Noname manuscript No.
(will be inserted by the editor)

A full second order model for multiscale texture analysis

Maïtine Bergounioux · Loïc Piffet

Received: date / Accepted: date

Abstract We presented 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 remainder term (0 order). For highly textured images the model gives a two-scale texture decomposition: u can be viewed as a "macro-texture" (larger scale) which oscillations are not too large and w is the "micro-texture" (very oscillating) that contains the 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 model is the Rudin-Osher-Fatemi denoising model (see [1, 16]). This model involves a regularization term that preserves discontinuities, what a classical H^1 -Tychonov regularization method does not. The observed image to recover is split in two parts $u_d = u + v$ where u represents the oscillating component (noise or texture) and 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 ([2,3,13]). The

M. Bergounioux
Université d'Orléans, UFR Sciences, Math., Labo. MAPMO, UMR 6628, Route de Chartres, BP 6759 45067 Orléans cedex 2,
Tel.: +33 238 -417316
Fax: +33 238 -417205
E-mail: maitine.bergounioux@univ-orleans.fr

L. Piffet
Université d'Orléans, UFR Sciences, Math., Labo. MAPMO, UMR 6628, Route de Chartres, BP 6759 45067 Orléans cedex 2,
E-mail: loic.piffet@univ-orleans.fr

regularization term involves only the so-called cartoon component v, while the remainder term $u := u_d - v$ represents the noise to be minimized.

A lot of people have investigated such decomposition models based on variational formulation , considering that an image can be decomposed into many components, each component describing a particular property of the image ([4,5,17–19,22] and references therein for example).

In [9] we have presented a (partial) second order model where the (first order) classical total variation term has been replaced by a second order total variation term (with the appropriate functional framework), that we have called ROF2. The use of such a model allows to get rid of the "staircasing effect" that appears with the ROF model in denoising processes.

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). 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 in three components: a smooth (continuous) part v, a "cartoon" (piecewise constant) 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 [15] 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 hope that u is piecewise constant so that its jump set gives the image contours. For highly textured images as the one of example (a) in figure 1, we shall see that the model gives a two-scale texture decomposition: u can be viewed as a "macro-texture" (larger scale) which oscillations are not too large and w is the "micro-texture" (very oscillating) that contains 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 gives 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 [9,11], 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

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

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

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

where

$$div(\xi) = (div(\xi_1), div(\xi_2), \ldots, div(\xi_n)), \tag{2}$$

with

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

We give thereafter important properties of these spaces which proofs can be found in [2,3,11,20] 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 $\dfrac{\partial u}{\partial x_i} \in BV(\Omega)$ 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 operator Φ_2 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 as well:

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 $q \leq \dfrac{n}{n-1}$, with continuous embedding. Moreover the embedding is compact if $q < \frac{n}{n-1}$.*
In particular

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

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 [9], we have studied the following restoration variational model

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

which only involves a second order term. We have shown that we can get rid of the staircasing effect while restoring noisy data. However, contours where not preserved as well as wanted and the resulting image was a bit blurred. To overcome this difficulty, we now consider a full second order model and we focus on texture extraction. More precisely, we assume that the image we want to recover from the data u_d can be decomposed as $u_d = w + u + v$ where u, v and w are functions that characterize different parts of u_d. In the sequel $u_d \in L^2(\Omega)$.

Components belong to different functional spaces: v is the (smooth) second order part and belongs to $BV^2(\Omega)$, u is $BV(\Omega)$ component and $w \in L^2(\Omega)$ is

the remainder term. In other words we expect v to be the smooth "colored" part of the image, u to be a $BV(\Omega)\backslash BV^2(\Omega)$ function which derivative is a measure supported by the contours and $w := u_d - u - v \in L^2$ is the noise and/or small textures.

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

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

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

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

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

- the first term $\|u_d - u - v\|^2_{L^2(\Omega)}$ is the fitting data term,
- $\Phi_1(u)$ is a standard total variation term widely used in image restoration (introduced in [16]),
- $\Phi_2(v)$ penalizes the second order total variation of component v,
- $\frac{\delta_1}{2}\|v\|^2_{L^2(\Omega} + \delta_2\Phi_1(v)$ 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. We could replace this penalization term by $\|v\|^2_{H^1}$ or $\|v\|_{W^{1,1}(\Omega)}$. We chose an intermediate penalization term between these two possibilities. The $W^{1,1}(\Omega)$ norm would have involved the L^1-norm of v which is non differentiable. As we already deal with the L^1 norm of ∇v (namely $\Phi_1(v)$ in our case) which is non differentiable as well, this would have added numerical technical difficulties. On the other hand the use of H^1-norm leads to the Laplacian operator on the dual of $H^1(\Omega)$ which is not easy to handle. Moreover, as we shall see in the sequel, we may set $\delta_2 = 0$ once the problem is discretized. In addition, we have noticed a null effect of δ_1 when performing numerical tests, so that we finally choose $\delta_1 = \delta_2 = 0$
  The δ-part is not useful (and not justified) from the modelling point of view. It is only necessary to prove existence and uniqueness of solutions.

