A PROJECTED GRADIENT ALGORITHM FOR IMAGE RESTORATION UNDER HESSIAN MATRIX-NORM REGULARIZATION

Stamatios Lefkimmiatis and Michael Unser

Biomedical Imaging Group, EPFL, CH-1015 Lausanne, Switzerland

Email:[stamatis.lefkimmiatis,michael.unser]@epfl.ch

ABSTRACT

We have recently introduced a class of non-quadratic Hessian-based regularizers as a higher-order extension of the total variation (TV) functional. These regularizers retain some of the most favorable properties of TV while they can effectively deal with the staircase effect that is commonly met in TV-based reconstructions. In this work we propose a novel gradient-based algorithm for the efficient minimization of these functionals under convex constraints. Furthermore, we validate the overall proposed regularization framework for the problem of image deblurring under additive Gaussian noise.

Index Terms— Linear inverse problems, image restoration, Hessian matrix norms, mixed-norm regularization.

1. INTRODUCTION

Artifacts degrading the quality of recorded images are mainly caused by blurring, which is a perturbation due to the imaging process (i.e. diffraction, aberrations etc.) and random noise that is intrinsic to the detection process. This image degradation can often be a major obstacle preventing image analysis and information extraction. To alleviate these effects, image restoration can serve as a desirable pre-processing technique.

Image deblurring belongs to the general family of inverse problems and it amounts to estimating an image f from the measurements y. The most-commonly used image-observation model involves linear measurements and can be formulated as

$$y = \mathcal{A}f + w\,, \tag{1}$$

where $\mathcal{A}$ is a linear blur operator and w is the measurement noise. The recovery of f from y is an *ill-posed problem* [1], due to the presence of noise and the blurring operator $\mathcal{A}$ that is usually ill-conditioned or non-invertible. To obtain a reasonable estimate of f one must thus reformulate the image restoration problem by taking into account the image formation and acquisition process as well as any available prior information about the properties of the image to be restored.

A common approach for restoring f is to form an objective function which quantifies the quality of a given estimate and has the form

$$\mathcal{J}\left(f\right) = \mathcal{J}_{\text{data}}\left(f\right) + \tau\mathcal{R}\left(f\right)\,. \tag{2}$$

The first term is known as *data fidelity* and measures the consistency between the estimation and the measurements, while the second one is the *regularization* term that constrains the set of plausible solutions. The *regularization parameter* $\tau \geq 0$ balances the contribution of the two terms. Image deblurring can then be cast as the minimization of (2).

This work was supported (in part) by the Hasler Foundation and the Indo-Swiss Joint Research Program.

A parameter that needs special treatment in this problem, since it significantly affects the quality of the restored image, is the form of the regularizer to be chosen. Several regularization approaches have been proposed for image deblurring, with a state-of-the art method based on minimizing the total-variation (TV) semi-norm [2]. Its success and wide use for the past two decades can be mainly attributed to its ability to produce results with well-preserved and sharp edges. In addition, TV is convex, thus permitting the design of efficient optimization methods.

Despite the popularity that TV regularization enjoys, it has been widely reported (cf. [3] for instance) that if it is applied, under the presence of noise, to signals not necessarily piecewise constant then it leads to the well-known *staircase* effect. Indeed, TV favors vanishing first-order derivatives and thus yields solutions belonging to the class of piecewise-constant functions. This effect can be highly undesirable especially in applications such as biomedical imaging where image interpretation can be severely obstructed.

To overcome this spurious effect induced by the TV norm, there is a growing interest in the recent literature for regularization techniques involving higher-order differential operators [3, 4, 5]. The motivation behind this attempt is to potentially restore a wider class of images avoiding staircase effects while preserving image sharpness. These regularizers often involve second-order operators, because vanishing second-order derivatives lead to piecewise-linear results that better fit smooth intensity changes.

In this paper, in Section 2 we briefly review the class of second-order regularizers that we recently introduced in [6] as an extension of TV. In Section 3 we propose a projected-gradient method for their efficient minimization under additional convex constraints. Then in Section 4, we perform a comparative performance study, considering image-deblurring experiments, and show that our regularizers are well-suited for restoring a larger class of images than merely piecewise-constant.

