Research Article

Mixed Higher Order Variational Model for Image Recovery

Pengfei Liu,[1] Liang Xiao,[1] and Liancun Xiu[2]

[1] *School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China*
[2] *Nanjing Center of China Geological Survey, Nanjing, Jiangsu 210016, China*

Correspondence should be addressed to Liang Xiao; xiaoliang@mail.njust.edu.cn

A novel mixed higher order regularizer involving the first and second degree image derivatives is proposed in this paper. Using spectral decomposition, we reformulate the new regularizer as a weighted L_1-L_2 mixed norm of image derivatives. Due to the equivalent formulation of the proposed regularizer, an efficient fast projected gradient algorithm combined with monotone fast iterative shrinkage thresholding, called, FPG-MFISTA, is designed to solve the resulting variational image recovery problems under majorization-minimization framework. Finally, we demonstrate the effectiveness of the proposed regularization scheme by the experimental comparisons with total variation (TV) scheme, nonlocal TV scheme, and current second degree methods. Specifically, the proposed approach achieves better results than related state-of-the-art methods in terms of peak signal to ratio (PSNR) and restoration quality.

1. Introduction

During image acquisition, the observed images are mainly degraded by blurring and random noise, which are tightly linked to imaging system. In order to attenuate the undesired effects, image recovery plays an integral part of modern imaging sciences in such areas, including medical imaging, microscopy, astronomical imaging, remote sensing, and photography [1–5].

Image recovery which amounts to estimating a desired image from its degraded measurements is mathematically an inverse problem [6]. To cope with the ill posedness of image recovery, the standard approach is to use image a priori to constrain the solutions so that many variational regularization approaches have been successfully proposed for image recovery. As we know, one of the most popular regularizers is the total variation (TV) seminorm which was firstly introduced by Rudin et al. for image denoising [7] and turned out to be very efficient for preserving image edges. Thus, TV has been widely and effectively used in many other image applications [8–11] and provides state-of-the-art result. However, TV-based reconstruction methods often give rise to the undesired staircase effect, since TV favors piecewise constant solutions. Thus, the main challenge

for image recovery is to reduce the staircase effect while preserving edges sharpness as far as possible. As is known to us, piecewise-vanishing second degree derivatives yield piecewise-linear solutions which better fit smooth intensity changes. To alleviate the staircase effect, several regularizers involving second degree image derivatives were introduced mainly for image denoising over past few years [12–15] and provide better performance than TV. Recently, many novel higher order regularization methods which constitute valid extensions of TV are effectively introduced for more image processing tasks [16–18] so that they minimize the staircase effect and provide state-of-the-art results. However, the second-order regularizers [16–18] which are the mixed norms of second degree directional image derivatives tend to suppress larger second degree directional derivatives thus leading to blurring across the sharp edges. Thus, the reconstructed images with better preserved edges and less staircase effect are more desirable. To this end, as reported in [19–23], the regularizers that combine first and second degree image derivatives have been proposed so as to provide better edge preservation while reducing the staircase effect in smooth regions of image; thus, we are motivated to investigate the regularizers involving derivatives of different degrees in this paper.

In this paper, our main contributions are twofolds. First, we reinterpret the isotropic second degree total variation (ISDTV) regularizer [16] which is the L_1-L_2 mixed norms of the second degree directional derivatives and give its equivalent formulation. We then propose a novel mixed higher order regularizer involving the first and second degree derivatives of the image and reformulate the new regularizer as a weighted L_1-L_2 mixed norm of image derivatives in spectral decomposition framework. The new regularizer which inherits the desirable properties of TV turns out to be well suited for image recovery. Second, for the proposed variational image recovery model, a fast projected gradient algorithm combined with monotone fast iterative shrinkage thresholding (FPG-MFISTA) which is derived from majorization-minimization (MM) framework has been utilized to solve the minimization problem under additional convex constraints. Finally, we compare our proposed schemes with TV [24], nonlocal TV [25], and current second degree methods [18, 26], which currently provide state-of-the-art results.

The rest of the paper is organized as follows. In Section 2, we first present the mathematic formulation of image recovery problem and review the ISDTV regularizer. We then propose a novel regularizer at the end of Section 2. In Section 3, we propose our variational regularization-based image recovery model and give a detailed description of the corresponding minimization algorithm. In Section 4, we assess the performance of our approach on both natural and biomedical images and provide systematic comparisons with TV, nonlocal TV, and current second degree methods. Finally, we conclude the paper in Section 5.

2. Mixed Higher Order Total Variation for Image Recovery

2.1. Regularization for Image Recovery Problem. We first consider the recovery of a continuously differentiable image f : $\Omega \to \mathfrak{R}$ from its degraded observation g, where $\Omega \subset \mathfrak{R}^2$ is the spatial support of the image. The most commonly used image observation model can be linearly formulated as

$$g = Af + e, \tag{1}$$

where A is a linear blur operator and e is Gaussian white noise with standard deviation σ. As we know, the recovery of f from its observation g is an ill-posed problem, due to the fact that the blur operator A is ill-conditioned. To stabilize the recovery process, some available a priori knowledge about the desired image is utilized. Thus, a common variational regularization approach is employed to formulate image recovery as an optimization problem:

$$\widehat{f} = \underset{f}{\operatorname{argmin}} \left\{ C(f) = \frac{1}{2} \int_{\Omega} |g - Af|^2 \, dxdy + \lambda J(f) \right\}, \tag{2}$$

where the first term is known as data fidelity term, which measures the consistency between estimation and observation, while the second one is called regularization term which provides some a priori knowledge of the desired image, and the regularization parameter $\lambda \geq 0$ provides a balance between the two terms.

2.2. Isotropic Second Degree Total Variation and Equivalent Formulation. Hu and Jacob have introduced the isotropic higher degree total variation (I-HDTV) scheme [16], which inherits the desirable properties of TV, such as, convexity and invariance to rotations and translations. As detailed in [16], the first degree directional derivative of f along the unit vector $\mathbf{s}_0(\theta) = (\cos\theta, \sin\theta)^{\mathrm{T}}$ at coordinate (x, y) denoted by $f_{\theta,1}(x, y)$ is defined as

$$f_{\theta,1}(x, y) = \underbrace{[\cos\theta, \sin\theta]}_{\mathbf{s}_0^{\mathrm{T}}(\theta)} \underbrace{\begin{bmatrix} \dfrac{\partial f(x, y)}{\partial x} \\ \dfrac{\partial f(x, y)}{\partial y} \end{bmatrix}}_{\mathbf{g}_0(x, y)}, \tag{3}$$

where $\mathbf{g}_0(x, y)$ denotes the gradient of f at coordinate (x, y) and $(\cdot)^{\mathrm{T}}$ denotes transpose operator. Specially, as proved in [16], the standard TV is reinterpreted as the L_1-L_2 mixed norms of the first degree directional derivative, formulated as

$$\begin{aligned} \mathrm{TV}(f) &= \int_{\Omega} \|\mathbf{g}_0(x, y)\|_2 dxdy \\ &= \sqrt{2} \int_{\Omega} \|f_{\theta,1}(x, y)\|_{L_2[0,2\pi]} dxdy \\ &= \sqrt{2} \int_{\Omega} \sqrt{\frac{1}{2\pi} \int_0^{2\pi} |f_{\theta,1}(x, y)|^2 d\theta} \, dxdy, \end{aligned} \tag{4}$$

where $\|\cdot\|_2$ denotes the Euclidean norm.

Now, we consider the isotropic second degree total variation (ISDTV) of image f, which is defined as

$$\begin{aligned} J_2(f) &= \int_{\Omega} \|f_{\theta,2}(x, y)\|_{L_2[0,2\pi]} dxdy \\ &= \int_{\Omega} \sqrt{\frac{1}{2\pi} \int_0^{2\pi} |f_{\theta,2}(x, y)|^2 d\theta} \, dxdy, \end{aligned} \tag{5}$$

where $f_{\theta,2}(x, y)$ is the second degree directional derivative of image f at coordinate (x, y), defined as

$$f_{\theta,2}(x, y) = \mathbf{s}^{\mathrm{T}}(\theta) \mathbf{g}(x, y), \tag{6}$$

with

$$\mathbf{s}(\theta) = \left(\cos^2\theta, 2\sin\theta\cos\theta, \sin^2\theta \right)^{\mathrm{T}}, \tag{7}$$

$$\mathbf{g}(x, y) = \left(\frac{\partial^2 f(x, y)}{\partial x^2}, \frac{\partial^2 f(x, y)}{\partial x \partial y}, \frac{\partial^2 f(x, y)}{\partial y^2} \right)^{\mathrm{T}}. \tag{8}$$

Based on the work of [16], we derive an equivalent formulation of the integrand of (5) in following proposition, whose proof is provided in Appendix A.

Proposition 1. *Let* $\mathbf{W} = (1/(2\sqrt{2})) \begin{bmatrix} 1 & 0 & -1 \\ \sqrt{2} & 0 & \sqrt{2} \\ 0 & 2 & 0 \end{bmatrix}$; *then the* $L_2[0, 2\pi]$ *norm of the second degree directional derivative of f at coordinate (x, y) is equal to* $\|\mathbf{W}\mathbf{g}(x, y)\|_2$, *where* $\|\cdot\|_2$ *stands for the Euclidean norm.*

Therefore, according to the differential operator U defined in Table 1, the ISDTV regularizer in (5) can be equivalently expressed as

$$J_2(f) = \int_\Omega \|\mathbf{W}Uf(x,y)\|_2 dxdy. \tag{9}$$

2.3. Mixed Higher Order Total Variation and Equivalent Formulation. Based on the discussion in above Section 2.2, TV is reinterpreted as the L_1-L_2 mixed norms of first degree directional derivative of the image and can well preserve image edges, and ISDTV is defined as the L_1-L_2 mixed norms of second degree directional derivatives of the image and can effectively preserve image features sharpness and minimize the staircase effect. As we know, the recovered image with well-preserved edges and features while not suffering from severely staircase effect is more desirable in image recovery. Furthermore, as detailed in [19–23], the regularizers that combine first and second degree derivatives of the image are reported to well preserve edges and reduce the staircase effect at the same time, and it enables us to extend TV and ISDTV to mixed higher order case. To this end, we thus propose a new regularizer which combines first and second degree derivatives of the image in this subsection.

