(Φ, Φ^*) image decomposition models and minimization algorithms

Triet M. Le
triet.le@yale.edu
Mathematics
Yale University
New Haven, CT, USA

Linh H. Lieu
llieu@math.ucdavis.edu
Mathematics
UC Davis
Davis, CA, USA

Luminita A. Vese
lvese@math.ucla.edu
Mathematics
UCLA
Los Angeles, CA, USA

UCLA CAM Report 05-13, March 2005

Abstract

We propose in this paper minimization algorithms for image restoration using dual functionals and dual norms. In order to extract a clean image u from a degraded version $f = Ku + n$ (where f is the observation, K is a blurring operator and n represents additive noise), we impose a standard regularization penalty $\Phi(u) = \int \phi(|Du|)dx < \infty$ on u, where ϕ is positive, increasing and has at most linear growth at infinity. However, on the residual $f - Ku$ we impose a dual penalty $\Phi^*(f - Ku) < \infty$, instead of the more standard $\|f - Ku\|_{L^2}^2$ fidelity term. In particular, when ϕ is convex, homogeneous of degree one, and with linear growth (for instance the total variation of u), we recover the (BV, BV^*) decomposition of the data f, as suggested by Y. Meyer [32]. Practical minimization methods are presented, together with theoretical, experimental results and comparisons to illustrate the validity of the proposed models. Moreover, we also show that by a slight modification of the associated Euler-Lagrange equations, we obtain well-behaved approximations and improved results.

Keywords: image restoration; image decomposition; BV duality, functional minimization.

1 Introduction

Let Ω be an open, bounded and connected subset of $\mathbb{R}^2$, with Lipschitz boundary $\partial\Omega$. For two dimensional images, Ω is in general the interior of a rectangle in the plane. Assume $f : \Omega \to \mathbb{R}$ is a given image. Since at every pixel the light intensity has finite energy, it is natural to assume that $f \in L^\infty(\Omega)$. And since $L^\infty(\Omega) \subset L^2(\Omega)$ for bounded Ω, it is not too restrictive to assume that $f \in L^2(\Omega)$. Let us assume the linear degradation model $f = Ku + n$, where $f, u : \Omega \to \mathbb{R}$ are the degraded and the clean unknown images respectively, $K : L^2(\Omega) \to L^2(\Omega)$ is a linear and continuous operator, and n represents additive noise of zero mean. The problem of recovery of the unknown image u, given f and given this degradation model, is known to be an ill-posed problem. Therefore, regularization techniques as a-priori smoothness on the unknown u are usually imposed in a minimization approach, of the form

$$\inf_u E(u) = R(u) + \lambda F(f - Ku), \tag{1}$$

where the first term $R(u)$ acts as a regularization term (usually depending on spatial derivatives of the unknown u), $F(f - Ku)$ acts as a fidelity term, and $\lambda \geq 0$ is a tuning parameter. The behavior of functionals R and F is chosen function of the a-priori smoothness assumptions on u and function of the noise statistics. The standard case is when $R(u)$ depends on the gradient Du of u (and possibly on its discontinuity set S_u), and if additive Gaussian noise of zero mean is observed, then $F(u) = \|f - Ku\|^2_{L^2(\Omega)}$. These cases include the Mumford and Shah model [35] for computing optimal piecewise-smooth approximations u of f and the models of D. Geman, S. Geman and collaborators [21], [22], [23], [24] in the non-convex case, or the total variation minimization of Rudin, Osher and Fatemi [38], [39] in the convex case. Other related models and analysis in a variational approach include Acar-Vogel [2], Chambolle-Lions [13], Aubert and collaborators [7], [41]. In the PDE approach, we mention the Perona-Malik equation [37] ($\lambda = 0$), as well as the anisotropic smoothing Catté et al. [12], L. Alvarez et al. [4], and subsequent papers.

More recently, Y. Meyer on one side [32], and D. Mumford - B. Gidas [34] on the other side advocated the use of generalized functions as distributions in dual spaces for modeling images with oscillations, such as natural images, noise, texture, oscillatory patterns; thus proposing spaces such as $H^{-s}(\Omega)$ [34] and spaces that approximate the dual $BV^*(\Omega)$ of the space $BV(\Omega)$ [32]. Such oscillatory images are better modeled if weaker (dual) norms are considered as penalty or assumption, instead of the more standard $\|\cdot\|^2_{L^2(\Omega)}$ fidelity penalty.

Here, we follow the approach suggested by Y. Meyer of using duality to obtain weaker norms to represent the oscillatory component $v = f - Ku$. When analyzing the Rudin-Osher-Fatemi model [38] in the book manuscripts by Y. Meyer [32] and Andreu-Vaillo, Caselles and Mazón [1], a dual functional $\Phi^*(v)$ of the total variation $\Phi(u) = |Du|(\Omega)$ appears in the characterization of minimizers, applied to the residual term $v := f - Ku$. We therefore propose in this paper minimization models of the form

$$\inf_u \Big\{\Phi(u) + \lambda\Phi^*(f - Ku)\Big\}.$$

We will consider in particular the penalty $u \in BV(\Omega)$, or more generally of the form $\phi(|Du|)(\Omega) < \infty$, with ϕ convex and of linear growth ($\phi(t) = |t|$, $\phi(t) = \sqrt{1+t^2}$, $\phi(t) = \log\cosh(|t|)$, see [16] for the notion of convex functions of measures), as well as non-convex potentials that we write as $\Phi(u) = \int_\Omega \phi(|Du|)dx$ for functions in $W^{1,1}(\Omega)$. In the convex, well-posed case, these include the total variation minimization proposed by L. Rudin, S. Osher, and E. Fatemi [38]. Related recent work is proposed by J.-F. Aujol and A. Chambolle [8], and by S. Levine [28], where the authors use the duality given by the Legendre-Fenchel transform to solve cartoon and texture decomposition models. However, our approach proposed here is different.

In [42], [43], [36], and Aujol et al. [10], other approximations to the $(BV, G = \dot{W}^{-1,\infty})$ model of Y. Meyer have been previously proposed. The use of the dual norm of the total variation has also appeared in S. Kindermann et al.'s work [26], in a slightly different framework.

In the context of modeling oscillatory components by generalized functions, we refer to Y. Meyer [32], D. Mumford and B. Gidas [34], and to [42], [43], [36]. Recently, in [27], the authors propose practical methods for solving approximations to Meyer's $(BV, F = \mathrm{div}(BMO) = B\dot{M}O^{-1})$ decomposition model, while in [29], the authors generalize the models [42], [43], [36], [17] by proposing a (BV, H^{-s}) decomposition model and theoretical results based on duality. Finally, in Garnett et al. [20] and Aujol-Chambolle [8], the $(BV, E = \dot{B}^{-1}_{\infty,\infty})$ decomposition model proposed by Y. Meyer is also analyzed in theory and in practice for modeling decompositions into cartoon and texture.

For more properties and notations regarding characterization of minimizers by duality for the Rudin-Osher-Fatemi model, we refer the readers to the book manuscripts [32] and [1]. A preliminary, short version of this work has been presented at the 2006 SPIE Electronic Imaging Conference [14].

The outline of the paper is as follows: in the first part of Section 2 we give motivations and we present the general (Φ, Φ^*) minimization model and properties, together with the associated Euler-Lagrange equations. Section 2.1 is devoted to the particular case when Φ is defined to be the total variation or the BV semi-norm. Here we further analyze the model of the form (BV, BV^*), we show existence of minimizers and provide a regularized version for which Uzawa's method is convergent. Section 3 is devoted to variants of the proposed models in practice and to several experimental results and comparisons for image decomposition, denoising and deblurring.

2 Description of the general model and properties

Following the terminology from [1], let E be a normed space, and let E^* be its dual space. Let $\Phi : E \to [0, \infty]$ be any function. Let us define $\Phi^* : E^* \to [0, \infty]$, by

$$\Phi^*(v) = \sup\Big\{\frac{\langle v, u\rangle}{\Phi(u)} : \ u \in E\Big\},$$

with the convention that $\frac{0}{0} = 0$, $\frac{0}{\infty} = 0$. Here, $\langle v, u\rangle = v(u)$ denotes the duality pairing.

Note that $\Phi^*(v) \geq 0$ for any $v \in E^*$. Note also that the supremum is attained on the set of $u \in E$ such that $\langle v, u\rangle \geq 0$. We have the following Cauchy-Schwartz inequality

$$\langle v, u\rangle \leq \Phi^*(v)\Phi(u) \ \text{ if } \ \Phi(u) > 0.$$

Let $K : E \mapsto E$ be a linear and continuous operator, with linear and continuous adjoint $K^* : E^* \mapsto E^*$ (we will assume later that $E{=}E^*$; also, usually, in inverse problems such as deblurring, we may need to assume that $K\chi_\Omega = \chi_\Omega$). We wish to propose here image decomposition, denoising and deblurring models of the form

$$\inf_{u \in E} \Big\{\Phi(u) + \lambda\Phi^*(f - Ku)\Big\}, \tag{2}$$

where $f \in E \subset E^*$ is a given data and $\lambda > 0$ is a tuning parameter. We will choose Φ so that $\Phi(u) < \infty$ implies u is a piecewise-smooth function, with homogeneous regions and sharp boundaries (sometimes called "cartoon" component); Φ^* will have the role of attracting oscillatory components of zero mean, $v := f - Ku$ from f. Thus v could represent additive noise (when K is the identity or for general blurring operator K), or v could represent the texture component (if K is the identity).

To motivate the above model and its relation with the more standard one,

$$\inf_{u \in E} \Big\{\Phi(u) + \lambda\|f - Ku\|^2\Big\}, \tag{3}$$

we first need some preliminary remarks (valid in the case when Φ is convex, lower semi-continuous, and positive homogeneous of degree one, in particular if Φ is a smoothing (semi) norm on E). We follow [1], [32].