2. SECOND-ORDER REGULARIZATION

The TV norm, assuming that image f is smooth, is defined as

$$\text{TV}(f) = \int_{\Omega} \left\|\nabla f\left(\mathbf{x}\right)\right\|_2 \mathrm{d}\mathbf{x}\,, \tag{3}$$

where $\left\|\nabla f\right\|_2$ is the Euclidean norm of the image gradient and $\Omega \subset \mathbb{R}^2$. As we observed in [6], and one can easily verify by direct calculations, the 2-D TV functional can be equivalently written as

$$\text{TV}(f) = \frac{1}{h\left(q\right)} \int_{\Omega} \left\|\mathrm{D}_{\theta} f\left(\mathbf{x}\right)\right\|_{L_q([0,\,2\pi])} \mathrm{d}\mathbf{x}, \tag{4}$$

where $h\left(q\right) = \left\|\cos\left(\theta\right)\right\|_{L_q([0,\,2\pi])}$ and D_{θ} is the first directional derivative in the direction of the unit-norm vector $\mathbf{u}_{\theta} = (\cos\theta,\,\sin\theta)$, defined as $\mathrm{D}_{\theta} f = \langle \nabla f,\,\mathbf{u}_{\theta}\rangle$. Therefore, according

Table 1. Definition of Continuous Differential Operators.

$\Delta = \partial_{xx} + \partial_{yy}$	$\Delta_e = (\partial_{xx} + \partial_{yy},\, 0)$
$\mathcal{U} = (\partial_{xx} - \partial_{yy},\, 2\partial_{xy})$	$\mathcal{V} = \left(\partial_{xx}, \sqrt{2}\partial_{xy},\, \partial_{yy}\right)$

to (4), TV can be interpreted as a mixed L_1-L_q norm where the L_1-norm acts on the image domain Ω, while the L_q-norm acts on the angular orientation of the directional derivative, $\theta \in [0,\, 2\pi]$.

In two dimensions, a natural way to compute second derivatives is to use the Hessian operator $\mathcal{H}_f = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$ where $f_{ij}(\mathbf{x}) = \frac{\partial^2}{\partial i\, \partial j} f(\mathbf{x})$. While the use of second derivatives for constructing a regularizer is straightforward in the 1-D setting, there are plenty of possible choices in 2-D. To obtain second-order regularizers that promote invariances and at the same time qualify as valid extensions of TV, we recently introduced in [6] a novel class that extends Definition (4) to the second-order case. This is accomplished by increasing the order of differentiation using second-order bidirectional derivatives, and by defining our second-order regularizers as

$$\mathcal{R}(f) = \frac{1}{h^2(q)} \int_{\Omega} \left\| \mathrm{D}^2_{\theta,\phi} f(\mathbf{x}) \right\|_{L_q(\mathcal{S})} \mathrm{d}\mathbf{x}\,, \tag{5}$$

where $\mathcal{S} = [0,\, 2\pi] \times [0,\, 2\pi]$ and $\mathrm{D}^2_{\theta,\phi}$ is the second directional derivative in the directions dictated by the angles θ and ϕ, defined as $\mathrm{D}^2_{\theta,\phi} f(\mathbf{x}) = \mathrm{D}_{\theta}(\mathrm{D}_{\phi} f)(\mathbf{x}) = \mathbf{u}_{\theta}^{\mathsf{T}} \mathcal{H}_f(\mathbf{x})\, \mathbf{v}_{\phi}$, with $\mathbf{v}_{\phi} = (\cos\phi,\, \sin\phi)$. Note that contrary to the 2-D TV, the resulting functionals in (5) are not equivalent for different choices of L_q-norms. In this paper, we consider two members of this family that arise by employing the mixed norms L_1-L_∞ and L_1-L_2. Interestingly, as we prove in [6], these regularizers involve the spectral and the Frobenius matrix-norms, respectively, and are defined as:

$$\mathcal{R}_{\mathrm{S}}(f) = \int_{\Omega} \left\| \mathrm{D}^2_{\theta,\phi} f(\mathbf{x}) \right\|_{L_\infty(\mathcal{S})} \mathrm{d}\mathbf{x} = \int_{\Omega} \left\| \mathcal{H}_f(\mathbf{x}) \right\|_2 \mathrm{d}\mathbf{x}\,, \tag{6}$$

$$\mathcal{R}_{\mathrm{F}}(f) = \frac{1}{\pi} \int_{\Omega} \left\| \mathrm{D}^2_{\theta,\phi} f(\mathbf{x}) \right\|_{L_2(\mathcal{S})} \mathrm{d}\mathbf{x} = \int_{\Omega} \left\| \mathcal{H}_f(\mathbf{x}) \right\|_F \mathrm{d}\mathbf{x}\,. \tag{7}$$

It is also worthwhile to note that these regularizers are convex, homogeneous, rotation and translation invariant.

