

A full second order variational model for multiscale texture analysis

Maitine Bergounioux · Loïc Piffet

Received: 15 June 2011 / Published online: 20 April 2012
© Springer Science+Business Media, LLC 2012

Abstract We present a second order image decomposition model to perform denoising and texture extraction. We look for the decomposition $f = u + v + w$ where u is a first order term, v a second order term and w the (0 order) remainder term. For highly textured images the model gives a two-scale texture decomposition: u can be viewed as a *macro-texture* (larger scale) whose oscillations are not too large and w is the *micro-texture* (very oscillating) that may contain noise. We perform mathematical analysis of the model and give numerical examples.

Keywords Second order total variation · Image reconstruction · Denoising · Texture analysis · Variational method

1 Introduction

The most famous variational model for image denoising is the Rudin-Osher-Fatemi denoising model (see [1, 19]). This model contains a regularization term based on the total variation which preserves discontinuities, which a classical H^1 -Tychonov regularization method does not do. The observed image to recover is split into two parts $u_d = u + v$: u represents the oscillating component (noise or texture) while v is the smooth part. So we look for the solution as $u + v$ with $v \in BV(\Omega)$ and $u \in L^2(\Omega)$, where $BV(\Omega)$ is the functions of bounded variation space defined on an open set Ω [2, 3, 16]. The regularization term deals with the so-called *cartoon* component v ,

M. Bergounioux (✉) · L. Piffet
UFR Sciences, Math., Labo. MAPMO, UMR 7349, Université d'Orléans, Route de Chartres,
BP 6759, 45067 Orléans cedex 2, France
e-mail: maitine.bergounioux@univ-orleans.fr

L. Piffet
e-mail: loic.piffet@univ-orleans.fr

while the remainder term $u := u_d - v$ represents the noise to be minimized:

$$\text{(ROF)} \quad \min\{\|u\|_{L^2} + \Phi_1(v) \mid u + v = u_d, u \in L^2(\Omega), v \in BV(\Omega)\}.$$

Here $\|u\|_{L^2}$ is the L^2 -norm of u and $\Phi_1(v)$ stands for the total variation of v (we recall the definition in next section).

A lot of people have investigated such variational decomposition models, assuming that an image can be decomposed into many components, each component describing a particular property of the image ([4, 5, 20–22, 26] and references therein for example).

In [9] we have presented a second order model where the (first order) classical total variation term Φ_1 was replaced by a second order total variation term Φ_2 (using the appropriate functional framework that we called ROF2. The use of such a model allows to get rid of the *staircasing effect* that appears with the ROF denoising model, since solutions are piecewise constant. However, we have noticed that while the staircasing effect disappeared, the resulting image was slightly blurred (the blur effect was not as important as if we had used a classical low-pass filter however) since solutions are now piecewise affine. To remove this undesirable effect due to the partial ROF2, we decide to use a full second order model. More precisely, we assume that the image can be split into three components: a smooth (continuous) part v , a piecewise constant cartoon part u and an oscillating part w that should involve noise and/or fine textures. Such decompositions have already been investigated by Aujol and al. [4, 5, 7]. These authors use the Meyer space of oscillating functions [18] instead of the $BV^2(\Omega)$ space. We present these spaces in the sequel. However, the model we propose here is different: the oscillating part of the image is not penalized but a priori included in the remainder term $w = u_d - u - v$, while v is the smooth part (in $BV^2(\Omega)$) and u belongs to $BV(\Omega)$: we expect that u is a piecewise constant function so that its jump set gives the image contours. For highly textured images as the one of example (a) in Fig. 1, we shall see that the model gives a two-scale texture decomposition: u can be viewed as a *macro-texture* (larger scale) whose oscillations are not too large and w is the *micro-texture* (quite oscillating) that may contain the noise as well.

Therefore, we look for components u , v and w that belong to different spaces: u belongs to $BV(\Omega)$ (and if possible not to $W^{1,1}(\Omega)$), $v \in BV^2(\Omega)$ and $w \in L^2(\Omega)$. This last component $w = u_d - u - v$ lies in the same space as the observed image u_d .

The paper is organized as follows. We present the model and briefly give main properties of the spaces of functions of first and second order bounded variation. We give existence, partial uniqueness results and optimality conditions. Section 3 is devoted to the discretization and numerical process. We present some results in the last section.

2 The mixed second order model

2.1 Spaces $BV(\Omega)$ and $BV^2(\Omega)$

Let Ω be an open bounded subset of \mathbb{R}^n , $n \geq 2$ (practically $n = 2$) smooth enough (with the cone property and Lipschitz for example). Following [2, 3, 6] and [9, 14],

we recall the definitions and main properties of the spaces of functions of first and second order bounded variation. The space $BV(\Omega)$ is the classical space of functions of bounded variation defined by

$$BV(\Omega) = \{u \in L^1(\Omega) \mid \Phi_1(u) < +\infty\},$$

where

$$\Phi_1(u) := \sup \left\{ \int_{\Omega} u(x) \operatorname{div} \xi(x) \, dx \mid \xi \in C_c^1(\Omega), \|\xi\|_{\infty} \leq 1 \right\}. \tag{1}$$

The space of functions with bounded Hessian, denoted $BV^2(\Omega)$, is the space of $W^{1,1}(\Omega)$ functions such that $\Phi_2(u) < +\infty$, where

$$W^{1,1}(\Omega) = \{u \in L^1(\Omega) \mid \nabla u \in L^1(\Omega)\}.$$

Here ∇u stands for the first order derivative of u (in the sense of distributions) and

$$\Phi_2(u) := \sup \left\{ \int_{\Omega} \langle \nabla u, \operatorname{div}(\xi) \rangle_{\mathbb{R}^n} \mid \xi \in C_c^2(\Omega, \mathbb{R}^{n \times n}), \|\xi\|_{\infty} \leq 1 \right\} < \infty, \tag{2}$$

where

$$\operatorname{div}(\xi) = (\operatorname{div}(\xi_1), \operatorname{div}(\xi_2), \dots, \operatorname{div}(\xi_n)),$$

with

$$\forall i, \xi_i = (\xi_i^1, \xi_i^2, \dots, \xi_i^n) \in \mathbb{R}^n \quad \text{and} \quad \operatorname{div}(\xi_i) = \sum_{k=1}^n \frac{\partial \xi_i^k}{\partial x_k}.$$

We give thereafter important properties of these spaces. Proofs can be found in [2, 3, 14, 23] for example.

Theorem 1 (Banach properties)

- The space $BV(\Omega)$, endowed with the norm $\|u\|_{BV(\Omega)} = \|u\|_{L^1} + \Phi_1(u)$, is a Banach space. The derivative in the sense of distributions of every $u \in BV(\Omega)$ is a bounded Radon measure, denoted Du , and $\Phi_1(u) = \int_{\Omega} |Du|$ is the total variation of u .
- The space $BV^2(\Omega)$ endowed with the following norm

$$\|f\|_{BV^2(\Omega)} := \|f\|_{W^{1,1}(\Omega)} + \Phi_2(f) = \|f\|_{L^1} + \|\nabla f\|_{L^1} + \Phi_2(f), \tag{3}$$

where Φ_2 is given by (2) is a Banach space.

Theorem 2 (Structural properties of the derivative) *Let Ω be an open subset of \mathbb{R}^n with Lipschitz boundary.*

- For every $u \in BV(\Omega)$, the Radon measure Du can be decomposed into $Du = \nabla u \, dx + D^s u$, where $\nabla u \, dx$ is the absolutely continuous part of Du with respect of the Lebesgue measure and $D^s u$ is the singular part.

- A function u belongs to $BV^2(\Omega)$ if and only if $u \in W^{1,1}(\Omega)$ and $\frac{\partial u}{\partial x_i} \in BV(\Omega)$ for $i \in \{1, \dots, n\}$. In particular

$$\Phi_2(u) \leq \sum_{i=1}^n \Phi_1\left(\frac{\partial u}{\partial x_i}\right) \leq n\Phi_2(u).$$

We get lower semi-continuity results for the semi-norms Φ_1 and Φ_2 :

Theorem 3 (Semi-continuity)

- The mapping $u \mapsto \Phi_1(u)$ is lower semi-continuous (denoted in short *lsc*) from $BV(\Omega)$ to \mathbb{R}^+ for the $L^1(\Omega)$ -topology.
- The mapping $u \mapsto \Phi_2(u)$ is lower semi-continuous from $BV^2(\Omega)$ endowed with the strong topology of $W^{1,1}(\Omega)$ to \mathbb{R}^+ .