Firstly, with the definitions of $\mathbf{s}_0(\theta)$, $\mathbf{g}_0(x,y)$, $\mathbf{s}(\theta)$, and $\mathbf{g}(x,y)$ in (3), (7), and (8), we define

$$\mathbf{S}(\theta) = \left(\cos\theta, \sin\theta, \cos^2\theta, 2\sin\theta\cos\theta, \sin^2\theta\right)^\mathrm{T},$$

$$\mathbf{G}(x,y)$$

$$= \left(\frac{\partial f(x,y)}{\partial x}, \frac{\partial f(x,y)}{\partial y}, \frac{\partial^2 f(x,y)}{\partial x^2}, \frac{\partial^2 f(x,y)}{\partial x\partial y}, \frac{\partial^2 f(x,y)}{\partial y^2}\right)^\mathrm{T}. \tag{10}$$

Based on the definitions in (10), we define the new regularizer as

$$J_{1,2}(f) = \int_\Omega \left\|\mathbf{S}^\mathrm{T}(\theta)\mathbf{G}(x,y)\right\|_{L_2[0,2\pi]} dxdy$$

$$= \int_\Omega \sqrt{\frac{1}{2\pi}\int_0^{2\pi} \left|\mathbf{S}^\mathrm{T}(\theta)\mathbf{G}(x,y)\right|^2 d\theta}\, dxdy. \tag{11}$$

Definition (11) leads to a convex regularizer which is also homogeneous, rotation, and translation invariant. Moreover, we have

$$\mathbf{S}^\mathrm{T}(\theta)\mathbf{G}(x,y) = \mathbf{s}_0^\mathrm{T}(\theta)\mathbf{g}_0(x,y) + \mathbf{s}^\mathrm{T}(\theta)\mathbf{g}(x,y)$$

$$= f_{\theta,1}(x,y) + f_{\theta,2}(x,y); \tag{12}$$

namely, the regularizer in (11) involves the first and second degree directional derivatives of the image which are related to TV and ISDTV, respectively, and takes full advantage of the ability to preserve edges and minimize the staircase effect. Thus, the resulting regularizer constitutes a valid extension of TV and ISDTV. For simplicity, we term the new regularizer in (11) as mixed higher order total variation (MHOTV).

Based on the following proposition whose proof is given in Appendix B, we give an equivalent formulation of the MHOTV regularizer.

TABLE 1: Definition of differential operators.

Vectorial operators
$U = \left(\partial_{xx}, \partial_{xy}, \partial_{yy}\right)^\mathrm{T}$
$V = \left(\partial_x, \partial_y, \partial_{xx}, \partial_{xy}, \partial_{yy}\right)^\mathrm{T}$

Proposition 2. *Let* $\mathbf{R} = (1/(2\sqrt{2}))\begin{bmatrix} 0 & 0 & 1 & 0 & -1 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 & \sqrt{2} \\ 0 & 0 & 0 & 2 & 0 \end{bmatrix}$; *then the integrand of* (11) *is equal to* $\|\mathbf{R}\mathbf{G}(x,y)\|_2$, *where* $\|\cdot\|_2$ *stands for the Euclidean norm.*

Therefore, according to the differential operator V defined in Table 1, we further express our MHOTV regularizer as

$$J_{1,2}(f) = \int_\Omega \|\mathbf{R}Vf(x,y)\|_2 dxdy. \tag{13}$$

With the equivalent formulations of the ISDTV and MHOTV regularizers defined in (9) and (13), which can be also understood as weighted L_1-L_2 mixed norm of image derivatives, thus the two equivalent formulations are more preferable for the minimization algorithm we will propose in Section 3.

3. FPG-MFISTA: The Efficient Algorithm for the Proposed Model

Using the ISDTV and MHOTV regularizers described in (9) and (13), we present a unified variational regularization image recovery model to reconstruct the desired image from its noisy observation, which is formulated as

$$\widehat{f} = \underset{f}{\operatorname{argmin}}\left\{C(f) = \frac{1}{2}\int_\Omega |g - Af|^2 dxdy + \lambda J(f)\right\}$$

$$J(f) \in \{J_2(f), J_{1,2}(f)\}, \tag{14}$$

where the last expression in (14) means that the regularization term can be either ISDTV or MHOTV regularizer.

3.1. Discrete Formulation for ISDTV- and MHOTV-Based Image Recovery Model. Now we consider the discrete formulation of proposed model (14). In order to simplify our analysis, we assume that the image with size $m \times n$ is stacked in a vector of size $N = m \times n$. Then, the objective function in (14) can be discretized as

$$C(\mathbf{f}) = \frac{1}{2}\|\mathbf{g} - \mathbf{A}\mathbf{f}\|_2^2 + \lambda J(\mathbf{f}), \tag{15}$$

where $\mathbf{A} \in \mathfrak{R}^{N \times N}$ is the blurring matrix and $\mathbf{g}, \mathbf{f} \in \mathfrak{R}^N$ are the vectorized versions of the observed image and the image to be recovered, respectively. To discretize the differential operators defined in Table 1, we assume reflexive boundary conditions for the images and use forward finite differences to approximate the partial derivatives [27].

To simplify our analysis, we will introduce some auxiliary definitions and notations as follows.

First, we define operators $\mathbf{U} : \mathfrak{R}^N \rightarrow \mathfrak{R}^{3 \times N}$ and $\mathbf{V} : \mathfrak{R}^N \rightarrow \mathfrak{R}^{5 \times N}$ to refer to the discrete versions

of the operators U and V defined in Table 1, respectively. For simplicity, we denote space $X_j = \mathfrak{R}^{j \times N} = \{\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) \mid \mathbf{u}_i \in \mathfrak{R}^j, \forall i = 1, 2, \ldots, N\}$ and equip the space X_j with the inner product $\langle \cdot, \cdot \rangle_{X_j}$ and norm $\|\cdot\|_{X_j}$. Let $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) \in X_j$ and $\mathbf{v} = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N) \in X_j$; we define them as

$$\langle \mathbf{u}, \mathbf{v} \rangle_{X_j} = \sum_{i=1}^{N} \mathbf{u}_i^{\mathrm{T}} \mathbf{v}_i,$$

$$\|\mathbf{u}\|_{X_j} = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle_{X_j}}. \tag{16}$$

Next, let us define the adjoint operators of $\mathbf{U}$ and $\mathbf{V}$ which are the discrete operators $\mathbf{U}^{\mathrm{T}} : X_3 \rightarrow \mathfrak{R}^N$ and $\mathbf{V}^{\mathrm{T}} : X_5 \rightarrow \mathfrak{R}^N$, respectively. Thus, let $\mathbf{p} \in X_3$, $\mathbf{q} \in X_5$ and $\mathbf{f} \in \mathfrak{R}^N$; we have

$$\langle \mathbf{p}, \mathbf{Uf} \rangle_{X_3} = \langle \mathbf{U}^{\mathrm{T}} \mathbf{p}, \mathbf{f} \rangle_2,$$

$$\langle \mathbf{q}, \mathbf{Vf} \rangle_{X_5} = \langle \mathbf{V}^{\mathrm{T}} \mathbf{q}, \mathbf{f} \rangle_2, \tag{17}$$

where $\langle \cdot, \cdot \rangle_2$ stands for the inner product in Euclidean space $\mathfrak{R}^N$.

Furthermore, with the l_1-l_2 mixed norm of a vector field $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) \in X_j$ defined as $\|\mathbf{u}\|_{1,2} = \sum_{i=1}^{N} \|\mathbf{u}_i\|_2$, we thus write the discrete versions of the ISDTV and MHOTV regularizers defined in (9) and (13) as

$$J_2(\mathbf{f}) = \sum_{i=1}^{N} \|\mathbf{W}(\mathbf{Uf})_i\|_2 = \|\mathbf{WUf}\|_{1,2}, \tag{18}$$

$$J_{1,2}(\mathbf{f}) = \sum_{i=1}^{N} \|\mathbf{R}(\mathbf{Vf})_i\|_2 = \|\mathbf{RVf}\|_{1,2}, \tag{19}$$

where $(\cdot)_i$ denotes the ith column element of the corresponding vector filed.

Therefore, we can write the discrete formulation of proposed model (14) as

$$\widehat{\mathbf{f}} = \arg\min_{\mathbf{f}} \left\{ C(\mathbf{f}) = \frac{1}{2} \|\mathbf{g} - \mathbf{Af}\|_2^2 + \lambda \|\mathbf{Hf}\|_{1,2} \right\}$$

$$\mathbf{H} \in \{\mathbf{WU}, \mathbf{RV}\}, \tag{20}$$

with the last expression $\mathbf{H} \in \{\mathbf{WU}, \mathbf{RV}\}$ in (20) meaning that the regularization term can be either ISDTV or MHOTV regularizer.

3.2. Majorization-Minimization (MM).

In the following, we focus on the minimization of model (20). Due to the fact that the ISDTV and MHOTV regularizers are both nondifferentiable, we solve model (20) by a majorization-minimization (MM) approach [28–30]. As described in the MM framework, an iterative algorithm for solving the proposed model (20) at kth iterate $\mathbf{f}^{(k)}$ involves the following two steps:

(1) constructing the surrogate functional $L(\mathbf{f}; \mathbf{f}^{(k)})$ that satisfies the following properties:

$$C(\mathbf{f}) \leq L(\mathbf{f}; \mathbf{f}^{(k)}), \quad \forall \mathbf{f}; \quad C(\mathbf{f}^{(k)}) = L(\mathbf{f}^{(k)}; \mathbf{f}^{(k)}), \tag{21}$$

(2) solving the next iterate $\mathbf{f}^{(k+1)} = \operatorname{argmin}_{\mathbf{f}} L(\mathbf{f}; \mathbf{f}^{(k)})$.

