Image restoration and decomposition via bounded total variation and negative Hilbert-Sobolev spaces

Linh Lieu and Luminita Vese

Department of Mathematics, University of California, Los Angeles
520 Portola Plaza, Los Angeles, CA 90095-1555
Email Addresses : llieu@math.ucla.edu, lvese@math.ucla.edu

UCLA C.A.M. Report 05-33, May 2005

Abstract

In this paper we present a new class of models for image restoration and decomposition by functional minimization. Following Meyer's ideas in a total variation minimization framework of L. Rudin, S. Osher, and E. Fatemi, our model decomposes a given (degraded or textured) image f into a sum $u+v$, where $u \in BV$ is a function of bounded variation (the cartoon component of f), and the noisy (or textured) component v is modeled by tempered distributions belonging to the negative Hilbert-Sobolev space H^{-s}. The proposed models can be seen as generalizations of Osher-Sole-Vese's model and have been motivated also by Mumford-Gidas. We present proofs for existence and uniqueness of solution to our model, as well as characterizations of the solution. We also give a numerical algorithm for solving the minimization problem. And finally, we present various numerical results on denoising, deblurring, and decompositions of both real and synthetic images.

Keywords: functional minimization, functions of bounded variation, negative Hilbert-Sobolev spaces, image restoration, image decomposition, image deblurring, image analysis.

1. Introduction and motivations

Image restoration and decomposition are of important interests in image processing. In the restoration case, one seeks to recover a 'true' image u from an observed, often noisy and/or blurry, image f. Many well-known models for image restoration work by decomposing the observed image f into a sum of two components, e.g. $f = u + v$, where u is a piece-wise smooth geometric component representing the objects in the image and v is the oscillatory component representing the noise. One classical example of such models is proposed by Rudin, Osher, and Fatemi [19] :

$$\inf_{u \in BV(\Omega)} \Big\{ F(u) = \lambda \int_\Omega |\nabla u| + \frac{1}{2} \int_\Omega |f-u|^2 dxdy \Big\}. \tag{1}$$

The first term on the right hand side of (1) is called the *Total Variation (TV) or regularization term*, and the second the *fidelity (fitting) term*. λ is a positive (*tuning*) parameter. The infimum is taken over all functions in $BV(\Omega)$ (functions of bounded variations). The solution to (1) is the geometric component u in the decomposition $f = u + v$. Here, the v component is written as $f - u$. Existence and uniqueness results of this minimization problem can be found in [1], [2], [8], [23].

The ROF model performs very well for removing noise while preserving edges. However, it fails to separate well oscillatory components from high-frequencies components. For example, edges of an object are high-frequencies components and noise are oscillatory components. Both the u and v components in the ROF model contain high-frequencies components. That is, the v component not only contains oscillatory components, it also contains high-frequencies components. One can see shadows of edges in the noise/textured component v, even when the parameter λ is not so small. Alternatively, to remedy this situation, Yves Meyer suggested to replace the L^2-norm in the fidelity term of (1) with a weaker norm more appropriate for modeling textured or oscillatory patterns. He proposed the following minimization problem [16] :

$$\inf_{u\in BV(\Omega)} \Big\{ E(u) = \int_\Omega |\nabla u| + \lambda \parallel f - u \parallel_* \Big\}, \tag{2}$$

where the $\| \cdot \|_*$-norm is defined as follows:

Definition 1.1. Let G denote the Banach space consisting of all generalized functions $f(x,y)$ which can be written as

$$f(x,y) = div\big(\vec{g}(x,y)\big), \quad \vec{g} = (g_1, g_2), \quad g_1, g_2 \in L^\infty(\Omega),\ \vec{g}\cdot\vec{n} = 0 \text{ on } \partial\Omega$$

induced by the norm $\parallel f \parallel_*$ defined as

$$\parallel f \parallel_* = \inf_{f=div(\vec{g})} \parallel \sqrt{g_1(x,y)^2 + g_2(x,y)^2} \parallel_\infty .$$

The space G is the dual to the space $W_0^{1,1}(\Omega) := \{f \ :\ \nabla f \in L^1(\Omega)^2,\ f \equiv 0 \ on\ \partial\Omega\}$. In [16], Meyer also introduces two other spaces, denoted by F and E. The space F is defined as G, but instead of $L^\infty(\Omega)$, g_1, g_2 belong to the John and Nirenberg space $BMO(\Omega)$ (and considered in practice by Le and Vese [15]). Similarly for the space E, where g_1, g_2 belong to the Besov space $B_\infty^{-1,\infty}(\Omega)$. Observe that E coincides with $\dot{B}_\infty^{-1,\infty}(\Omega)$. Oscillatory patterns have smaller norms in G, F and E than in $L^2(\Omega)$ (see [16] for details). Thus, in a minimization problem such as (1) and (2), these spaces are more appropriate for decomposition of an image into a piece-wise smooth (or cartoon) part and a oscillatory (or textured) part than the space $L^2(\Omega)$.

The difficulty with the minimization problem (2) is that it cannot be solved directly due to the form of the $*-norm$. That is to say, the associated Euler-Lagrange equation cannot be expressed directly with respect to u. For a practical alternative, Vese and Osher [24] have introduced the following minimization problem as an approximation to (2):

$$\inf_{u\in BV(\Omega),\vec{g}=(g_1,g_2)\in L^p(\Omega)^2} G_p(u,\vec{g}), \text{ where}$$

$$G_p(u,\vec{g}) = \int_\Omega |\nabla u| + \lambda \int_\Omega |f - (u + div\vec{g})|^2 dxdy + \mu\Big[\int_\Omega \Big(\sqrt{g_1^2 + g_2^2}\Big)^p dxdy\Big]^{\frac{1}{p}}, \tag{3}$$

where $\lambda, \mu > 0$ are tuning parameters, and $p \geq 1$. Here, the unknowns are u, g_1, g_2. The first term in (3) ensures that $u \in BV(\Omega)$, the second term ensures that $v = f - u \approx div(\vec{g})$, and the last term ensures that v is approximately in the space $G_p(\Omega)$, which is given by

$$G_p(\Omega) = \{v = div(\vec{g}), \vec{g} = (g_1, g_2), g_1, g_2 \in L^p(\Omega),\ \vec{g}\cdot\vec{n} = 0 \text{ on } \partial\Omega\}$$

induced by the norm

$$\parallel v \parallel_{G_p(\Omega)} = \inf_{v=div(\vec{g}),\vec{g}\in L^p(\Omega)^2} \|\sqrt{g_1^2 + g_2^2}\|_p.$$

To see that (3) is an approximation to (2), take $\lambda \to \infty$ and $p \to \infty$, then in the limit, $f - u = div(\vec{g})$ a.e., for all those $\vec{g}$ with smallest $L^\infty(\Omega)$ norm, and the middle term disappear and the last term becomes $\| f - u \|_*$.

The space $G_p(\Omega)$ above can be identified with the space $W^{-1,p}(\Omega)$, the dual to the Sobolev space $W_0^{1,p'}(\Omega)$, where $\frac{1}{p} + \frac{1}{p'} = 1$, and the norm $\| v \|_{G_p(\Omega)}$ is a dual norm to the Sobolev norm $\| \nabla u \|_{p'}$. Thus, in the case $p = 2$, the $v = div\vec{g}$ in (3) corresponds to $v \in H^{-1}(\Omega)$, the dual of the space $H_0^1(\Omega)$. Indeed, for $v \in H^{-1}(\Omega)$, there is a unique $P \in H_0^1(\Omega)$ such that

$$\| v \|^2_{H^{-1}(\Omega)} := \inf_{g_1, g_2 \in L^2(\Omega), v = div\vec{g}} \int_\Omega (g_1^2 + g_2^2) dxdy = \int_\Omega |\nabla P|^2 dxdy,$$

that is to say, $v = div(\nabla P) = \triangle P$, or equivalently, $P = \triangle^{-1} v$.

Limiting to the case $p = 2$ and using the inverse Laplacian, Osher, Solé, Vese [20] have modified the minimization problem (3) to a new problem corresponding to the case $\lambda = \infty$ in (3). They proposed the following:

$$\inf_{u \in BV(\Omega)} \Big\{ \int_\Omega |\nabla u| + \lambda |f - u|^2_{H^{-1}(\Omega)} \Big\}, \tag{4}$$

where $|v|^2_{H^{-1}(\Omega)} = \int_\Omega |\nabla \triangle^{-1} v|^2 dxdy$, which is a semi-norm dual to the semi-norm $|u|_{H^1(\Omega)} := \int_\Omega |\nabla u|^2$ of $H^1(\Omega)$. This model yields an exact decomposition of $f = u + v$, with $u \in BV(\Omega)$ and $v \in H^{-1}(\Omega)$. Solving (4) leads to solving an equivalent non-linear fourth order Euler-Lagrange PDE with a severe CFL condition (for details see [20]). Numerical experiments show that the models (3) and (4) separates texture from objects better than the ROF model.

Mumford and Gidas, taking a statistical approach, have proposed in [17] several stochastic models for analysis of generic images. In this paper, they pointed out that Gaussian white noise lives in $\cap_{\varepsilon > 0} H^{-1-\varepsilon}_{loc}$, where H^{-s}_{loc} (for any $s \geq 0$) is the Hilbert-Sobolev space of negative degree of differentiability.

Motivated by the works of Meyer [16], Osher-Solé-Vese [20] and Mumford-Gidas [17], we propose in this paper a new convex minimization problem (or rather, a class of convex minimization problems), in which we decompose a given image into $f = u + v$ where u is of bounded variation and v belongs to one of Hilbert-Sobolev spaces of negative degree of differentiability.

The precise formulation of our proposed model is given in Section 2. As we shall show below, solving the proposed convex minimization problem with respect to the unknown u leads to solving a second-order Euler-Lagrange PDE, instead of a fourth-order PDE. Moreover, the new model is more general than (1) and (4). In fact, it recovers (1) when we set $s = 0$, and it becomes an equivalent form of (4) when we make simple modification in the form of the norm H^{-s}.

The plan of this paper is as follows: in Section 2, we give a detailed description of the proposed model. Existence and uniqueness of solutions and characterizations of minimizers are given in Sections 3 and 4, respectively. In Section 5 we give detailed explanation to our numerical algorithm and implementation of our model. Then in Section 6, we give numerical results and comparisons of these with other models.