The Hessian spectral norm can be alternatively defined as

$$\left\| \mathcal{H}_f(\mathbf{x}) \right\|_2 = \max_{i=1,2} \left(|\lambda_i f(\mathbf{x})| \right)\,, \tag{8}$$

where $\lambda_i f(\mathbf{x})$ are the eigenvalues of the Hessian matrix of f at coordinates $\mathbf{x}$. These two eigenvalues are given by

$$\lambda_{1,2} f(\mathbf{x}) = \frac{1}{2}\Big(\Delta f(\mathbf{x}) \pm \left\| \mathcal{U} f(\mathbf{x}) \right\|_2 \Big)\,, \tag{9}$$

where the associated differential operators are defined in Table 1. Using the identity,

$$\max\left(|\alpha + \beta|\,, |\alpha - \beta| \right) = |\alpha| + \beta\,, \forall \beta \geq 0\,, \tag{10}$$

and combining it with (6), (8) and (9), we finally express our regularizer involving the spectral norm of the Hessian matrix, in an easier-to-minimize form

$$\mathcal{R}_{\mathrm{S}}(f) = \frac{1}{2} \int_{\Omega} \left(|\Delta f(\mathbf{x})| + \left\| \mathcal{U} f(\mathbf{x}) \right\|_2 \right) \mathrm{d}\mathbf{x}\,. \tag{11}$$

Based on (11) we can also interpret this functional as an equally weighted compound regularizer whose first term corresponds to the L_1-norm of the Laplacian.

Regarding the second regularizer, since the Frobenius norm of a matrix is equal to the Euclidian norm of its vectorized version, we equivalently write (7) as

$$\mathcal{R}_{\mathrm{F}}(f) = \int_{\Omega} \left\| \mathcal{V} f(\mathbf{x}) \right\|_2 \mathrm{d}\mathbf{x}\,, \tag{12}$$

where the differential operator $\mathcal{V}$ is defined in Table 1. This last form is more preferable for the description of the minimization algorithm we propose in Section 3.

3. OBJECTIVE FUNCTION MINIMIZATION

3.1. Majorization-Minimization Approach

Hereafter, we will consider the discrete formulation of the image-restoration problem and we will use bold-faced symbols to distinguish between the continuous and discrete domains. Assuming Gaussian noise degrading the measurements, the data term is quadratic and the objective function reads as

$$\mathcal{J}(\mathbf{f}) = \frac{1}{2} \left\| \mathbf{y} - \mathbf{A}\mathbf{f} \right\|_2^2 + \tau \mathcal{R}(\mathbf{f})\,, \tag{13}$$

where $\mathbf{A} \in \mathbb{R}^{N \times N}$ is the convolution matrix describing the blurring operation, and $\mathbf{y}\,, \mathbf{f} \in \mathbb{R}^N$ are the N-dimensional rasterized observed and unknown images, respectively, with $N = m \times n$.

We minimize (13) following a majorization-minimization (MM) approach (cf. [7, 8, 9] for instance). Specifically, we upper bound the data term of our objective function using the following majorizer

$$Q(\mathbf{f}, \mathbf{f}_0) = \frac{1}{2} \left\| \mathbf{y} - \mathbf{A}\mathbf{f} \right\|_2^2 + d(\mathbf{f}, \mathbf{f}_0)\,, \tag{14}$$