Under the MM framework, we find the optimal solution of model (20) by iteratively minimizing a sequence of surrogate functionals $\{L(\mathbf{f}; \mathbf{f}^{(k)})\}$ that upper bound the objective function $C(\mathbf{f})$.

Therefore, to obtain the surrogate functionals $\{L(\mathbf{f}; \mathbf{f}^{(k)})\}$, we first upper bound the data fidelity term of $C(\mathbf{f})$ at kth iterate $\mathbf{f}^{(k)}$ using the following majorizer:

$$Q(\mathbf{f}; \mathbf{f}^{(k)}) = \frac{\alpha}{2} \|\mathbf{f} - \mathbf{z}\|_2^2 + c, \tag{22}$$

where $\mathbf{z} = \mathbf{f}^{(k)} + (1/\alpha)\mathbf{A}^{\mathrm{T}}(\mathbf{g} - \mathbf{Af}^{(k)})$ and c is a constant which is independent of $\mathbf{f}$. To make sure that $Q(\mathbf{f}; \mathbf{f}^{(k)})$ is valid, we choose $\alpha > \|\mathbf{A}^{\mathrm{T}}\mathbf{A}\|$ to make $\alpha\mathbf{I} - \mathbf{A}^{\mathrm{T}}\mathbf{A}$ positive definite. Thus, the whole resulting majorizer of $C(\mathbf{f})$ at kth iterate $\mathbf{f}^{(k)}$ can be written as

$$L(\mathbf{f}; \mathbf{f}^{(k)}) = Q(\mathbf{f}; \mathbf{f}^{(k)}) + \lambda \|\mathbf{Hf}\|_{1,2}$$

$$= \frac{\alpha}{2} \|\mathbf{f} - \mathbf{z}\|_2^2 + \lambda \|\mathbf{Hf}\|_{1,2} + c. \tag{23}$$

Then, the minimization problem (20) is converted to iteratively minimizing (23) with regard to $\mathbf{f}$, which can be formulated as

$$\mathbf{f}^{(k+1)} = \operatorname{argmin}_{\mathbf{f}} \left\{ L(\mathbf{f}; \mathbf{f}^{(k)}) = \frac{\alpha}{2} \|\mathbf{f} - \mathbf{z}\|_2^2 + \lambda \|\mathbf{Hf}\|_{1,2} + c \right\}$$

$$\mathbf{H} \in \{\mathbf{WU}, \mathbf{RV}\}. \tag{24}$$

As we can see, the resulting minimization problem (24) is much simpler and its solution can be alternatively interpreted as the solution of a denoising problem with $\mathbf{z}$ being the noisy observation. In the following, we will present a primal-dual approach based on the monotone fast iterative shrinkage thresholding algorithm (MFISTA) [24] to seek the optimal solution of (24).

3.3. FPG-MFISTA.

Considering that $\mathbf{f}$ is a digital image whose intensities are finite and bounded, we can use a convex closed set to constrain the range of image intensities. Based on that, the resulting denoising subproblem (24) can be equivalently formulated as

$$\mathbf{f}^{(k+1)} = \operatorname{argmin}_{\mathbf{f}} \frac{\alpha}{2} \|\mathbf{f} - \mathbf{z}\|_2^2 + \lambda \|\mathbf{Hf}\|_{1,2} + I_S(\mathbf{f})$$

$$= \operatorname{argmin}_{\mathbf{f} \in S} \frac{\alpha}{2} \|\mathbf{f} - \mathbf{z}\|_2^2 + \lambda \|\mathbf{Hf}\|_{1,2}, \tag{25}$$

where $\mathbf{H} \in \{\mathbf{WU}, \mathbf{RV}\}$ means that the regularization term can be either ISDTV or MHOTV regularizer, $S = \{\mathbf{f} \in \mathfrak{R}^N \mid \mathbf{f}_i \in [l, u], \forall i = 1, \ldots, N\}$ is a convex closed set which enforces bounded constraints on the solution, and $I_S(\mathbf{f})$ is the indicator function of S that takes the value 0 for $\mathbf{f} \in S$ and ∞ otherwise.

Next, we present a primal-dual approach which can efficiently compute the solution of the constrained problem

(25). Without loss of generality, here we just describe all the steps of the minimization algorithm for model (25) involving the proposed MHOTV regularizer, which corresponds to the case that $\mathbf{H} = \mathbf{RV}$.

Therefore, by denoting $\mu = \lambda/\alpha$, the corresponding MHOTV based denoising subproblem can be further rewritten as

$$\mathbf{f}^{(k+1)} = \arg\min_{\mathbf{f} \in S} \frac{1}{2}\|\mathbf{f} - \mathbf{z}\|_2^2 + \mu\|\mathbf{RVf}\|_{1,2}. \tag{26}$$

As referred in [31], $\|\cdot\|_{\infty,2}$ is the dual norm of $\|\cdot\|_{1,2}$, and then the l_1-l_2 mixed norm of the vector field $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) \in X_j$ can be rewritten as

$$\|\mathbf{u}\|_{1,2} = \max_{\boldsymbol{\omega} \in X}\langle\mathbf{u}, \boldsymbol{\omega}\rangle_{X_j}, \tag{27}$$

where $X = \{\boldsymbol{\omega} = (\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots, \boldsymbol{\omega}_N) \in X_j \mid \|\boldsymbol{\omega}_i\|_2 \leq 1, \forall i = 1, \ldots, N\}$ denotes the l_∞-l_2 unit norm ball. Using (27), we rewrite model (26) as

$$\begin{aligned}\mathbf{f}^{(k+1)} &= \arg\min_{\mathbf{f} \in S} \frac{1}{2}\|\mathbf{f} - \mathbf{z}\|_2^2 + \mu\max_{\boldsymbol{\omega} \in X}\langle\mathbf{RVf}, \boldsymbol{\omega}\rangle_{X_5} \\ &= \arg\min_{\mathbf{f} \in S} \frac{1}{2}\|\mathbf{f} - \mathbf{z}\|_2^2 + \mu\max_{\boldsymbol{\omega} \in X}\langle\mathbf{f}, \mathbf{V}^T\mathbf{R}^T\boldsymbol{\omega}\rangle_2;\end{aligned} \tag{28}$$

thus, the resulting formulation (28) makes us to solve the following minimax problem:

$$\min_{\mathbf{f} \in S}\max_{\boldsymbol{\omega} \in X}\left\{E(\mathbf{f}, \boldsymbol{\omega}) = \frac{1}{2}\|\mathbf{f} - \mathbf{z}\|_2^2 + \mu\langle\mathbf{f}, \mathbf{V}^T\mathbf{R}^T\boldsymbol{\omega}\rangle_2\right\}. \tag{29}$$

Since the cost function $E(\mathbf{f}, \boldsymbol{\omega})$ is strictly convex in $\mathbf{f}$ and concave in $\boldsymbol{\omega}$, we can exchange the order of the minimum and maximum, and the saddle-point $(\mathbf{f}^{(k+1)}, \boldsymbol{\omega}^{(k)})$ of (29) is attained as reported in [32]. Therefore,

$$\min_{\mathbf{f} \in S}\max_{\boldsymbol{\omega} \in X} E(\mathbf{f}, \boldsymbol{\omega}) = E\left(\mathbf{f}^{(k+1)}, \boldsymbol{\omega}^{(k)}\right) = \max_{\boldsymbol{\omega} \in X}\min_{\mathbf{f} \in S} E(\mathbf{f}, \boldsymbol{\omega}). \tag{30}$$

As we know, $E(\mathbf{f}, \boldsymbol{\omega})$ can be equivalently written as

$$\begin{aligned}E(\mathbf{f}, \boldsymbol{\omega}) &= \frac{1}{2}\left\|\mathbf{f} - \left(\mathbf{z} - \mu\mathbf{V}^T\mathbf{R}^T\boldsymbol{\omega}\right)\right\|_2^2 \\ &\quad + \frac{1}{2}\|\mathbf{z}\|_2^2 - \frac{1}{2}\left\|\mathbf{z} - \mu\mathbf{V}^T\mathbf{R}^T\boldsymbol{\omega}\right\|_2^2.\end{aligned} \tag{31}$$

Based on (30) and (31), we can obtain the saddle-point $(\mathbf{f}^{(k+1)}, \boldsymbol{\omega}^{(k)})$ of (29) by solving

$$\begin{aligned}\mathbf{f}^{(k+1)} &= \arg\min_{\mathbf{f} \in S}\left\{\max_{\boldsymbol{\omega} \in X} \frac{1}{2}\left\|\mathbf{f} - (\mathbf{z} - \mu\mathbf{V}^T\mathbf{R}^T\boldsymbol{\omega})\right\|_2^2\right. \\ &\quad \left. + \frac{1}{2}\|\mathbf{z}\|_2^2 - \frac{1}{2}\left\|\mathbf{z} - \mu\mathbf{V}^T\mathbf{R}^T\boldsymbol{\omega}\right\|_2^2\right\} \\ &= P_S\left(\mathbf{z} - \mu\mathbf{V}^T\mathbf{R}^T\boldsymbol{\omega}^{(k)}\right),\end{aligned} \tag{32}$$

where P_S denotes the orthogonal projection onto the convex set S, while $\boldsymbol{\omega}^{(k)}$ is the maximizer of the dual problem, formulated as

$$\begin{aligned}\boldsymbol{\omega}^{(k)} = \arg\max_{\boldsymbol{\omega} \in X}\left\{h(\boldsymbol{\omega}) = \frac{1}{2}\left\|P_S\left(\mathbf{z} - \mu\mathbf{V}^T\mathbf{R}^T\boldsymbol{\omega}\right)\right.\right. \\ \left.\left. - (\mathbf{z} - \mu\mathbf{V}^T\mathbf{R}^T\boldsymbol{\omega})\right\|_2^2\right. \\ \left. + \frac{1}{2}\|\mathbf{z}\|_2^2 - \frac{1}{2}\left\|\mathbf{z} - \mu\mathbf{V}^T\mathbf{R}^T\boldsymbol{\omega}\right\|_2^2\right\}.\end{aligned} \tag{33}$$