One related work has been done by Daubechies and Teschke [9], [10] via wavelets approach. In [9], [10], the authors modified the Vese-Osher and Osher-Solé-Vese energies in (3) and (4) (by replacing $BV(\Omega)$ with the space $B_1^1(L^1(\Omega))$ in the regularizing term, and limiting themselves to the case $p = 2$, whereby they arrived to a new minimization problem which gives a decomposition $f \approx u + v$, with $u \in B_1^1(L^1(\Omega))$ and $v \in H^{-1}(\Omega)$. Other interesting related work for image decomposition and cartoon and texture separation using generalized functions and dual spaces

are by Aujol and collaborators [5], [6], [4], [7], Starck, Elad and Donoho [22], Tadmor, Nezzar and Vese [21], Esedoglu and Osher [13], among others. A related preliminary work inspired from the OSV model[20] and also using the Fourier Transform for computations and variants is by Roudenko [18].

2. Description of the new model

2.1. The space $H^s(\mathbb{R}^n)$ [11].

Definition 2.1. For any $s \in \mathbb{R}$, the space $H^s(\mathbb{R}^n)$ consists of tempered distributions g such that $(1+|\xi|^2)^{s/2}\hat{g} \in L^2(\mathbb{R}^n)$, where $\hat{g}$ is the Fourier transform of g.

The space $H^s(\mathbb{R}^n)$ is a Hilbert space equipped with the inner product

$$< f, g >_s = \int (1+|\xi|^2)^s \hat{f}\bar{\hat{g}} d\xi$$

and the associated norm $\| f \|_s = \sqrt{< f, f >_s}$.

When $s = m$ is an integer, then $H^s(\mathbb{R}^n)$ is the same as the Sobolev Space $H^m(\mathbb{R}^n)$ with equivalent norms. The dual to $H^{-s}(\mathbb{R}^n)$ is the space $H^s(\mathbb{R}^n)$.

Observe that if $s_1 > s_2 \geq 0$, then $\| f \|_{s_1} > \| f \|_{s_2}$. Thus we have the following continuous embeddings (injections) of Hilbert spaces

$$\cdots \subset H^{s_1}(\mathbb{R}^n) \subset H^{s_2}(\mathbb{R}^n) \subset \cdots \subset H^0(\mathbb{R}^n) = L^2(\mathbb{R}^n) \subset \cdots \subset H^{-s_2}(\mathbb{R}^n) \subset H^{-s_1}(\mathbb{R}^n) \cdots$$

2.2. Extending from a bounded domain to $\mathbb{R}^2$.

Often when working with images, we are dealing with functions defined only on a bounded domain. On the other hand, the H^s norm is defined using the Fourier Transform which is defined for functions given on the whole space. To resolve this, we will consider extending a function given on a bounded domain $\Omega \subset \mathbb{R}^2$ by zeros to the whole $\mathbb{R}^2$ space. It is clear that extension by zeros is a continuous embedding of $L^2(\Omega)$ into $L^2(\mathbb{R}^2)$. Moreover, it is an embedding of $BV(\Omega)$ into $BV(\mathbb{R}^2)$ [14]. In addition, since $\partial\Omega$ is Lipschitz, Poincaré-Wirtinger Inequality [14] implies that $BV(\Omega)$ is continuously embedded into $L^2(\Omega)$. We thus have the following continuous embeddings: $BV(\Omega) \subset L^2(\Omega) \subset L^2(\mathbb{R}^2) \subset H^{-s}(\mathbb{R}^2)$.

Henceforth in this paper, when the functions are defined only on $\Omega \subset \mathbb{R}^2$, $\partial\Omega$ is Lipschitz, we will use extension by zeros to compute the H^{-s} norm. We will consider a function $f : \Omega \to \mathbb{R}$ to be in $H^{-s}(\mathbb{R}^2)$, $s > 0$, if after extending by zeros to $\mathbb{R}^2$, the extended function, which we will also call f, belongs to $H^{-s}(\mathbb{R}^2)$.

Note that in our Euler-Lagrange equation given later, for the unknown $u \in BV(\Omega)$, we will first impose $\frac{Du}{|Du|} \cdot \vec{n} = 0$ on $\partial\Omega$ (in the divergence operator computed in the spatial domain) and $u = 0$ outside $\overline{\Omega}$ when we work with the Fourier Transform $\hat{u}$ of u.

2.3. The new model.

Hereafter, s will always be positive, and we will write L^2 and H^{-s} in place of $L^2(\mathbb{R}^2)$ and $H^{-s}(\mathbb{R}^2)$. Consider the following new model for image denoising and decomposition:

$$\inf_{u \in BV(\Omega)} F(u) := \lambda |u|_{BV(\Omega)} + \| f - Ku \|^2_{-s}, \tag{5}$$

where Ω is a bounded domain in $\mathbb{R}^2$ with Lipschitz boundary, $\lambda > 0$ is the tuning parameter, $|u|_{BV(\Omega)} := \int_\Omega |Du|$ is the regularizing (total variation) term, $\| f - Ku \|^2_{-s} = \int_{\mathbb{R}^2} (1+|\xi|^2)^{-s} |\hat{f} -$

$\hat{Ku}|^2\ d\xi$ is the fidelity term. We assume that the operator K is an injective continuous linear operator from L^2 into L^2 such that $K1_\Omega \neq 0$.

Observe that the regularizing term is convex and the fidelity term is strictly convex, thus the energy $F(u)$ is convex. Secondly, since the H^{-s}-norm is bounded by the L^2-norm, the operator K is also continuous in the H^{-s}-norm on L^2 as a subspace of H^{-s}. Hence by Hahn-Banach Theorem K can be extended to a continuous linear operator $K : H^{-s} \to H^{-s}$.

Remarks 2.1. (i) Taking K to be the identity operator and $s = 0$ and applying Parseval Identity $\int |g|^2 dx = \int |\hat{g}|^2 d\xi$, (5) becomes exactly (1), the ROF model.

(ii) We can obtain an equivalent form of (4) from (5) by setting $s = 1$ and writing $\| f-u \|_{H_0^{-s}} := \int |\xi|^{-2s} |\hat{f} - \hat{u}|^2 d\xi$, that is, we change the H^{-s} norm into a semi-norm by removing 1 from the first term in the integrand.

2.4. **The Euler-Lagrange Equation.** We can formally compute the Euler-Lagrange equation for this minimization problem as follows (recall that here $u = 0$ outside $\overline{\Omega}$ when $\hat{u}$ appears, and we will also impose $\frac{Du}{|Du|} \cdot \vec{n} = 0$ on $\partial\Omega$, where $\vec{n}$ is the unit outward normal to $\partial\Omega$).

Take any test function $\varphi \in C^\infty(\Omega)$. Define

$$(6) \quad g(\varepsilon) \ = \ F(u+\varepsilon\varphi) \ := \ \lambda \int_\Omega \big|Du + \varepsilon D\varphi\big| \ + \ \int_{\mathbb{R}^2} \frac{(\hat{f} - \hat{Ku} - \varepsilon\hat{K\varphi})(\bar{\hat{f}} - \bar{\hat{K}}u - \varepsilon\bar{\hat{K}}\varphi)}{(1+|\xi|^2)^s} d\xi.$$

Recall the two identities $\int v\hat{w}dx = \int \hat{v}w dx$ and $\int v\bar{\hat{w}}dx = \int \bar{\hat{v}}w dx$, for any $w(x) \in \mathbb{R}$. Applying these to (6), we obtain

$$\begin{aligned} g'(0) = \ & \lambda \int_\Omega \frac{Du}{|Du|} \cdot D\varphi \ - \ \int_{\mathbb{R}^2} \frac{\bar{\hat{f}} - \bar{\hat{K}}u}{(1+|\xi|^2)^s} \hat{K\varphi} \ - \ \int_{\mathbb{R}^2} \frac{\hat{f} - \hat{Ku}}{(1+|\xi|^2)^s} \bar{\hat{K}}\varphi \\ = \ & \lambda \int_{\partial\Omega} \varphi \frac{Du}{|Du|} \cdot \nu \ ds - \lambda \int_\Omega div\big(\frac{Du}{|Du|}\big)\varphi - \int_{\mathbb{R}^2} \mathfrak{F}\Big(\frac{\bar{\hat{f}} - \bar{\hat{K}}u}{(1+|\xi|^2)^s}\Big) \ K\varphi \\ & - \ \overline{\int_{\mathbb{R}^2} \mathfrak{F}\Big(\frac{\bar{\hat{f}} - \overline{\hat{Ku}}}{(1+|\xi|^2)^s}\Big) \ K\varphi}. \end{aligned}$$

We adopt the convention that $\frac{Du}{|Du|} \cdot \vec{n} = 0$ on $\partial\Omega$. Then the first integral vanishes. The last two integrals are conjugates of each other. So we obtain

$$\begin{aligned} g'(0) = \ & -\lambda \int_\Omega div\big(\frac{Du}{|Du|}\big)\varphi \ - \ 2\int_{\mathbb{R}^2} \mathbf{Re}\Big\{\mathfrak{F}\Big(\frac{\bar{\hat{f}} - \bar{\hat{K}}u}{(1+|\xi|^2)^s}\Big)\Big\} \ K\varphi \\ = \ & -\lambda \int_\Omega div\big(\frac{Du}{|Du|}\big)\varphi \ - \ 2\int_{\mathbb{R}^2} K^*\Big(\mathbf{Re}\Big\{\mathfrak{F}\Big(\frac{\bar{\hat{f}} - \bar{\hat{K}}u}{(1+|\xi|^2)^s}\Big)\Big\}\Big) \ \varphi \end{aligned}$$

for all $\varphi \in C^\infty(\Omega)$, where K^* is the adjoint of K. Thus, the Euler-Lagrange equation is

$$(7) \quad \begin{cases} \lambda div\Big(\frac{Du}{|Du|}\Big) \ + \ K^*\Big(2\ \mathbf{Re}\Big\{\mathfrak{F}\Big(\frac{\bar{\hat{f}} - \bar{\hat{K}}u}{(1+|\xi|^2)^s}\Big)\Big\}\Big) & = \ 0 \quad \text{in } \Omega \\ \frac{Du}{|Du|} \cdot \vec{n} & = \ 0 \quad \text{on } \partial\Omega, \\ u & = \ 0 \quad \text{outside } \overline{\Omega} \end{cases}$$

Remarks 2.2. (i) When $s = 0$ and K is the identity operator, the Euler-Lagrange equation in (7) above is the same as that of the ROF model (because $f - u = \mathfrak{F}(\bar{\hat{f}} - \bar{\hat{u}})$ when f and u are real-valued). This proves that the obtained PDE is consistent with the energy (5), (see Remark 2.1(i) above).