Finally, we have embedding results:

Theorem 4 Embedding results Assume $n \geq 2$. Then

- $BV(\Omega) \subset L^2(\Omega)$ with continuous embedding, if $n = 2$,
- $BV(\Omega) \subset L^p(\Omega)$ with compact embedding, for every $p \in [1, 2)$, if $n = 2$,
- $BV^2(\Omega) \hookrightarrow W^{1,q}(\Omega)$ with $q \leq \frac{n}{n-1}$, with continuous embedding. Moreover the embedding is compact if $q < \frac{n}{n-1}$. In particular

$$BV^2(\Omega) \hookrightarrow L^q(\Omega), \quad \forall q \in [1, \infty[, \text{ if } n = 2.$$

In the sequel, we set $n = 2$ and Ω is a bounded, open, Lipschitz subset of \mathbb{R}^2 , so that $BV^2(\Omega) \subset H^1(\Omega)$ with continuous embedding and $BV^2(\Omega) \subset W^{1,1}(\Omega)$ with compact embedding.

2.2 The variational model

In a previous work [9], we have studied the following restoration variational model

$$\inf \left\{ \frac{1}{2} \|u_d - v\|_{L^2(\Omega)}^2 + \mu \Phi_2(v) + \delta \|v\|_{W^{1,1}(\Omega)} \mid v \in BV^2(\Omega) \right\}.$$

which only involves a second order term Φ_2 . The motivation was to get rid of the staircasing effect while restoring noisy data. Indeed, the original ROF model provides piecewise constant solutions [12, 24]. We inferred that the use of a second order penalization term leads to piecewise affine solutions so that there is no staircasing any longer. However, we observed that the contours were not kept as well as we wanted and that the resulting image was slightly blurred. To overcome this difficulty, we now consider a full second order model involving both first and second order penalization terms. Furthermore, we concentrate us on the extraction of the texture; indeed denoising can be handled in a similar way, considering that noise is a very fine texture.

Specifically, we assume that the image we want to rebuild from the data can be decomposed as $u_d = w + u + v$ where u, v and w are functions that characterize

the various structures of u_d . In the sequel $u_d \in L^2(\Omega)$. These functions belong to different functional spaces:

- v is the (smooth) second order part and belongs to $BV^2(\Omega)$. It is expected to describe the dynamics of the image.
- u is the $BV(\Omega)$ -component. We conjecture that $u \in BV(\Omega) \setminus BV^2(\Omega)$, and could be piecewise constant, so that its derivative is a measure supported by the contours.
- The remainder term $w := u_d - u - v \in L^2$ is the noise and/or fine textures.

We consider the following cost functional defined on $BV(\Omega) \times BV^2(\Omega)$:

$$F_{\lambda, \mu, \delta}(u, v) = \frac{1}{2} \|u_d - u - v\|_{L^2(\Omega)}^2 + \lambda \Phi_1(u) + \mu \Phi_2(v) + \frac{\delta_1}{2} \|v\|_{L^2(\Omega)}^2 + \delta_2 \Phi_1(v),$$

where $\lambda, \mu, \delta_i > 0$. We are looking for a solution to the optimization problem

$$(\mathcal{P}_{\lambda, \mu, \delta}) \quad \inf\{F_{\lambda, \mu, \delta}(u, v) \mid (u, v) \in BV(\Omega) \times BV^2(\Omega)\}.$$

Let us comment the different terms of the cost functional $F_{\lambda, \mu, \delta}$:

- the first term $\|u_d - u - v\|_{L^2(\Omega)}^2$ is the fitting data term,
- $\Phi_1(u)$ is a standard total variation term widely used in image restoration,
- $\Phi_2(v)$ penalizes the second order total variation of component v as in [9],
- $\frac{\delta_1}{2} \|v\|_{L^2(\Omega)}^2 + \delta_2 \Phi_1(v)$ is a penalization term which is necessary to get a priori estimates on minimizing sequences and obtain existence results.

It is a theoretical tool and $\delta_i > 0, i = 1, 2$ can be chosen as small as wanted.

Remark 1 We could replace the last penalization term by $\|v\|_{H^1}^2$ or $\|v\|_{W^{1,1}(\Omega)}$. We have chosen an intermediate penalization term between these two possibilities. Indeed, using the $W^{1,1}(\Omega)$ norm would add another non differentiable penalization $\|v\|_{L^1}^2$. As we already deal with $\Phi_1(v)$ which is not differentiable either, this would add numerical technical difficulties.

We could get rid of this δ -penalization term. Indeed, we may use Poincaré-Wirtinger inequalities to get the appropriate a priori estimates and deduce existence of solutions (see [8]). However, this impose to change the functional framework and introduce additional constraints. For example, this requires that the BV -component u has 0 mean-value and/or that the BV^2 -component v vanishes on the boundary. Moreover, we would not have strict convexity any longer so that uniqueness of solutions cannot be ensured any more. This will be precisely investigated in a future work.

The δ -part is not useful (and unjustified) from the modelling point of view. It is only necessary to prove existence and uniqueness of the solution. Besides, as we shall see it after, we can fix $\delta_2 = 0$ once the problem is discretized. Furthermore, we noticed that δ_1 has no influence during the execution of the numeric tests, so that we finally chose $\delta_1 = \delta_2 = 0$.

We first give an existence and uniqueness result for problem $(\mathcal{P}_{\lambda, \mu, \delta})$.

Theorem 5 *Assume that $\lambda > 0, \mu > 0$ and $\delta_i > 0, i = 1, 2$. Problem $(\mathcal{P}_{\lambda, \mu, \delta})$ has a unique solution (u, v) .*

Proof We first prove existence. Let $(u_n, v_n) \in BV(\Omega) \times BV^2(\Omega)$ be a minimizing sequence, i.e.

$$\lim_{n \rightarrow +\infty} F_{\lambda, \mu, \delta}(u_n, v_n) = \inf\{F_{\lambda, \mu, \delta}(u, v) \mid (u, v) \in BV(\Omega) \times BV^2(\Omega)\} < +\infty.$$

The sequence $(v_n)_{n \in \mathbb{N}}$ is L^2 -bounded (with the δ_1 -term) and L^1 -bounded since Ω is bounded. The sequence $\Phi_1(v_n)$ is bounded thanks to the δ_2 -term and $\Phi_2(v_n)$ is bounded as well. Therefore $(v_n)_{n \in \mathbb{N}}$ is bounded in $BV^2(\Omega)$.

As $(v_n)_{n \in \mathbb{N}}$ is $L^1(\Omega)$ -bounded and $(u_n + v_n)_{n \in \mathbb{N}}$ is $L^2(\Omega)$ -bounded, this yields that $(u_n)_{n \in \mathbb{N}}$ is $L^1(\Omega)$ -bounded. As $(\Phi_1(u_n))_{n \in \mathbb{N}}$ is bounded then $(u_n)_{n \in \mathbb{N}}$ is bounded in $BV(\Omega)$.

With the compactness result of Theorem 4, we infer that $(v_n)_{n \in \mathbb{N}}$ strongly converges (up to a subsequence) in $W^{1,1}(\Omega)$ to $v^* \in BV^2(\Omega)$ and Theorem 3 gives the following:

$$\Phi_2(v^*) \leq \liminf_{n \rightarrow +\infty} \Phi_2(v_n).$$

Similarly, the compactness embedding of $BV(\Omega)$ in $L^1(\Omega)$ (Proposition 2) gives the existence of a subsequence still denoted $(u_n)_{n \in \mathbb{N}}$ and $u^* \in BV(\Omega)$ such that u_n strongly converges in $L^1(\Omega)$ to u^* , and

$$\Phi_1(u^*) \leq \liminf_{n \rightarrow +\infty} \Phi_1(u_n).$$

Finally

$$F_{\lambda, \mu, \delta}(u^*, v^*) \leq \liminf_{n \rightarrow +\infty} F_{\lambda, \mu, \delta}(u_n, v_n) = \min_{(u, v) \in BV(\Omega) \times BV^2(\Omega)} F_{\lambda, \mu, \delta}(u, v).$$

The pair (u^*, v^*) is a solution to $(\mathcal{P}_{\lambda, \mu, \delta})$. Uniqueness is straightforward with the strict convexity of $F_{\lambda, \mu, \delta}$. □

It is easy to see that (u^*, v^*) is a solution to $(\mathcal{P}_{\lambda, \mu, \delta})$ if and only if

$$\begin{aligned} u^* &= \operatorname{argmin} \left\{ \frac{1}{2} \|u_d - v^* - u\|^2 + \lambda \Phi_1(u), u \in BV(\Omega) \right\}, \\ v^* &= \operatorname{argmin} \left\{ \frac{1}{2} \|u_d - u^* - v\|^2 + \frac{\delta_1}{2} \|v\|_{L^2}^2 + \delta_2 \Phi_1(v) + \mu \Phi_2(v), v \in BV^2(\Omega) \right\}. \end{aligned} \tag{4}$$

and we may derive optimality conditions in a standard way:

Theorem 1 (u^*, v^*) is a solution to $(\mathcal{P}_{\lambda, \mu, \delta})$ if and only if

$$u_d - u^* - v^* \in \lambda \partial \Phi_1(u^*), \tag{5a}$$

$$u_d - u^* - (1 + \delta_1)v^* \in \mu \partial \Phi_2(v^*) + \delta_2 \partial \Phi_1(v^*). \tag{5b}$$

Here $\partial \Phi(w)$ denotes the subdifferential of Φ at w (see [3, 15] for example). The proof is straightforward since Φ_1 and Φ_2 are convex and continuous and variables u and v can be decoupled.