Since the dual function $h(\boldsymbol{\omega})$ in (33) is differentiable and has Lipschitz continuous gradient, as referred in [24], then we can compute the gradient of $h(\boldsymbol{\omega})$ as

$$\nabla h(\boldsymbol{\omega}) = \mu\mathbf{RV}P_S\left(\mathbf{z} - \mu\mathbf{V}^T\mathbf{R}^T\boldsymbol{\omega}\right). \tag{34}$$

To accelerate the convergence, we use Nesterov's iterative method [33] to solve model (33), which is a gradient-based approach that exhibits convergence rates of one order higher than the standard gradient ascent method. Furthermore, we employ the operator P_X to return the orthogonal projection onto the l_∞-l_2 unit norm ball X, where this operation is performed by projecting independently each component $\boldsymbol{\omega}_i$, $i = 1, 2, \ldots, N$ of $\boldsymbol{\omega}$ onto the l_2 unit norm ball.

Therefore, as described above, we first find the maximizer of the dual problem (33) and then obtain the optimal solution of our primal problem (26) through (32).

In the same way, by using (28)–(33) and replacing $\mathbf{R}$ and $\mathbf{V}$ with $\mathbf{W}$ and $\mathbf{U}$, respectively, we can obtain the solution of model (25) involving the ISDTV regularizer.

Finally, a detailed description of the fast projected gradient algorithm combined with monotone fast iterative shrinkage thresholding (FPG-MFISTA) for ISDTV and MHOTV regularization-based image recovery is given in Algorithm 1, while the corresponding denoising subproblem (24) can be efficiently solved by a fast projected gradient algorithm described in Algorithm 2.

4. Experimental Results

To evaluate the performance of our proposed ISDTV and MHOTV regularization schemes in the context of image deblurring, we provide systematical experimental comparisons with five state-of-the-art methods, namely, TV regularization method [24], nonlocal TV regularization method [25], Hessian Frobenius-norm and Hessian Spectral-norm regularization methods [18], and the Hessian Nuclear-norm regularization method [26]. Finally, the quality of the reconstructed image is evaluated in terms of peak signal-to-noise ratio (PSNR), which is measured in decibels (dB) and defined as

$$\text{PSNR} = 10\log_{10}\left(\frac{\max_{1 \leq i \leq N}\left|\mathbf{f}_i^{(0)}\right|^2}{(1/N)\left\|\mathbf{f}^{(0)} - \widehat{\mathbf{f}}\right\|_2^2}\right), \tag{35}$$

where $\widehat{\mathbf{f}} \in \mathfrak{R}^N$ and $\mathbf{f}^{(0)} \in \mathfrak{R}^N$ are the vectorized versions of the reconstructed image and the original image, respectively.

Step 1. Input $\mathbf{g}$, $\mathbf{A}$, $\lambda > 0$, $\alpha > \left\|\mathbf{A}^{\mathrm{T}}\mathbf{A}\right\|$, ε ($\varepsilon > 0$ is a tolerance), *Maxiter*
      (Number of iterations). Set $k = 0$. Initialize $\mathbf{v}^{(1)} = \mathbf{f}^{(0)}$, $t_1 = 1$, $c_1 = C(\mathbf{f}^{(0)})$.
Step 2. Compute

$$\mathbf{s}^{(k)} \longleftarrow \mathrm{FPG} \left(\mathbf{v}^{(k)} + \frac{1}{\alpha}\mathbf{A}^{\mathrm{T}}\left(\mathbf{g} - \mathbf{A}\mathbf{v}^{(k)}\right), \frac{\lambda}{\alpha}\right);$$

$$t_{k+1} \longleftarrow \frac{1 + \sqrt{1 + 4t_k^2}}{2}; \quad c_{k+1} = C\left(\mathbf{s}^{(k)}\right);$$

if $c_{k+1} > c_k$ **then**
    $c_{k+1} = c_k;$ $\mathbf{f}^{(k+1)} \longleftarrow \mathbf{f}^{(k)};$
else
    $\mathbf{f}^{(k+1)} \longleftarrow \mathbf{s}^{(k)};$
end

$$\mathbf{v}^{(k+1)} \longleftarrow \mathbf{f}^{(k+1)} + \frac{t_k}{t_{k+1}}\left(\mathbf{s}^{(k)} - \mathbf{f}^{(k+1)}\right) + \frac{t_k - 1}{t_{k+1}}\left(\mathbf{f}^{(k+1)} - \mathbf{f}^{(k)}\right).$$

Step 3. Terminate if $\left\|\mathbf{f}^{(k+1)} - \mathbf{f}^{(k)}\right\|_2 / \left\|\mathbf{f}^{(k)}\right\|_2 < \varepsilon$. Otherwise, set $k = k + 1$ and go to Step 2
Step 4. Output $\widehat{\mathbf{f}}$-An optimal solution of (20).

ALGORITHM 1: FPG-MFISTA for ISDTV and MHOTV regularizer-based image recovery.

Step 1. Input $\mathbf{z}$, $\mathbf{H}$, $\mu > 0$, $\gamma > \left\|\mathbf{H}^{\mathrm{T}}\mathbf{H}\right\|$, ε_0 ($\varepsilon_0 > 0$ is a tolerance), *Inneriter*
      (Number of iterations). Set $k = 0$. Initialize $\mathbf{u}^{(1)} = \boldsymbol{\omega}^{(0)} = \mathbf{0}$, $t_1 = 1$.
Step 2. Compute

$$\boldsymbol{\omega}^{(k)} \longleftarrow P_X \left(\mathbf{u}^{(k)} + \frac{1}{\mu\gamma}\mathbf{H}P_S\left(\mathbf{z} - \mu\mathbf{H}^{\mathrm{T}}\mathbf{u}^{(k)}\right)\right);$$

$$t_{k+1} \longleftarrow \frac{1 + \sqrt{1 + 4t_k^2}}{2};$$

$$\mathbf{u}^{(k+1)} \longleftarrow \boldsymbol{\omega}^{(k)} + \frac{t_k - 1}{t_{k+1}}\left(\boldsymbol{\omega}^{(k)} - \boldsymbol{\omega}^{(k-1)}\right);$$

Step 3. Terminate if $\left\|\boldsymbol{\omega}^{(k+1)} - \boldsymbol{\omega}^{(k)}\right\|_2 / \left\|\boldsymbol{\omega}^{(k)}\right\|_2 < \varepsilon_0$. Otherwise, set $k = k + 1$ and go to Step 2
Step 4. Output $\mathbf{f}^{(k)} = P_S(\mathbf{z} - \mu\mathbf{H}^{\mathrm{T}}\boldsymbol{\omega}^{(k-1)})$—An optimal solution of (25).

ALGORITHM 2: FPG $(\mathbf{z}, \mu)$ (fast projected gradient) algorithm for image denoising under ISDTV ($\mathbf{H} = \mathbf{WU}$) or MHOTV ($\mathbf{H} = \mathbf{RV}$) regularizer.

4.1. Experiment Setting. For the image deblurring experiments, our test images contain six natural images and six biomedical cell images which are shown in Figures 1 and 2, respectively, where the cell images are part of the biomedical image database [34], and they are all converted to grayscale. In our experiments, the intensities of all the test images are initially normalized to the range [0, 1], and we use three kinds of point spread function (PSF) to produce the blurred images, which are a 9×9 Gaussian PSF of standard deviation $\sigma_b = 6$, a 9×9 uniform one, and a 19×19 motion one, respectively. Furthermore, the Gaussian noise of two different levels corresponding to a blurred SNR (BSNR) of $\{20, 30\}$ dB is added to generate the eventually degraded images, and the BSNR is defined as

$$\mathrm{BSNR} = \frac{\mathrm{var}(\mathbf{Af})}{\sigma_n^2}, \tag{36}$$

where var($\mathbf{Af}$) is the variance of blurred image and σ_n is the standard deviation of Gaussian noise. In addition, we constrain the intensities of image recovered by all the seven methods to lie in the convex closed set $S = \{\mathbf{f} \in \mathfrak{R}^N \mid \mathbf{f}_i \in$ $[0, 1], \forall i = 1, \ldots, N\}$. For the sake of fair comparisons, the regularization parameter λ for each regularization method is chosen to give the best PSNR performance. Finally, we set $\varepsilon = 10^{-5}$, *Maxiter* $= 100$, $\varepsilon_0 = 10^{-3}$, and *Inneriter* $= 10$ in our experiments.

4.2. Recovery on Natural Images. In Table 2, we provide systematic comparative results for all the regularization schemes on the natural images under different degradation conditions. As we can see from Table 2, regarding comparisons among the seven regularization schemes, the best PSNR results, on average, are achieved for the proposed MHOTV regularizer, the nonlocal TV regularizer provides better results than TV regularizer for most of the cases, and ISDTV regularizer outperforms TV and Hessian Spectral-norm regularizers for most of the cases, while the Hessian Frobenius-norm regularizer always performs similar results with ISDTV scheme, and the PSNR results of Hessian Nuclear-norm regularizer are almost always a little better than that of Hessian Spectral-norm and Frobenius-norm

TABLE 2: PSNR (dB) results of the seven regularization methods on natural images for three blurring kernels and two different noise levels.