(ii) When the operator K is convolution with a kernel k, (e.g. K is convolution with a Gaussian kernel), then owing to the identity $\widehat{K\varphi} = \hat{k}\hat{\varphi}$, equation (7) becomes

$$\lambda div\Big(\frac{Du}{|Du|}\Big) \;+\; 2\;\mathbf{Re}\Big\{\mathfrak{F}\Big(\frac{\bar{\hat{f}} - \bar{\hat{k}}\bar{\hat{u}}}{(1+|\xi|^2)^s}\hat{k}\Big)\Big\} \;=\; 0.$$

In fact, $\mathfrak{F}\Big(\frac{\bar{\hat{f}}-\bar{\hat{k}}\bar{\hat{u}}}{(1+|\xi|^2)^s}\hat{k}\Big)$ is real-valued, hence we get

$$\lambda div\Big(\frac{Du}{|Du|}\Big) \;+\; 2\;\mathfrak{F}\Big(\frac{\bar{\hat{f}} - \bar{\hat{k}}\bar{\hat{u}}}{(1+|\xi|^2)^s}\hat{k}\Big) \;=\; 0. \tag{8}$$

(iii) If we integrate both sides of (8) over $\mathbb{R}^2$, imposing that $u = 0$ outside of $\overline{\Omega}$ and $\frac{Du}{|Du|}\cdot\vec{n} = 0$ on $\partial\Omega$ ($\vec{n}$ is the outward normal to $\partial\Omega$), then the first term vanishes, and we get

$$\int_{\mathbb{R}^2}\mathfrak{F}\Big(\frac{\bar{\hat{f}} - \bar{\hat{k}}\bar{\hat{u}}}{(1+|\xi|^2)^s}\hat{k}\Big) = 0, \text{ implying that } \mathfrak{F}\mathfrak{F}\Big(\frac{\bar{\hat{f}} - \bar{\hat{k}}\bar{\hat{u}}}{(1+|\xi|^2)^s}\hat{k}\Big)(0) = 0.$$

If we write $w(\xi) = \frac{\hat{f}(\xi)-\hat{k}(\xi)\hat{u}(\xi)}{(1+|\xi|^2)^s}\bar{\hat{k}}(\xi),$ then $\bar{w}(\xi) = w(-\xi)$. Then

$$\begin{aligned}\hat{\bar{w}}(x) &= \int_{\mathbb{R}^2} w(-\xi)e^{-2\pi i\xi\cdot x}d\xi, \text{ (make a change of variables } -\xi\to\xi),\\ &= \int_{\mathbb{R}^2} w(\xi)e^{2\pi i\xi\cdot x}d\xi\\ &= \check{w}(x)\end{aligned}$$

Therefore, $\hat{\bar{w}}(0) = w(0) = \big(\hat{f}(0) - \hat{k}(0)\hat{u}(0)\big)\bar{\hat{k}}(0) = 0$. We impose that $\hat{k}(0) \neq 0$, that is, k has nonzero mean, then $\hat{f}(0) - \hat{u}(0) = 0$, which means the component $v = f - u$ has zero mean.

3. Existence and Uniqueness of Solutions

We now prove existence and uniqueness of minimizers for the proposed model, adapting the techniques from [23], [8], [1] to the (BV, H^{-s}) case.

Theorem 3.1. *Given $\Omega \subset \mathbb{R}^2$ bounded domain with Lipschitz boundary, $f \in H^{-s}$, $f = 0$ outside of $\bar{\Omega}$, $\lambda > 0$, and $K : L^2 \to L^2$ is an injective continuous linear operator such that $K1_\Omega \neq 0$. Then the minimization problem*

$$\inf_{u\in BV(\Omega)} F(u) = \lambda|u|_{BV(\Omega)} \;+\; \parallel f - Ku \parallel^2_{-s}, \quad s > 0,$$

has a unique solution in $BV(\Omega)$.

Proof. Let $u_n \in BV(\Omega)$ be a minimizing sequence. Then there exists a constant $M > 0$ for which $|u_n|_{BV(\Omega)} \;\leq\; M$ and $\parallel f - Ku_n \parallel^2_{-s} \leq\; M$ for all n. By the Poincaré-Wirtinger inequality, we

have a constant $C > 0$ which depends only on Ω such that

$$\| u_n - \frac{1}{|\Omega|}\int_\Omega u_n \|_{L^2(\Omega)} \le C|u_n|_{BV(\Omega)} \le CM, \text{ for all } n. \tag{9}$$

Claim 3.1. *There exists a constant $C' > 0$ such that $\frac{1}{|\Omega|}\left|\int_\Omega u_n\right| \le C', \ \forall n.$*

(We will postpone the proof to this claim until later).

By the Claim, $\| u_n \|_{L^2(\Omega)} \le M', \ \forall n$, implying $\| u_n \|_{L^1(\Omega)} \le M', \ \forall n$. Hence

$$\| u_n \|_{BV(\Omega)} := \| u_n \|_{L^1(\Omega)} + |u_n|_{BV(\Omega)} \le M'', \ \forall n.$$

Therefore there exist a subsequence, still denoted u_n, and a $u \in BV(\Omega)$ such that $u_n \to u$ in $L^1(\Omega)$ and

$$|u|_{BV(\Omega)} \le \liminf_{n\to\infty} |u_n|_{BV(\Omega)}. \tag{10}$$

Moreover, by passing to a subsequence if necessary, $u_n \rightharpoonup u$ weakly in $L^2(\Omega)$. After extending u_n, u by zeros to $\mathbb{R}^2$, we still have $u_n \rightharpoonup u$ weakly in L^2. Since K is a continuous linear operator from L^2 to L^2, $\int_{\mathbb{R}^2} Ku_n(x)\varphi(x)dx = \int_{\mathbb{R}^2} u_n(x)K^*\varphi(x)dx, \quad \forall \varphi \in L^2$. Therefore, $Ku_n \rightharpoonup Ku$ weakly in L^2.

To show $Ku_n \rightharpoonup Ku$ weakly in H^{-s}, we recall that for any $\varphi \in H^{-s}$, $\frac{\hat\varphi(\xi)}{(1+|\xi|^2)^{s/2}} \in L^2(\mathbb{R}^2;\mathbb{C})$, therefore $\frac{\hat\varphi(\xi)}{(1+|\xi|^2)^{s}} \in L^2(\mathbb{R}^2;\mathbb{C})$. So $\mathfrak{F}\Big(\ \frac{\hat\varphi(\xi)}{(1+|\xi|^2)^s}\ \Big) \in L^2(\mathbb{R}^2;\mathbb{C})$. Hence (by subsequently taking the real part and then the imaginary part and combining them afterwards)

$$\int_{\mathbb{R}^2} \overline{\mathfrak{F}\Big(\ \frac{\hat\varphi(\xi)}{(1+|\xi|^2)^s}\ \Big)}(x)Ku_n(x)dx \to \int_{\mathbb{R}^2} \overline{\mathfrak{F}\Big(\ \frac{\hat\varphi(\xi)}{(1+|\xi|^2)^s}\ \Big)}(x)Ku(x)dx \quad \text{as } n\to\infty.$$

Applying $\int v\bar{\hat w} = \int \bar{\hat v} w$, we thus have for any $\varphi \in H^{-s}$,

$$\int_{\mathbb{R}^2} \frac{\hat\varphi\,\overline{\widehat{Ku_n}}}{(1+|\xi|^2)^s}d\xi \to \int_{\mathbb{R}^2} \frac{\hat\varphi\,\overline{\widehat{Ku}}}{(1+|\xi|^2)^s}d\xi \quad \text{as } n\to\infty.$$

Therefore, $f - Ku_n \rightharpoonup f - Ku$ weakly in H^{-s} and by the lower semi-continuity of the H^{-s}-norm, we get

$$\| f - Ku \|^2_{-s} \le \liminf_{n\to\infty} \| f - Ku_n \|^2_{-s}.$$

Hence,

$$F(u) := \lambda|u|_{BV(\Omega)} + \| f - Ku \|^2_{-s} \le \liminf_{n\to\infty} \Big(\lambda|u_n|_{BV(\Omega)} + \| f - Ku_n \|^2_{-s}\Big).$$

So u is a minimizer.

For uniqueness, suppose $u, v \in BV(\Omega)$ are 2 minimizers : $F(u) = F(v) = \inf_{w\in BV(\Omega)} F(w)$. Suppose $Ku \neq Kv$. We have

$$\begin{aligned} \| f - \frac{1}{2}Ku - \frac{1}{2}Kv \|^2_{-s} &= \frac{1}{4}\| f - Ku \|^2_{-s} + \frac{1}{2}\mathbf{Re} < f - Ku, f - Kv >_{-s} + \frac{1}{4}\| f - Kv \|^2_{-s} \\ &< \frac{1}{2}\| f - Ku \|^2_{-s} + \frac{1}{2}\| f - Kv \|^2_{-s}. \end{aligned}$$

Then $\| f-\frac{1}{2}Ku-\frac{1}{2}Kv \|^2_{-s}+\lambda|\frac{1}{2}(u+v)|_{BV(\Omega)} \leq \| f-\frac{1}{2}Ku-\frac{1}{2}Kv \|^2_{-s}+\lambda\frac{1}{2}(|u|_{BV(\Omega)}+|v|_{BV(\Omega)}) < \frac{1}{2}\| f-Ku \|^2_{-s}+\frac{1}{2}\| f-Kv \|^2_{-s}+\lambda\frac{1}{2}(|u|_{BV(\Omega)}+|v|_{BV(\Omega)}) = inf_{w\in BV(\Omega)}F(w)$. This implies $F(\frac{1}{2}(u+v)) < F(u)=F(v)$, which cannot be so if u and v are minimizers.