Definition 1. *Let $u \in E$. We say that $v^* \in E^*$ is a subgradient of Φ at u, and we write $v^* \in \partial\Phi(u)$, if and only if $\Phi(u)$ is finite and $\Phi(v) \geq \Phi(u) + \langle v^*, v - u\rangle$, for any $v \in E$.*

Theorem 1. *[1] Assume that Φ is convex, lower semi-continuous and positive homogeneous of degree 1. Then $v^* \in \partial\Phi(u)$ if and only if $\phi^*(v^*) \leq 1$ and $\langle v^*, u\rangle = \Phi(u)$.*

Proof. If $v^* \in \partial\Phi(u)$ then $\Phi(v) \geq \Phi(u) + \langle v^*, v - u\rangle, \forall v \in E$. By choosing $v = 0$ and $v = 2u$ (using $\Phi(0) = 0$), we easily obtain that $\langle v^*, u\rangle = \Phi(u)$. Furthermore, we obtain $\langle v^*, v\rangle \leq \Phi(v)$, for any $v \in E$, thus $\Phi^*(v^*) \leq 1$. Conversely, if $\Phi^*(v^*) \leq 1$ (which implies $\Phi(v) \geq \langle v^*, v\rangle$ for any $v \in E$) and $\Phi(u) = \langle v^*, u\rangle$, we have $\Phi(v) - \Phi(u) - \langle v^*, v - u\rangle = \Phi(v) - \langle v^*, v\rangle \geq 0$, thus $v^* \in \partial\Phi(u)$. $\square$

Remark 1. *Let us assume that E is a Hilbert space, with inner product $\langle \cdot, \cdot \rangle$ that induces the norm $\| \cdot \|$ in (3), and that Φ is convex, l.s.c. and positive homogeneous of degree 1. If u is a minimizer of (3), then this is equivalent with $2\lambda K^*(f - Ku) \in \partial\Phi(u)$, which by Theorem 1 gives $2\lambda\Phi^*(K^*(f - Ku)) \leq 1$ and $2\lambda\langle K^*(f - Ku), u\rangle = \Phi(u)$. Thus, for instance if K is the identity operator, then model (3), although it imposes that $\|f - u\|$ must be finite and even small, we also obtain that $\Phi^*(f - u) \leq \frac{1}{2\lambda} < \infty$ and this weaker condition, that does not penalize oscillations, could be an alternative constraint on the oscillatory component $f - u$, or more generally on $f - Ku$. This is in the spirit of Y. Meyer's ideas [32], as we will see later.*

For applications to image processing, we work with regularizations Φ as in [21], [22], [38], [7], [41]. We first define a continuous, even potential function $\phi : \mathbb{R} \to [0, \infty)$, increasing on $[0, \infty)$ (possibly non-convex), and with at most linear growth at infinity. In the convex case, $\phi(|Du|)$ is well defined for $u \in BV(\Omega)$, convex and lower semi-continuous, as a convex function of measures (see Demengel-Témam [16]), assuming two constants $a > 0$, $b \geq 0$ exist, such that $a|x| - b \leq \phi(x) \leq a|x| + b$ for any $x \in \mathbb{R}^2$. In the non-convex case (well defined and well-posed only in the discrete setting), we assume only $0 \leq \phi(x) \leq a|x| + b$ and we formally work with $u \in W^{1,1}(\Omega) \subset BV(\Omega) \subset L^2(\Omega)$ and the distributional gradient Du, as a function in $L^1(\Omega) \times L^1(\Omega)$.

We consider (2) by letting $E = E^* = L^2(\Omega)$, $\langle u, v\rangle = \int_\Omega u(x)v(x)dx$ for $u, v \in L^2(\Omega)$, and defining

$$\Phi(u) = \begin{cases} \int_\Omega \phi(|Du|), & \text{if } u \in BV(\Omega), \\ +\infty, & \text{if } u \in L^2(\Omega) \setminus BV(\Omega), \end{cases} \quad \text{for convex } \phi,$$

and

$$\Phi(u) = \begin{cases} \int_\Omega \phi(|Du|)dx, & \text{if } u \in W^{1,1}(\Omega), \\ +\infty, & \text{if } u \in L^2(\Omega) \setminus W^{1,1}(\Omega), \end{cases} \quad \text{for non-convex } \phi.$$

We propose here the following general denoising, deblurring and decomposition model,

$$\inf_u \left\{ \int_\Omega \phi(|Du|)dx + \lambda \left(\sup_{w,\ \int_\Omega \phi(|Dw|)dx \neq 0} \frac{\int_\Omega (f - Ku)w dx}{\int_\Omega \phi(|Dw|)dx} \right) \right\}, \tag{4}$$

as an alternative to the more standard one

$$\inf_u \int_\Omega \phi(|Du|)dx + \lambda\|f - Ku\|^2_{L^2(\Omega)}. \tag{5}$$

In practice, to computationally minimize the functional in (4), we formally apply Uzawa's method [18]: we define

$$L(u, w) = \int_\Omega \phi(|Du|)dx + \lambda \frac{\int_\Omega (f - Ku)w dx}{\int_\Omega \phi(|Dw|)dx} \text{ for } u, w \in W^{1,1}(\Omega),\ \int_\Omega \phi(|Dw|)dx \neq 0,$$

and problem (4) formally becomes

$$\inf_u \sup_w L(u, w).$$

The main minimization steps formally are: start with initial estimates $u = u^n$, $w = w^n$, $n = 0$. Then for $n \geq 0$, let

$$\begin{aligned} u^{n+1} &:= \arg\min_u L(u, w^n), \\ w^{n+1} &:= \arg\max_w L(u^{n+1}, w). \end{aligned}$$

To solve the above alternating minimization and maximization problems, we shall drive to steady state the following evolutionary coupled system in the unknowns $u = u^{n+1}$, $w = w^{n+1}$ (making ϕ differentiable):

$$\frac{\partial u}{\partial t} = K^* w^n + \frac{\int_\Omega \phi(|Dw^n|)dx}{\lambda}\text{div}\Big(\phi'(|Du|)\frac{Du}{|Du|}\Big), \quad \frac{\partial u}{\partial \vec{n}}\Big|_{\partial\Omega}(t,x) = 0, \quad t > 0 \tag{6}$$

$$\frac{\partial w}{\partial t} = f - Ku^{n+1} + \frac{\int_\Omega (f - Ku^{n+1})wdx}{\int_\Omega \phi(|Dw|)dx}\text{div}\Big(\phi'(|Dw|)\frac{Dw}{|Dw|}\Big), \quad \frac{\partial w}{\partial \vec{n}}\Big|_{\partial\Omega}(t,x) = 0 \quad t > 0 \tag{7}$$

where $\vec{n}$ denotes the exterior unit normal to $\partial\Omega$. Equation (6) in u is a gradient descent and equation (7) in w a gradient ascent.

Remark 2. *In practice, we regularize ϕ at the origin if necessary: for instance, if $\phi(t) = |t|^p$, $0 < p \leq 1$, then we substitute it by $\phi(t) = \sqrt{\epsilon^2 + |t|^2}^p$, with very small $\epsilon > 0$, and this also insures that $\int_\Omega \phi(|Dw|)dx \neq 0$ for any $w \in W^{1,1}(\Omega)$.*

Remark 3. *In practice, for the maximization process (7) in w, for fixed u^{n+1}, we start with an initial guess w such that $\int_\Omega (f - Ku^{n+1})wdx \geq 0$; then by selecting the correct time-step discretization $\triangle t$ to compute a new w, we insure that the energy in w, $\frac{\int_\Omega (f-Ku^{n+1})wdx}{\int_\Omega \phi(|Dw|)dx}$, is increasing. Therefore the condition $\int_\Omega (f - Ku^{n+1})wdx \geq 0$ remains satisfied. If $\triangle t$ is not optimally selected at every iteration, then we can add the absolute value, thus working with the equivalent formulation $\sup_w \frac{|\int_\Omega (f-Ku)wdx|}{\int_\Omega \phi(|Dw|)dx}$.*

For comparison purposes, we also recall here the gradient descent for formally minimizing (5),

$$\frac{\partial u}{\partial t} = 2\lambda K^*(f - Ku) + \text{div}\Big(\phi'(|Du|)\frac{Du}{|Du|}\Big), \quad \frac{\partial u}{\partial \vec{n}}\Big|_{\partial\Omega}(t,x) = 0, \quad t > 0. \tag{8}$$

Experimental results and comparisons will be shown with the above models for image decomposition, denoising and deblurring. We will also consider a rescaling of the descent speed in u, that makes the differential operators better behaved in practice and speeds up the convergence.

We consider in the next subsection the case when ϕ defines the total variation for BV functions, directly related and equivalent with Y. Meyer's proposal.

2.1 The BV case

In the particular case $E = E^* = L^2(\Omega)$ and $\Phi(u) = \int_\Omega |Du|$ (total variation of u) if $u \in BV(\Omega)$ ($+\infty$ otherwise), models (3) and (5) become the Rudin-Osher-Fatemi models [38], [39]. When $K = I$, the ROF model has been further analyzed by Y. Meyer in [32] obtaining a characterization of minimizers similar with the one in Remark 1, but with $\Phi^*(\cdot) = \|\cdot\|_*$ defined in a different, equivalent way. Moreover, it is suggested in [32] a more refined (BV, G) decomposition model, as an alternative to the ROF model: $\inf_{u \in BV(\Omega)}\{\int_\Omega |Du| + \lambda\|f - u\|_*\}$. As justified in [32] and as we will see later, the norm $\|\cdot\|_*$ does not penalize oscillations, thus it is a more refined texture norm than the $L^2(\Omega)$ norm for modeling oscillations in images. In this paper we propose an alternative way for solving in practice the (BV, G) Meyer's model as an exact decomposition (with generalizations), by directly using the definition of Φ^*. This is different from the earlier approximations to the (BV, G) model from [42], [43], [9], [10], [6].

For equivalent definitions and properties of BV functions, we refer the reader to [19], [5], [1], among other references.

Definition 2. *A function $u \in L^1(\Omega)$ has bounded variation in Ω if*

$$\int_\Omega |Du| := \sup\left\{ \int_\Omega u\, div\phi dx : \quad \phi \in C_c^1(\Omega, \mathbb{R}^2),\ |\phi| \le 1 \right\} < \infty.$$

We write $u \in BV(\Omega)$ to denote the space of functions of bounded variation, and $|u|_{BV(\Omega)} := \int_\Omega |Du|$ is called the total variation of u.