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 at 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.

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

The sequence $(v_n)_{n\in\mathbb{N}}$ is bounded in $BV^2(\Omega)$: indeed

- it is bounded in L^2 (with the δ_1-term) and in L^1 since Ω is bounded
- $\Phi_1(v_n)$ is bounded thanks to the δ_2-term, and

– $\Phi_2(v_n)$ is bounded as well.

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:

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

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

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

Finally

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

The pair (u^*,v^*) is a solution to $(\mathcal{P}_{\lambda,\mu,\delta})$.
Uniqueness is straightforward with the strict convexity of $F_{\lambda,\mu,\delta}$ due to the δ_1-term and the data fitting term.

■

Remark 1 *We can get rid of the 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, we require that the BV component u has 0 mean value and/or that the BV^2 component v vanishes on the boundary. Moreover, we do not have strict convexity any longer so that uniqueness of solutions is not ensured any more.. This will be precisely investigated in a future work.*

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

$$u^*=\mathrm{argmin}\{\frac{1}{2}\|\mathrm{u_d}-\mathrm{v}^*-\mathrm{u}\|^2+\lambda\Phi_1(\mathrm{u})\ ,\mathrm{u}\in\mathrm{BV}(\Omega)\}\,, \tag{4}$$

$$v^*=\mathrm{argmin}\{\ \frac{1}{2}\|\mathrm{u_d}-\mathrm{u}^*-\mathrm{v}\|^2+\frac{\delta_1}{2}\|\mathrm{v}\|^2_{\mathrm{L}^2}+\delta_2\Phi_1(\mathrm{v})+\mu\Phi_2(\mathrm{v}),\mathrm{v}\in\mathrm{BV}^2(\Omega)\}.$$

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 [3,12] 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

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

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 for example by the forward scheme:

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

where

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

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

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

where $\left\|(\nabla u)_{i,j}\right\|_{\mathbb{R}^2} = \left\|\left(\nabla u^1_{i,j}, \nabla u^2_{i,j}\right)\right\|_{\mathbb{R}^2} = \sqrt{\left(\nabla u^1_{i,j}\right)^2 + \left(\nabla u^2_{i,j}\right)^2}$.
The discrete divergence operator div is the opposite of the adjoint operator of the gradient operator ∇:

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

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:

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

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

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

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

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

Theorem 6 *Assume $\lambda \geq 0, \mu \geq 0, \delta_2 \geq 0$ and $\delta_1 > 0$. Problem $\mathbf{P}_{\lambda,\mu,\delta}$ 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

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

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 $(\mathcal{P}_{\lambda,\mu,\delta})$ if and only if*

$$u_d - u^* - v^* \in \lambda \partial J_1(u^*), \tag{10a}$$

$$u_d - u^* - (1+\delta)v^* \in \mu \partial J_2(v^*). \tag{10b}$$

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

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

$$u^* = u_d - v^* - \Pi_{\lambda K_1}(u_d - v^*)\ , \tag{11a}$$

$$v^* = \frac{1}{1+\delta}\left(u_d - u^* - \Pi_{\mu K_2}(u_d - u^*)\right). \tag{11b}$$

where K_1 and K_2 are the following convex closed subsets :

$$K_1 = \{\text{div } p \mid p \in X^2,\ \|p_{i,j}\|_{\mathbb{R}^2} \leq 1\ \forall i = 1,\dots,N, j = 1,\dots,M\} \tag{12a}$$

$$K_2 = \{H^* p \mid p \in X^4,\ \|p_{i,j}\|_{\mathbb{R}^4} \leq 1,\ \forall i = 1,\dots,N, j = 1,\dots,M\}. \tag{12b}$$

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

Proof.- Following [9,10], relations (10) are equivalent to

$$u^* \in \partial J_1^* \left(\frac{u_d - u^* - v^*}{\lambda} \right) = \partial \iota_{K_1} \left(\frac{u_d - u^* - v^*}{\lambda} \right), \tag{13a}$$

$$v^* \in \partial J_2^* \left(\frac{u_d - u^* - (1+\delta)v^*}{\mu} \right) = \partial \iota_{K_2} \left(\frac{u_d - u^* - (1+\delta)v^*}{\mu} \right), \tag{13b}$$

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

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

for every $c > 0$. Then relation (13a) with $c = \lambda$ is equivalent to

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

since $\Pi_K\left(\frac{u}{c}\right) = \frac{1}{c}\Pi_{cK}\left(u\right)$. Similarly (13b) with $c = \frac{\mu}{1+\delta}$ is equivalent to