Thus, $Ku=Kv$. Since K is injective, $u=v$. □

Proof. of Claim 3.1 (we follow [1], [23])

Denote by $w_n = \left(\frac{1}{|\Omega|}\left|\int_\Omega u_n\right|\right)\chi_\Omega$ and $v_n = u_n - w_n$, where u_n is a minimizing sequence as above. Then clearly $w_n, v_n \in BV(\Omega)$. Moreover, we have from (9)

$$\| v_n \|_{-s} \leq \| v_n \|_{L^2} = \| u_n - \frac{1}{|\Omega|}\int_\Omega u_n \|_{L^2(\Omega)} \leq C,$$

for some constant $C>0$ independent of n, for all n. Thus,

$$\begin{aligned}
M &\geq \| f-Ku_n \|^2_{-s} \\
&= \| f-Kv_n-Kw_n \|^2_{-s} \\
&= \| f-Kv_n \|^2_{-s} + \| Kw_n \|^2_{-s} - 2\,\mathbf{Re} < f-Kv_n, Kw_n >_{-s} \\
&\geq \| Kw_n \|^2_{-s} - 2\| f-Kv_n \|_{-s}\| Kw_n \|_{-s} \\
&= \| Kw_n \|_{-s}\left(\| Kw_n \|_{-s} - 2\| f-Kv_n \|_{-s}\right). \\
&\geq \| Kw_n \|_{-s}\left(\| Kw_n \|_{-s} - 2\left(\| f \|_{-s} + \| K \|\| v_n \|_{-s}\right)\right).
\end{aligned}$$

Denote by $x_n = \| Kw_n \|_{-s}$ and $a_n = \| f \|_{-s} + \| K \|\| v_n \|_{-s}$. Then

$$x_n(x_n-2a_n) \leq M, \text{ and } 0 \leq a_n \leq \| f \|_{-s} + \| K \| \cdot C = M', \ \forall n.$$

Hence $0 \leq x_n \leq a_n + \sqrt{a_n^2+M} \leq M''$. This implies

$$\frac{1}{|\Omega|}\left|\int_\Omega u_n\right| \| K\chi_\Omega \|_{-s} = \| Kw_n \|_{-s} \leq M'', \ \forall n.$$

Therefore, $\frac{1}{|\Omega|}\left|\int_\Omega u_n\right| \leq C', \ \forall n.$ □

Remark 3.1. The above existence and uniqueness result also holds for other regularizing functionals on $BV(\Omega)$ instead of the total variation. For example, $|u|_{BV(\Omega)}$ can be substituted by $\int_\Omega \phi(Du)$ defined in the sense of convex functions of measures (see [12], [23]), where $\phi : \mathbb{R}^2 \to [0,\infty)$ is continuous, even, convex, and satisfying $\phi(0)=0$, $a|x|-b \leq \phi(x) \leq a|x|+b$ for some constants $a>0$ and $b \geq 0$, and any $x \in \mathbb{R}^2$. Examples are $\phi(Du) = |u_{x_1}| + |u_{x_2}|$, $\phi(Du) = \sqrt{\alpha+|Du|^2}$, $\phi(Du) = \log\cosh(\alpha+|Du|^2)$, with $\alpha>0$.

4. Characterization of Solutions

In this section we present some theoretical results characterizing the minimizer of the new model. Similar calculations have been made by Y. Meyer in [16] for the ROF model (see also [3] for related discussion).

We begin by defining a semi-norm $\| \cdot \|_*$ on H^{-s} dual to the $|\cdot|_{BV}$ norm as follows:

$$\| f \|_* = \sup_{u\in BV(\Omega),\ |u|_{BV(\Omega)}\neq 0} \frac{|\mathbf{Re} < f, u >_{-s}|}{|u|_{BV(\Omega)}}, \quad f \in H^{-s}. \tag{11}$$

$\| \cdot \|_*$ in (11) is indeed a semi-norm since for any $f, g \in H^{-s}$ and $u \in BV(\Omega)$, $\mathbf{Re} < f+g, u >_{-s} = \mathbf{Re} < f, u >_{-s} + \mathbf{Re} < g, u >_{-s}$, so $|\mathbf{Re} < f+g, u >_{-s}| \leq |\mathbf{Re} < f, u >_{-s}$

$| + |\mathbf{Re} < g, u >_{-s} |$. This implies the triangle inequality. Moreover, it is clear that for any $\lambda \in \mathbb{R}$, $\| \lambda f \|_* = |\lambda| \| f \|_*$.

Lemma 4.1. *If $f \in H^{-s}$ is such that $\|f\|_* < \infty$, then $\mathbf{Re} < f, 1_\Omega >_{-s} = 0$.*

Proof. Let $u \in BV(\Omega)$ be such that $|u|_{BV(\Omega)} \neq 0$. Then for any constant $c \in \mathbb{R}$

$$\frac{|\mathbf{Re} < f, u + c >_{-s} |}{|u + c|_{BV(\Omega)}} = \frac{|\mathbf{Re} < f, u + c >_{-s} |}{|u|_{BV(\Omega)}} = \frac{|\mathbf{Re}(< f, u >_{-s} + < f, c >_{-s})|}{|u|_{BV(\Omega)}} \leq \|f\|_* < \infty.$$

This implies that there is a constant C such that $|\mathbf{Re} < f, c >_{-s} | \leq C < \infty$ for any $c \in \mathbb{R}$. Then $|c||\mathbf{Re} < f, 1_\Omega >_{-s} | \leq C < \infty$ for any $c \in \mathbb{R}$, therefore $|\mathbf{Re} < f, 1_\Omega >_{-s} |$ must be zero. □

Remark 4.1. The above Lemma implies that if $\|f\|_* < \infty$ is defined by (11), then $|\mathbf{Re} < f, u >_{-s} | \leq |u|_{BV(\Omega)} \|f\|_*$ for any $u \in BV(\Omega)$.

Theorem 4.1. *Fix $\lambda > 0$. Let u be a minimizer of (5) and set $v := f - Ku$. Then the following holds,*

(I) $\| K^* f \|_* \leq \frac{\lambda}{2}$ *if and only if* $u = 0, v = f$.

(II) *Suppose* $\| K^* f \|_* > \frac{\lambda}{2}$. *Then $u \in BV(\Omega), v = f - Ku$ is minimizer if and only if* $\| K^* v \|_* \leq \frac{\lambda}{2}$ *and* $\mathbf{Re} < K^* v, u >_{-s} = \frac{\lambda}{2} |u|_{BV}$. *In both cases, if in addition $|u|_{BV(\Omega)} \neq 0$ then $\|K^* v\|_* = \frac{\lambda}{2}$.*

Proof. (I) The minimizer is $u = 0, v = f$ if and only if $\forall h \in BV(\Omega)$ and $\forall \varepsilon \in \mathbb{R}, \varepsilon \neq 0$,

$$\lambda |\varepsilon h|_{BV(\Omega)} + \| f - \varepsilon K h \|_{-s}^2 \geq \| f \|_{-s}^2 .$$

Expanding the term $\| f - \varepsilon K h \|_{-s}^2$, we get

$$\lambda |\varepsilon| |h|_{BV(\Omega)} - 2\varepsilon \, \mathbf{Re} < f, Kh >_{-s} + \varepsilon^2 \| Kh \|_{-s}^2 \geq 0.$$

If $\varepsilon > 0$: $\lambda |\varepsilon| |h|_{BV(\Omega)} + \varepsilon^2 \| Kh \|_{-s}^2 \geq 2|\varepsilon| \mathbf{Re} < f, Kh >_{-s}$.
If $\varepsilon < 0$: $\lambda |\varepsilon| |h|_{BV(\Omega)} + \varepsilon^2 \| Kh \|_{-s}^2 \geq -2|\varepsilon| \mathbf{Re} < f, Kh >_{-s}$.
Hence,

$$\lambda |\varepsilon| |h|_{BV(\Omega)} + \varepsilon^2 \| Kh \|_{-s}^2 \geq 2|\varepsilon| \big| \mathbf{Re} < f, Kh >_{-s} \big|.$$

Dividing both sides by $|\varepsilon|$ and taking $|\varepsilon| \to 0$, we obtain

$$\lambda |h|_{BV(\Omega)} \geq \big| \mathbf{Re} < f, Kh >_{-s} \big| = \big| \mathbf{Re} < K^* f, h >_{-s} \big|, \text{ for all } h \in BV(\Omega). \tag{12}$$

Hence, $\| K^* f \|_* \leq \frac{\lambda}{2}$.

Conversely, if $\| K^* f \|_* \leq \frac{\lambda}{2}$, then by the remark after the Lemma, the last inequality (12) holds for any $h \in BV(\Omega)$. From here, add $\| f \|_{-s}^2$ to both sides of (12) and $\| Kh \|_{-s}^2$ to the left hand side, we obtain

$$\lambda |h|_{BV(\Omega)} + \| f \|_{-s}^2 - 2 \, \mathbf{Re} < f, Kh >_{-s} + \| Kh \|_{-s}^2 \geq \| f \|_{-s}^2 .$$

That is,

$$\lambda |h|_{BV(\Omega)} + \| f - Kh \|_{-s}^2 \geq \| f \|_{-s}^2 .$$

This is true for all $h \in BV(\Omega)$, thus $u = 0$ is the minimizer.

(II) We have $u \in BV(\Omega), v = f - Ku$ is minimizer if and only if $\forall h \in BV(\Omega), \forall \varepsilon \in \mathbb{R}$:

$$\lambda |u + \varepsilon h|_{BV(\Omega)} + \| v - \varepsilon K h \|_{-s}^2 \geq \lambda |u|_{BV(\Omega)} + \| v \|_{-s}^2 .$$

Then,

$$\lambda |\varepsilon| |h|_{BV(\Omega)} - 2\varepsilon \, \mathbf{Re} < v, Kh >_{-s} + \varepsilon^2 \| Kh \|_{-s}^2 \geq 0,$$

for any $\varepsilon \in \mathbb{R}$, hence

$$\lambda|\varepsilon||h|_{BV(\Omega)} + \varepsilon^2 \parallel Kh \parallel^2_{-s} \geq \ 2|\varepsilon||\mathbf{Re} < K^*v, h >_{-s}|.$$

Divide both sides by $|\varepsilon|$ and taking $\varepsilon \to 0$, we get

$$\frac{\lambda}{2}|h|_{BV(\Omega)} \ \geq \ |\mathbf{Re} < K^*v, h >_{-s}|, \quad \forall h \in BV(\Omega).$$

Hence, $\parallel K^*v \parallel_* \leq \ \frac{\lambda}{2}$.

Repeat the same calculation with specifically $h = u$ and $-1 < \varepsilon < 0$ (so that $1 + \varepsilon \ > \ 0$), we get

$$\lambda(1+\varepsilon)|u|_{BV(\Omega)} + \parallel v \parallel^2_{-s} -2\varepsilon \ \mathbf{Re} < v, Ku >_{-s} + \varepsilon^2 \parallel Ku \parallel^2_{-s} \geq \ \lambda|u|_{BV(\Omega)} + \parallel v \parallel^2_{-s}.$$

$$\Longrightarrow \quad \lambda\varepsilon|u|_{BV(\Omega)} - 2\varepsilon \ \mathbf{Re} < K^*v, u >_{-s} + \varepsilon^2 \parallel Ku \parallel^2_{-s} \geq \ 0.$$

Since $\varepsilon < 0$, this is the same as

$$-\lambda|\varepsilon||u|_{BV(\Omega)} + 2|\varepsilon|\mathbf{Re} < K^*v, u >_{-s} + \varepsilon^2 \parallel Ku \parallel^2_{-s} \geq \ 0.$$

Divide both sides by $|\varepsilon|$ and then let $\varepsilon \nearrow 0$, we obtain

$$2 \ \mathbf{Re} < K^*v, u >_{-s} \geq \ \lambda|u|_{BV(\Omega)}.$$

Since $|\mathbf{Re} < K^*v, u >_{-s} \ | \leq \ \frac{\lambda}{2}|u|_{BV(\Omega)}$, we get $|\mathbf{Re} < K^*v, u >_{-s}| = \frac{\lambda}{2}|u|_{BV(\Omega)}$. If $|u|_{BV(\Omega)} \neq 0$, then $\parallel K^*v \parallel_* = \frac{\lambda}{2}$.