3 The discretized problem

This section is devoted to numerical analysis of the previous model. We first (briefly) present the (standard) discretization process.

3.1 Discretization process

We assume that the image is rectangular with size $N \times M$. We note $X := \mathbb{R}^{N \times M} \simeq \mathbb{R}^{NM}$ endowed with the usual inner product and the associated Euclidean norm

$$\langle u, v \rangle_X := \sum_{1 \leq i \leq N} \sum_{1 \leq j \leq M} u_{i,j} v_{i,j}, \quad \|u\|_X := \sqrt{\sum_{1 \leq i \leq N} \sum_{1 \leq j \leq M} u_{i,j}^2}. \tag{6}$$

We set $Y = X \times X$. It is classical to define the discrete total variation with finite difference schemes as following (see for example [6]): the discrete gradient of the numerical image $u \in X$ is $\nabla u \in Y$ computed by the forward scheme for example:

$$(\nabla u)_{i,j} = ((\nabla u)_{i,j}^1, (\nabla u)_{i,j}^2), \tag{7}$$

where

$$(\nabla u)_{i,j}^1 = \begin{cases} u_{i+1,j} - u_{i,j} & \text{if } i < N \\ 0 & \text{if } i = N, \end{cases} \quad \text{and} \quad (\nabla u)_{i,j}^2 = \begin{cases} u_{i,j+1} - u_{i,j} & \text{if } j < M \\ 0 & \text{if } j = M. \end{cases}$$

The (discrete) total variation corresponding to $\Phi_1(u)$ is given by

$$J_1(u) = \frac{1}{NM} \sum_{1 \leq i \leq N} \sum_{1 \leq j \leq M} \|(\nabla u)_{i,j}\|_{\mathbb{R}^2}, \tag{8}$$

where

$$\|(\nabla u)_{i,j}\|_{\mathbb{R}^2} = \left\| \left(\nabla u_{i,j}^1, \nabla u_{i,j}^2 \right) \right\|_{\mathbb{R}^2} = \sqrt{\left(\nabla u_{i,j}^1 \right)^2 + \left(\nabla u_{i,j}^2 \right)^2}.$$

The discrete divergence operator $-\text{div}$ is the adjoint operator of the gradient operator ∇ :

$$\forall (p, u) \in Y \times X, \quad \langle -\text{div } p, u \rangle_X = \langle p, \nabla u \rangle_Y.$$

To define a discrete version of the second order total variation Φ_2 we have to introduce the discrete Hessian operator. For any $v \in X$, the Hessian matrix of v , denoted Hv is identified to a X^4 vector:

$$(Hv)_{i,j} = \left((Hv)_{i,j}^{11}, (Hv)_{i,j}^{12}, (Hv)_{i,j}^{21}, (Hv)_{i,j}^{22} \right).$$

We refer to [9] for the detailed expressions of these quantities. The discrete second order total variation corresponding to $\Phi_2(v)$ writes

$$J_2(v) = \frac{1}{NM} \sum_{1 \leq i \leq N} \sum_{1 \leq j \leq M} \|(Hv)_{i,j}\|_{\mathbb{R}^4}, \tag{9}$$

with

$$\|(Hv)_{i,j}\|_{\mathbb{R}^4} = \sqrt{(Hv_{i,j}^{11})^2 + (Hv_{i,j}^{12})^2 + (Hv_{i,j}^{21})^2 + (Hv_{i,j}^{22})^2}.$$

The discretized problem stands

$$\inf_{(u,v) \in X \times X} \frac{1}{2} \|u_d - u - v\|_X^2 + \lambda J_1(u) + \mu J_2(v) + \frac{\delta_1}{2} \|v\|_X^2 + \delta_2 J_1(v). \tag{10}$$

Theorem 6 Assume $\lambda \geq 0, \mu \geq 0, \delta_2 \geq 0$ and $\delta_1 > 0$. Problem (10) has a unique solution.

Proof The proof is obvious since the cost functional is strictly convex and coercive because of the data-fitting term and the δ_1 -term. □

In the sequel we set $\lambda > 0, \mu > 0, \delta := \delta_1 > 0$ and $\delta_2 = 0$.

3.2 Numerical realization and algorithm

Let (u^*, v^*) be the unique solution to

$$(\mathbf{P}_{\lambda,\mu,\delta}) \quad \inf_{(u,v) \in X \times X} \frac{1}{2} \|u_d - u - v\|_X^2 + \lambda J_1(u) + \mu J_2(v) + \frac{\delta}{2} \|v\|_X^2.$$

Using the subdifferential properties and decoupling u^* and v^* gives the following necessary and sufficient optimality conditions:

Proposition 1 (u^*, v^*) is a solution to $(\mathbf{P}_{\lambda,\mu,\delta})$ if and only if

$$u_d - u^* - v^* \in \lambda \partial J_1(u^*), \tag{11a}$$

$$u_d - u^* - (1 + \delta)v^* \in \mu \partial J_2(v^*). \tag{11b}$$

We can perform an explicit computation to get the following result:

Theorem 2 (u^*, v^*) is a solution to $(\mathbf{P}_{\lambda,\mu,\delta})$ if and only if

$$u^* = u_d - v^* - \Pi_{\lambda K_1}(u_d - v^*), \tag{12a}$$

$$v^* = \frac{1}{1 + \delta} (u_d - u^* - \Pi_{\mu K_2}(u_d - u^*)), \tag{12b}$$

where K_1 and K_2 are the following convex closed subsets:

$$K_1 = \{\operatorname{div} p \mid p \in X^2, \|p_{i,j}\|_{\mathbb{R}^2} \leq 1, \forall i = 1, \dots, N, j = 1, \dots, M\}, \tag{13a}$$

$$K_2 = \{H^* p \mid p \in X^4, \|p_{i,j}\|_{\mathbb{R}^4} \leq 1, \forall i = 1, \dots, N, j = 1, \dots, M\}, \tag{13b}$$

and Π_{K_i} denotes the orthogonal projection on K_i .

Proof Following [9, 13], relations (11a), (11b) are equivalent to

$$u^* \in \partial J_1^* \left(\frac{u_d - u^* - v^*}{\lambda} \right) = \partial \iota_{K_1} \left(\frac{u_d - u^* - v^*}{\lambda} \right), \tag{14a}$$

$$v^* \in \partial J_2^* \left(\frac{u_d - u^* - (1 + \delta)v^*}{\mu} \right) = \partial \iota_{K_2} \left(\frac{u_d - u^* - (1 + \delta)v^*}{\mu} \right), \tag{14b}$$

where J^* is the Fenchel-Legendre transform of J , and ι_K is the indicatrix function of K :

$$\iota_K(x) = \begin{cases} 0 & \text{if } x \in K \\ +\infty & \text{else.} \end{cases}$$

Let Π_K be the orthogonal projection on a closed convex set K . Recall that

$$\lambda \in \partial \iota_K(u) \iff \lambda = c \left(u + \frac{\lambda}{c} - \Pi_K \left(u + \frac{\lambda}{c} \right) \right) \iff u = \Pi_K \left(u + \frac{\lambda}{c} \right),$$

for every $c > 0$. Then relation (14a) with $c = \lambda$ is equivalent to

$$u_d - v^* - u^* = \lambda \Pi_{K_1} \left(\frac{u_d - v^*}{\lambda} \right) = \Pi_{\lambda K_1} (u_d - v^*),$$

since $\Pi_K(\frac{\mu}{c}) = \frac{1}{c} \Pi_{cK}(u)$. Similarly (14b) with $c = \frac{\mu}{1+\delta}$ is equivalent to

$$\frac{u_d - u^* - (1 + \delta)v^*}{\mu} = \Pi_{K_2} \left(\frac{u_d - u^*}{\mu} \right) = \frac{1}{\mu} \Pi_{\mu K_2} (u_d - u^*). \quad \square$$