where $d(\mathbf{f}, \mathbf{f}_0) = 0.5\,(\mathbf{f} - \mathbf{f}_0)^{\mathsf{T}} \left[\alpha \mathbf{I} - \mathbf{A}^{\mathsf{T}}\mathbf{A} \right] (\mathbf{f} - \mathbf{f}_0)$. In order $Q(\mathbf{f}, \mathbf{f}_0)$ to be a valid majorizer, we need to ensure that $d(\mathbf{f}, \mathbf{f}_0) \geq 0$, $\forall \mathbf{f} \neq \mathbf{f}_0$, with equality iff $\mathbf{f} = \mathbf{f}_0$. This is satisfied if $\alpha \mathbf{I} - \mathbf{A}^{\mathsf{T}}\mathbf{A}$ is positive definite, which implies that $\alpha > \left\| \mathbf{A}^{\mathsf{T}}\mathbf{A} \right\|$. The upper-bounded version of (13) can then be expressed as

$$\begin{aligned} \tilde{\mathcal{J}}(\mathbf{f}, \mathbf{f}_0) &= Q(\mathbf{f}, \mathbf{f}_0) + \tau \mathcal{R}(\mathbf{f}) \\ &= \frac{\alpha}{2} \left\| \mathbf{f} - \mathbf{z} \right\|_2^2 + \tau \mathcal{R}(\mathbf{f}) + c\,, \end{aligned} \tag{15}$$

where $\mathbf{z} = \mathbf{f}_0 + \alpha^{-1}\mathbf{A}^{\mathsf{T}}(\mathbf{y} - \mathbf{A}\mathbf{f}_0)$ and c is a constant. Then we iteratively minimize (15) w.r.t $\mathbf{f}$, setting $\mathbf{f}_0$ to the previous iteration's solution. As we can see, the new objective function does not involve the term $\left\| \mathbf{A}\mathbf{f} \right\|_2^2$ anymore, which turns the minimization task into a much simpler one: the minimizer of (15) is the solution of a denoising problem. Since the convergence of this method can be slow, to speed it up we employ the FISTA algorithm [9], which exhibits state-of-the-art convergence rates. Description of our approach using the monotonic version of FISTA [10] is given in Algorithm 1.

3.2. Denoising Step

The MM formulation of (13) relies on solving the problem

$$\arg\min_{\mathbf{f} \in \mathcal{C}} \frac{1}{2} \left\| \mathbf{f} - \mathbf{z} \right\|_2^2 + \tau \mathcal{R}(\mathbf{f})\,, \tag{16}$$

where $\mathcal{C}$ is a convex set representing additional convex constraints on the solution such as box or positivity constraints, and $\mathcal{R}(\mathbf{f})$ is one of the two studied regularizers. First we describe the proposed minimization algorithm for the regularizer involving the Hessian spectral

Algorithm 1 : Image reconstruction under Hessian-based norm regularization based on MFISTA.

Input: $\mathbf{y}$, $\mathbf{A}$, $\tau > 0$, $\alpha > \left\|\mathbf{A}^{\mathsf{T}}\mathbf{A}\right\|$.
Initialization: $\mathbf{v}_0 = \mathbf{f}_0$, $t_0 = 1$, $c_0 = \mathcal{J}(\mathbf{f}_0)$.
while *stopping criterion is not satisfied* **do**
  $\mathbf{s}_n \leftarrow \texttt{denoise}\left(\mathbf{v}_n + \alpha^{-1}\mathbf{A}^{\mathsf{T}}(\mathbf{y} - \mathbf{A}\mathbf{v}_n), \frac{\tau}{\alpha}\right)$;
  $t_{n+1} \leftarrow \frac{1+\sqrt{1+4t_n^2}}{2}$;
  $c_{n+1} = \mathcal{J}(\mathbf{s}_n)$;
  if $c_{n+1} > c_n$ **then**
    $c_{n+1} = c_n$;
    $\mathbf{f}_{n+1} \leftarrow \mathbf{f}_n$;
    $\mathbf{v}_{n+1} \leftarrow \mathbf{f}_n + \frac{t_n}{t_{n+1}}(\mathbf{s}_n - \mathbf{f}_n)$;
  else
    $\mathbf{f}_{n+1} \leftarrow \mathbf{s}_n$;
    $\mathbf{v}_{n+1} \leftarrow \mathbf{s}_n + \left(\frac{t_n - 1}{t_{n+1}}\right)(\mathbf{s}_n - \mathbf{f}_n)$;
  end
  $n \leftarrow n + 1$;
end
return $\mathbf{f}_n$;

matrix-norm and then we show how it differentiates for the other one involving the Hessian Frobenius matrix-norm.