Conversely, suppose $u \in BV(\Omega)$ and $v = f - Ku$ satisfy $\parallel K^*v \parallel_* \leq \frac{\lambda}{2}$ (with equality if $|u|_{BV(\Omega)} \neq 0$) and $\mathbf{Re} < K^*v, u >_{-s} = \frac{\lambda}{2}|u|_{BV(\Omega)}$. Then $\forall h \in BV(\Omega)$ and $\varepsilon \in \mathbb{R}$,

$$\begin{aligned}
&\lambda|u + \varepsilon h|_{BV(\Omega)} + \parallel v - \varepsilon Kh \parallel^2_{-s} \\
&\geq 2 \ \mathbf{Re} < K^*v, u + \varepsilon h >_{-s} + \parallel v \parallel^2_{-s} -2\varepsilon \ \mathbf{Re} < v, Kh >_{-s} + \varepsilon^2 \parallel Kh \parallel^2_{-s} \\
&= \parallel v \parallel^2_{-s} + 2 \ \mathbf{Re} < K^*v, u >_{-s} + \varepsilon^2 \parallel Kh \parallel^2_{-s} \\
&\geq \parallel v \parallel^2_{-s} + \lambda|u|_{BV(\Omega)}.
\end{aligned}$$

Therefore, u is minimizer. □

Remark 4.2. Note that if $\mathbf{Re} < f, K1_\Omega >_{-s} = 0$ (can be simply obtained by subtracting a constant from f), then in part (II) above we always have $|u|_{BV(\Omega)} \neq 0$ (because if u is a constant minimizer, then u must be zero in this case, but this cannot hold in part (II)).

Remark 4.3. The above characterization of minimizers holds if the total variation $|u|_{BV(\Omega)}$ is substituted by another functional Φ on $BV(\Omega)$ that is convex, lower semi-continuous and positive homogeneous of degree 1.

5. Numerical approximation of the model

We describe here the algorithm that we use in our numerical experiments. We work only with the case when the operator K is the convolution with a kernel k, i.e. the functional that we will minimize is

$$\inf_{u \in BV(\Omega)} F(u) = \lambda \int_\Omega |Du|dx + \int_{\mathbb{R}^2} \frac{|\hat{f} - \hat{k}\hat{u}|^2}{(1 + |\xi|^2)^s} d\xi. \tag{13}$$

However, since the Total Variation is not differentiable at 0, we approximate our functional by the following regularized functional

$$\inf_{u_\varepsilon \in BV(\Omega)} F_\varepsilon(u_\varepsilon) = \lambda \int_\Omega \sqrt{\varepsilon^2 + |Du_\varepsilon|^2} dx + \int_{\mathbb{R}^2} \frac{|\hat{f} - \hat{k}\hat{u}_\varepsilon|^2}{(1+|\xi|^2)^s} d\xi \tag{14}$$

Remark 5.1. It can be shown that the regularized minimization problem in (14) has a unique solution for each $\lambda > 0$ and $\varepsilon > 0$ by repeating the steps we did for the original minimization problem. In addition, following the arguments in [8], it can be shown that as ε converges to zero, the unique solution of (14) converges to the solution of (13).

Applying gradient descent method to minimize (14), we arrive to the following non-linear PDE:

$$\begin{cases} \frac{\partial u}{\partial t} = \lambda div\Big(\frac{\nabla u}{\sqrt{\varepsilon^2+|\nabla u|^2}}\Big) \; + \; 2\,\mathfrak{F}\Big(\frac{\bar{\hat{f}}-\bar{\hat{k}}\bar{\hat{u}}}{(1+|\xi|^2)^s}\hat{k}\Big) & \text{in } \Omega \\ \frac{\nabla u}{\sqrt{\varepsilon^2+|\nabla u|^2}} \cdot \vec{n} \; = \; 0 & \text{on } \partial\Omega \\ u = 0 \text{ outside } \overline{\Omega}. \end{cases} \tag{15}$$

To proceed with the discrete numerical algorithm for solving (15), let us simplify things a little by assuming that the initial discrete image f is of $M \times M$ pixels, and that it has been sampled from its continuous version at uniformly spaced points starting at $(0,0)$, i.e. $f_{i,j} = f(i\triangle x, j\triangle x)$ for $i, j = 0, 1, \cdots, M-1$, where $\triangle x$ is to be determined.

5.1. **The "force" term.** Computing the force term $\mathfrak{F}\Big(\frac{\bar{\hat{f}}-\bar{\hat{k}}\bar{\hat{u}}}{(1+|\xi|^2)^s}\hat{k}\Big)$ in the PDE (15) requires the Discrete Fourier Transform (DFT), which is defined to be

$$\hat{f}_{m,n} = \frac{1}{\sqrt{M}} \sum_{i,j=0}^{M-1} f_{i,j} e^{-2\pi\sqrt{-1}(im+jn)/M}, \quad \text{for } m, n = 0, 1, \cdots, M-1.$$

And the Inverse Discrete Fourier Transform (IDFT) is

$$\check{f}_{i,j} = \frac{1}{\sqrt{M}} \sum_{i,j=0}^{M-1} f_{i,j} e^{2\pi\sqrt{-1}(im+jn)/M}, \quad \text{for } m, n = 0, 1, \cdots, M-1.$$

The DFT array $\hat{f}_{m,n}$ is, indeed, as taken from its continuous counterpart at frequencies $(m\triangle\xi, n\triangle\xi)$, for $m, n = 0, 1, ..., M-1$. The definitions of the DFT and IDFT imply that $\triangle x$ and $\triangle\xi$ are inversely related by

$$\triangle\xi = \frac{1}{M\triangle x}.$$

Therefore, to give our numerical computations a balance weight between the spatial and the Fourier frequency terms, we choose $\triangle x = \frac{1}{\sqrt{M}}$ and $\triangle\xi = \frac{1}{\sqrt{M}}$.

When computing the force term, we first multiply $(-1)^{i+j}$ to $f_{i,j}$ and $k_{i,j}$, $i, j = 0, 1, \cdots, M-1$, before taking the DFT. This shifts the origin of the frequency domain to the center of the image. Thus, as m varies from 0 to $M-1$, and n from 0 to $M-1$, the coefficient $\hat{f}_{m,n}$ corresponds to the Fourier coefficient at frequency $((m - \frac{M}{2})\triangle\xi, (n - \frac{M}{2})\triangle\xi)$. Hence, we evaluate the weight function $\frac{1}{(1+|\xi|^2)^s}$ at points $\xi_1, \xi_2 = -\frac{M\triangle\xi}{2}, -\frac{(M-1)\triangle\xi}{2}, \cdots, \frac{(M-1)\triangle\xi}{2}$.

5.2. **The curvature term.** For the discrete gradient, we shall use the following usual notations:

$$\nabla^{+,+}u = (\nabla_x^+ u, \nabla_y^+ u), \qquad \nabla^{+,-}u = (\nabla_x^+ u, \nabla_y^- u)$$
$$\nabla^{-,+}u = (\nabla_x^- u, \nabla_y^+ u), \qquad \nabla^{-,-}u = (\nabla_x^- u, \nabla_y^- u)$$

where

$$\nabla_x^+ u = u_{i+1,j} - u_{i,j}, \qquad \nabla_x^- = u_{i,j} - u_{i-1,j}$$
$$\nabla_y^+ u = u_{i,j+1} - u_{i,j}, \qquad \nabla_y^- = u_{i,j} - u_{i,j-1}.$$

Since the dual operators to $\nabla^{+,+}$, $\nabla^{+,-}$, $\nabla^{-,+}$, $\nabla^{-,-}$ are, respectively, the operators $div^{-,-}$, $div^{-,+}$, $div^{+,-}$, $div^{+,+}$, we can write the regularized curvature term in one of four ways:

$$div\Big(\frac{\nabla u}{\sqrt{\varepsilon^2 + |\nabla u|^2}}\Big) \approx div^{\alpha*,\beta*}\Big(\frac{\nabla^{\alpha,\beta} u}{\sqrt{\varepsilon^2 + |\nabla^{\alpha,\beta} u|^2}}\Big) \tag{16}$$

where $div^{\alpha*,\beta*}$ denotes the dual operator of $\nabla^{\alpha,\beta}$, with $\alpha, \beta = +, -$.

To make our numerical scheme rotationally invariant, we use all four approximations to the gradient operator by alternating them with each iteration [3]. For example, if u^n was computed using $\nabla^{+,+}$, then we use $\nabla^{+,-}$ to compute u^{n+1}, $\nabla^{-,+}$ to compute u^{n+2}, and $\nabla^{-,-}$ to compute u^{n+3}, and then repeat.

We note here that we compute the curvature term for the ROF model in the same way.

5.3. **Numerical Algorithm.** We solve (15) with the following iterative semi-implicit scheme [23]:

1. u^0 is arbitrarily given (we can take $u^0 = f$)
2. Once u^n is calculated, compute the forcing term $F^n = \mathfrak{F}\Big(\frac{\bar{\hat{f}} - \bar{\hat{k}}\bar{\hat{u}}^n}{1+|\xi|^2}\hat{k}\Big)$
3. Compute $u_{i,j}^{n+1}$, for $i, j = 1, 2, \cdots, M-2$ as the solution of the linear discrete equation:
   $u_{i,j}^{n+1} = u_{i,j}^n + \triangle t\Big(\frac{\lambda}{\triangle x} div^{\alpha*,\beta*}\Big(\frac{\nabla^{\alpha,\beta} u^{n+1}}{\sqrt{\varepsilon^2 + |\nabla^{\alpha,\beta} u^n|^2}}\Big)_{i,j} + 2F_{i,j}^n\Big)$
   with $\varepsilon = \varepsilon' \triangle x$, some $\varepsilon' > 0$ small, and the boundary conditions
   $u_{0,j}^{n+1} = u_{1,j}^{n+1}, \quad u_{M-1,j}^{n+1} = u_{M-2,j}^{n+1}, \quad u_{i,0}^{n+1} = u_{i,1}^{n+1}, \quad u_{i,M-1}^{n+1} = u_{i,M-2}^{n+1}$.