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

■

We may write relations (11) as a fixed point equation $(u, v) = G(u, v)$, where

$$\begin{aligned} G : X^2 &\to X^2 \\ (u, v) &\mapsto \begin{pmatrix} u_d - v - \Pi_{\lambda K_1}\left(u_d - v\right) \\ \frac{1}{1+\delta}\left(u_d - u - \Pi_{\mu K_2}\left(u_d - u\right)\right) \end{pmatrix}. \end{aligned} \tag{14}$$

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

$$G_\alpha(u, v) = \begin{pmatrix} (1-\alpha)u + \alpha(u_d - v - \Pi_{\lambda K_1}\left(u_d - v\right)) \\ (1-\alpha)v + \frac{\alpha}{1+\delta}(u_d - u - \Pi_{\mu K_2}\left(u_d - u\right)) \end{pmatrix}. \tag{15}$$

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

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

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 de G.*

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

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

If $\alpha \in (0\,1/2)$, then $\max\left(|1-\alpha|, 2\alpha, \frac{2\alpha}{1+\delta}\right) < 1$, and G_α is a contraction. Therefore, the sequence $(u_{n+1}, v_{n+1}) = G_\alpha(u_n, v_n)$ converges to G_α fixed point. 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

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

3. **Stopping test.**

We can prove similarly that for $\alpha \in (0, 1/2)$, the sequence given by $\mathcal{A}$ converges to G fixed point.

4 Numerical results

We have tested the model and performed numerical experimentation on images of Figure 1. Image (a) is a picture of an old wall that 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 20485× 20485 pixels, pixel size is 0.28 μm 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 [14].

We have computed the total variation J_1 and the second order total variation J_2 that we report in Table 1.

We have also computed and the G-norm[1] for every image. Recall (see [15]) that the G-space is the dual space of $\mathcal{BV}$ (the closure of the Schwartz class in $BV(\Omega)$) :

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

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

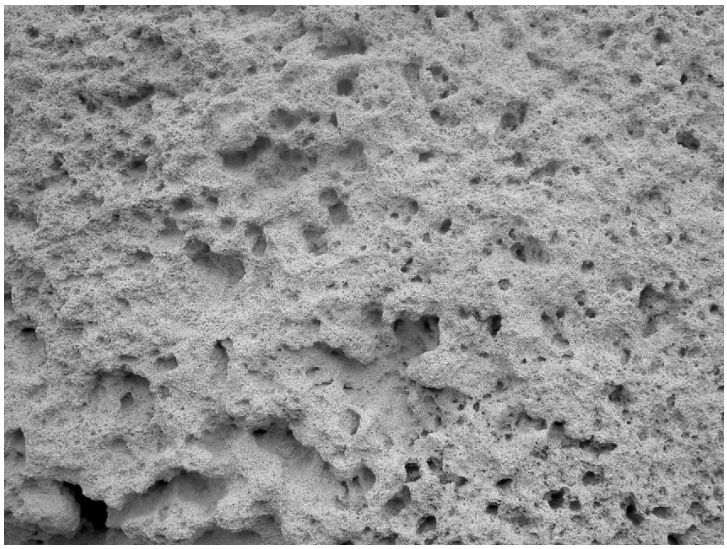
$$\|f\|_G := \inf\{\|\sqrt{\varphi_1^2 + \varphi_2^2}\|_\infty \mid f = \text{div}\ \varphi\ \}\ .$$

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.

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

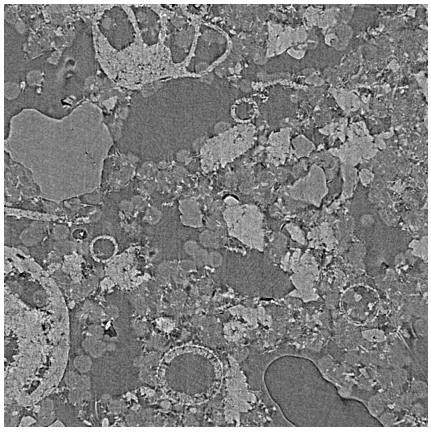
Table 1 Total variation J_1, second order total variation J_2 and G-norm for tests images - the last column is the "normalized" G-norm: we have divided by the L^2-norm.

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



(a) Wall

(b) Butterfly



(c) Tuffeau

Fig. 1 Examples

The projections in step 2. have been computed using a Nesterov-type algorithm inspired by [21] 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) :

Image	λ	μ	$\|u\|_G$	$\|v\|_G$	$\|w\|_G$	$J_1(u)$	$J_2(v)$	$\|w\|_{L^2}$	$\bar{u} \cdot 10^7$
	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
(a)	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
	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
(b)	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
	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
(c)	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

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)

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.
We present thereafter some results[2] 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".