We may write relations (12a), (12b) as a fixed point equation $(u, v) = G(u, v)$, where

$$G : X^2 \rightarrow X^2$$

$$(u, v) \mapsto \left(\begin{array}{l} u_d - v - \Pi_{\lambda K_1} (u_d - v) \\ \frac{1}{1+\delta} (u_d - u - \Pi_{\mu K_2} (u_d - u)) \end{array} \right). \tag{15}$$

Let us introduce G_α defined by

$$G_\alpha(u, v) = \begin{pmatrix} u \\ v \end{pmatrix} + \alpha \left(G(u, v) - \begin{pmatrix} u \\ v \end{pmatrix} \right).$$

We get

$$G_\alpha(u, v) = \begin{pmatrix} (1 - \alpha)u + \alpha(u_d - v - \Pi_{\lambda K_1} (u_d - v)) \\ (1 - \alpha)v + \frac{\alpha}{1+\delta}(u_d - u - \Pi_{\mu K_2} (u_d - u)) \end{pmatrix}. \tag{16}$$

This leads to the following fixed-point algorithm:

Algorithm \mathcal{A}_0

1. **Initialization step.** Choose u_0 and v_0 (for example $u_0 = 0$ and $v_0 = u_d$) and $0 < \alpha < 1/2$.
2. **Iteration.** Define the sequences $((u_n, v_n))_n$ as

$$\begin{cases} u_{n+1} = u_n + \alpha (u_d - v_n - \Pi_{\lambda K_1} (u_d - v_n) - u_n) \\ v_{n+1} = v_n + \frac{\alpha}{1 + \delta} (u_d - u_n - \Pi_{\mu K_2} (u_d - u_n) - (1 + \delta)v_n). \end{cases}$$

3. **Stopping test.**

Theorem 7 For every $\alpha \in (0, 1/2)$, the sequence (u_n, v_n) converges to the (unique) fixed point of G .

Proof It is sufficient to prove that $G_\alpha = (G_\alpha^1, G_\alpha^2)$ is strictly contracting. Let be $\alpha > 0$. For every $(u_1, v_1), (u_2, v_2) \in X^2$, we have

$$\begin{aligned} & \left\| G_\alpha^1(u_1, v_1) - G_\alpha^1(u_2, v_2) \right\|_X + \left\| G_\alpha^2(u_1, v_1) - G_\alpha^2(u_2, v_2) \right\|_X \\ & \leq |1 - \alpha| \|u_1 - u_2\|_X + 2\alpha \|v_1 - v_2\|_X + |1 - \alpha| \|v_1 - v_2\|_X + \frac{2\alpha}{1 + \delta} \|u_1 - u_2\|_X \\ & \leq \max \left(|1 - \alpha|, 2\alpha, \frac{2\alpha}{1 + \delta} \right) (\|u_1 - u_2\|_X + \|v_1 - v_2\|_X). \end{aligned}$$

If $\alpha \in (0, 1/2)$, then $\max(|1 - \alpha|, 2\alpha, \frac{2\alpha}{1 + \delta}) < 1$, and G_α is contracting. Therefore, the sequence $(u_{n+1}, v_{n+1}) = G_\alpha(u_n, v_n)$ converges to the fixed point of G_α . Moreover G and G_α have the same fixed points. □

For the numerical realization a (standard) relaxed version of the algorithm is used.

Algorithm (\mathcal{A})

1. **Initialization step.** Choose u_0 and v_0 (for example $u_0 = 0$ and $v_0 = u_d$) and $0 < \alpha < 1/2$.
2. **Iteration.** Define the sequences $((u_n, v_n))_n$ as

$$\begin{cases} u_{n+1} = u_n + \alpha (u_d - u_n - v_n - \Pi_{\lambda K_1} (u_d - v_n)) \\ v_{n+1} = v_n + \frac{\alpha}{1 + \delta} (u_d - u_{n+1} - (1 + \delta)v_n - \Pi_{\mu K_2} (u_d - u_{n+1})). \end{cases}$$

3. **Stopping test.** We can prove similarly that for any $\alpha \in (0, 1/2)$, the sequence generated by (\mathcal{A}) converges to the fixed point of G .

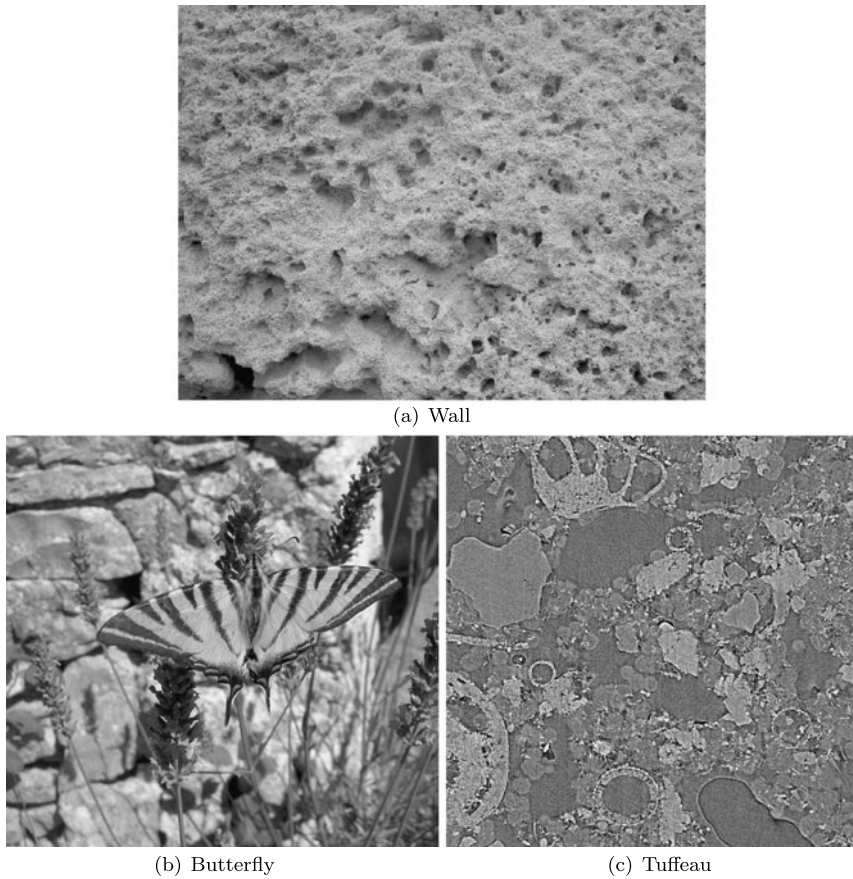


Fig. 1 Examples

4 Numerical results

We have performed numerical experimentation on the three (natural) images of Fig. 1:

- Image (a) is a picture of an old damaged wall which can be considered as pure texture.
- Image (b) involves both sharp contours and small details.
- The third image (c) is a micro-tomographic image of tuffeau (stone): the image is one slice extracted from a 3D tomographic image of a tuffeau sample. The image is 2048×2048 pixels, the pixel size is $0.28 \mu\text{m}$ with resin, silica (opal sphere), air bubble in the resin (caused by the impregnation process), silica (quartz crystal), calcite and phyllosilicate. The segmentation of such an image is a hard challenge [17].

We have computed the total variation J_1 and the second order total variation J_2 of every image and reported them in Table 1.