As we already showed in Sec. 2, we can decompose the Hessian spectral matrix-norm into a linear combination of two vector norms. Thus, we write (16) as

$$\arg\min_{\mathbf{f}\in\mathcal{C}} \frac{1}{2}\|\mathbf{f} - \mathbf{z}\|_2^2 + \frac{\tau}{2}\left(\|\boldsymbol{\Delta}\mathbf{f}\|_1 + \|\mathbf{U}\mathbf{f}\|_{1,2}\right), \tag{17}$$

where $\|\mathbf{u}\|_{1,2} \triangleq \sum_{k=1}^{N}\|\mathbf{u}_i\|_2$ is the ℓ_1-ℓ_2 mixed-norm of $\mathbf{u} \in \mathbb{R}^{N\times 2}$. By introducing the differential operator $\boldsymbol{\Delta}_e$ (see Table 1 for its continuous definition) we write (17) in the equivalent form

$$\arg\min_{\mathbf{f}\in\mathcal{C}} \frac{1}{2}\|\mathbf{f} - \mathbf{z}\|_2^2 + \frac{\tau}{2}\left(\|\boldsymbol{\Delta}_e\mathbf{f}\|_{1,2} + \|\mathbf{U}\mathbf{f}\|_{1,2}\right). \tag{18}$$

Since $\|\,\|_{\infty,2}$ is the dual norm of $\|\,\|_{1,2}$ [11] we express the latter as

$$\|\mathbf{u}\|_{1,2} \triangleq \max_{\boldsymbol{\omega}\in\mathcal{B}} \langle\boldsymbol{\omega}, \mathbf{u}\rangle, \tag{19}$$

where $\mathcal{B} = \left\{\boldsymbol{\omega} = (\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_N) \,|\, \|\boldsymbol{\omega}_k\|_2 \leq 1,\ \forall k = 1, \ldots, N\right\}$ is the ℓ_∞-ℓ_2 unit-norm ball. Using (19) we rewrite (18) as

$$\begin{aligned}&\arg\min_{\mathbf{f}\in\mathcal{C}} \frac{1}{2}\|\mathbf{f} - \mathbf{z}\|_2^2 + \frac{\tau}{2}\left(\max_{\boldsymbol{\omega}_1\in\mathcal{B}}\langle\boldsymbol{\omega}_1, \boldsymbol{\Delta}_e\mathbf{f}\rangle + \max_{\boldsymbol{\omega}_2\in\mathcal{B}}\langle\boldsymbol{\omega}_2, \mathbf{U}\mathbf{f}\rangle\right)\\ &= \arg\min_{\mathbf{f}\in\mathcal{C}} \frac{1}{2}\|\mathbf{f} - \mathbf{z}\|_2^2 + \frac{\tau}{2}\max_{\boldsymbol{\omega}_1,\boldsymbol{\omega}_2\in\mathcal{B}}\left\langle\boldsymbol{\Delta}_e^{\mathsf{T}}\boldsymbol{\omega}_1 + \mathbf{U}^{\mathsf{T}}\boldsymbol{\omega}_2, \mathbf{f}\right\rangle,\end{aligned} \tag{20}$$

where $^{\mathsf{T}}$ denotes the transpose operation. This naturally leads us to the following minimax problem

$$\min_{\mathbf{f}\in\mathcal{C}}\max_{\boldsymbol{\omega}\in\mathcal{B}} \mathcal{L}(\mathbf{f}, \boldsymbol{\omega}) \tag{21}$$

where

$$\mathcal{L}(\mathbf{f}, \boldsymbol{\omega}) = \frac{1}{2}\|\mathbf{f} - \mathbf{z}\|_2^2 + \tau\left\langle\mathbf{U}_\alpha^{\mathsf{T}}\boldsymbol{\omega}, \mathbf{f}\right\rangle, \tag{22}$$