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

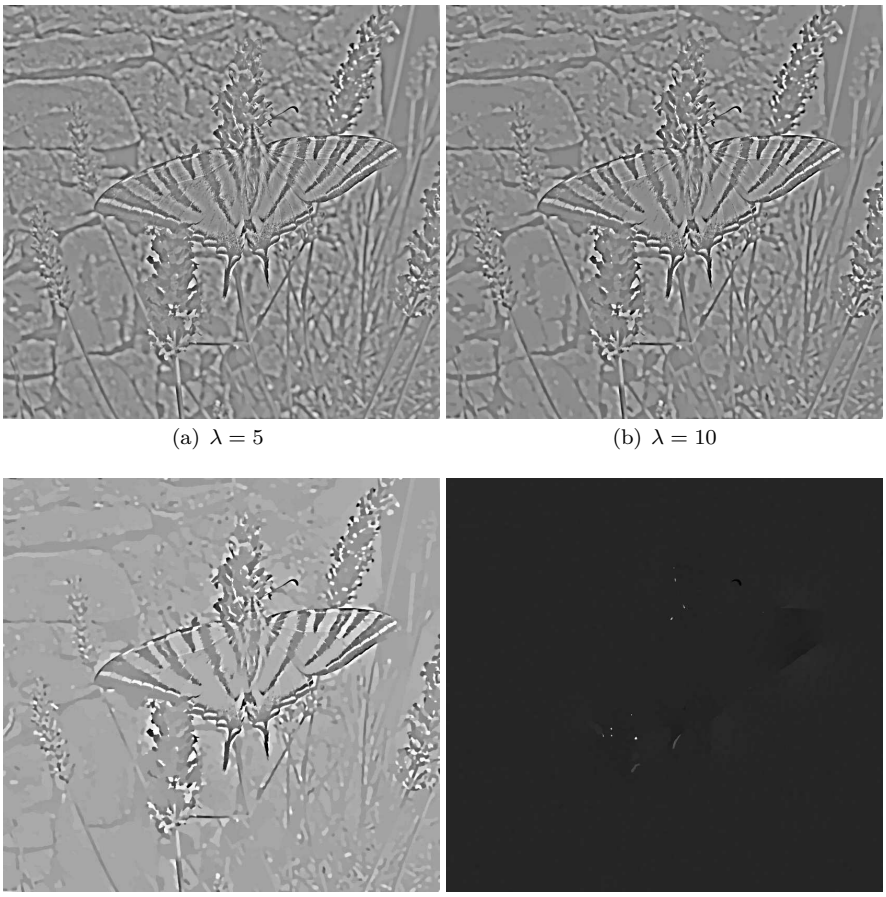
4.1 Sensitivity with respect to λ

(a) $\lambda = 5$ $(\rho = 10)$ (b) $\lambda = 10$ $(\rho = 5)$

(c) $\lambda = 20$ $(\rho = 2.5)$ (d) $\lambda = 50$ $(\rho = 1)$

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

We can see that the ratio $\rho := \dfrac{\lambda}{\mu}$ is significant : indeed if $\mu >> \lambda$ the second-order term is more weighted than the first order one and the BV^2 component 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.