Table 1 Total variation J_1 , second order total variation J_2 and G -norm for tests images—the last column is the “normalized” G -norm

Image u_d	$J_1(u_d)$	$J_2(u_d)$	$\ u_d\ _G$	$\ u_d\ _{L^2}$	$\ u_d\ _G/\ u_d\ _{L^2}$
Wall (a)	23.27	43.07	7.62	0.5277	14.4408
Butterfly (b)	10.27	11.14	12.11	0.5463	22.1659
Tuffeau (c)	27.62	43.79	2.76	0.4978	5.555

Table 2 G -norm and total variation of components for different parameters λ, μ — \bar{u} is the mean value of u —We recall that $\|u_d\|_G = 7.62$ for image (a) (wall), $\|u_d\|_G = 12.11$ for image (b) (butterfly) and $\|u_d\|_G = 2.76$ for image (c) (tuffeau)

Image	λ	μ	$\ u\ _G$	$\ v\ _G$	$\ w\ _G$	$J_1(u)$	$J_2(v)$	$\ w\ _{L^2}$	$\bar{u} \cdot 10^7$
(a)	2	6	0.376	8.769	0.333	16.80	1.293	3.276	2.46
	5	10	0.510	8.905	0.316	11.24	1.230	6.994	3.39
	10	20	0.630	8.803	0.324	6.218	1.383	11.13	-0.17
	50	100	1.476	7.905	0.452	0.160	1.632	19.88	6.87
(b)	2	6	0.262	12.11	0.225	4.018	2.402	1.996	-0.43
	5	10	0.316	12.00	0.272	2.519	2.461	3.683	0.56
	5	50	0.903	11.437	0.418	6.415	0.626	3.391	5.27
	10	5	0.245	11.90	0.206	0.004	5.05	3.882	-1.78
	10	20	0.423	11.78	0.314	1.908	2.231	5.792	4.03
	50	50	0.825	11.59	0.284	0.043	2.712	11.18	-7.88
	50	100	1.196	11.37	0.363	0.386	2.038	13.63	-2.54
(c)	2	6	0.277	3.089	0.258	20.20	3.68	3.3485	12.69
	10	20	0.460	4.365	0.299	12.61	1.199	10.98	-10.88
	20	50	0.906	6.220	0.281	8.012	0.875	16.58	-43.36
	50	100	1.082	5.232	0.345	1.730	1.502	25.90	-4.07

We have also computed and the G -norm¹ for every image. Recall (see [18]) that the G -space is the dual space of \mathcal{BV} (the closure of the Schwartz class in $BV(\Omega)$):

$$G := \{f \mid \exists \varphi = (\varphi_1, \varphi_2) \in (L^\infty(\mathbb{R}^2))^2 \mid f = \operatorname{div} \varphi\},$$

and the $\|\cdot\|_G$ norm is defined as

$$\|f\|_G := \inf \left\{ \left\| \sqrt{\varphi_1^2 + \varphi_2^2} \right\|_\infty \mid f = \operatorname{div} \varphi \right\}.$$

Though the G -norm cannot measure image oscillations (non oscillating images may have a small G -norm) it may be an indicator on oscillations amplitude. The more the function is oscillating, the smaller its G -norm is.

¹We are very grateful to Pierre Weiss who provided the codes to compute the G -norm efficiently.



Fig. 2 BV^2 component— v — $\mu = 50$ — $\rho := \frac{\lambda}{\mu}$

The projections in step 2 of algorithm have been computed using a Nesterov-type algorithm inspired by [25] which has been adapted for the projection on K_2 . The stopping criterion is based on the difference between two consecutive iterates that should be less than 10^{-3} coupled with a maximal number of iterations (here 175). We give thereafter the values of the G -norm for the components $u, v, w := u_d - u - w$ and different pairs (λ, μ) (see Table 2).

We note that the G -norm of the BV^2 component v (few oscillations) and the L^2 component w (many oscillations) are independent on the choice of λ and μ . This is not the case for the BV component u . Moreover, though the amplitude of the BV component may be quite large (for example, if $\lambda = 2$ and $\mu = 6$ we get $\max u \simeq 139$ and $\min u \simeq -86$ for image (b)) we note (at least numerically) that the mean value of the BV component is always null. The same holds for the remainder term (though it is less significant since it is much smaller). This confirms that the BV^2 component involves all the image dynamic information as contrast, luminance and so on.

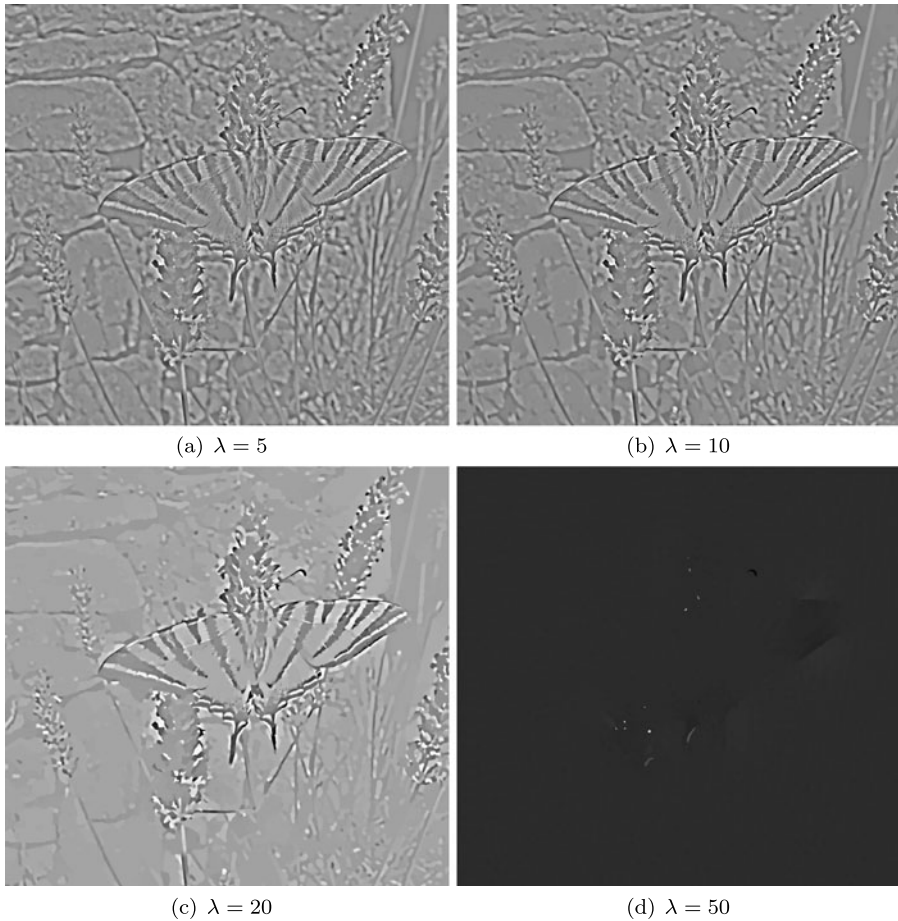


Fig. 3 BV component— $u - \mu = 50$

We present thereafter some results² for different values of λ and μ . All numerical tests have been performed with $\delta = 0$ since we noticed this parameter has no influence on the results. We use MATLAB[®] software. We do not report on CPU time since our numerical codes have not been optimized.

In what follows images have been contrasted or equalized to be more “readable”.

4.1 Sensitivity with respect to λ

We can see that the ratio $\rho := \frac{\lambda}{\mu}$ is significant: indeed if $\mu \gg \lambda$ the second-order term is more weighted than the first order one and the BV^2 component (see Fig. 2) has a small second derivative. This means that there are less and less details as the ratio ρ grows and the resulting image is more and more blurred.

²Many others examples can be found at <http://web.me.com/maitine.bergounioux/PagePro/Publications.html>.

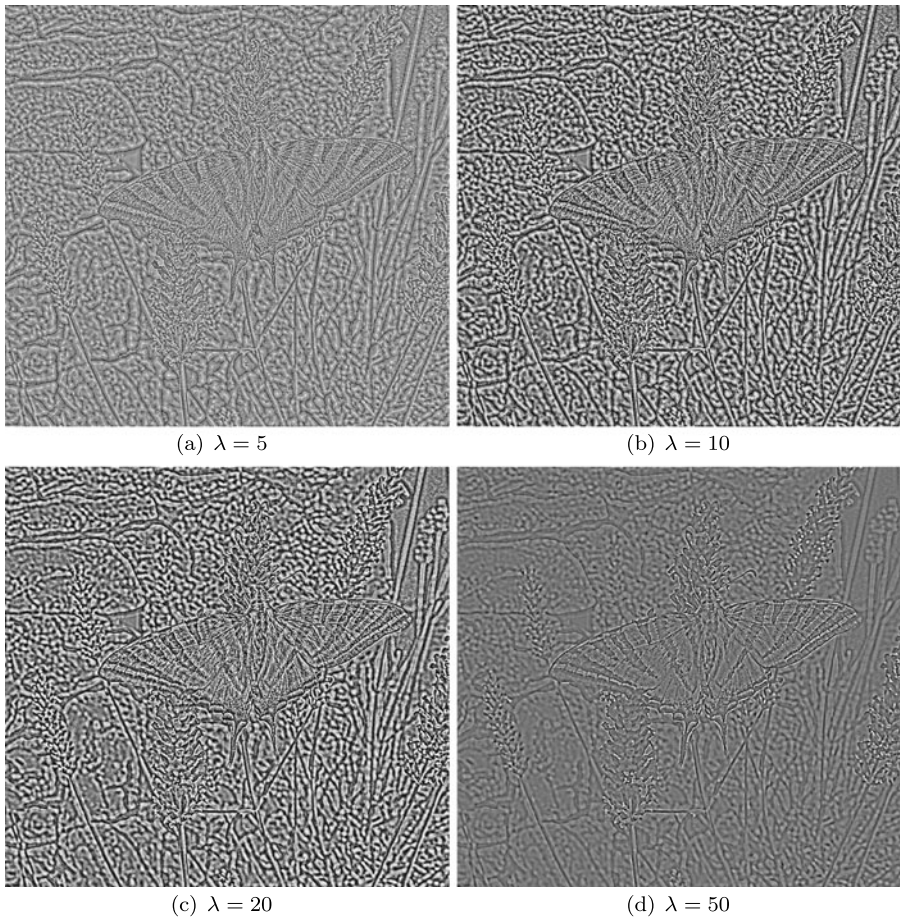


Fig. 4 L^2 component— $w = u_d - u - v - \mu = 50$ —images after histogram equalization

The ratio ρ is less significant for the BV component u which is sensible to the λ parameter (see Fig. 3). One sees that the larger λ is, the more u is piecewise constant. This is consistent with the fact that the optimal value for $\Phi_1(u)$ should be smaller as λ grows.