Blur	Gaussian		Uniform		Motion		Gaussian		Uniform		Motion	
BSNR (dB)	20	30	20	30	20	30	20	30	20	30	20	30
	Barbara						Boat					
TV	23.22	24.12	23.26	24.15	23.34	24.72	26.83	29.12	26.97	29.35	26.23	28.79
Nonlocal TV	**23.42**	24.21	**23.45**	24.22	23.28	25.02	26.94	29.25	27.07	29.46	25.87	28.67
Hessian Spectral	23.32	24.18	23.35	24.22	23.32	24.68	26.91	29.11	27.00	29.30	26.05	28.47
Hessian Frobenius	23.36	24.23	23.38	24.27	23.37	24.78	27.01	29.27	27.12	29.47	26.15	28.60
Hessian Nuclear	23.38	24.26	23.40	24.29	23.41	24.85	27.07	29.34	27.19	29.57	26.19	28.65
ISDTV	23.35	24.22	23.37	24.26	23.36	24.76	26.98	29.23	27.09	29.43	26.13	28.58
MHOTV	**23.42**	**24.28**	**23.45**	**24.32**	**23.50**	**24.88**	**27.15**	**29.46**	**27.27**	**29.67**	**26.48**	**28.96**
	Face						Fingerprint					
TV	26.13	28.83	26.11	28.73	24.97	28.43	23.47	26.37	23.31	26.26	22.43	25.30
Nonlocal TV	26.90	29.28	26.73	29.32	24.44	27.77	**24.86**	26.88	24.57	26.72	23.03	25.84
Hessian Spectral	26.84	29.61	26.74	29.49	25.70	29.23	24.68	27.07	24.53	27.04	23.75	26.30
Hessian Frobenius	26.98	29.77	26.89	29.69	25.79	29.39	24.79	27.13	**24.62**	27.11	23.79	**26.36**
Hessian Nuclear	27.03	29.79	26.96	29.69	25.83	29.46	24.75	27.07	24.60	27.06	**23.80**	26.32
ISDTV	26.95	29.82	26.87	29.67	25.78	29.44	24.74	27.09	24.58	27.06	23.76	26.32
MHOTV	**27.12**	**30.01**	**27.00**	**29.84**	**25.86**	**29.58**	24.69	**27.27**	24.43	**27.16**	23.45	26.23
	Lena						Peppers					
TV	25.49	27.40	25.54	27.56	25.59	28.37	26.21	29.35	26.54	29.49	24.77	27.19
Nonlocal TV	25.57	27.49	25.67	27.57	25.07	28.17	26.53	28.64	26.82	28.98	24.24	27.64
Hessian Spectral	25.34	27.10	25.41	27.20	25.25	28.16	26.19	28.98	26.46	29.08	25.12	27.97
Hessian Frobenius	25.44	27.22	25.50	27.32	25.39	28.33	26.35	29.14	26.64	29.25	25.23	28.15
Hessian Nuclear	25.51	27.26	25.56	27.37	25.51	28.43	26.48	29.23	26.77	29.37	25.31	28.16
ISDTV	25.42	27.19	25.48	27.29	25.36	28.29	26.29	29.09	26.59	29.21	25.19	28.11
MHOTV	**25.78**	**27.60**	**25.83**	**27.73**	**25.81**	**28.73**	**26.69**	**29.65**	**27.01**	**29.75**	**25.56**	**28.66**

FIGURE 1: Set of natural images (left to right): Barbara, Boat, Face, Fingerprint, Lena, and Peppers.
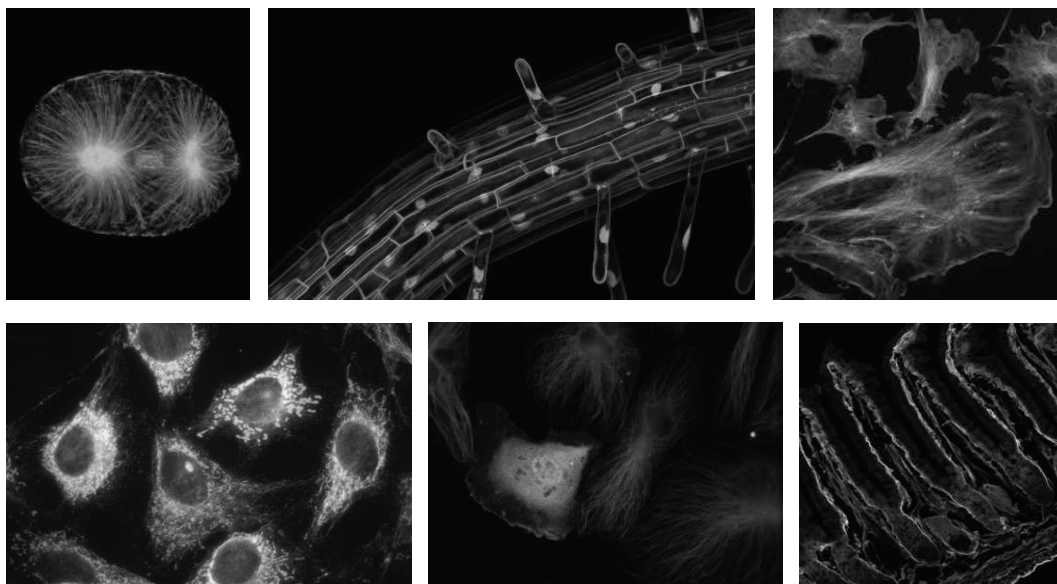

FIGURE 2: Set of biomedical cell images (left to right): Fluorescent cell 1, Fluorescent cell 2, Fluorescent cell 3, CIL 248, CIL 7437, and CIL 12293.
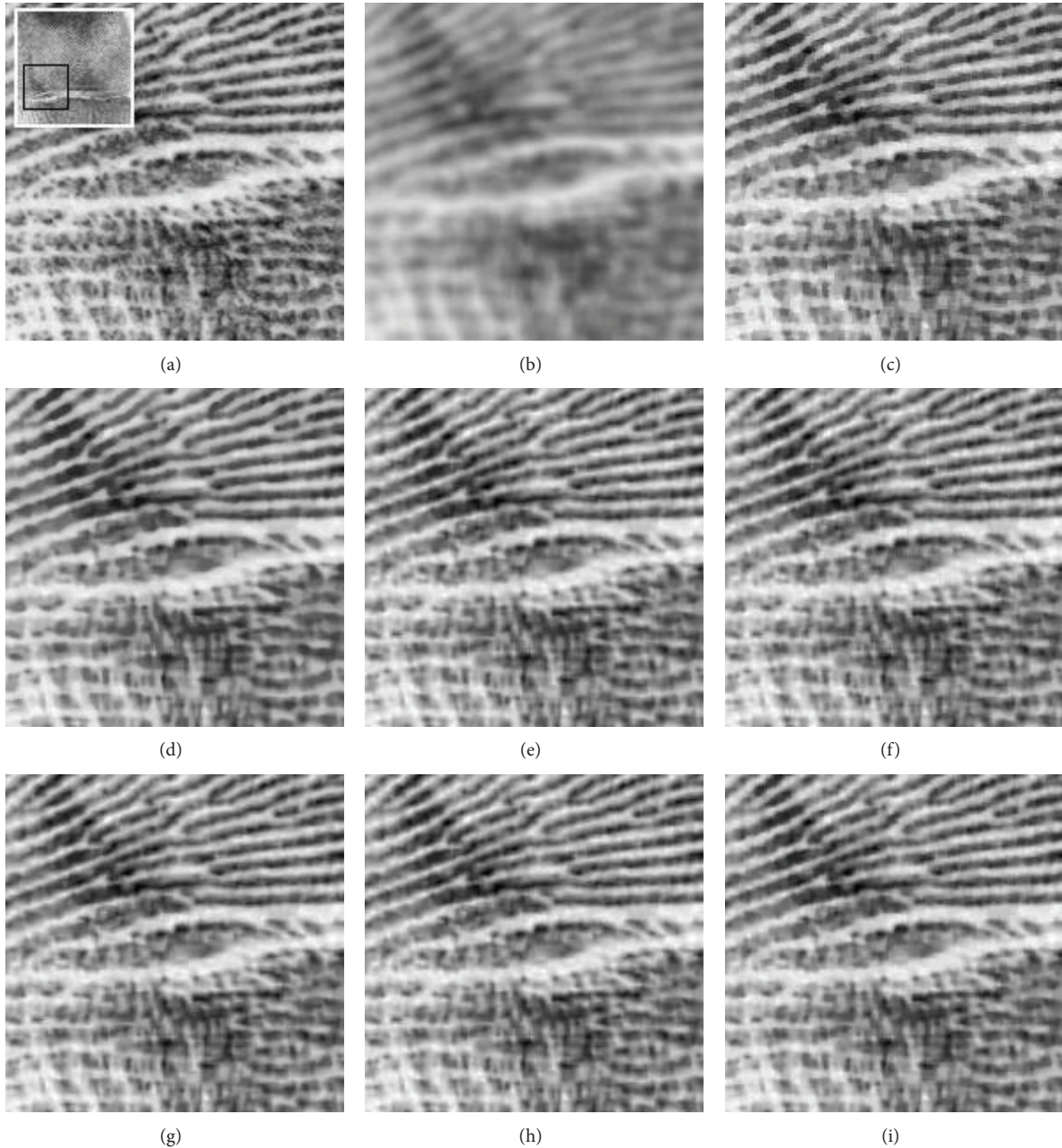
FIGURE 3: Results of different methods on the Fingerprint image degraded by Gaussian blurring and noise corresponding to a BSNR level of 30 dB: (a) original image, (b) degraded image (PSNR = 18.56 dB), (c) TV solution (PSNR = 26.37 dB), (d) nonlocal TV solution (PSNR = 26.88 dB), (e) Hessian Frobenius-norm solution (PSNR = 27.13 dB), (f) Hessian Spectral-norm solution (PSNR = 27.07 dB), (g) Hessian Nuclear-norm solution (PSNR = 27.07 dB), (h) ISDTV solution (PSNR = 27.09 dB), and (i) MHOTV solution (PSNR = 27.27 dB).

regularizers. Thus, we can conclude that our MHOTV-based regularization scheme is proven to be a valid approach for image recovery.