with $\boldsymbol{\omega} = \left[\boldsymbol{\omega}_1^{\mathsf{T}}\,\boldsymbol{\omega}_2^{\mathsf{T}}\right]^{\mathsf{T}}$ and $\mathbf{U}_\alpha = 0.5\left[\boldsymbol{\Delta}_e^{\mathsf{T}}\,\mathbf{U}^{\mathsf{T}}\right]^{\mathsf{T}}$. Since the function $\mathcal{L}(\mathbf{f}, \boldsymbol{\omega})$ is strictly convex in $\mathbf{f}$ and concave in $\boldsymbol{\omega}$ we can exchange the order of the minimum and maximum [12] and get

$$\begin{aligned}\max_{\boldsymbol{\omega}\in\mathcal{B}}\min_{\mathbf{f}\in\mathcal{C}} &\frac{1}{2}\left\|\mathbf{f} - \left(\mathbf{z} - \tau\mathbf{U}_\alpha^{\mathsf{T}}\boldsymbol{\omega}\right)\right\|_2^2\\ &+\frac{1}{2}\|\mathbf{z}\|_2^2 - \frac{1}{2}\left\|\mathbf{z} - \tau\mathbf{U}_\alpha^{\mathsf{T}}\boldsymbol{\omega}\right\|_2^2.\end{aligned} \tag{23}$$

Algorithm 2 : $\texttt{denoise}(\mathbf{z}, \tau)$ – denoising algorithm under Hessian-based norm regularization. For the different regularizers $\boldsymbol{\Psi} = \mathbf{U}_\alpha$ (spectral norm) or $\boldsymbol{\Psi} = \mathbf{V}$ (Frobenius norm).

Input: $\mathbf{z}$, $\boldsymbol{\Psi}$, $\tau > 0$, $\gamma \geq \left\|\boldsymbol{\Psi}^{\mathsf{T}}\boldsymbol{\Psi}\right\|$, $\mathcal{P}_\mathcal{C}$.
Initialization: $\mathbf{v}_1 = \mathbf{0}$, $t_1 = 1$.
Output: $\mathbf{f}^*$ – approximate optimal solution of (16).
while *stopping criterion is not satisfied* **do**
  $\boldsymbol{\omega}_n \leftarrow \mathcal{P}_\mathcal{B}\left(\mathbf{v}_n + \frac{1}{\gamma\tau}\boldsymbol{\Psi}\mathcal{P}_\mathcal{C}\left(\mathbf{z} - \tau\boldsymbol{\Psi}^{\mathsf{T}}\mathbf{v}_n\right)\right)$;
  $t_{n+1} \leftarrow \frac{1+\sqrt{1+4t_n^2}}{2}$;
  $\mathbf{v}_{n+1} \leftarrow \boldsymbol{\omega}_n + \left(\frac{t_n - 1}{t_{n+1}}\right)(\boldsymbol{\omega}_n - \boldsymbol{\omega}_{n-1})$;
  $n \leftarrow n + 1$;
end
return $\mathcal{P}_\mathcal{C}\left(\mathbf{z} - \tau\boldsymbol{\Psi}^{\mathsf{T}}\boldsymbol{\omega}_{n-1}\right)$;

The inner minimization problem is solved exactly by

$$\mathbf{f} = \mathcal{P}_\mathcal{C}\left(\mathbf{z} - \tau\mathbf{U}_\alpha^{\mathsf{T}}\boldsymbol{\omega}\right), \tag{24}$$

leading to the dual formulation

$$\begin{aligned}\max_{\boldsymbol{\omega}\in\mathcal{B}}\Bigg(g(\boldsymbol{\omega}) \;\triangleq\; &\frac{1}{2}\left\|\mathcal{P}_\mathcal{C}\left(\mathbf{z} - \tau\mathbf{U}_\alpha^{\mathsf{T}}\boldsymbol{\omega}\right) - \left(\mathbf{z} - \tau\mathbf{U}_\alpha^{\mathsf{T}}\boldsymbol{\omega}\right)\right\|_2^2\\ &+\frac{1}{2}\|\mathbf{z}\|_2^2 - \frac{1}{2}\left\|\mathbf{z} - \tau\mathbf{U}_\alpha^{\mathsf{T}}\boldsymbol{\omega}\right\|_2^2\Bigg),\end{aligned} \tag{25}$$