Moreover, if λ is large enough then $u = 0$ (Fig. 3(d)). Indeed we have noticed that the optimal solution (u^*, v^*) satisfies (4). This means that u^* is the solution to the classical Rudin-Osher-Fatemi problem

$$u^* = \operatorname{argmin} \left\{ \frac{1}{2} \|f - u\|^2 + \lambda \Phi_1(u), u \in BV(\Omega) \right\}$$

with $f := u_d - v^*$. With a result by Meyer ([18], Lemma 3, p. 42) we conclude that $u^* = 0$ if $\lambda > \|u_d - v^*\|_G$. Figure 4 shows the L^2 component.

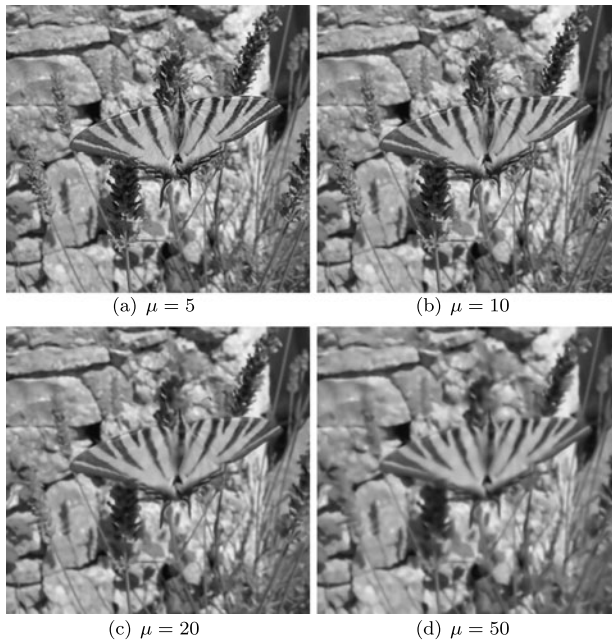


Fig. 5 BV^2 component— $v-\lambda = 10$

4.2 Sensitivity with respect to μ

The same comments hold: the ratio ρ is the significant quantity with respect to the behavior of the BV^2 component (see Fig. 5).

If $\mu \ll \lambda$ then the BV component u is constant (this is consistent with the fact that λ is large enough). Once again, if μ grows (while λ is fixed) the BV component is “less” piecewise constant. The effect of μ on the remainder term w seems more significant than the effect of λ , see Figs. 6 and 7.

4.3 Decomposition with 3 components

We present (Figs. 8, 9, 10 and 11) the three components together for image (a) and different values of λ and μ . This image may be considered as pure texture. We clearly see that the BV^2 component involves the image dynamic, the BV component u extracts a macro-texture and the remainder term w a micro-structure. The scaling between u and w is tuned via parameters λ and μ (the ratio $\rho := \frac{\lambda}{\mu}$ has no influence).

4.4 Denoising and texture extraction

We end with image (c) (Figs. 12, 13 and 14). It is quite difficult to perform segmentation of such an image. Indeed, the image is noisy and there are texture areas (due to the micritic calcite part). The denoising process should preserve the texture which involves physical information. As we want to recover the vacuum area we have to

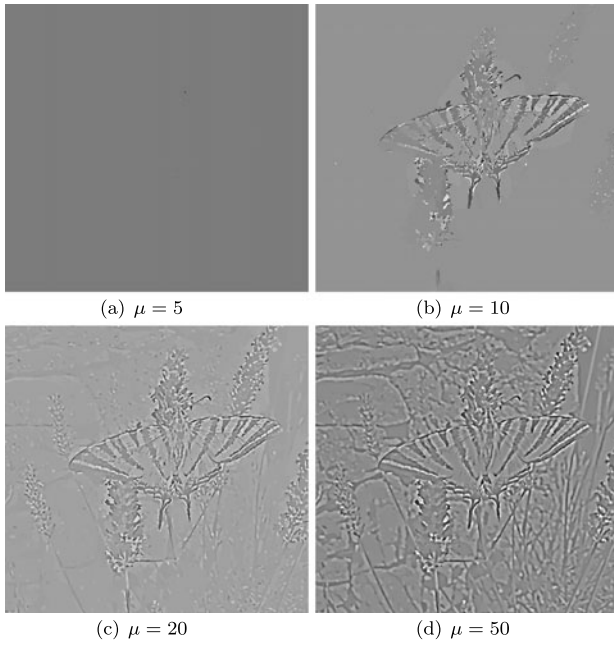


Fig. 6 BV component— $u-\lambda = 10$

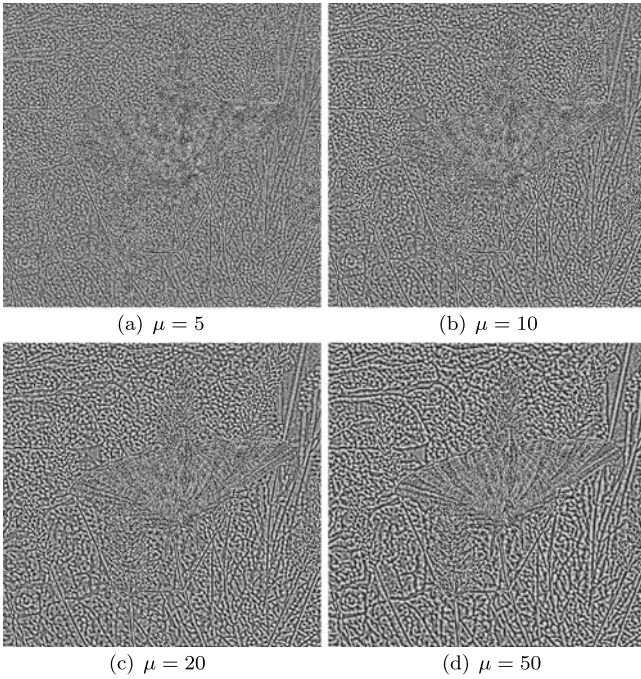


Fig. 7 L^2 component— $w = u_d - u - v - \lambda = 10$ —Histogram equalization has been performed