(a) $\lambda = 5$ (b) $\lambda = 10$

(c) $\lambda = 20$ (d) $\lambda = 50$

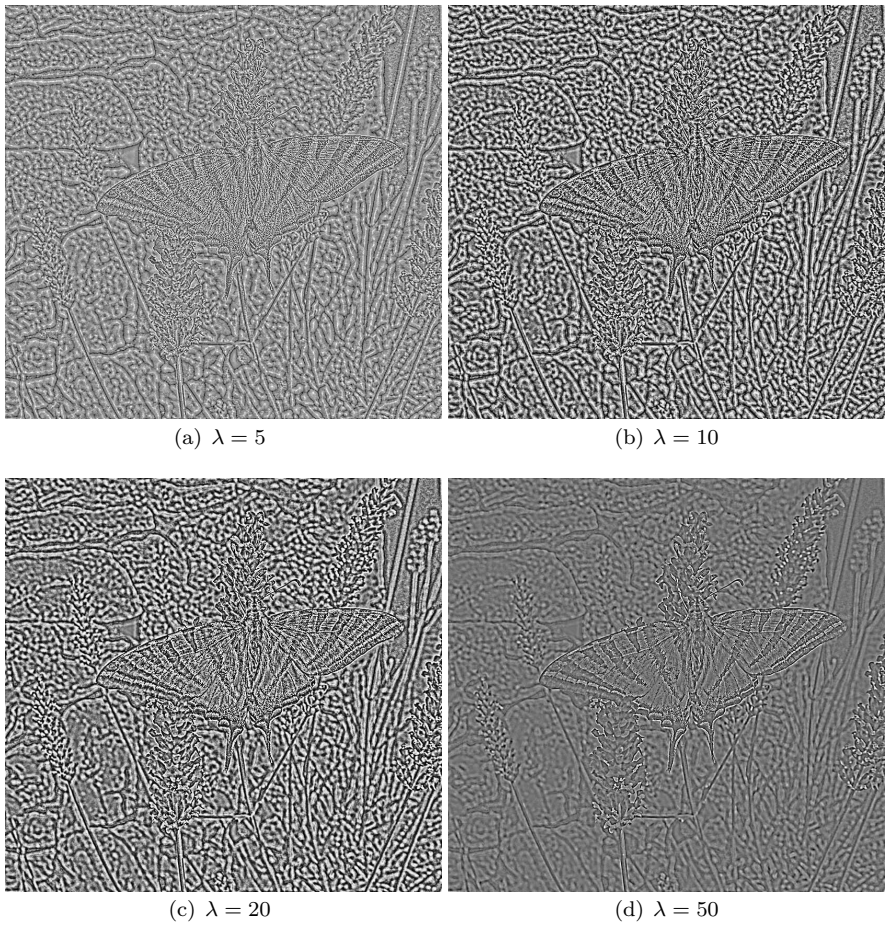
Fig. 3 BV component - u - $\mu = 50$

The ratio ρ is less significant for the BV component u which is sensible to the λ parameter. 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

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

with $f := u_d - v^*$. With a result by Meyer ([15], Lemma 3, p.42) we conclude that $u^* = 0$ if $\lambda > \|u_d - v^*\|_G$.



(a) $\lambda = 5$ (b) $\lambda = 10$

(c) $\lambda = 20$ (d) $\lambda = 50$

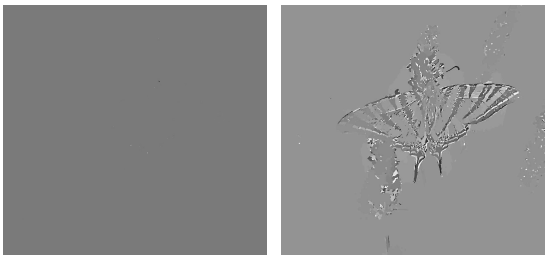
Fig. 4 L^2 component - $w = u_d - u - v$- $\mu = 50$ - images after histogram equalization