where $\mathcal{P}_\mathcal{C}$ is the orthogonal projection onto the convex set $\mathcal{C}$. Contrary to our initial problem, the dual one is smooth and its gradient

$$\nabla g(\boldsymbol{w}) = \tau\mathbf{U}_\alpha\mathcal{P}_\mathcal{C}\left(\mathbf{z} - \tau\mathbf{U}_\alpha^{\mathsf{T}}\boldsymbol{\omega}\right) \tag{26}$$

is well defined. To obtain (26) we use the property that the gradient of a function $s(\boldsymbol{\omega}) = \|\boldsymbol{\omega} - \mathcal{P}_\mathcal{C}(\boldsymbol{\omega})\|_2^2$ is equal to $\nabla s(\boldsymbol{\omega}) = 2(\boldsymbol{\omega} - \mathcal{P}_\mathcal{C}(\boldsymbol{\omega}))$ [10]. The solution of (17) can thus be obtained by first solving the dual problem (25) and then using (24). Since (25) has not a closed-form solution ($\mathbf{U}_\alpha$ has not a stable inverse), we employ Nesterov's iterative method [13] for smooth functions and the operator $\mathcal{P}_\mathcal{B}$ that returns the orthogonal projection onto the convex set $\mathcal{B}$. This method exhibits a state-of-the-art convergence rate $O\left(\frac{1}{n^2}\right)$, where n is the number of iterations. The solution of (16) under the Hessian Frobenius-norm regularizer is obtained in exactly the same way, using formulae (21)–(26) and replacing $\mathbf{U}_\alpha$ with $\mathbf{V}$, which is the discretized version of $\mathcal{V}$. Algorithm 2 describes in detail the proposed denoising approach for both the spectral- and the Frobenius-norm-based regularization.

4. EXPERIMENTS

To validate the effectiveness of the studied second-order regularization framework and the efficiency of the proposed minimization algorithm, we provide experimental results for the task of image deblurring. We compare our results with TV [10] for a set of four standard test-images. In our experiments we use two different *point spead functions* (psf), a uniform and a Gaussian one, both of support 9×9. The standard deviation for the Gaussian psf is $\sigma_b = 6$. In Table 2 we provide numerical results in terms of ISNR for three BSNRs (BSNR $= \text{var}\left[\mathbf{A}\mathbf{f}\right]/\sigma^2$) corresponding to different levels of Gaussian noise. In addition, since the intensities of the original images are in the range of $[0\,1]$, we constrain the solutions in all cases to lie in the convex set $\mathcal{C} = \{\mathbf{f}\,|\mathbf{f}_k \in [0\,1],\ \forall k = 1, \ldots, N\}$. For the sake of fairness among comparisons, the reported results for each regularizer are obtained using the individualized regularization parameter τ

Table 2. ISNR results for the three regularizers under comparison.

Image/ BSNR		Uniform blur: 9×9			Gaussian blur: 9×9		
		TV	Spec.	Frob.	TV	Spec.	Frob.
Boat	20 dB	3.83	3.86	**3.98**	3.45	3.49	**3.59**
	25 dB	4.82	4.86	**5.01**	4.40	4.45	**4.58**
	30 dB	5.99	5.96	**6.13**	5.54	5.52	**5.67**
Fl. Cells	20 dB	3.51	4.21	**4.28**	3.21	3.92	**4.00**
	25 dB	4.32	5.06	**5.15**	3.95	4.72	**4.81**
	30 dB	5.43	6.15	**6.25**	5.05	5.76	**5.87**
Fingerpr.	20 dB	5.19	6.45	**6.54**	5.00	6.22	**6.31**
	25 dB	6.55	7.70	**7.79**	6.30	7.38	**7.47**
	30 dB	8.07	8.85	**8.92**	7.82	8.49	**8.56**
Lena	20 dB	3.79	3.94	**4.04**	3.54	3.70	**3.80**
	25 dB	4.38	4.58	**4.69**	4.06	4.28	**4.39**
	30 dB	5.30	5.48	**5.60**	4.91	5.12	**5.24**

that gives the best ISNR performance. For the discretization of the second-order differential operators, we use forward finite differences with Neumann boundary conditions (reflexive boundaries). Finally, regarding the minimization of the objective functions, we use the proposed algorithm and the method in [10] for our second-order regularizers and TV, respectively, with a stopping criterion set to either reaching a relative normed difference of 10^{-5} between two successive estimates, or a maximum of 100 MFISTA iterations. We use 10 inner iterations to solve the corresponding denoising problem.

