), 503–516.

[6] T. Chen and H. R. Wu, Space variant median filters for the restoration of impulse noise corrupted images, *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 48 (2001), 784–789.

[7] X. Chen and W. Zhou, Convergence of the reweighted $\ell_1$ minimization algorithm for $\ell_2$-$\ell_p$ minimization, *Comput. Optim. Appl.*, 59 (2014), 47–61.

[8] M. Hintermüller and T. Wu, Nonconvex TV$^k$-models in image restoration: Analysis and a trust-region regularization-based superlinearly convergent solver, *SIAM Journal on Imaging Sciences*, 6 (2013), 1385–1415.

[9] H. Hwang and R. A. Haddad, Adaptive median filters: new algorithms and results *IEEE Transactions on Image Processing*, 4 (1995), 499–502.

[10] D. Krishnan and R. Fergus, Fast image deconvolution using hyper-laplacian priors, in *Advances in Neural Information Processing Systems 22* (eds. Y. Bengio, D. Schuurmans, J. Lafferty, C. Williams and A. Culotta), Curran Associates, Inc., 2009, 1033–1041.

[11] Z. Lin, R. Liu and Z. Su, Linearized alternating direction method with adaptive penalty for low-rank representation, in *Advances in Neural Information Processing Systems* (eds.
[12] X. Liu, Alternating minimization method for image restoration corrupted by impulse noise, *Multimedia Tools and Applications*, 76 (2017), 12505–12516.

[13] 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*, 12 (2003), 1579–1590.

[14] M. Nikolova, A variational approach to remove outliers and impulse noise, *J. Math. Imaging Vision*, 20 (2004), 99–120, Special issue on mathematics and image analysis.

[15] S. Oh, H. Woo, S. Yun and M. Kang, Non-convex hybrid total variation for image denoising, *Journal of Visual Communication and Image Representation*, 24 (2013), 332–344.

[16] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer-Verlag, 1998.

[17] Y. Wang, J. Yang, W. Yin and Y. Zhang, A new alternating minimization algorithm for total variation image reconstruction, *SIAM Journal on Imaging Sciences*, 1 (2008), 248–272.

[18] J. Yang, Y. Zhang and W. Yin, An efficient TVL1 algorithm for deblurring multichannel images corrupted by impulse noise, *SIAM J. Sci. Comput.*, 31 (2009), 2842–2865.

Received May 2016; revised April 2017.

E-mail address: yoonmojung@skku.edu
E-mail address: iamlogin@yonsei.ac.kr
E-mail address: yswmathedu@skku.edu