4.2 Sensitivity with respect to μ

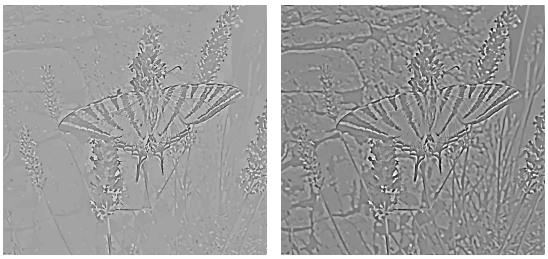
(a) $\mu = 5$ (b) $\mu = 10$

(c) $\mu = 20$ (d) $\mu = 50$

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

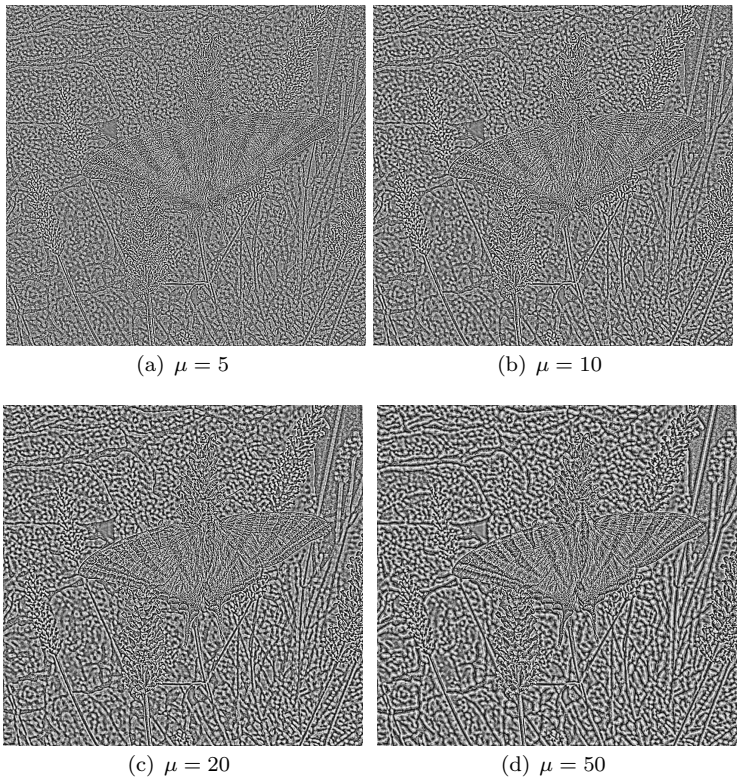


(a) $\mu = 5$ (b) $\mu = 10$



(c) $\mu = 20$ (d) $\mu = 50$

Fig. 6 BV component - u - $\lambda = 10$ -



(a) $\mu = 5$ (b) $\mu = 10$

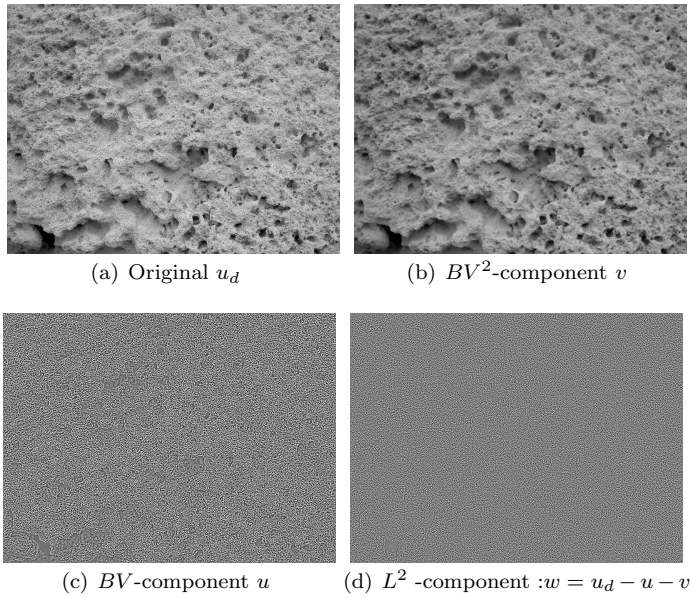
(c) $\mu = 20$ (d) $\mu = 50$

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

The same comments hold : the ratio ρ is the significant quantity with respect to the behaviour of the BV^2 component.
If $\mu << \lambda$ 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.

4.3 Decomposition with 3 components

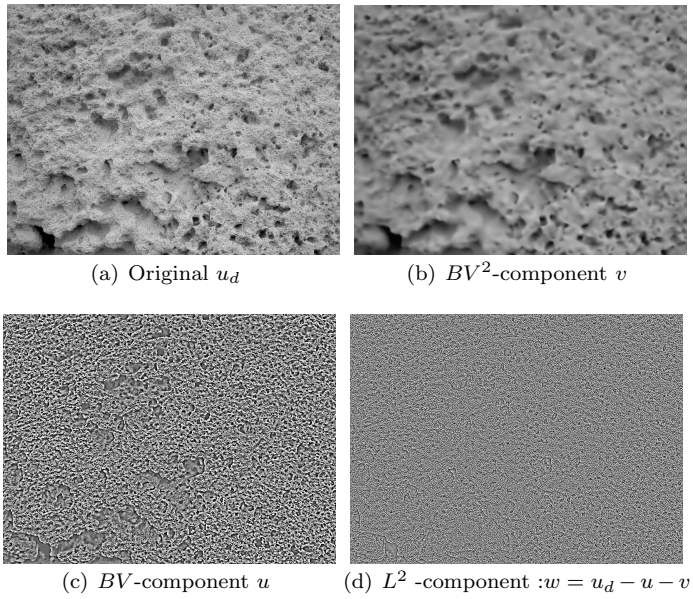
We present the 3 components altogether on image (a) for some values of λ and μ.This image may be considered as pure texture. We clearly see that the BV^2 component involves the image dynamic, the BV component u extracts a macro-texture and the remainder term w a micro-structure. The scaling between u and w is tuned via parameters λ and μ (the ratio $\rho := \dfrac{\lambda}{\mu}$ has no influence).



(a) Original u_d (b) BV^2-component v

(c) BV-component u (d) L^2 -component :$w = u_d - u - v$

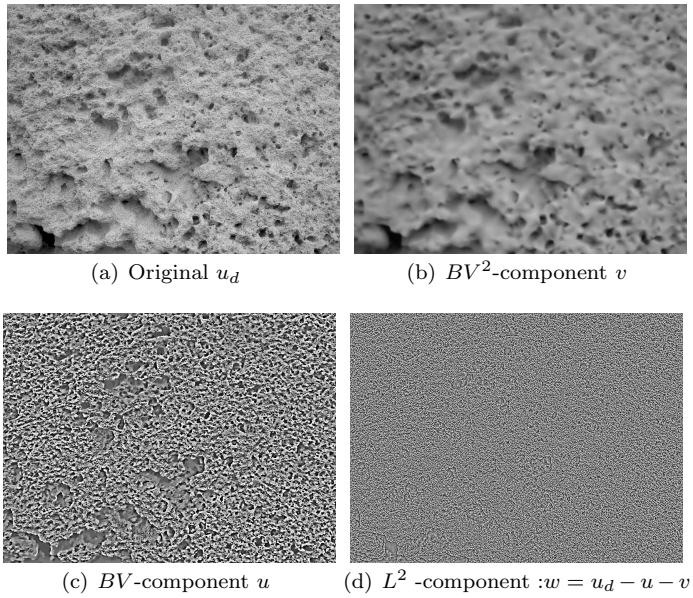
Fig. 8 Wall for $\lambda = 1$ and $\mu = 1$ - $\rho = 1$