With reference to the results of Table 2, the reconstruction under our second-order regularization framework, using either the spectral- or the Frobenius-norm, yields quantitatively better results than TV for the majority of images and different combinations of psfs and noise levels. The efficacy of our approach can also be visually appreciated from the representative deblurring example shown in Fig. 1. In this example, we can ascertain that while TV regularization introduces staircase artifacts mixing structural details of the image, our solutions avoid this problem without compromising the quality of the reconstruction.

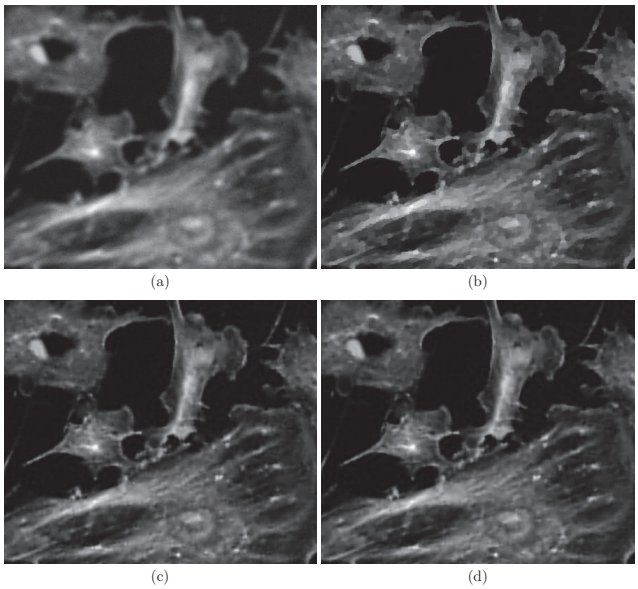

Fig. 1. Results on Fluorescent cell image. Close up of (a) Blurred image (Gaussian psf, BSNR=20 dB), (b) TV reconstruction (ISNR=3.21 dB), (c) Spectral-norm reconstruction (ISNR=3.92 dB), (d) Frobenius-norm reconstruction (ISNR=4.00 dB). The details of this figure are better seen in the electronic version of this paper.

5. CONCLUSIONS

We have proposed an efficient constrained gradient-based minimization algorithm for image restoration under our recently introduced second-order regularization framework. We have also demonstrated that our regularizers are more adapted than TV for the class of images that consist mostly of ridges and smooth transition of intensities, both from a qualitative and quantitative point of view. In particular, our regularizers circumvent staircase effects as well fine-scale-structure deformations that occur with TV and at the same time can restore edges in a satisfactory way.

6. REFERENCES

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

[2] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, pp. 259–268, 1992.

[3] M. Lysaker and X.-C. Tai, "Iterative image restoration combining total variation minimization and a second-order functional," *Int. J. Computer Vision*, vol. 66, pp. 5–18, 2006.

[4] G. Steidl, "A note on the dual treatment of higher-order regularization functionals," *Computing*, vol. 76, pp. 135–148, 2006.

[5] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM J. Imaging Sci.*, vol. 3, no. 3, pp. 492–526, 2010.

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

[7] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.*, vol. 57, pp. 1413–1457, 2004.

[8] M.A.T. Figueiredo, J.M. Bioucas-Dias, and R.D. Nowak, "Majorization–minimization algorithms for wavelet-based image restoration," *IEEE Trans. Image Processing*, vol. 16, pp. 2980–2991, 2007.

[9] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sciences*, vol. 2, pp. 183–202, 2009.

[10] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Processing*, vol. 18, pp. 2419–2434, 2009.

[11] S. Suvrit, "Fast projections onto $\ell_{1,q}$-norm balls for grouped feature selection," in *Proceedings of the European conference on Machine learning and knowledge discovery in databases (ECML PKDD'11)*, 2011, vol. 3, pp. 305–317.

[12] R. T. Rokcafellar, *Convex Analysis*, Princeton, NJ: Princeton Univ. Press, 1970.

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