$BV(\Omega)$ becomes a Banach space endowed with the norm $\|u\|_{BV(\Omega)} = \|u\|_{L^1(\Omega)} + |u|_{BV(\Omega)}$. If $u \in W^{1,1}(\Omega)$, then $|u|_{BV(\Omega)} = \int_\Omega |\nabla u| dx$. Let us also recall the Poincaré-Wirtinger inequality in two dimensions: there is a constant $C > 0$ such that for any $u \in BV(\Omega)$,

$$\|u - u_\Omega\|_{L^2(\Omega)} \le C \int_\Omega |Du|,$$

where $u_\Omega = \frac{\int_\Omega u(x)dx}{|\Omega|}$ denotes the mean of u over Ω.

To consider the total variation of u, we define

$$\Phi(u) = \begin{cases} \int_\Omega |Du| \text{ if } u \in BV(\Omega), \\ +\infty \text{ if } u \in L^2(\Omega) \setminus BV(\Omega). \end{cases} \tag{9}$$

Definition 3. *For any $v \in L^2(\Omega)$, we define*

$$\Phi^*(v) = \sup_{\substack{w \in BV(\Omega) \\ |w|_{BV(\Omega)} \neq 0}} \frac{\int_\Omega vw dx}{|w|_{BV(\Omega)}} \Bigg(= \sup_{\substack{w \in BV(\Omega) \\ |w|_{BV(\Omega)} \neq 0}} \frac{|\int_\Omega vw dx|}{|w|_{BV(\Omega)}} \Bigg) \le \infty. \tag{10}$$

Let $X := \left\{ v \in L^2(\Omega) : \ \int_\Omega v(x)dx = 0 \right\}$.

Lemma 1. *Let $v \in L^2(\Omega)$ such that $\Phi^*(v) < \infty$. Then $v \in X$.*

Proof. Take any fixed $w \in BV(\Omega)$ with $|w|_{BV(\Omega)} \neq 0$. Then for any constant c,

$$\frac{\int_\Omega v(w+c)dx}{|w+c|_{BV(\Omega)}} = \frac{\int_\Omega v(w+c)dx}{|w|_{BV(\Omega)}} = \frac{\int_\Omega vw dx}{|w|_{BV(\Omega)}} + c\frac{\int_\Omega v dx}{|w|_{BV(\Omega)}} \le \Phi^*(v) < \infty.$$

If $\int_\Omega v(x)dx \neq 0$, then by letting $|c| \to \infty$ for c of the same sign with $\int_\Omega v(x)dx$ we obtain a contradiction. Therefore, v must be of zero mean in $L^2(\Omega)$, thus $v \in X$. □

The converse property is also true.

Lemma 2. *Let $v \in X$. Then $\Phi^*(v) < \infty$.*

Proof. For any $w \in BV(\Omega)$ with $|w|_{BV(\Omega)} \neq 0$, and since v is of zero mean, we have:

$$\frac{\int_\Omega v(x)w(x)dx}{|w|_{BV(\Omega)}} = \frac{\int_\Omega v(x)(w(x) - w_\Omega)dx}{|w|_{BV(\Omega)}} \le \frac{\int_\Omega |v(x)||w(x) - w_\Omega|dx}{|w|_{BV(\Omega)}}$$

$$\le \frac{\|v\|_{L^2(\Omega)} \|w - w_\Omega\|_{L^2(\Omega)}}{|w|_{BV(\Omega)}} \le C\|v\|_{L^2(\Omega)} < \infty,$$

where we have used the Cauchy-Schwartz and Poincaré-Wirtinger inequalities. Since this holds for any $w \in BV(\Omega)$ with $|w|_{BV(\Omega)} \neq 0$, we obtain $\Phi^*(v) \le C\|v\|_{L^2(\Omega)} < \infty$. □

Remark 4. *Any $v \in X$ can be identified with a linear and continuous (bounded) form from $\dot{BV}^*(\Omega) \cap L^2(\Omega)$. Here $\dot{BV}(\Omega)$ is the homogeneous version of $BV(\Omega)$ endowed with the norm $\|\cdot\|_{\dot{BV}(\Omega)} := |\cdot|_{BV(\Omega)}$ (or the quotient space $BV(\Omega)/P_0$ where functions different by a constant are identified), and $\dot{BV}^*(\Omega)$ is the space of distributions dual of $\dot{BV}(\Omega)$. We also have that for $v \in X$,*

$$\Phi^*(v) := \sup_{w \in BV(\Omega),\ |w|_{BV(\Omega)} \neq 0} \frac{\int_\Omega vw dx}{|w|_{BV(\Omega)}} = \sup_{w \in BV(\Omega),\ |w|_{BV(\Omega)} \leq 1} \int_\Omega vw dx.$$

Note that not all elements of $\dot{BV}^*(\Omega)$ can be expressed in this way (for instance, see T. De Pauw [15] for a characterization of the dual of $SBV(\Omega)$, of special functions of bounded variation). We also note that we have $\dot{BV}(\Omega)^* \cap L^2(\Omega) = \dot{W}^{1,1}(\Omega)^* \cap L^2(\Omega)$. In conclusion, the general model (2), when Φ is defined as in (9), can be called a $(\dot{BV}, \dot{BV}^*)$ model, and this model is equivalent with the (BV, G) model proposed by Y. Meyer, as we will see next.

We introduce now the space $G(\Omega)$ [32], [6], [1].

Definition 4. *We denote by $G(\Omega)$ the subset of $L^2(\Omega)$ defined by*

$$G(\Omega) = \{v \in L^2(\Omega),\ \textit{there is } \vec{g} \in L^\infty(\Omega, \mathbb{R}^2),\ v = -div\vec{g} \textit{ in } \mathcal{D}'(\Omega),\ \vec{g} \cdot \vec{n}|_{\partial\Omega} = 0\}.$$

Also, let us define, for $v \in G(\Omega)$,

$$\|v\|_* := \inf \left\{ \|\, |\vec{g}|\, \|_\infty\!:\ |\vec{g}| = \sqrt{g_1(x)^2 + g_2(x)^2},\ v = -\mathrm{div}(\vec{g}) \textit{ in } \mathcal{D}'(\Omega),\ \vec{g} \in L^\infty(\Omega)^2,\ \vec{g} \cdot \vec{n}|_{\partial\Omega} = 0 \right\}. \tag{11}$$

Following [1] and [6], it is possible to show that $\|\cdot\|_*$ is convex, lower semi-continuous and positive homogeneous of degree 1. Moreover, if $\|v\|_* < \infty$, then it is easy to show that the infimum in (11) is attaint [3]. We also have the following facts:

(i) $X = \{v \in L^2(\Omega) :\ \Phi^*(v) < \infty\}$;

(ii) $G(\Omega) = X$;

(iii) For any $v \in X = G(\Omega)$, we have $\|v\|_* = \Phi^*(v)$.

The proof of (i) is the direct consequence of Lemmas 1 and 2.

The proof of (ii) is shown by Aubert-Aujol [6]; one inclusion is obvious, while the other inclusion is the direct implication of a result by Bourgain-Brezis [11].

The proof of (iii) is shown in Andreu-Vaillo et al. [1] by interesting techniques related to the framework from the previous section.

In the original work of Y. Meyer and A. Haddad [32], [25], the case $\Omega = \mathbb{R}^2$ is considered, and it is shown that the space of distributions $G = \{v = -\mathrm{div}\vec{g},\ \vec{g} \in L^\infty(\mathbb{R}^2, \mathbb{R}^2)\}$, endowed with the norm $\|v\|_G = \inf\{\|\vec{g}\|_{L^\infty(\mathbb{R}^2,\mathbb{R}^2)}, v = -\mathrm{div}\vec{g}\}$, is isometrically isomorphic to the dual space of $(\{u \in BV(\mathbb{R}^2), Du \in L^1(\mathbb{R}^2)^2\}, |\cdot|_{BV(\mathbb{R}^2)})$.

Example 1. *Consider the sequence of oscillatory functions of zero mean on $\Omega = (0, 2\pi)$, $f_n(x) = \cos(nx) \in X$. Then it is easy to verify that $\|f_n\|_* = \frac{1}{n} \to 0$ as $n \to \infty$ (similarly, it is easy to verify in this case that $\sup_{\{w \in BV(0,2\pi),\ |w|_{BV(0,2\pi)} \leq 1\}} \int_0^{2\pi} \cos(nx) w(x) dx = \frac{1}{n}$). Therefore, more and more oscillations are not penalized by the $\|\cdot\|_*$ norm (which coincides with $\Phi^*(\cdot)$ in the case of the total variation). Therefore, small $\|\cdot\|_*$ norm or small $\Phi^*(\cdot)$ encourages oscillations. However, $\|f_n\|_{L^2(0,2\pi)} = \sqrt{\pi} > 0$ and does not converge to zero as the number of oscillations becomes larger and larger. Thus, it is natural to use $\|f - Ku\|_*$ or $\Phi^*(f - Ku)$ as a penalty for oscillatory components such as noise or texture, in a minimization approach.*

Theorem 2. *Let $f \in L^2(\Omega)$ and $K : L^2(\Omega) \to L^2(\Omega)$ be a linear and continuous operator, with adjoint K^*, such that $K\chi_\Omega = K^*\chi_\Omega = \chi_\Omega$, and let $\lambda > 0$. Then there exist at least one solution u of the minimization problem*

$$\inf_{u \in BV(\Omega)} F(u) = \Phi(u) + \lambda \Phi^*(f - Ku), \tag{12}$$

with Φ defined in (9).