Furthermore, we show the representative Fingerprint, Face, and Peppers deblurring examples in Figures 3, 4, and 5, which are degraded by Gaussian blurring, uniform blurring, and motion blurring with different Gaussian noise levels, respectively.

From these figures, we can clearly see that the images recovered by TV regularization method suffer from heavy blocky artifacts appearing in the smoother regions of the image and result in some loss of details, the nonlocal TV regularization method preserves the sharp edges but smoothes some moderate gradient areas, the Hessian Spectral-norm, Hessian Frobenius-norm, Hessian Nuclear-norm, and ISDTV regularization methods are very competitive, and they can also preserve some sharp edges and minimize the staircase effect, while our MHOTV regularization method provides the best image quality which can effectively minimize the staircase effect and better preserve edges

Figure 4: Results of different methods on the Face image degraded by uniform blurring and noise corresponding to a BSNR level of 20 dB: (a) original image, (b) degraded image (PSNR = 21.77 dB), (c) TV solution (PSNR = 26.11 dB), (d) nonlocal TV solution (PSNR = 26.73 dB), (e) Hessian Frobenius-norm solution (PSNR = 26.89 dB), (f) Hessian Spectral-norm solution (PSNR = 26.74 dB), (g) Hessian Nuclear-norm solution (PSNR = 26.96 dB), (h) ISDTV solution (PSNR = 26.87 dB), and (i) MHOTV solution (PSNR = 27.00 dB).

sharpness. For example, in Figure 3, since the Fingerprint image is mainly composed of smooth transitions and ridge features, in order to highlight the differences, we present the details of the corresponding zooming versions of the region (marked by the black box in Figure 3(a)) for all the seven methods in Figures 3(c)–3(i). Comparing the results of these seven methods, we observe that the TV solution shown in Figure 3(c) results in some loss of fine ridge details and produces the staircase effect, the nonlocal TV solution shown in Figure 3(d) smoothes some ridge details, and the

remaining five methods all preserve fine ridge details and edges better than TV and nonlocal TV methods, while our MHOTV solution provides the best performance in terms of both PSNR and image quality.

4.3. Recovery on Biomedical Cell Images.

In this section, we mainly talk about the performance of the proposed MHOTV regularization scheme on biomedical images, which are closely related to the practical applications. To evaluate

FIGURE 5: Results of different methods on the Peppers image degraded by motion blurring and noise corresponding to a BSNR level of 30 dB: (a) original image, (b) degraded image (PSNR = 18.98 dB), (c) TV solution (PSNR = 27.19 dB), (d) nonlocal TV solution (PSNR = 27.64 dB), (e) Hessian Frobenius-norm solution (PSNR = 28.15 dB), (f) Hessian Spectral-norm solution (PSNR = 27.97 dB), (g) Hessian Nuclear-norm solution (PSNR = 28.16 dB), (h) ISDTV solution (PSNR = 28.11 dB), and (i) MHOTV solution (PSNR = 28.66 dB).

the effectiveness of our approach, we provide deblurring experiments for all the methods on the set of biomedical cell images, and their comparative results are given in Table 3.

With the PSNR comparisons in Table 3, we can clearly observe that the proposed MHOTV regularization approach provides the best PSNR results for most of the cases, the ISDTV method always performs better than nonlocal TV, TV, and Hessian Spectral-norm methods, and the nonlocal TV method provides the better results than TV method for most of the cases, while the Hessian Frobenius-norm method always provides the comparatively similar results to ISDTV method, and the PSNR results of Hessian Nuclear-norm method are almost always a little better than that of Hessian Spectral-norm and Frobenius-norm methods. Meanwhile, we show the representative Fluorescent cell 3 and CIL 12293 deblurring examples in Figures 6 and 7, respectively.
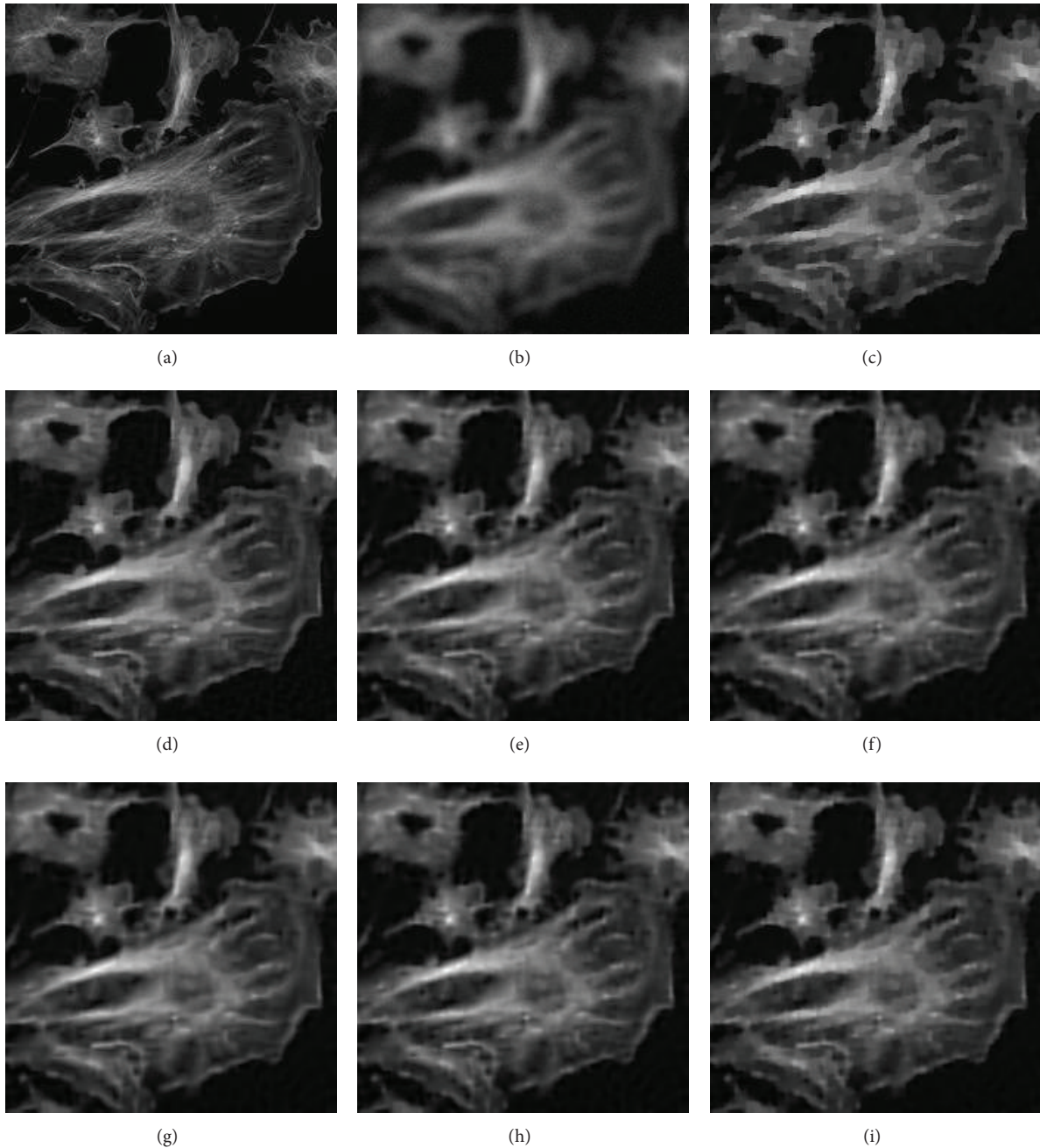
(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(i)

FIGURE 6: Results of different methods on the Fluorescent cell 3 image degraded by Gaussian blurring and noise corresponding to a BSNR level of 20 dB: (a) original image, (b) degraded image (PSNR = 24.40 dB), (c) TV solution (PSNR = 27.15 dB), (d) nonlocal TV solution (PSNR = 27.47 dB), (e) Hessian Frobenius-norm solution (PSNR = 27.53 dB), (f) Hessian Spectral-norm solution (PSNR = 27.47 dB), (g) Hessian Nuclear-norm solution (PSNR = 27.53 dB), (h) ISDTV solution (PSNR = 27.52 dB), and (i) MHOTV solution (PSNR = 27.67 dB).

While carefully inspecting the recovered images shown in Figures 6 and 7, we obviously find that the advantage of our MHOTV approach consists in better preserving the ridges and filament-like features sharpness without introducing severe block artifacts, and the nonlocal TV, ISDTV, and the three Hessian-norm regularization methods can preserve some fine ridge features and alleviate the staircase effect, as opposed to TV which introduces the heavy block artifacts which tends to produce the oversmoothed image

ridge and filament-like features. Therefore, we conclude that our MHOTV regularization approach can be effectively and successfully applied in the process of biomedical imaging.