(a) Original u_d (b) BV^2-component v

(c) BV-component u (d) L^2 -component :$w = u_d - u - v$

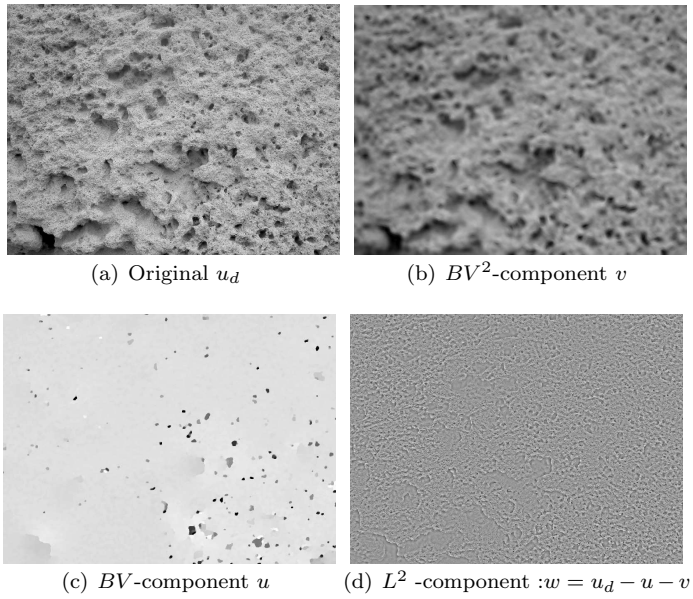
Fig. 9 Wall for $\lambda = 2$ and $\mu = 6$ - $\rho = 0.33$