Proof. Let $\bar{u} = f_\Omega \chi_\Omega$, where f_Ω is the mean of f over Ω. Then $F(\bar{u}) = \lambda \Phi^*(f - f_\Omega) < \infty$, since $f - f_\Omega$ is of zero mean. Therefore $0 \leq \inf_u F(u) < \infty$, and there must be a minimizing sequence $u_n \in BV(\Omega)$ such that $\inf_{u \in BV(\Omega)} F(u) = \lim_{n\to\infty} F(u_n)$. We obtain that $|u_n|_{BV(\Omega)} \leq M$, for all $n \geq 0$. Also, $f - Ku_n \in L^2(\Omega)$ must be of zero mean, since $\Phi^*(f - Ku_n) < \infty$. This implies that $\int_\Omega f dx = \int_\Omega Ku_n dx = \int_\Omega u_n K^* \chi_\Omega dx = \int_\Omega u_n dx$, thus each u_n has the same mean equal with f_Ω. By Poincaré-Wirtinger inequality, we deduce that $\|u_n\|_{L^2(\Omega)}$ are uniformly bounded. Since Ω is bounded, we also deduce that $\|u_n\|_{L^1(\Omega)}$ are uniformly bounded, therefore the sequence u_n is uniformly bounded in $BV(\Omega)$. Then there is a $u \in BV(\Omega)$ and a subsequence u_n converging to u strongly in $L^1(\Omega)$, weakly in $L^2(\Omega)$ and weakly in $BV - w^*(\Omega)$. We also have that $f - Ku_n$ converges to $f - Ku$ weakly in $L^2(\Omega)$. In conclusion, by the lower semi-continuity of $|\cdot|_{BV(\Omega)}$ and of $\Phi^*(\cdot)$, we deduce that

$$F(u) \leq \liminf_{n\to\infty} F(u_n) = \inf_{w \in BV(\Omega)} F(w),$$

thus existence of minimizers in $BV(\Omega)$. □

Remark 5. *If working on the entire plane as in [32], it is possible to show that there is no uniqueness of minimizers [33].*

Theorem 3. *Let $v \in X$. The maximization problem*

$$\sup_{w \in BV(\Omega),\ |w|_{BV(\Omega)} \leq 1} \int_\Omega v(x) w(x) dx =: \Phi^*(v) < \infty$$

has at least one solution w.

Proof. A maximizing sequence w_n with finite energy must exist, since the supremum is finite. We therefore have $|w_n|_{BV(\Omega)} \leq 1$. It is also sufficient to consider the restriction that w_n have zero mean. Then by Poincaré-Wirtinger inequality, we deduce that w_n are uniformly bounded in $L^2(\Omega)$, then in $L^1(\Omega)$ and finally in $BV(\Omega)$. Then there is a $w \in BV(\Omega)$ and a subsequence w_n converging to w strongly in $L^1(\Omega)$, weakly in $L^2(\Omega)$ and $|w|_{BV(\Omega)} \leq \liminf_{n\to\infty} |w_n|_{BV(\Omega)} \leq 1$. We obtain that $\int_\Omega v w_n dx \to \int_\Omega v w dx$, thus the limit w is a solution of the maximization process. □

Remark 6. *In the related parallel work by Kindermann, Osher and Xu [26], the maximization in Thm. 3 is also proposed. An interesting theoretical result of Strang [40] is used for the numerical computation, that says that it is sufficient to consider maximizers $w \in BV(\Omega)$ among characteristic functions of sets with finite perimeter.*

2.1.1 A convergent Uzawa's algorithm of regularized BV version

In practice, we use the gradient descent and ascent methods as given in (6)-(7) for the (BV, BV^*) model. However, we cannot directly apply Proposition 1.1, page 189 from Ekeland-Témam [18] to show the convergence of Uzawa's algorithm in this case. We show in this paragraph that the general convergence result from [18] can be applied, if we regularize $|u|_{BV(\Omega)}$ in the minimization process in u by making the problem strictly elliptic (although this does not directly give a practical optimization method): we work on (a subspace of) $H^1(\Omega)$ substituting $|u|_{BV(\Omega)}$ by $F_1(u)+\epsilon F_2(u)$, where $F_1(u)=\int_\Omega \sqrt{\epsilon+|Du|^2}dx$ (convex and Gâteaux-differentiable) and $F_2(u)=\|Du\|^2_{L^2(\Omega)}$ is the H^1 semi-norm; then we can apply the result from Ekeland-Témam [18].

Recall that $f\in L^2(\Omega)$ with $f_\Omega=0$ (without loss of generality), and consider the Hilbert spaces $V=(\{u\in H^1(\Omega), u_\Omega=0\}, |u|_{H^1(\Omega)})$ and $Z=L^2(\Omega)$. Let $\mathcal{A}=V\subset V$ (as a closed convex set of V), and the closed, convex and bounded set $\mathcal{B}=\{w\in BV(\Omega):\ w_\Omega=0,\ |w|_{BV(\Omega)}\leq 1\}\subset L^2(\Omega)$. Let small fixed $\epsilon>0$, $F_1, F_2: V\to\mathbb{R}$ defined by $F_1(u)=\int_\Omega\sqrt{\epsilon+|Du|^2}dx$ and $F_2(u)=\|Du\|^2_{L^2(\Omega)}$. Define the mapping $\Psi: V\to Z$ by $\Psi(u)=f-Ku$, where K is linear and continuous (bounded) on $L^2(\Omega)$, such that $K\chi_\Omega=K^*\chi_\Omega=\chi_\Omega$. Finally, define

$$L(u,w)=F_1(u)+\epsilon F_2(u)+\lambda\int_\Omega (f-Ku)w dx = J(u)+\lambda\langle w,\Psi(u)\rangle_Z,$$

and consider the optimization problem

$$\inf_{u\in\mathcal{A}}\Big\{\sup_{w\in\mathcal{B}} L(u,w)\Big\}. \tag{13}$$

Our aim is to approximate the solution u of the optimization problem (13), thus we want to approximate a saddle point (u,w) of L, that must satisfy

$$J(u)+\lambda\langle w,\Psi(u)\rangle_Z\leq J(v)+\lambda\langle w,\Psi(v)\rangle_Z,\ \forall\ v\in\mathcal{A},$$

$$J(u)+\lambda\langle w,\Psi(u)\rangle_Z\geq J(u)+\lambda\langle \nu,\Psi(u)\rangle_Z,\ \forall\ \nu\in\mathcal{B}.$$

The last relation implies

$$\langle \nu-w,\Psi(u)\rangle_Z\leq 0,\ \forall \nu\in\mathcal{B},$$

which is equivalent with ([18] page 40)

$$w=\Pi_{\mathcal{B}}(w+\rho\Psi(u)),\ \forall\rho>0,$$

where $\Pi_{\mathcal{B}}$ is the projection in $Z=L^2(\Omega)$, of $\mathcal{B}$.

Uzawa's algorithm constructs two sequences of elements $u^n\in\mathcal{A}$, $w^n\in\mathcal{B}$, defined as: start with any $w^0\in\mathcal{B}$, then we calculate u^0, then w^1, u^1, etc.

- w^n being known, we determine u^n as the element of $\mathcal{A}$ which minimizes $J(v)+\lambda\langle w^n,\Psi(v)\rangle_Z$.
- Then we define $w^{n+1}=\Pi_{\mathcal{B}}(w^n+\rho_n\Psi(u^n))$, where $\rho_n>0$ will be chosen later on.

Theorem 4. *Under the assumptions in this subsection 2.1.1, the above Uzawa's algorithm is convergent in the following sense: $u^n\to u$ in V, where u is solution of problem (13) provided the ρ_n satisfy $0<\rho_*\leq\rho_n\leq\rho'_*$, ρ'_* sufficiently small.*

Proof. We follow the steps from [18]. From the assumptions, we have that $\mathcal{A}\subset V$ is a non-empty closed convex set (actually equal with V), and $\mathcal{B}\subset Z=L^2(\Omega)$ is non-empty, closed, convex and bounded. Both functionals $F_1(u)$ and $F_2(u)$ are Gâteaux-differentiable, therefore J is also. Moreover, since F_1 is convex, then F'_1 is "monotone", in the sense that

$$\langle F'_1(u)-F'_1(v),u-v\rangle_V\geq 0\ \forall u,v\in\mathcal{A}.$$

Also, F_2 satisfies

$$\langle F'_2(u)-F'_2(v),u-v\rangle\geq(=)2\|u-v\|^2_V,\ \forall u,v\in\mathcal{A}.$$

Combining these two inequalities, we obtain

$$\langle J'(u)-J'(v),u-v\rangle_V\geq 2\epsilon\|u-v\|^2_V,\ \forall u,v\in\mathcal{A}.$$

For any $w \in \mathcal{B}$, the mapping $u \to \int_\Omega (f - Ku)w dx$ is convex and l.s.c. on $\mathcal{A}$. Also, Ψ is lipschitzian from $\mathcal{A}$ to Z, since $\|\Psi(u) - \Psi(v)\|_Z = \|Ku - Kv\|_{L^2(\Omega)} \leq \|K\|\|u - v\|_{L^2(\Omega)} \leq C\|K\|\|D(u - v)\|_{L^2(\Omega)} = C\|K\|\|u - v\|_V$, due to Poincaré inequality. We have thus shown that all the assumptions necessary in the general Proposition 1.1, pages 189-190 from [18] are satisfied, thus we conclude the required convergence. □

3 Experimental results

In our experimental results for image denoising, image deblurring and cartoon and texture separation, we have applied the above introduced models, discretized by finite differences. We have also considered the case when Φ is not convex. Moreover, with a minor modification of the gradient descent, our nonlinear evolution PDE's in the unknown u are numerically better behaved, giving faster convergence and improved results. These are presented in the next section.

3.1 (BV, BV^*) image denoising and decomposition results

We first consider the time-dependent system (6)-(7), with $K = I$ and $\phi(|t|) = |t|$ (corresponding to total variation), that can be made differentiable by working with $\phi(|t|) = \sqrt{\epsilon^2 + |t|^2}$ instead (this is what we call the (BV, BV^*) model).