To clarify the notations in step 3 of our algorithm, assume that we are solving at pixel (i, j) the equation

$$u_{i,j}^{n+1} = u_{i,j}^n + \triangle t\Big(\frac{\lambda}{(\triangle x)} div^{-,-}\Big(\frac{\nabla^{+,+} u^{n+1}}{\sqrt{\varepsilon^2 + |\nabla^{+,+} u^n|^2}}\Big)_{i,j} + 2F_{i,j}^n\Big).$$

Let $d_{i,j} = (\sqrt{\varepsilon^2 + |\nabla^{+,+} u^n|^2})_{i,j}$, which is known since u^n is already computed. Then

$$div^{-,-}\Big(\frac{\nabla^{+,+} u^{n+1}}{\sqrt{\varepsilon^2 + |\nabla^{+,+} u^n|^2}}\Big)_{i,j} = \frac{u_{i+1,j}^n - u_{i,j}^{n+1}}{d_{i,j}} - \frac{u_{i,j}^{n+1} - u_{i-1,j}^n}{d_{i-1,j}} + \frac{u_{i,j+1}^n - u_{i,j}^{n+1}}{d_{i,j}} - \frac{u_{i,j}^{n+1} - u_{i,j-1}^n}{d_{i,j-1}}$$

Basically, we set all terms at the current pixel (i, j) in the curvature term to be unknown. Setting $c_{i,j} = \frac{\lambda \triangle t}{d_{i,j}(\triangle x)}$, then we have

$$(1 + c_{i-1,j} + 2c_{i,j} + c_{i,j-1}) u_{i,j}^{n+1} = u_{i,j}^n + c_{i-1,j} u_{i-1,j}^n + c_{i,j} u_{i+1,j}^n + c_{i,j} u_{i,j+1}^n + c_{i,j-1} u_{i,j-1}^n + 2\triangle t F_{i,j}^n.$$

Hence, $u_{i,j}^{n+1}$ is obtained by dividing the coefficient $(1 + c_{i-1,j} + 2c_{i,j} + c_{i,j-1})$ to both sides of the equation. We remark here that the semi-implicit scheme above approaches the steady state equation much faster than an explicit scheme, and hence, an advantage.

5.4. **The blurring kernel.** In our implementation, we perform blurring directly in the Fourier frequency domain. Since the Fourier Transform of a Gaussian is again a Gaussian, to blur an image, we multiply (pixel-wise) a Gaussian kernel to the DFT of the image, then we take the inverse transform. For a Gaussian kernel we used the form

$$k(\xi_1, \xi_2) = exp\Big(- \frac{\xi_1^2 + \xi_2^2}{2/\alpha^2}\Big).$$

Thus, in our numerical computation, we never compute a discrete convolution.

6. Numerical results for image restoration

In this section we present numerical results obtained by applying our proposed new model to image denoising, deblurring and decomposition. For comparison, we also present results from the ROF model in [19] performed on the same images. In our implementation, s is kept as a parameter. We will show numerical results obtained with various values of s.

For denoising, the parameter λ was chosen so that the best residual-mean-squared-error ($RMSE$) is obtained. For the $RMSE$ and SNR, we used the expressions

$$RMSE = \frac{\sqrt{\sum \big(u_{i,j} - u_{org,i,j}\big)^2}}{MN}, \qquad SNR = \frac{\sum \big(u_{i,j} - \frac{\sum u_{i,j}}{MN}\big)^2}{\sum \big(u_{i,j} - u_{org,i,j} - \frac{\sum (u_{i,j} - u_{org,i,j})}{MN}\big)^2}.$$

In Figure 2, we show the denoising results obtained from our proposed new model with H^{-1} and H^{-0} norms in the fidelity and the ROF model performed on a synthetic piece-wise constant image with additive Gaussian white noise of $\sigma = 30$ (shown in Fig. 1). In Figure 3, we show more results on the same image from our proposed model using $s = 0.5$ and $s = 2$, respectively, for the H^{-s} norm in the fidelity term. We also show the results from our model using H_0^{-1} semi-norm for the fidelity term (as discussed in Remark 2.1, equivalent case with the model from [20]). The $RMSE$ and SNR, together with the value of the parameter λ, for all results in this experiment are shown in Table 1. Comparing the results, our model with $H^{-0.5}$ norm gives results with the best $RMSE$, while the H_0^{-1} semi-norm gives results with the best SNR. Visually, the results from our model with H^{-1} norm preserves best the edges in the u component. Overall, our proposed model performs much better than the ROF model, as expected.

In Figure 5 we show results from another denoising application. We applied our model using $H^{-0.5}$ and H^{-1} and the ROF model to a noisy image of a woman. The noise is additive white noise with standard deviation $\sigma = 20$ (see Fig. 4). The $RMSE$ and SNR of the original noisy image is 0.0762599 and 8.01743, respectively. After denoising, the best results from the ROF, $H^{-0.5}$, and H^{-1} have the SNR of 30.64769, 32.97592, 32.82501, and the $RMSE$ of 0.03548729, 0.03461537, 0.03472017, respectively. The values for λ that yield these results are $14, 0.4, 0.25$ for the ROF,$H^{-0.5}, H^{-1}$ model, respectively. Again, our model yields better results than the ROF model.

We also try texture removal with our H^{-1}-model and compare with ROF model. The results are shown in Figures 6, 7. In Figure 6, the texture image is synthetically created, and in Figure 7, the image is natural. We can see that the H^{-1}-norm models better the texture than the L^2-norm.

In Figures 8, 9, we show deblurring results on a synthetic image and a natural image of an office. The blur is done as described in Sect. 5.4 with $\alpha = 0.8$. For the result in Figure 8, the blurred image has $RMSE = 0.1016$. The improved image has $RMSE = 0.03618513$ using the parameter $\lambda = 0.0011$. For the result in Figure 9, the blurred image has $RMSE = 0.181955$,

Restoration Model	λ	$RMSE$	SNR
Original Image		0.224578	4.03064
ROF	55	0.0526239	51.99021
H^0	9.3	0.05204916	53.11084
$H^{-0.5}$	3	0.04953237	58.99322
H^{-1}	1.75	0.04959289	59.41552
H^{-2}	0.9	0.05290355	52.67199
H_0^{-1}	3.8	0.04977601	59.905

TABLE 1. $RMSE$ and SNR for the denoising results on the synthetic piece-wise constant image shown in Figures 2,3.

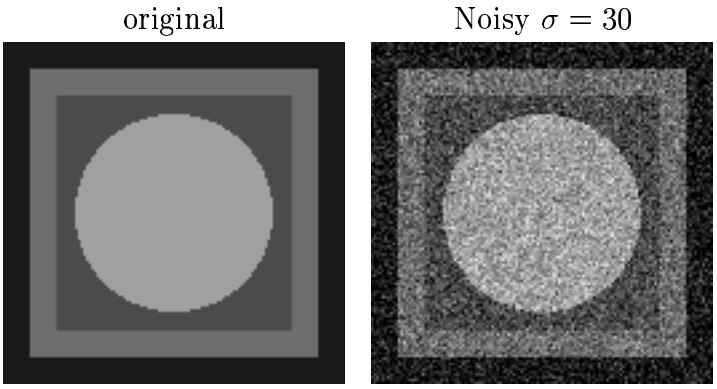

FIGURE 1. A synthetic image and its noisy version with additive Gaussian white noise with standard deviation 30 and zero mean.

and the improved image has $RMSE = 0.10373$ using $\lambda = 0.0004$. Visually, we see a significant improvement in the recovered image as compared to the degraded one.

Our last experiment is recovering from a blurred and noisy image. In Figure 10, we add white noise of standard deviation $\sigma = 10$ to the blurry image in Figure 8. The degradation has $RMSE = 0.121289$. The improved image is obtained with $\lambda = 0.0675$, and the $RMSE = 0.0607262$.

Acknowledgments. This work has been supported in part by an Alfred P. Sloan Fellowship, by the National Science Foundation (Grants NSF DMS 0312222 and ITR/AP 0113439) and by the National Institutes of Health through the NIH Roadmap for Medical Research, Grant U54 RR021813 entitled Center for Computational Biology (CCB).

REFERENCES

[1] R. ACAR AND C.R. VOGEL, *Analysis of bounded variation penalty methods of ill-posed problems*, Inverse Problems, 10 (1994), pp. 1217-1229.

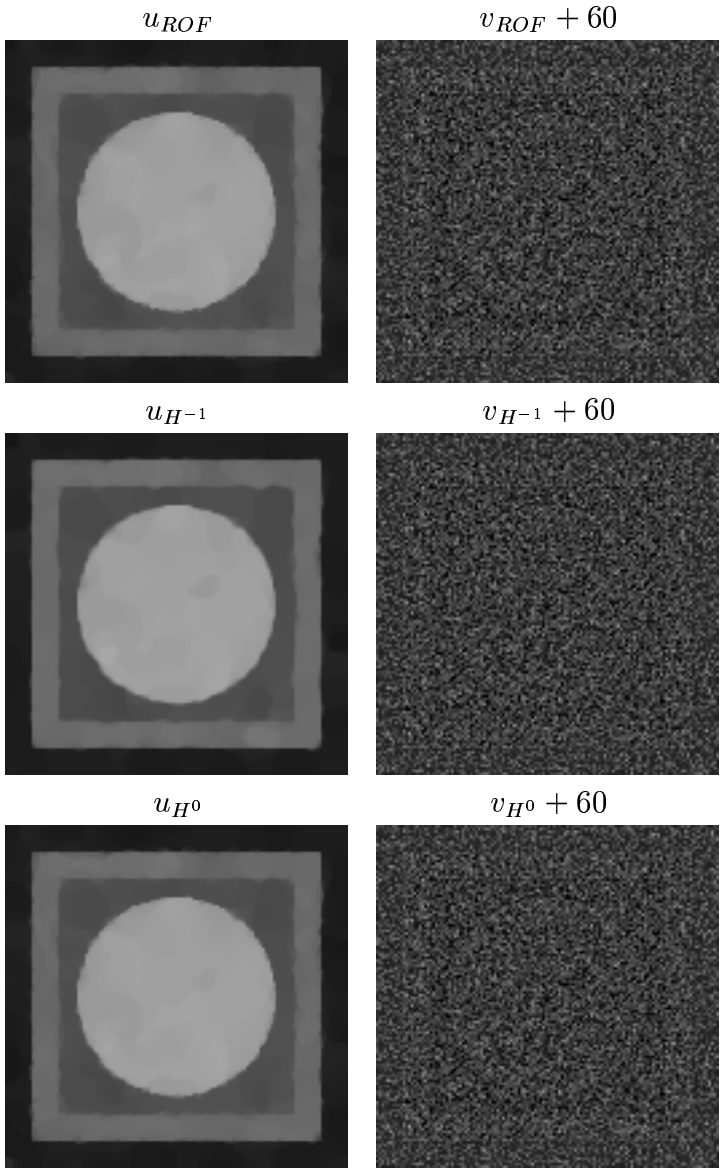

FIGURE 2. Comparison of results from our proposed model ($s = 0$ and $s = -1$) with the ROF model. Top: denoising results obtained from the ROF model. Middle: denoising results obtained with H^{-1}. Bottom: denoising results obtained with H^0.