(a) Original u_d (b) BV^2-component v

(c) BV-component u (d) L^2 -component :$w = u_d - u - v$

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



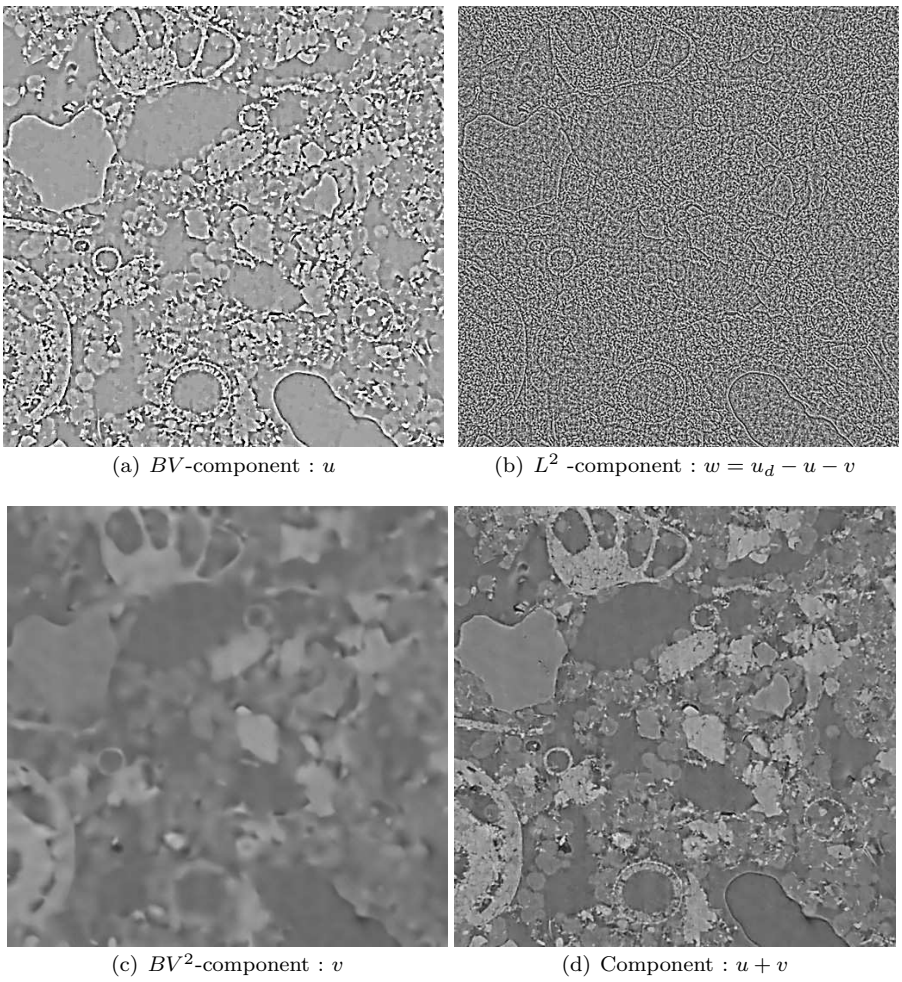
(a) Original u_d (b) BV^2-component v

(c) BV-component u (d) L^2 -component :$w = u_d - u - v$

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

4.4 Denoising and texture extraction

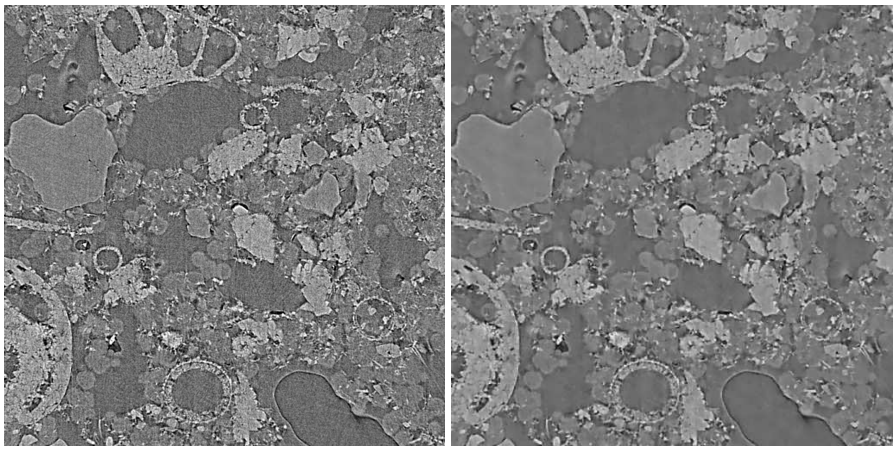
We end with image (c). It is quite difficult to perform segmentation of such an image. Indeed, the image is noisy and there are texture areas (due to the micritic calcite part). The denoising process should preserve the texture which involves physical information. As we want to recover the vacuum area we have to 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 component w from the "macro-texture" component u and perform a classical segmentation method on u.



(a) BV-component : u (b) L^2 -component : $w = u_d - u - v$

(c) BV^2-component : v (d) Component : $u + v$

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



(a) Original (noisy) image u_d (b) Component $:u+v$ (denoised image)

Fig. 13 Original and denoised image

5 Conclusion

This model seems promising to perform two-scale texture analysis. Indeed, the choice parameters λ and μ determines the scale of the "macro-texture". Once this part has been isolated, it possible to perform segmentation or statistical analysis.

There are many open questions that will 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 [20].

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 Problems 10, no. 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 (2000).
3. Attouch, H., Buttazzo, Michaille, G., *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*, Applied Mathematics and Optimization, 51, no 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*, Journal of Mathematical Imaging and Vision 22, no 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 147, Springer Verlag (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*, submitted , `http://hal.archives-ouvertes.fr/hal-00515451/fr/`,
9. Bergounioux, M., Piffet, L., , *A second-order model for image denoising and/or texture extraction*, Set Valued and Variational Analysis, Vol. 18, 3-4, pp. 277-306
10. Chambolle, A., *An algorithm for total variation minimization and applications*, Journal of Mathematical Imaging and Vision, 20 89–97 (2004)
11. Demengel, F., *Fonctions à hessien borné*, Annales de l'institut Fourier, tome 34, no 2, pp. 155-190 (1984)
12. Ekeland, I., Temam, R. , *Convex Analysis and Variational problems*, SIAM Classic in Applied Mathematics, 28, (1999)
13. Evans, L.C., Gariepy, R., *Measure theory and fine properties of functions*, CRC Press, 1992
14. Guillot, L., Le Trong, E., Rozenbaum, 0., Bergounioux, M., Rouet, J.L. , *A mixed model of active geodesic contours with gradient vector flows for X-ray microtomography segmentation,* Actes du colloque *Mathématiques pour l'image*, Bergounioux M. ed, PUO, 2009 ,
`http://hal.archives-ouvertes.fr/hal-00267007/fr/`
15. Meyer, Y. , *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, University Lecture Series, Vol. 22, AMS, 2001.
16. Osher, S., Fatemi, E, Rudin L., *Nonlinear total variation based noise removal algorithms*, Physica D 60, 259–268 (1992)
17. Osher, S., Sole, A., Vese L., *Image decomposition and restoration using total variation minimization and the H^1 norm*, SIAM Journal on Multiscale Modeling and Simulation, 1-3, 349–370 (2003)
18. Osher, S.,Vese, L., *Modeling textures with total variation minimization and oscillating patterns in image processing*, Journal of Scientific Computing 19, no 1-3, 553–572 (2003)
19. Osher, S. J, Vese, L. A.,, *Image denoising and decomposition with total variation minimization and oscillatory functions. Special issue on mathematics and image analysis*, J. Math. Imaging Vision, 20, no. 1-2, 7–18 (2004)
20. Piffet, L. , *Modèles variationnels pour l'extraction de textures 2D*, PhD Thesis, Orléans, 2010,
`http://tel.archives-ouvertes.fr/tel-00598289/fr/`
21. Weiss, P., Blanc-Féraud, L., Aubert, G., *Efficient schemes for total variation minimization under constraints in image processing*, SIAM Journal on Scientific Computing, 2009,
22. Yin, W., Goldfarb, D., Osher, S., *A comparison of three total variation based texture extraction models*, J. Vis. Commun. Image, 18, 240–252 (2007)