Moreover, instead of (7), we also consider another gradient descending method in u (or rescaling the descent speed, related with [30], [31]), for $K = I$,

$$\frac{\partial u}{\partial t} = h_{\epsilon,p}(|Du|)\left[\frac{\int_\Omega |Dw|\ dx}{\lambda}\text{div}\left(\frac{Du}{|Du|}\right) + w\right], \tag{14}$$

where $h_{\epsilon,p}(|t|) > 0$ and satisfies

$$h_{\epsilon,p}(|t|) = \begin{cases} \epsilon & \text{if } |t| \geq \epsilon, \\ |t|^p & \text{if } |t| < \epsilon, \text{ for some } p \geq 1. \end{cases} \tag{15}$$

This modification makes the divergence operator numerically better behaved when $|Du| = 0$, and also gives faster convergence and improved results. Thus, we will use "the modified (BV, BV^*) model" when we compute u using equation (14).

We will show comparisons with the ROF model [38],

$$\inf_u \mathcal{F}(u) = \int_\Omega |Du|\ dx + \lambda \int_\Omega |f - Ku|^2\ dx. \tag{16}$$

To minimize (16) in practice, we also consider two gradient descending methods,

$$\frac{\partial u}{\partial t} = \frac{1}{2\lambda}\text{div}\left(\frac{Du}{|Du|}\right) + K^*(f - Ku) \qquad \text{(called ROF model)} \tag{17}$$

$$\frac{\partial u}{\partial t} = h_{\epsilon,p}(|Du|)\left[\frac{1}{2\lambda}\text{div}\left(\frac{Du}{|Du|}\right) + (f - u)\right] \qquad \text{(called modified ROF model when } K = I). \tag{18}$$

Our numerical computations use $\Delta x = \Delta y = 1$. For the computations which involve $h_{\epsilon,p}(t)$, we use $\epsilon = 0.5$ and $p = 1$.

We also show w as a by-product of the algorithm. The computed w depends on the number of iterations and on the initial guess w^0. We have tested as initial guess w^0 the given data f (then the computed w may look like a smoothed version of the data f), a rescaled version of f with small values, or a random initial w^0. We have observed that the computed image u does not depend on the initial w^0, but the computed w at steady state may depend on the initial w^0. For visualization purposes only, in each figure, w has been rescaled to $[0, 255]$, and to $f - u$ or to $f - Ku$ a constant between $[100, 125]$ has been added.

Let $\bar{u}$ be the true image of size $N \times M$, and u be the recovered image. To quantify how good the recovered image is, and to choose the parameter λ, we use the root mean square error

$$rmse = \frac{\sqrt{\sum_{i,j} |\bar{u}_{i,j} - u_{i,j}|^2}}{NM}.$$

Remark 7. *In all our experimental results we have selected the parameter λ for each method such that the best, smallest rmse is obtained (since in our artificial tests, the true image $\bar{u}$ is known). As suggested by one of the referees, it would be possible, when the noise level is known, to apply the method from [8]; in other words, to numerically or experimentally link the variance σ^2 of the noise n, with the dual norm of the noise, $\Phi^*(n)$. We could consider random noise images of various variances, for which we could estimate numerically their norm Φ^*. Having such information, λ as a Lagrange multiplier could then be automatically selected (as in [1] for the ROF model), by imposing a desired value $\Phi^*(f - Ku)$.*

3.2 Non-convex potential Φ in a dual decomposition

Similarly, the decomposition model (4), with $\phi(t)$ non-convex function, can be minimized by the algorithm given in (6)-(7). Here we will consider two choices (see [21], [22], [23], [24] for non-convex regularizations):

$$\phi(t) = |t|^p, \ 0 < p < 1, \tag{19}$$

and

$$\phi(t) = \frac{|t|^q}{1 + \alpha|t|^q}, \ 0 < \alpha < 1, \ q \geq 1. \tag{20}$$

Through out our numerical computations, we use $p = 0.75$ in (19) and $\alpha = 0.001$ and $q = 2$ in (20). We will use "nonconvex 1" when we refer to (6)-(7) computed using (19), and "nonconvex 2" when we refer to (6)-(7) computed using (20). The experimental results with a non-convex potential for denoising are improved over the ROF model, and we no longer see "geometry" in the v component.

We first present next our experimental results and comparisons for image denoising and decomposition using both convex and non-convex potentials, and comparisons with the ROF model.

Figure 1 shows an image decomposition result into cartoon and texture, applied to a fingerprint data, obtained by the (BV, BV^*) model.

Figure 2 shows several images and their degraded versions, to be used in the restoration process; we also give here the $rmse$ for each degraded data f.

Figure 3 shows results and comparisons of denoising of the geometric synthetic image (b) using the proposed models. Similarly, all models are compared on the noisy square image (f) in Figure 4. For both such images, note that the ROF model gives smaller $rmse$ than by using the (BV, BV^*) model. However, as a by-product of the (BV, BV^*) model, we also show the obtained images w: somehow surprisingly, w is a better denoised image, having the smallest $rmse$ among all models (these happen only for the synthetic images). In these cases, the initial w was the noisy data f.

Figures 5 and 6 show the same experimental results, now applied to the noisy real images of Barbara (k) and Lena (h). In these cases, the (BV, BV^*) models gave slightly better results than using the ROF models. The use of the non-convex potentials in the general (Φ, Φ^*) models gave similar results with the ROF model.

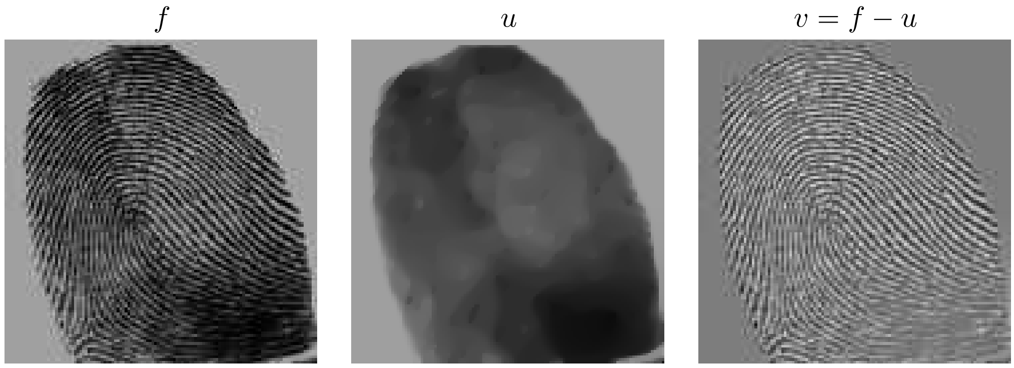


Figure 1: Exact image decomposition into cartoon and texture using the (BV, BV^*) minimization model. Left: real fingerprint data. Middle: cartoon component u. Right: oscillatory component $v = f - u$. $\lambda = 0.01$, 1000 iterations.

3.3 Image deblurring

We apply here the deblurring model (2) solved using (6)-(7), in the presence of blur, using $\phi(|t|) = |t|$ (corresponding to total variation, that can be made differentiable by working with $\phi(|t|) = \sqrt{\epsilon^2 + |t|^2}$ instead). The blur operator Ku is given by a convolution with a 5×5 blurring mask or kernel k of the form:

$$\frac{1}{273} \cdot \begin{array}{|c|c|c|c|c|} \hline 1 & 4 & 7 & 4 & 1 \\ \hline 4 & 16 & 26 & 16 & 4 \\ \hline 7 & 26 & 41 & 26 & 7 \\ \hline 4 & 16 & 26 & 16 & 4 \\ \hline 1 & 4 & 7 & 4 & 1 \\ \hline \end{array}$$

The equation in u from (6) is discretized using a semi-implicit scheme, as in [7], [41], while the discretization of the equation in w from (7) is done using a fully explicit scheme. We run 30 iterations in w, for every iteration in u. As before, we also have $\triangle x = \triangle y = 1$.

Below we give the number of iterations and rmse for our results, and comparisons with the Rudin-Osher model [39], for synthetic and real data. The results obtained using the new model are very similar with those obtained using the Rudin-Osher model (the new models did not produce an improvement in quality for deblurring). We also show the final w obtained as a by-product of the algorithm. The experimental results for deblurring are shown in Figures 7, 8 and 9.

4 Conclusion

We have considered here the exact decomposition model (BV, G) suggested by Y. Meyer [32], that is equivalent with a cartoon + texture $(BV, BV^* \cap L^2)$ model. Working directly with the duality definition for the texture norm, we have generalized this model and proposed several practical algorithms based on Uzawa's method for decomposition, denoising and deblurring of images. We have also continued the analysis of such variational models. Experimental results show that in some cases these exact decomposition models give slightly superior results to the more standard Rudin-Osher-Fatemi models; in other cases, no qualitative improvement has been obtained. Thus, although interesting mathematical and computational difficulties have been encountered, our proposed algorithm for the exact (BV, G) model does not provide a big improvement over the existing ones for image denoising or image deblurring.