[2] F. ANDREU-VAILLO, C. BALLESTER, V. CASELLES, AND J.M. MAZON, *Minimizing total variation flow*, C.R. Acad. Sci. Paris Sér. I Math., 331 (2000), pp. 867-872.
[3] F. ANDREU-VAILLO, V. CASELLES, AND J.M. MAZON, Parabolic Quasilinear Equations Minimizing Linear Growth Functionals, Birkhauser, 2004.
[4] G. AUBERT, J.-F. AUJOL, *Modeling very oscillating signals. Application to image processing* AMO 51 (2): 163-182, 2005.
[5] J.-F. AUJOL, G. AUBERT, L. BLANC-FERAUD, A. CHAMBOLLE, *Image decomposition application to SAR images*, LNCS 2695: 297-312, 2003.

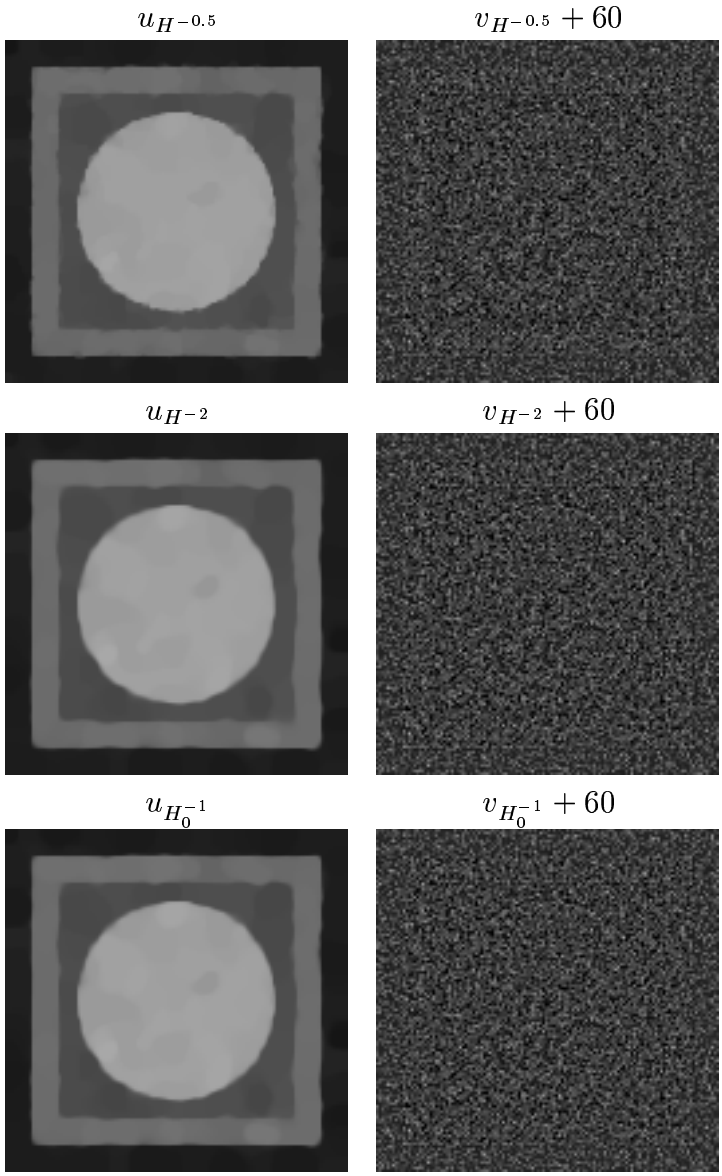

FIGURE 3. More results from our proposed model, using for the norm in the fidelity term $H^{-0.5}$ (row 1), H^{-2} (row 2), and H_0^{-1} semi-norm (row 3).

[6] J.-F. AUJOL, G. AUBERT, L. BLANC-FERAUD, A. CHAMBOLLE, *Image decomposition into a bounded variation component and an oscillating component*, Journal of Mathematical Imaging and Vision 22 (1): 71-88, 2005.

[7] J.-F. AUJOL, A. CHAMBOLLE, *Dual norms and image decomposition models*, IJCV 63(1): 85-104, 2005.

[8] A. CHAMBOLLE AND P.L. LIONS, *Image recovery via total variation minimization and related problems*, Numer. Math., 76 (1997), pp.167-188.

[9] I. DAUBECHIES, G. TESCHKE, *Wavelet-Based Image Decompositions by Variational Functionals*, Proc. SPIE Vol. 5266, p. 94-105, Wavelet Applications in Industrial Processing; Frederic Truchetet; Ed., Feb. 2004.

[10] I. DAUBECHIES AND G. TESCHKE, *Variational image restoration by means of wavelets: simultaneous decomposition, deblurring and denoising*, Appl. and Comp. Harm. Anal., to appear.

[11] R. DAUTRAY AND J.L. LIONS, Mathematical Analysis and Numerical Methods for Science and Technology, Vol. 2, Springer-Verlag, 1988.

original Noisy $\sigma = 20$

FIGURE 4. Lena image and its noisy version with additive white noise.

[12] F. DEMENGEL, R. TEMAM, *Convex Functions of a Measure and Applications*, Indiana University Mathematics Journal, Vol. 33, No. 5 (1984).
[13] S. ESEDOGLU, S.J. OSHER, *Decomposition of images by the anisotropic Rudin-Osher-Fatemi model*, Communications on Pure and Applied Mathematics, 57 (12): 1609-1626, 2004.
[14] L.C. EVANS AND R.F. GARIEPY, Measure Theory and Fine Properties of Functions, CRC Press, 1992.
[15] T. LE, L. VESE, *Image decomposition using total variation and div(BMO)*, UCLA CAM Report 04-36, 2004, to appear in MMS.
[16] Y. MEYER, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, Univ. Lecture Ser. 22, AMS, Providence, RI, 2002.
[17] D. MUMFORD AND B. GIDAS, *Stochastic models for generic images*, Quart. Appl. Math., 59 (2001), pp. 85-111.
[18] S. ROUDENKO, *Noise and Texture Detection in Image Processing*, Preprint, 2004.
[19] L. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), pp. 259-268.
[20] S. OSHER, A. SOLÉ, AND L. VESE, *Image decomposition and restoration using total variation minimization and the H^{-1} norm*, Multi. Model. Simul., Vol. 1, 3 (2003), pp. 349-370.
[21] E. TADMOR, S. NEZZAR, and L. VESE, *A multiscale image representation using hierarchical (BV, L2) decompositions*, Multiscale Modeling & Simulation 2(4):554-579, 2004.
[22] J.-L. STARCK, M. ELAD, and D.L. DONOHO, *Image Decomposition: Separation of Texture from Piece-Wise Smooth Content*, SPIE annual meeting, 3-8 August 2003, San Diego, California, USA.
[23] L. VESE, *A study in the BV space of a denoising-deblurring variational problem*, Appl. Math. Optm., 44 (2001), pp. 131-161.
[24] L. VESE AND S. OSHER, *Image denoising and decomposition with total variation minimization and oscillatory functions*, J. Math. Imaging Vision, 20 (2004), pp. 7-18.

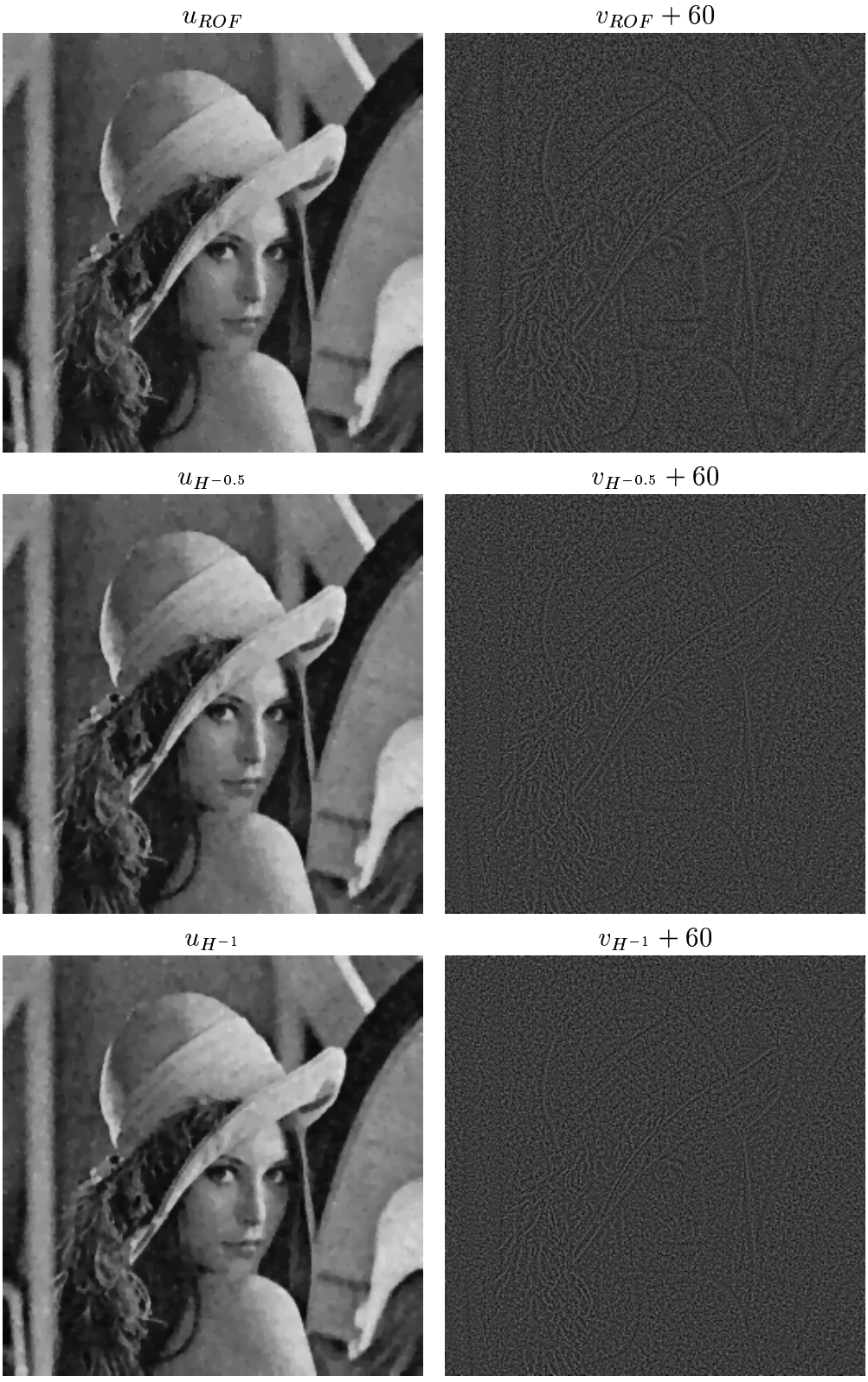

FIGURE 5. Denoising results on Lena image. Result of ROF model (top), result of our model with $H^{-0.5}$ (middle), result of our model with H^{-1} (bottom).