5. Conclusion

In this paper, we have proposed a novel MHOTV regularizer which combines the first and second degree derivatives

(a)

(a)   (b)   (c)

(d)   (e)   (f)

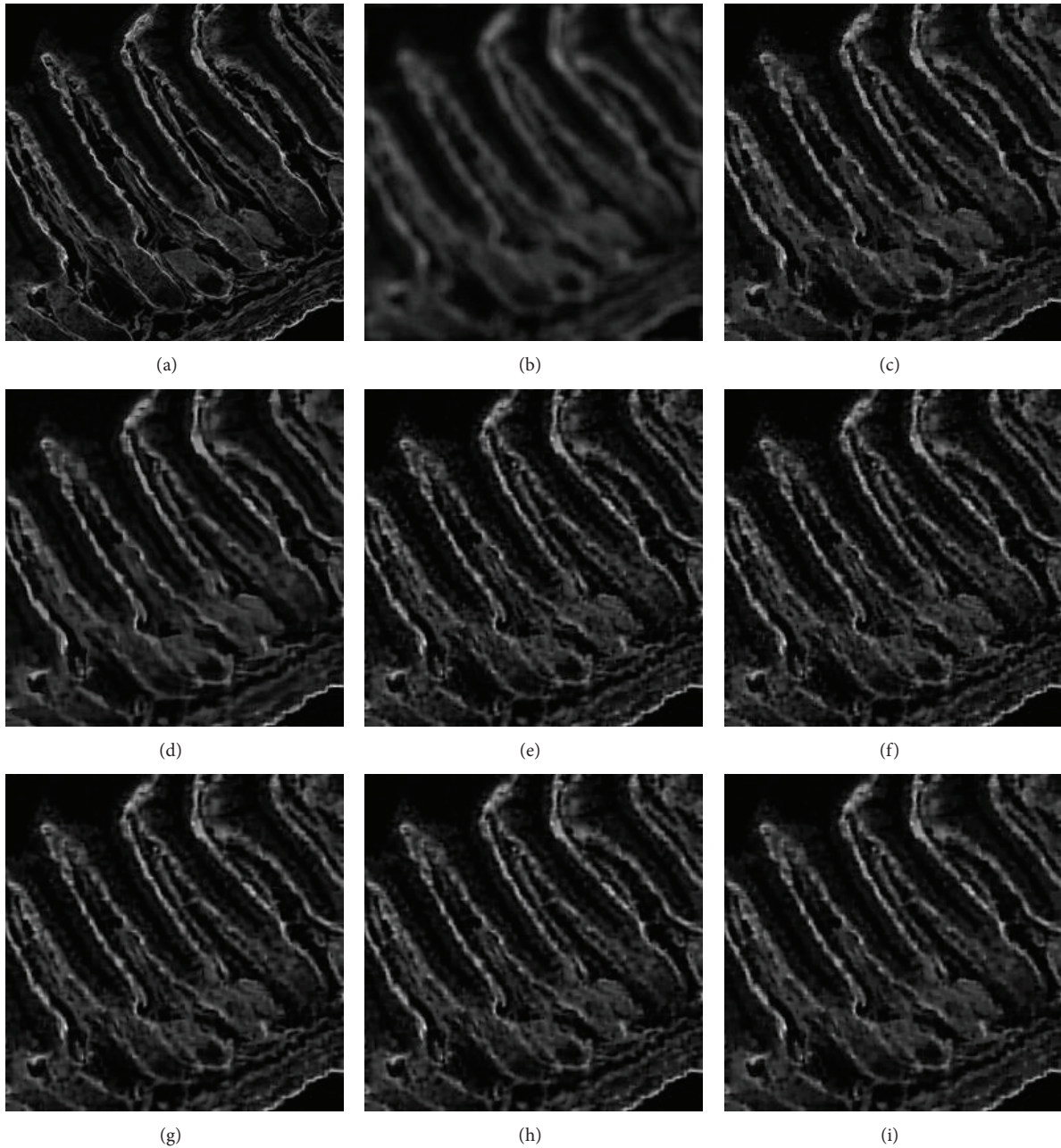(g)   (h)   (i)

FIGURE 7: Results of different methods on the CIL 12293 image degraded by uniform blurring and noise corresponding to a BSNR level of 20 dB: (a) original image, (b) degraded image (PSNR = 21.00 dB), (c) TV solution (PSNR = 25.04 dB), (d) nonlocal TV solution (PSNR = 24.95 dB), (e) Hessian Frobenius-norm solution (PSNR = 25.43 dB), (f) Hessian Spectral-norm solution (PSNR = 25.33 dB), (g) Hessian Nuclear-norm solution (PSNR = 25.44 dB), (h) ISDTV solution (PSNR = 25.48 dB), and (i) MHOTV solution (PSNR = 25.56 dB).

of the image. Within spectral decomposition framework, we reformulate both the ISDTV and MHOTV regularizers as weighted L_1-L_2 mixed norms of image derivatives. Furthermore, we present a unified variational model based on the ISDTV and MHOTV regularizers for image recovery. Based on the equivalent formulations of ISDTV and MHOTV regularizers, we then introduce an efficient fast projected gradient algorithm combined with monotone fast iterative shrinkage thresholding (FPG-MFISTA) to solve the

corresponding optimization problems under majorization-minimization framework. Finally, the performance of the proposed MHOTV regularization scheme was assessed by comparing the experimental results on both natural images and biomedical cell images against various state-of-the-art methods. Specifically, the proposed MHOTV regularization scheme performs improved results and can effectively minimize the staircase effect that is commonly met in TV scheme while better preserving edges sharpness.

TABLE 3: PSNR (dB) results of the seven regularization methods on biomedical cell images for two blurring kernels and two different noise levels.

Blur	Gaussian		Uniform		Gaussian		Uniform	
BSNR (dB)	20	30	20	30	20	30	20	30
	Fluorescent cell 1				Fluorescent cell 2			
TV	27.52	29.20	27.62	29.29	28.21	30.12	28.28	30.35
Nonlocal TV	27.44	29.22	27.46	29.23	28.40	30.12	28.41	30.28
Hessian Spectral	27.92	29.63	28.00	29.74	28.61	30.38	28.63	30.61
Hessian Frobenius	**27.97**	29.73	28.03	29.83	28.72	30.51	28.77	30.69
Hessian Nuclear	**27.97**	**29.75**	**28.04**	**29.88**	**28.74**	**30.57**	**28.82**	30.77
ISDTV	27.96	29.70	28.01	29.80	28.69	30.48	28.74	30.70
MHOTV	**27.97**	29.67	28.02	29.76	28.72	30.55	28.78	**30.78**
	Fluorescent cell 3				CIL 248			
TV	27.15	29.01	27.15	29.24	27.70	29.65	27.69	29.79
Nonlocal TV	27.47	29.18	27.48	29.25	27.89	29.68	27.80	29.76
Hessian Spectral	27.47	29.25	27.49	29.47	28.23	30.00	28.19	30.10
Hessian Frobenius	27.53	29.35	27.56	29.55	28.26	30.10	28.19	30.18
Hessian Nuclear	27.53	29.37	27.60	29.56	28.23	30.08	28.17	30.18
ISDTV	27.52	29.32	27.55	29.53	28.25	30.11	28.20	30.21
MHOTV	**27.67**	**29.50**	**27.68**	**29.73**	**28.35**	**30.25**	**28.30**	**30.38**
	CIL 7437				CIL 12293			
TV	34.61	36.31	34.66	36.42	24.95	26.81	25.04	27.26
Nonlocal TV	34.63	36.43	34.57	36.43	24.86	26.58	24.95	26.88
Hessian Spectral	35.10	36.79	35.08	36.88	25.31	26.97	25.33	27.37
Hessian Frobenius	35.13	36.85	35.11	36.96	25.40	27.03	25.43	27.44
Hessian Nuclear	35.10	36.83	35.09	36.92	25.42	27.00	25.44	27.40
ISDTV	35.12	36.83	35.10	36.94	25.44	27.04	25.48	27.46
MHOTV	**35.19**	**36.91**	**35.19**	**37.01**	**25.49**	**27.08**	**25.56**	**27.52**

Appendix

A. Proof of Proposition 1

Proof. Using the rotation steerability of directional derivatives [35, 36], we simplify the integrand of (5) as

$$
\sqrt{\frac{1}{2\pi} \int_0^{2\pi} \left| f_{\theta,2}(x,y) \right|^2 d\theta}
$$

$$
= \sqrt{\mathbf{g}^{\mathrm{T}}(x,y) \underbrace{\left(\frac{1}{2\pi} \int_0^{2\pi} \mathbf{s}(\theta) \mathbf{s}^{\mathrm{T}}(\theta) d\theta \right)}_{\mathbf{C}_2} \mathbf{g}(x,y)} \quad \text{(A.1)}
$$

$$
= \sqrt{\mathbf{g}^{\mathrm{T}}(x,y) \, \mathbf{C}_2 \mathbf{g}(x,y)}.
$$

Here, $\mathbf{C}_2$ is a matrix with entries $\mathbf{C}_2(i,j) = (1/2\pi) \int_0^{2\pi} \mathbf{s}_i(\theta) \mathbf{s}_j(\theta) d\theta$, $i, j = 1, \ldots, 3$, where $\mathbf{s}_i(\theta)$ denotes the ith element of $\mathbf{s}(\theta)$. Based on that, we have

$$
\mathbf{C}_2 = \frac{1}{8} \begin{bmatrix} 3 & 0 & 1 \\ 0 & 4 & 0 \\ 1 & 0 & 3 \end{bmatrix}. \quad \text{(A.2)}
$$