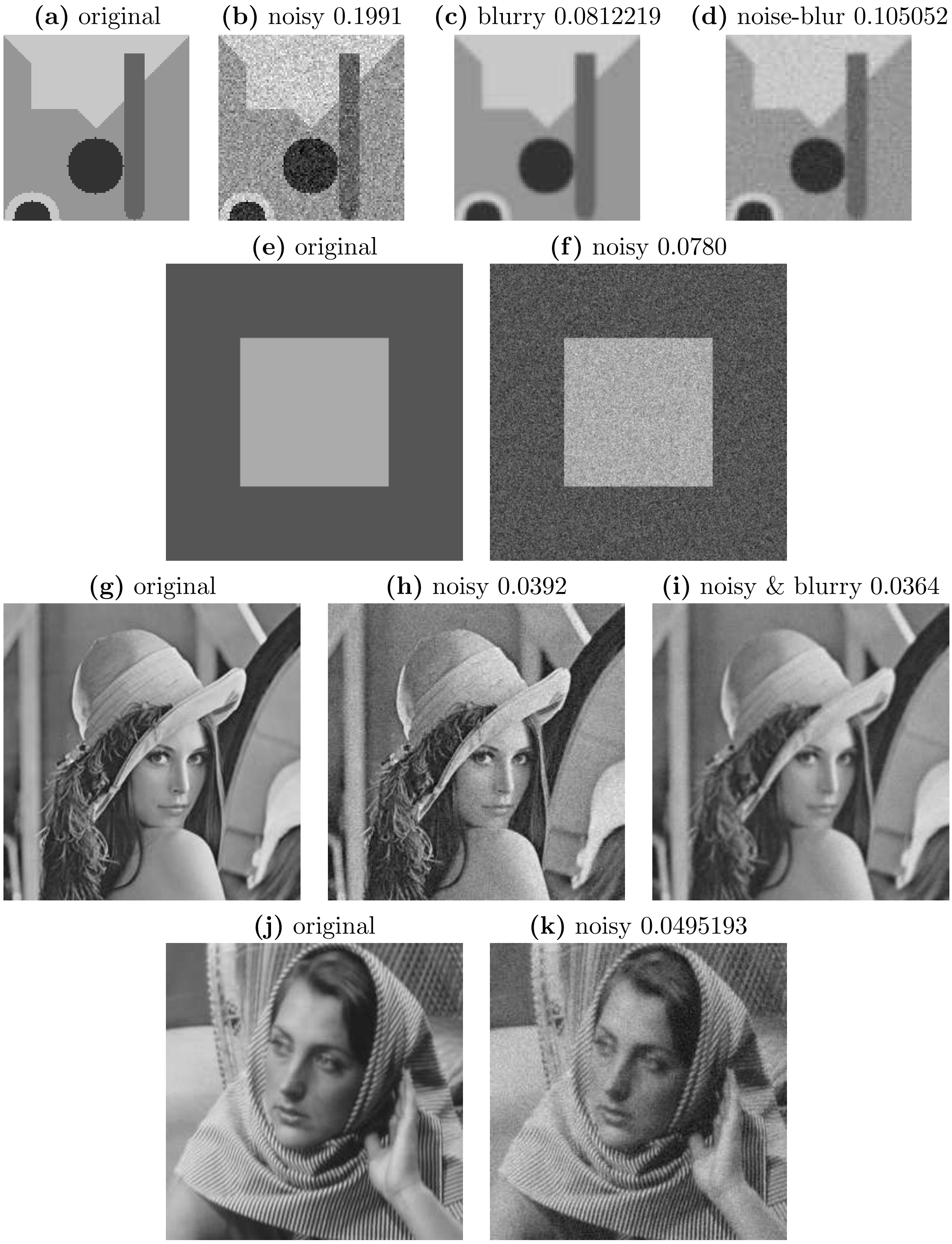


Figure 2: Original data images and their noisy and/or blurry versions f, together with the corresponding root mean square errors ($rmse$).

Acknowledgments: The authors would like to thank the organizers of the conference Mathematics and Image Analysis (MIA '06) for the invitation to participate to the conference and to contribute to this special issue. The authors would also like to thank the two unknown referees for their useful comments and suggestions that greatly helped to improve the presentation of the manuscript.

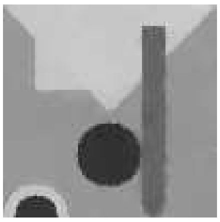
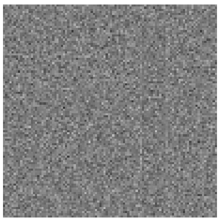
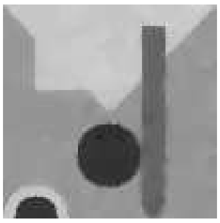
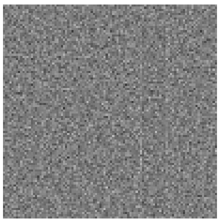
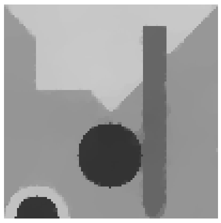
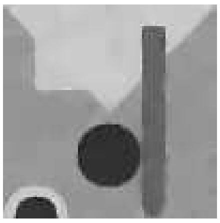
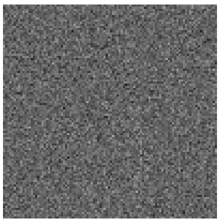
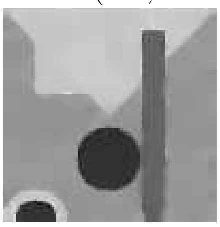
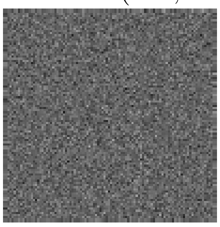
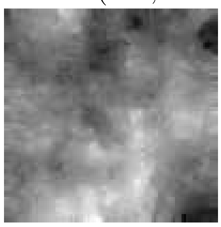
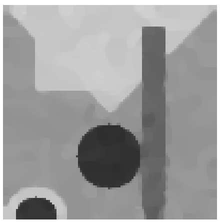
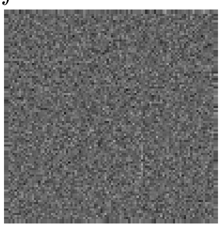
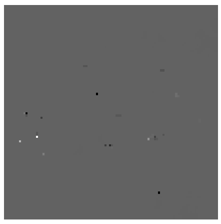
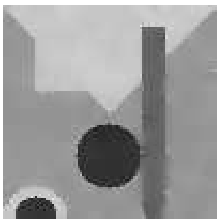
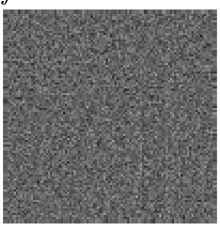
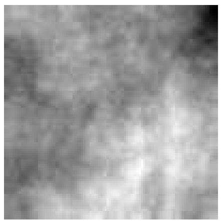
u ROF 	$f-u$ ROF 		$\lambda = 0.0265$ rmse 0.0514171086
u (BV, BV^*) 	$f-u$ (BV, BV^*) 	w (BV, BV^*) 	$\lambda = 0.05$ rmse u 0.0557836443 rmse w 0.0428309031
u mod. ROF 	$f-u$ mod. ROF 		$\lambda = 4.95$ rmse 0.05776094
u mod. (BV, BV^*) 	$f-u$ mod. (BV, BV^*) 	w mod. (BV, BV^*) 	$\lambda = 50$ rmse 0.05681891
u nonconv. 1 	$f-u$ nonconv. 1 	w nonconv. 1 	$\lambda = 8$ rmse 0.05111486
u nonconv. 2 	$f-u$ nonconv. 2 	w nonconv. 2 	$\lambda = 0.4$ rmse 0.04810694

Figure 3: Comparison for denoising the geometric data (b).

This work was supported in part by an Alfred P. Sloan fellowship, by the UCLA Institute for Pure and Applied Mathematics, and by the National Science Foundation grants ITR/ACI-0113439 ,

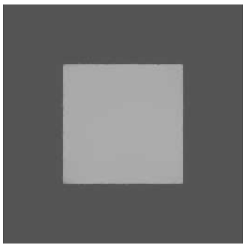
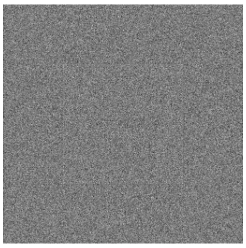
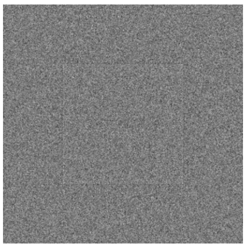
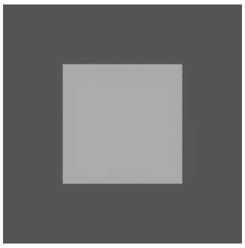
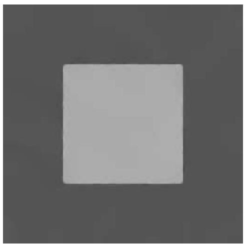
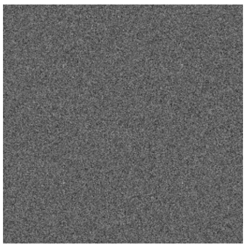
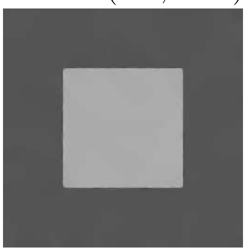
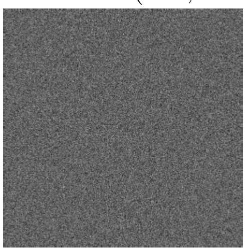
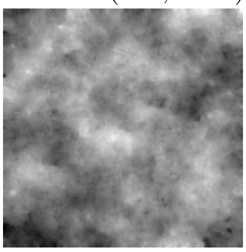
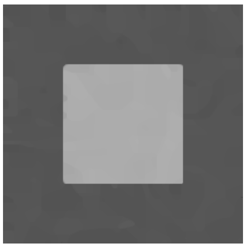
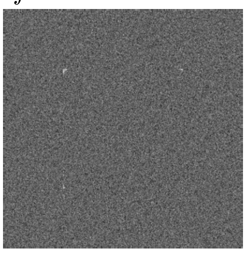
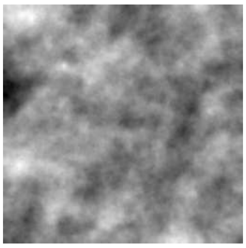
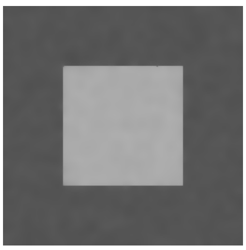
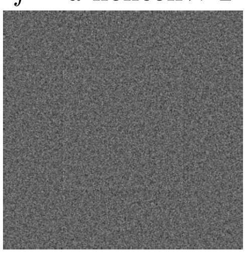
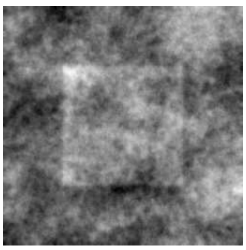
u ROF 	$f-u$ ROF 		$\lambda = 0.01$ rmse 0.00526068034
u (BV, BV^*) 	$f-u$ (BV, BV^*) 	w (BV, BV^*) 	$\lambda = 0.01$ rmse u 0.00643450115 rmse w 0.002649
u mod. ROF 	$f-u$ mod. ROF 		$\lambda = 3.825$ rmse 0.0058781
u mod. (BV, BV^*) 	$f-u$ mod. (BV, BV^*) 	w mod. (BV, BV^*) 	$\lambda = 150$ rmse 0.005767196
u nonconv. 1 	$f-u$ nonconv. 1 	w nonconv. 1 	$\lambda = 10$ rmse 0.00746253
u nonconv. 2 	$f-u$ nonconv. 2 	w nonconv. 2 	$\lambda = 1$ rmse 0.007092052

Figure 4: Comparison for denoising the square image (f).