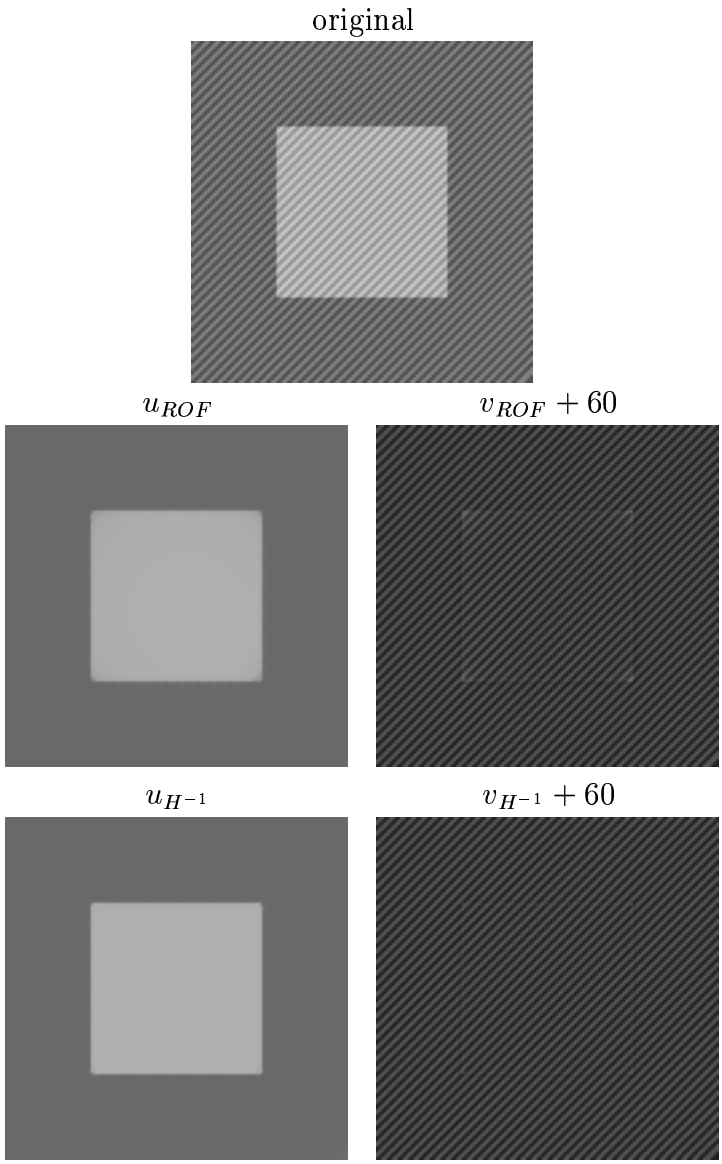

FIGURE 6. Decomposition of a synthetic textured image. Results from ROF model ($\lambda = 42$) and our model with H^{-1} ($\lambda = 0.5$).

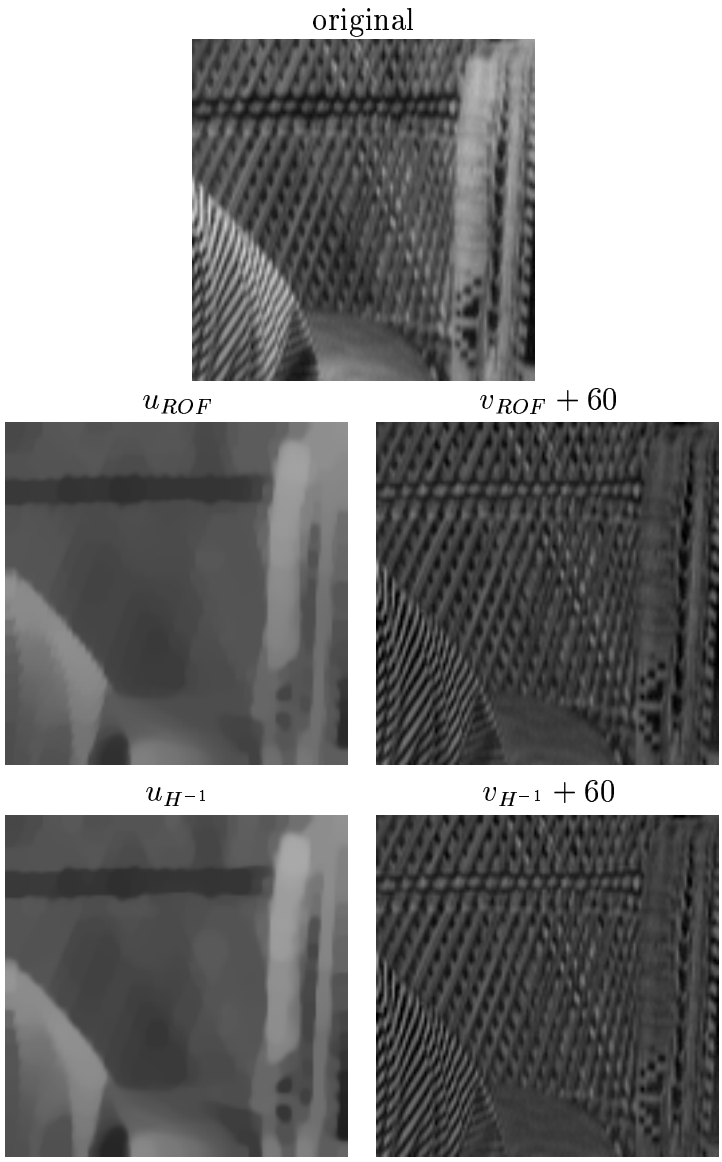

FIGURE 7. Decomposition of a natural textured image. Results from ROF model ($\lambda = 58$, middle) and our model with H^{-1} ($\lambda = 2.5$, bottom).

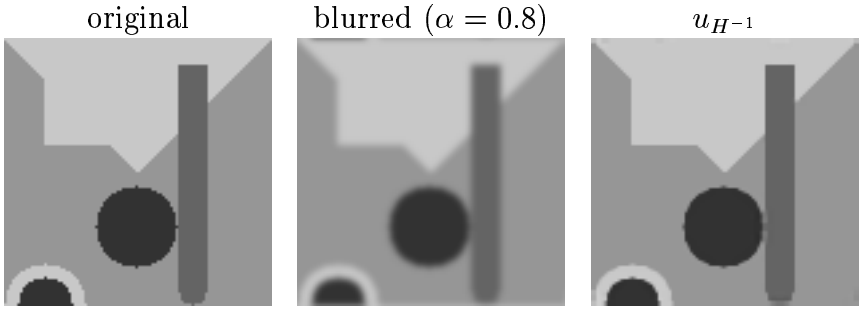

FIGURE 8. Deblurring on a synthetic image. Result is obtained from our model with H^{-1} ($\lambda = 0.0011$).

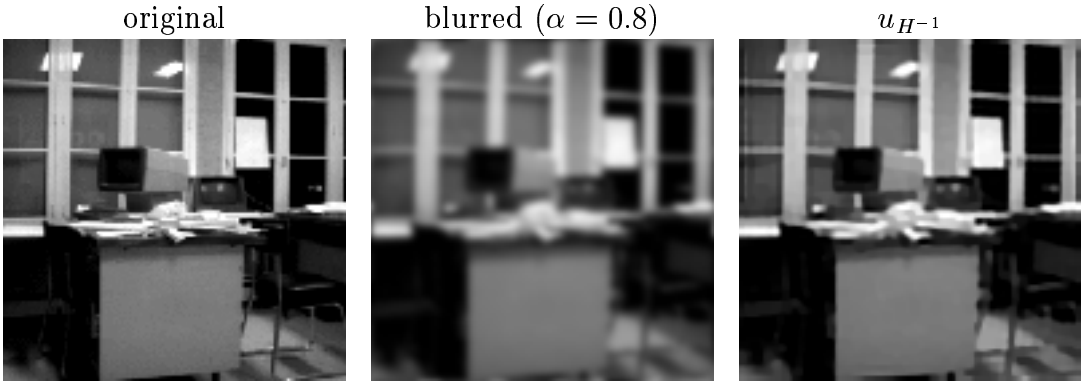

FIGURE 9. Deblurring on image of an office. Result is obtained from our model with H^{-1} ($\lambda = 0.0004$).

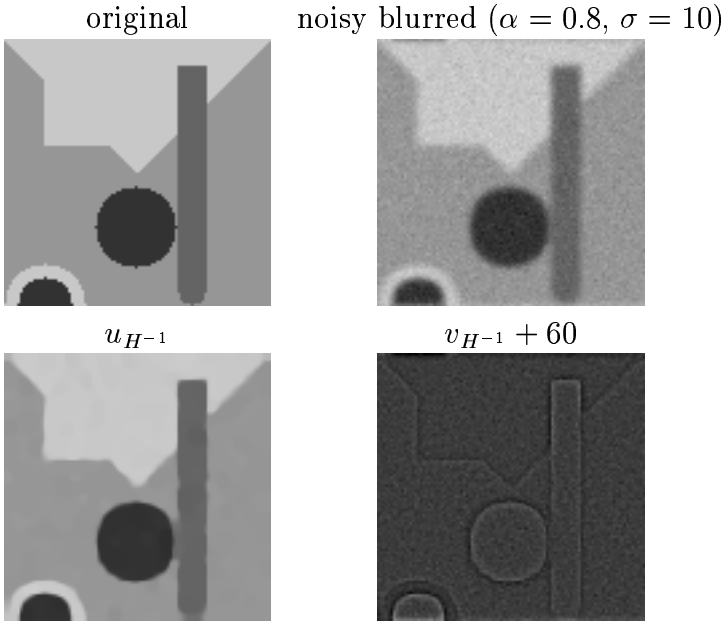

FIGURE 10. Denoising-deblurring result using our model with H^{-1} ($\lambda = 0.0675$).