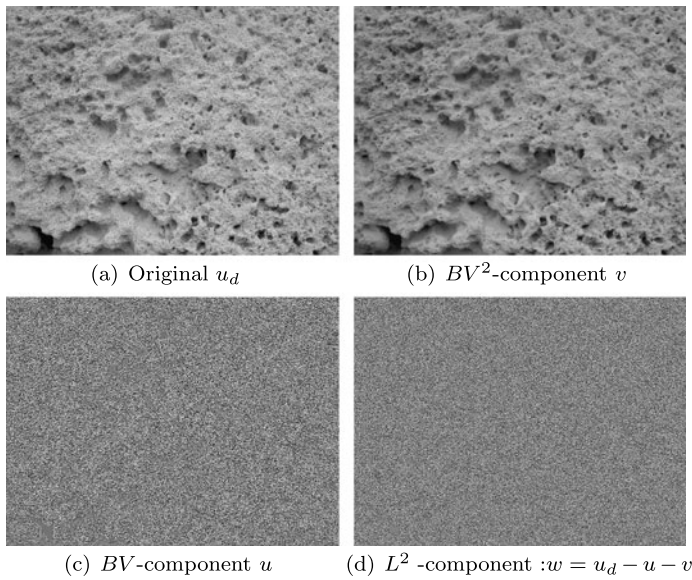


Fig. 8 Wall for $\lambda = 1$ and $\mu = 1 - \rho = 1$

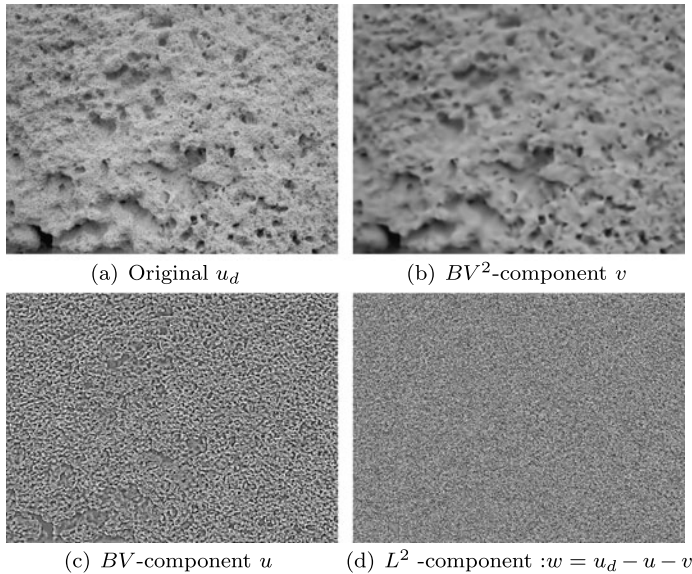


Fig. 9 Wall for $\lambda = 2$ and $\mu = 6 - \rho = 0.33$

perform a contour segmentation and if possible regions classification to recover the different physical components of the stone. The decomposition model we propose, can be used as a pre-processing treatment to separate the noise and fine texture com-

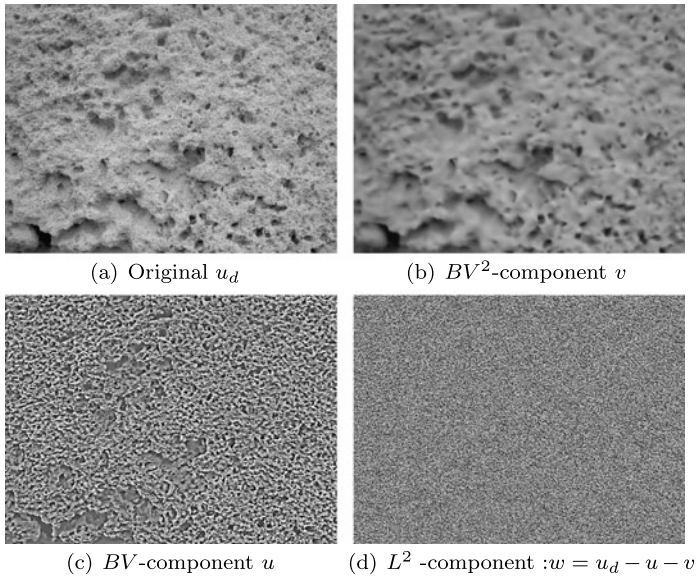


Fig. 10 Wall for $\lambda = 5$ and $\mu = 10 - \rho = 0.5$

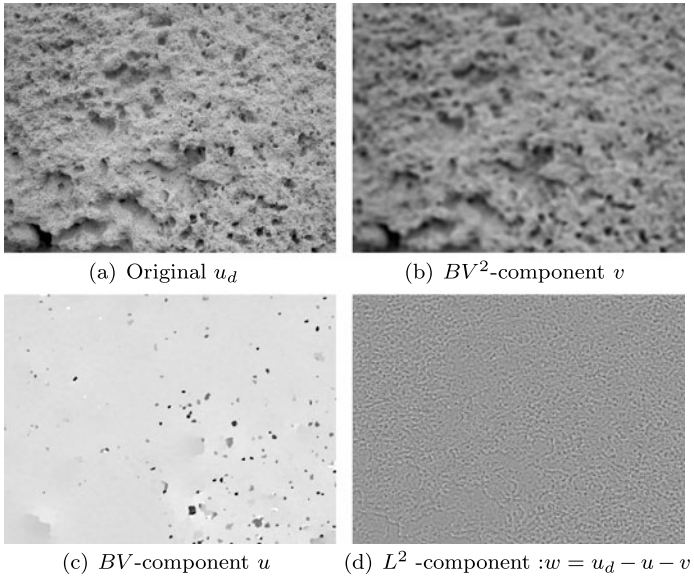


Fig. 11 Wall for $\lambda = 50$ and $\mu = 100 - \rho = 0.5$

ponent w from the macro-texture component u and perform a classical segmentation method on u .