DMS-0312222, DMS-0714945, and DMS-0809270.

References

[1] F. Andreu-Vaillo, V. Caselles and J.M. Mazon, *Parabolic Quasilinear Equations Minimizing Linear Growth Functionals*, Progress in Mathematics Vol. 223, Birkhäuser 2004.

[2] R. Acar, C.R. Vogel, *Analysis of Bounded Variation Penalty Methods for Ill-Posed Problems*, Inverse problems, 10(6), 1994.

[3] R.A. Adams, *Sobolev Spaces*, Academic Press 1975.

[4] L. Alvarez, P.-L. Lions, J.-M. Morel, *Image selective smoothing and edge detection by non-linear diffusion 2*, SINUM 29 (3): 845-866, 1992.

[5] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford University Press, 2000.

[6] G. Aubert, and J.-F. Aujol, *Modeling very oscillating signals. Application to image processing*, AMO 51(2): 163-182, 2005.

[7] G. Aubert and L. Vese, *A variational method in image recovery*, SIAM Journal on Numerical Analysis, 34 (5): 1948-1979, 1997.

[8] J.-F. Aujol, A. Chambolle, *Dual norms and image decomposition models*, IJCV 3(1): 85-104, 2005.

[9] J.-F. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle, *Image decomposition application to SAR images*, LNCS 2695: 297-312, 2003.

[10] J.-F. Aujol, G. Aubert, L. Blanc-Féraud, A. Chambolle, *Image decomposition into a bounded variation component and an oscillating component*, Journal of Mathematical Imaging and Vision, 22(1): 71-88, 2005.

[11] J. Bourgain and H. Brezis, *On the equation* $divY = f$ *and application to control of phases*, J. Amer. Math. Soc. 16: 393-426, 2003.

[12] F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll, *Image selective smoothing and edge-detection by nonlinear diffusion*, SINUM 29(1): 182-193, 1992.

[13] A. Chambolle, P.-L. Lions, *Image recovery via total variation minimization and related problems*, Numerische Mathematik 76(2): 167-188, 1997.

[14] G. Chung, T.M. Le, L.H. Lieu, N.M. Tanushev and L.A. Vese, *Computational methods for image restoration, image segmentation, and texture modeling*, Computational Imaging IV, edited by C.A. Bouman, E.L. Miller, I. Pollak, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 6065, pp. 60650J-1 – 60650J-15, 2006.

[15] T. De Pauw, *On SBV dual*, Indiana Univ. Math. J. 47(1): 99-121, 1998.

[16] F. Demengel and R. Témam, *Convex functions of a measure and applications*, Indiana University Mathematics Journal 33 (5): 673-709, 1984.

[17] I. Daubechies, G. Teschke, *Variational image restoration by means of wavelets: Simultaneous decomposition, deblurring, and denoising*, Applied and Computational Harmonic Analysis 19(1): 1-16, 2005.

[18] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*, SIAM, 1999.

[19] L.C. Evans, R.F. Gariepy *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, 1992.

[20] J.B. Garnett, T.M. Le, Y. Meyer and L.A. Vese, *Image Decompositions Using Bounded Variation and Homogeneous Besov Spaces*, Appl. Comput. Harmon. Anal. 23: 25-56, 2007.

[21] D. Geman and S. Geman, *Bayesian image analysis*, in Disordered Systems and Biological Organization, E. Bienenstock, F. Fogelman and G. Weisbuch Eds., Berlin, Springer Verlag, 1986, Vol. F20.

[22] D. Geman, G. Reynolds, *Constrained Restoration and the Recovery of Discontinuities*, IEEE Tr. on PAMI, 14(3): 367-383, 1992.

[23] D. Geman, C. Yang, *Nonlinear Image Recovery with Half-Quadratic Regularization*, IEEE TIP 4(7): 1995.

[24] S. Geman and D.E. McClure, *Statistical methods for tomographic image reconstruction*, in Proc. 46th Sess. Int. Stat. Inst. Bulletin ISI, 1987, vol. 52.

[25] A. Haddad, *Méthodes variationnelles en traitement d'image*, PhD Thesis, Ecole Normale Supérieure, Cachan, June 2005.

[26] S. Kindermann, S. Osher, and J. Xu, *Denoising by BV-Duality*, Journal of Scientific Computing 28(2-3): 411-444, 2006.

[27] T.M. Le and L.A. Vese, *Image Decomposition Using Total Variation and div*(BMO), Multiscale Modeling and Simulation 4(2): 390-423, 2005.

[28] S.E. Levine, *An Adaptive Variational Model for Image Decomposition*, in Energy Minimization Methods in Computer Vision and Pattern Recognition: 5th International Workshop, EMMCVPR 2005, St. Augustine, FL, USA, November 9-11, 2005. Proceedings. Editors: A. Rangarajan, B. Vemuri, A.L. Yuille, LNCS 3757: 382 - 397, 2005.

[29] L. Lieu and L. Vese, *Image Restoration and Decomposition via Bounded Total Variation and Negative Hilbert-Sobolev Spaces*, Applied Mathematics and Optimization 58: 167-193, 2008.

[30] A. Marquina, S. Osher, *A new time dependent model based on level set motion for nonlinear deblurring and noise removal*, LNCS 1682: 429-434, 1999.

[31] A. Marquina, S. Osher, *Explicit algorithms for a new time dependent model based on level set motion for nonlinear deblurring and noise removal*, SIAM Journal on Scientific Computing 22(2): 387-405, 2000.

[32] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, University Lecture Series, Col. 22, Amer. Math. Soc., 2001.

[33] Y. Meyer, Public Lecture, CMLA ENS-Cachan, Workshop "Contenu informatif des images numériques", 2004.

[34] D. Mumford and B. Gidas, *Stochastic models for generic images*, Quarterly of Applied Mathematics, 59(1): 85-111 2001.

[35] D. Mumford and J. Shah, *Optimal approximations by piecewise-smooth functions and associated variational problems*, Communications on Pure and Applied Mathematics 42(5): 577-685, 1989.

[36] S. Osher, A. Solé, L. Vese, *Image decomposition and restoration using total variation minimization and the H^{-1} norm*, Multiscale Modeling and Simulation 1(3): 349 - 370, 2003.

[37] P. Perona, and J. Malik, *Scale-space and edge detection using anisotropic diffusion*, IEEE Transactions on PAMI, 12(7): 629-639, 1990.

[38] L. Rudin, S. Osher, E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D 60: 259-268, 1992.

[39] L.I. Rudin and S. Osher, *Total variation based image restoration with free local constraints*, Proceedings of IEEE ICIP-94, Vol. 1: 31-35, 1994.

[40] G. Strang, *L^1 and L^∞ and approximation of vector fields in the plane*, in "Nonlinear Partial Differential Equations in Applied Science," H. Fujita, P. Lax, and G. Strang, eds., Lecture Notes in Num. Appl. Anal. 5 (1982) 273-288.

[41] L. Vese, *A study in the BV space of a denoising-deblurring variational problem*, Applied Mathematics and Optimization, 44(2): 131-161, 2001.

[42] L. Vese, S. Osher, *Modeling Textures with Total variation Minimization and Oscillating patterns in Image Processing*, Journal of Scientific Computing, 19(1-3): 553-572, 2003.

[43] L.A. Vese, and S.J. Osher, *Image Denoising and Decomposition with Total Variation Minimization and Oscillatory Functions*, Journal of Mathematical Imaging and Vision, 20: 7-18, 2004.

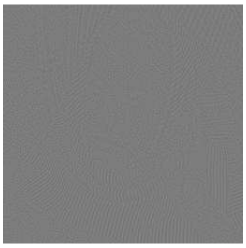
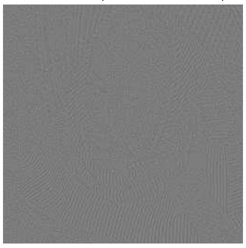
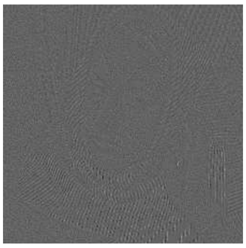
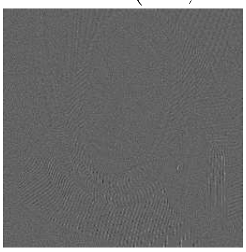
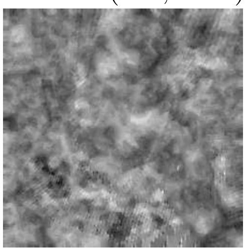
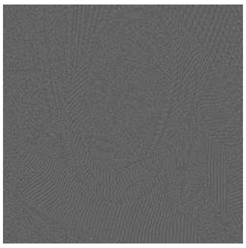
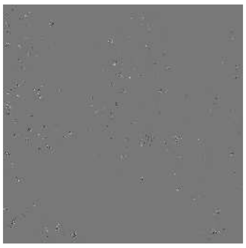
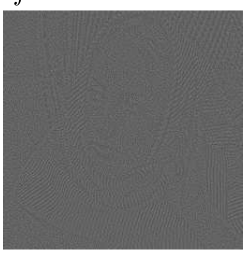
u ROF 	$f-u$ ROF 		$\lambda = 0.14$ rmse 0.0425563939
u (BV, BV^*) 	$f-u$ (BV, BV^*) 	w (BV, BV^*) 	$\lambda = 1500$ rmse 0.0415789001
u mod. ROF 	$f-u$ mod. ROF 		$\lambda = 4$ rmse 0.0474
u mod. (BV, BV^*) 	$f-u$ mod. (BV, BV^*) 	w mod. (BV, BV^*) 	$\lambda = 1500$ rmse 0.04524627
u nonconv. 1 	$f-u$ nonconv. 1 	w nonconv. 1 	$\lambda = 120$ rmse 0.04296311
u nonconv. 2 	$f-u$ nonconv. 2 	w nonconv. 2 	$\lambda = 9$ rmse 0.04297547

Figure 5: Comparison for denoising the Barbara image data (k).