Since $\mathbf{C}_2$ is a symmetric and positive definite matrix, we can use the spectral decomposition theorem to rewrite $\mathbf{C}_2$ as

$$
\mathbf{C}_2 = \mathbf{Q} \Lambda \mathbf{Q}^{\mathrm{T}}, \quad \text{(A.3)}
$$

where $\Lambda = (1/4) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ and $\mathbf{Q} = (1/\sqrt{2}) \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \\ -1 & 1 & 0 \end{bmatrix}$ are the diagonal matrix with the eigenvalues of $\mathbf{C}_2$ and the orthogonal matrix with the corresponding eigenvectors in its columns, respectively. Specially, let us define $\mathbf{W} = \Lambda^{1/2} \mathbf{Q}^{\mathrm{T}} = (1/(2\sqrt{2})) \begin{bmatrix} 1 & 0 & -1 \\ \sqrt{2} & 0 & \sqrt{2} \\ 0 & 2 & 0 \end{bmatrix}$; we then obtain that

$$
\sqrt{\frac{1}{2\pi} \int_0^{2\pi} \left| f_{\theta,2}(x,y) \right|^2 d\theta}
$$

$$
= \sqrt{\mathbf{g}^{\mathrm{T}}(x,y) \, \mathbf{C}_2 \mathbf{g}(x,y)}
$$

$$
= \sqrt{\mathbf{g}^{\mathrm{T}}(x,y) \, \mathbf{Q} \Lambda \mathbf{Q}^{\mathrm{T}} \mathbf{g}(x,y)}
$$

$$
= \sqrt{\mathbf{g}^{\mathrm{T}}(x,y) \, \mathbf{Q} \Lambda^{1/2} \underbrace{\Lambda^{1/2} \mathbf{Q}^{\mathrm{T}}}_{\mathbf{W}} \mathbf{g}(x,y)}
$$

$$
= \left\| \mathbf{W} \mathbf{g}(x,y) \right\|_2.
$$

$$
\text{(A.4)}
$$

□

B. Proof of Proposition 2

Proof. Using the rotation steerability of directional derivatives [35, 36], we can simplify the integrand of (11) as

$$\sqrt{\frac{1}{2\pi} \int_0^{2\pi} \left| \mathbf{S}^{\mathrm{T}}(\theta) \, \mathbf{G}(x, y) \right|^2 d\theta}$$

$$= \sqrt{\mathbf{G}^{\mathrm{T}}(x, y) \underbrace{\left(\frac{1}{2\pi} \int_0^{2\pi} \mathbf{S}(\theta) \, \mathbf{S}^{\mathrm{T}}(\theta) \, d\theta \right)}_{\mathbf{C}} \mathbf{G}(x, y)}$$

$$= \sqrt{\mathbf{G}^{\mathrm{T}}(x, y) \, \mathbf{C} \mathbf{G}(x, y)}, \tag{B.1}$$

where, $\mathbf{C}$ is a matrix with entries $\mathbf{C}(i, j) = (1/2\pi) \int_0^{2\pi} \mathbf{S}_i(\theta) \mathbf{S}_j(\theta) d\theta$, $i, j = 1, \ldots, 5$, where $\mathbf{S}_i(\theta)$ denotes the ith element of $\mathbf{S}(\theta)$. Based on that, we have

$$\mathbf{C} = \frac{1}{8} \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 1 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 1 & 0 & 3 \end{bmatrix}. \tag{B.2}$$

Since $\mathbf{C}$ is a symmetric and positive definite matrix, we can use the spectral decomposition theorem to reformulate $\mathbf{C}$ as

$$\mathbf{C} = \mathbf{B} \mathbf{D} \mathbf{B}^{\mathrm{T}}, \tag{B.3}$$

where $\mathbf{D} = (1/4) \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$ and $\mathbf{B} =$

$(1/\sqrt{2}) \begin{bmatrix} 0 & \sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} \\ -1 & 0 & 0 & 1 & 0 \end{bmatrix}$ are the diagonal matrix with the eigenvalues of $\mathbf{C}$ and the orthogonal matrix with the corresponding eigenvectors in its columns, respectively. Specially, we define

$$\mathbf{R} = \mathbf{D}^{1/2} \mathbf{B}^{\mathrm{T}} = \frac{1}{2\sqrt{2}} \begin{bmatrix} 0 & 0 & 1 & 0 & -1 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 & \sqrt{2} \\ 0 & 0 & 0 & 2 & 0 \end{bmatrix}. \tag{B.4}$$

Substituting (B.3) into (B.1), we obtain that

$$\sqrt{\frac{1}{2\pi} \int_0^{2\pi} \left| \mathbf{S}^{\mathrm{T}}(\theta) \, \mathbf{G}(x, y) \right|^2 d\theta}$$

$$= \sqrt{\mathbf{G}^{\mathrm{T}}(x, y) \, \mathbf{C} \mathbf{G}(x, y)}$$

$$= \sqrt{\mathbf{G}^{\mathrm{T}}(x, y) \, \mathbf{B} \mathbf{D} \mathbf{B}^{\mathrm{T}} \mathbf{G}(x, y)} \tag{B.5}$$

$$= \sqrt{\mathbf{G}^{\mathrm{T}}(x, y) \, \mathbf{B} \mathbf{D}^{1/2} \underbrace{\mathbf{D}^{1/2} \mathbf{B}^{\mathrm{T}}}_{\mathbf{R}} \mathbf{G}(x, y)}$$

$$= \left\| \mathbf{R} \mathbf{G}(x, y) \right\|_2.$$

□

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] O. Michailovich and D. Adam, "Deconvolution of medical images from microscopic to whole body images," in *Blind Image Deconvolution: Theory and Applications*, P. Campisi and K. Egiazarian, Eds., pp. 169–238, CRC Press, Boca Raton, Fla, USA, 2007.

[2] G. Sluder and D. E. Wolf, *Digital Microscopy*, Academic, New York, NY, USA, 3rd edition, 2007.

[3] C. R. Vogel, *Computational Methods for Inverse Problems*, SIAM, Philadelphia, Pa, USA, 2002.

[4] S. Twomey, *Introduction to the Mathematics of Inversion in Remote Sensing*, Dover, New York, NY, USA, 2002.

[5] M. Ben-Ezra and S. K. Nayar, "Motion-based motion deblurring," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 689–698, 2004.

[6] P. C. Hansen, J. G. Nagy, and D. P. O'Leary, *Deblurring Images: Matrices, Spectra, and Filtering*, SIAM, Philadelphia, Pa, USA, 2006.

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

[8] A. Marquina, "Nonlinear inverse scale space methods for total variation blind deconvolution," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 64–83, 2009.

[9] R. Choksi and Y. van Gennip, "Deblurring of one dimensional bar codes via total variation energy minimization," *SIAM Journal on Imaging Sciences*, vol. 3, no. 4, pp. 735–764, 2010.

[10] C. Louchet and L. Moisan, "Total variation as a local filter," *SIAM Journal on Imaging Sciences*, vol. 4, no. 2, pp. 651–694, 2011.

[11] M. K. Ng and W. Wang, "A total variation model for retinex," *SIAM Journal on Imaging Sciences*, vol. 4, no. 1, pp. 345–365, 2011.

[12] T. Chan, A. Marquina, and P. Mulet, "High-order total variation-based image restoration," *SIAM Journal on Scientific Computing*, vol. 22, no. 2, pp. 503–516, 2000.

[13] G. Steidl, S. Didas, and J. Neumann, "Splines in higher order TV regularization," *International Journal of Computer Vision*, vol. 70, no. 3, pp. 241–255, 2006.

[14] Y.-L. You and M. Kaveh, "Fourth-order partial differential equations for noise removal," *IEEE Transactions on Image Processing*, vol. 9, no. 10, pp. 1723–1730, 2000.

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

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

[16] Y. Hu and M. Jacob, "Higher degree total variation (HDTV) regularization for image recovery," *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2559–2571, 2012.

[17] S. Lefkimmiatis, A. Bourquard, and M. Unser, "Hessian-based norm regularization for image restoration with biomedical applications," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 983–995, 2012.

[18] S. Lefkimmiatis and M. Unser, "A projected gradient algorithm for image restoration under Hessian matrix-norm regularization," in *Proceedings of the IEEE International Conference on Image Processing*, pp. 3029–3032, 2012.

[19] F. Li, C. Shen, J. Fan, and C. Shen, "Image restoration combining a total variational filter and a fourth-order filter," *Journal of Visual Communication and Image Representation*, vol. 18, no. 4, pp. 322–330, 2007.

[20] M. Lysaker and X.-C. Tai, "Iterative image restoration combining total variation minimization and a second-order functional," *International Journal of Computer Vision*, vol. 66, no. 1, pp. 5–18, 2006.

[21] A. Chambolle and P.-L. Lions, "Image recovery via total variation minimization and related problems," *Numerische Mathematik*, vol. 76, no. 2, pp. 167–188, 1997.

[22] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 492–526, 2010.

[23] F. Knoll, K. Bredies, T. Pock, and R. Stollberger, "Second order total generalized variation (TGV) for MRI," *Magnetic Resonance in Medicine*, vol. 65, no. 2, pp. 480–491, 2011.

[24] 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.

[25] X. Zhang, M. Burger, X. Bresson, and S. Osher, "Bregmanized nonlocal regularization for deconvolution and sparse reconstruction," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 253–276, 2010.

[26] S. Lefkimmiatis, J. P. Ward, and M. Unser, "Hessian Schatten-norm regularization for linear inverse problems," *IEEE Transactions on Image Processing*, vol. 22, no. 5, pp. 1873–1888, 2013.

[27] J. E. Denis and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM, Philadelphia, Pa, USA, 1983.

[28] M. A. T. Figueiredo, J. M. Bioucas-Dias, and R. D. Nowak, "Majorization-minimization algorithms for wavelet-based image restoration," *IEEE Transactions on Image Processing*, vol. 16, no. 12, pp. 2980–2991, 2007.

[29] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[30] D. R. Hunter and K. Lange, "A tutorial on MM algorithms," *The American Statistician*, vol. 58, no. 1, pp. 30–37, 2004.

[31] S. Suvrit, "Fast projections onto $l_{1,q}$-norm balls for grouped feature selection," in *Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD '11)*, vol. 3, pp. 305–317, 2011.

[32] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, USA, 1970.

[33] Y. Nesterov, "A method for solving a convex programming problem with convergence rates O $(1/k^2)$," *Soviet Mathematics Doklady*, vol. 27, pp. 372–376, 1983.

[34] D. Brandner and G. Withers, *The Cell: An Image Library*, 2010, http://www.cellimagelibrary.org.

[35] E. P. Simoncelli and W. T. Freeman, "Steerable pyramid: a flexible architecture for multi-scale derivative computation," in *Proceedings of the IEEE International Conference on Image Processing*, vol. 3, pp. 444–447, October 1995.

[36] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 9, pp. 891–906, 1991.