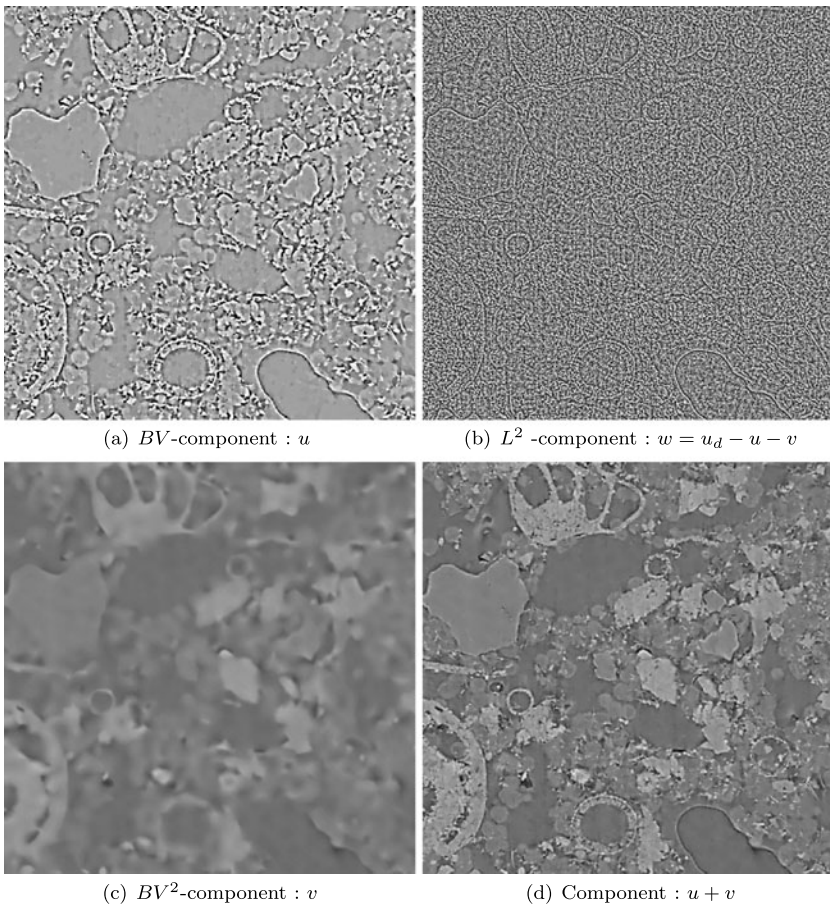


Fig. 12 Denoising and texture extraction— $\lambda = 10$, $\mu = 20$

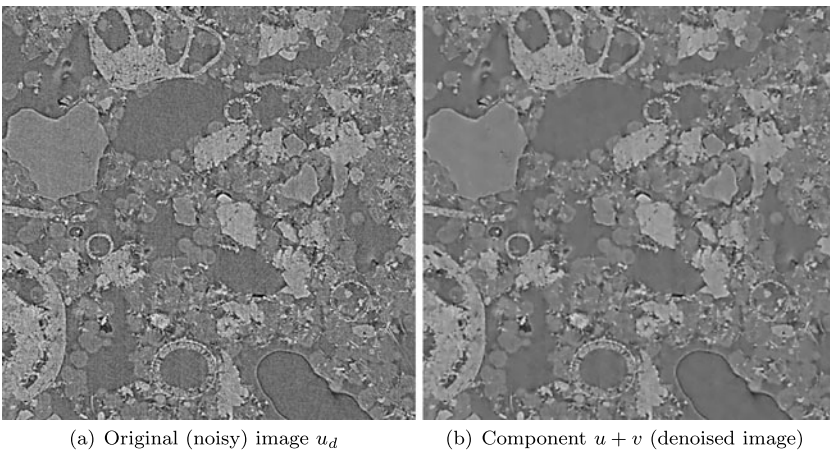


Fig. 13 Original and denoised image— $\lambda = 10$, $\mu = 20$

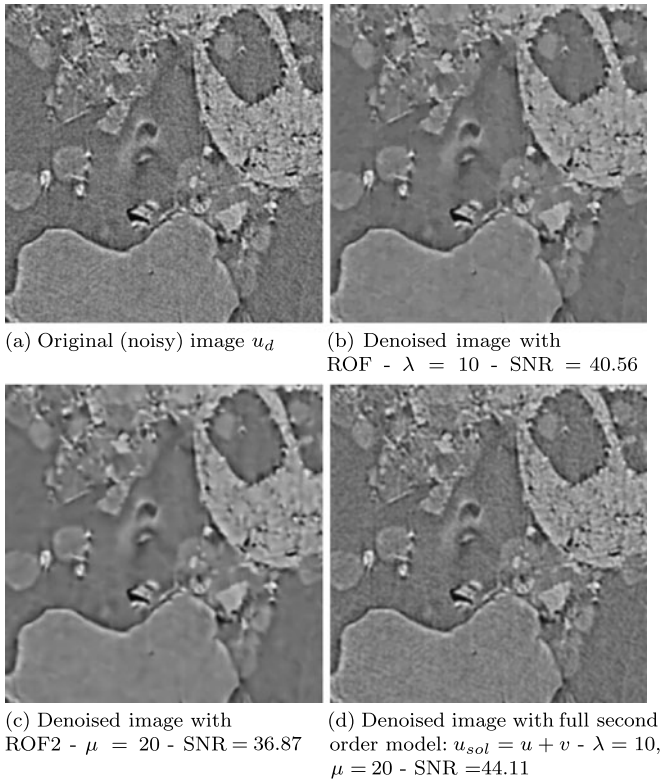


Fig. 14 Comparison of ROF, ROF2 and full second order model with SNR

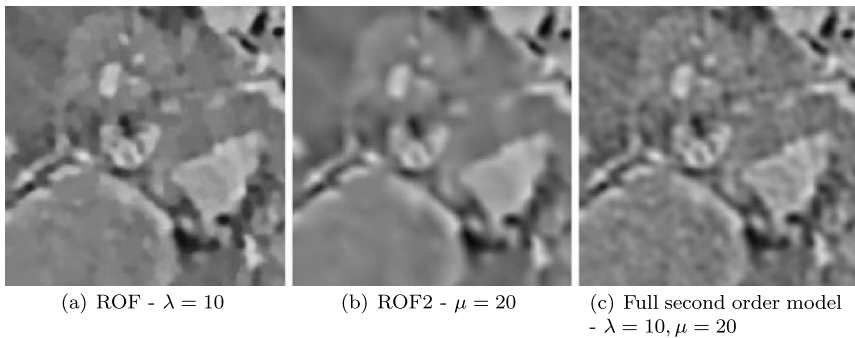


Fig. 15 Zoom for the comparison of ROF, ROF2 and full second order model

5 Conclusion

From the denoising point of view this model provides a good compromise between the ROF and the ROF2 [9] models as shown in Figs. 14 and 15. The signal to noise

ratio (SNR) is computed as

$$SNR(u_{sol}) = 20 \log_{10} \left(\frac{\|u_d\|_{L^2}}{\|u_d - u_{sol}\|_{L^2}} \right),$$

where u_{sol} is the computed solution and u_d the original noisy image.

Indeed, the solution to the ROF-model is piecewise constant: this why we observe staircasing effects. Similarly, the ROF2 model solution is piecewise affine so that we do not get staircasing any longer but the denoised image is slightly blurred. The use of the full second order makes it possible to keep all the advantages of the two precedents without having the disadvantages of them.

This model seems promising to perform two-scale texture analysis. This can be used to analyse the structure of radiographs in osteoporosis context [10, 11]. Indeed, the choice parameters λ and μ determines the scale of the macro-texture/micro-texture. Once this part has been isolated, it possible to perform segmentation or statistical analysis.

There are many open questions to be addressed in future works: existence an uniqueness results without penalization terms have to be investigated together with a sharp analysis of the continuous model. The comparison with existing models should be extensively performed as well: one can find many comparison tests in [23].

A last we have to improve the numerical process (both discretization and algorithm) to perform 3D tests in the future.

References

1. Acar, R., Vogel, C.R.: Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Probl.* **10**(6), 1217–1229 (1994)
2. Ambrosio, L., Fusco, N., Pallara, D.: *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford Mathematical Monographs. Oxford University Press, Oxford (2000)
3. Attouch, H., Buttazzo Michaille, G.: In: *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization*. MPS-SIAM Series on Optimization (2006)
4. Aubert, G., Aujol, J.F.: Modeling very oscillating signals, application to image processing. *Appl. Math. Optim.* **51**(2), 163–182 (2005)
5. Aubert, G., Aujol, J.F., Blanc-Feraud, L., Chambolle, A.: Image decomposition into a bounded variation component and an oscillating component. *J. Math. Imaging Vis.* **22**(1), 71–88 (2005)
6. Aubert, G., Kornprobst, P.: *Mathematical Problems in Image Processing, Partial Differential Equations and the Calculus of Variations*. Applied Mathematical Sciences, vol. 147. Springer, Berlin (2006)
7. Aujol, J.F.: *Contribution à l'analyse de textures en traitement d'images par méthodes variationnelles et équations aux dérivés partielles*. PhD Thesis, Nice (2004)
8. Bergounioux, M.: On Poincaré-Wirtinger inequalities in BV-spaces. *Control Cybern.* To appear, Issue 4 (2011)
9. Bergounioux, M., Piffet, L.: A second-order model for image denoising and/or texture extraction. *Set-Valued Var. Anal.* **18**(3–4), 277–306 (2010)
10. Biermé, H., Benhamou, C.L., Richard, F.: Parametric estimation for Gaussian operator scaling random fields and anisotropy analysis of bone radiograph textures. In: Pohl, K. (ed.) *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'09)*, London, pp. 13–24 (2009)
11. Biermé, H., Richard, F.: Analysis of texture anisotropy based on some Gaussian fields with spectral density. In: Bergounioux, M. (ed.) *Mathematical Image Processing*. Springer Proceedings in Mathematics, vol. 5, pp. 59–74. Springer, Berlin (2011)

12. Caselles, V., Chambolle, A., Novaga, M.: The discontinuity set of solutions of the TV denoising problem and some extensions. *SIAM Multiscale Model. Simul.* **6**(3), 879–894 (2007)
13. Chambolle, A.: An algorithm for total variation minimization and applications. *J. Math. Imaging Vis.* **20**, 89–97 (2004)
14. Demengel, F.: Fonctions à Hessien borné. *Ann. Inst. Fourier* **34**(2), 155–190 (1984)
15. Ekeland, I., Temam, R.: *Convex Analysis and Variational Problems*. SIAM Classic in Applied Mathematics, vol. 28 (1999)
16. Evans, L.C., Gariepy, R.: *Measure Theory and Fine Properties of Functions*. CRC Press, New York (1992)
17. Guillot, L., Le Trong, E., Rozenbaum, O., Bergounioux, M., Rouet, J.L.: A mixed model of active geodesic contours with gradient vector flows for X-ray microtomography segmentation. In: Bergounioux, B.M. (ed.) *Actes du colloque Mathématiques pour l'image* (2009). PUO <http://hal.archives-ouvertes.fr/hal-00267007/fr/>
18. Meyer, Y.: *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*. University Lecture Series, vol. 22. AMS, New York (2001)
19. Osher, S., Fatemi, E., Rudin, L.: Nonlinear total variation based noise removal algorithms. *Physica D* **60**, 259–268 (1992)
20. Osher, S., Sole, A., Vese, L.: Image decomposition and restoration using total variation minimization and the H^1 norm. *SIAM J. Multiscale Model. Simul.* **1–3**, 349–370 (2003)
21. Osher, S., Vese, L.: Modeling textures with total variation minimization and oscillating patterns in image processing. *J. Sci. Comput.* **19**(1–3), 553–572 (2003)
22. Osher, S.J., Vese, L.A.: Image denoising and decomposition with total variation minimization and oscillatory functions. *J. Math. Imaging Vis.* **20**(1–2), 7–18 (2004). Special issue on mathematics and image analysis
23. Piffet, L.: *Modèles variationnels pour l'extraction de textures 2D*. PhD Thesis, Orléans (2010)
24. Ring, W.: Structural properties of solutions of total variation regularization problems. *ESAIM, Math. Model. Numer. Anal.* **34**, 799–840 (2000)
25. Weiss, P., Blanc-Féraud, L., Aubert, G.: Efficient schemes for total variation minimization under constraints in image processing. *SIAM J. Sci. Comput.* (2009)
26. Yin, W., Goldfarb, D., Osher, S.: A comparison of three total variation based texture extraction models. *J. Vis. Commun. Image* **18**, 240–252 (2007)