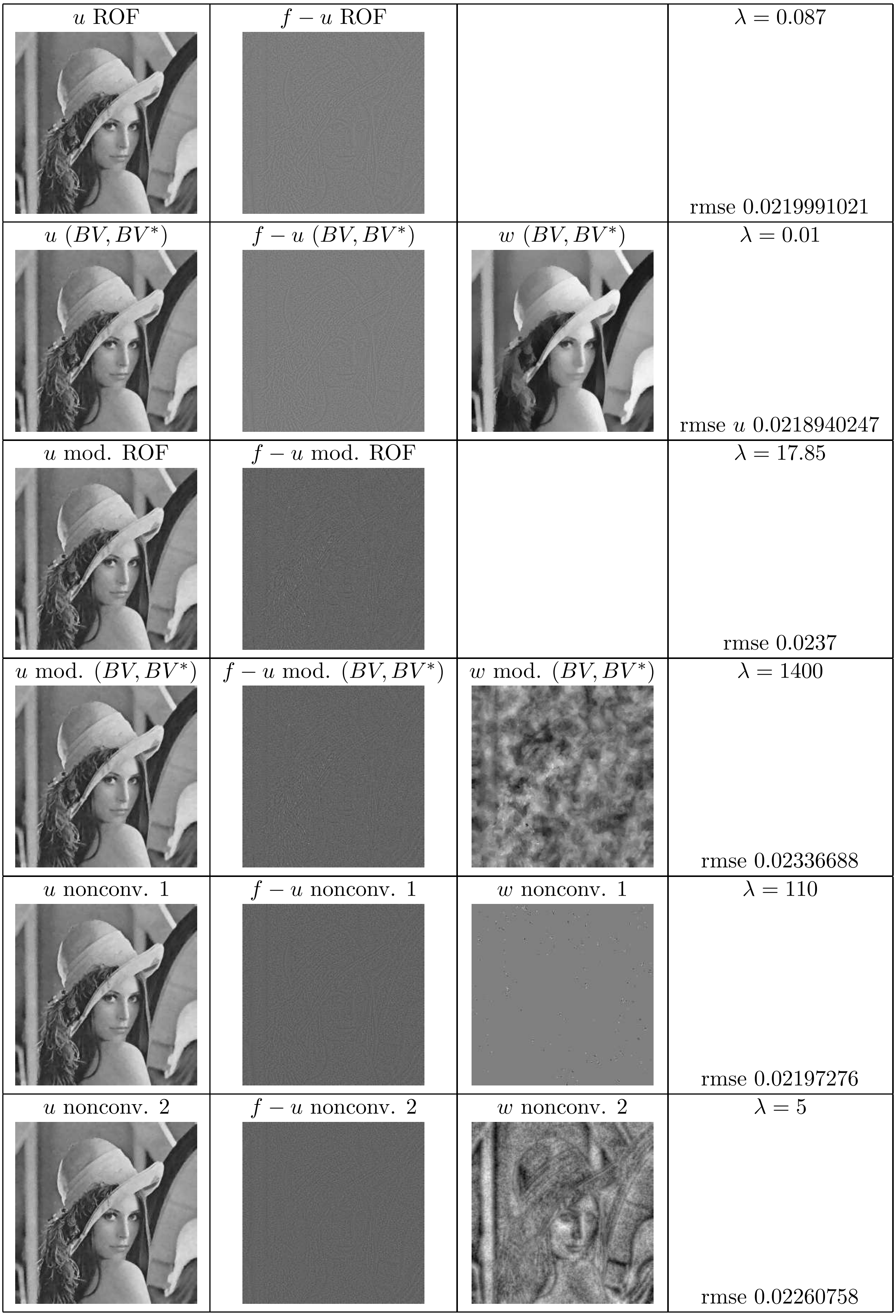


Figure 6: Comparison for denoising the Lena image data (h).

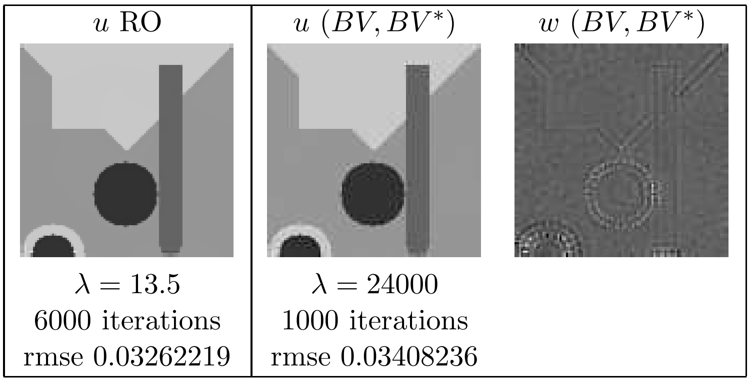


Figure 7: Deblurring of geometric image (c).

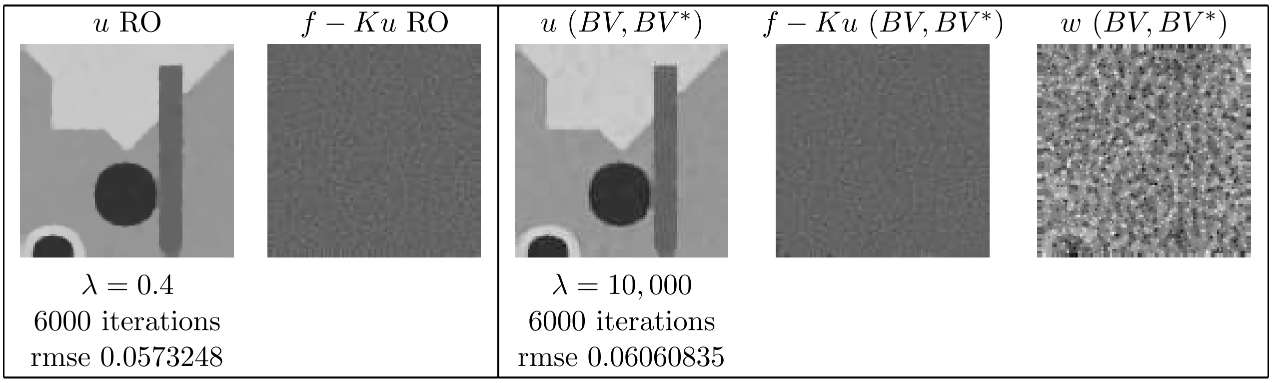


Figure 8: Denoising-deblurring of geometric image (d).

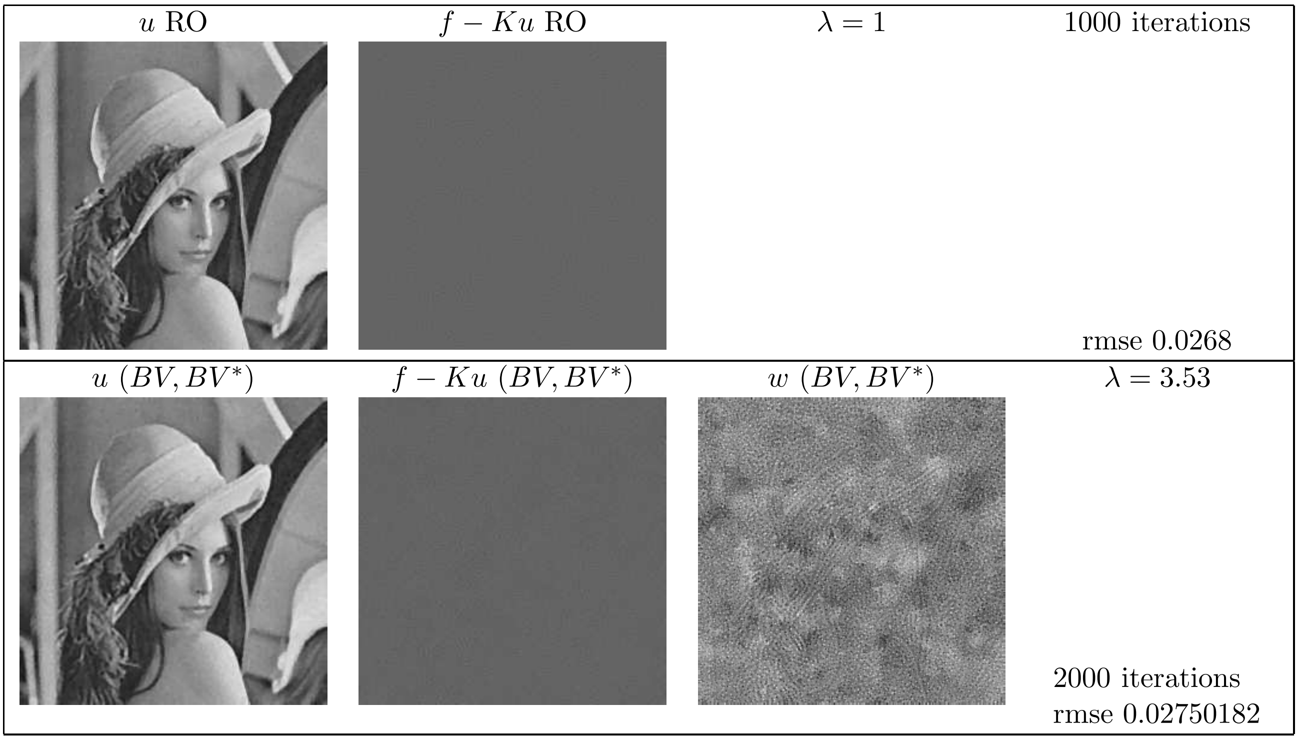


Figure 9: Denoising-deblurring of Lena image (i).