

Mathematical modeling of textures : application to color image decomposition with a projected gradient algorithm

Vincent Duval ¹ & Jean-François Aujol ² & Luminita A. Vese ³

¹ Institut Telecom, Telecom ParisTech, CNRS UMR 5141

vduval@enst.fr

² CMLA, ENS Cachan, CNRS, UniverSud

Jean-Francois.Aujol@cmla.ens-cachan.fr

³ UCLA, Mathematics Department

lvese@ucla.math.edu

Abstract: In this paper, we are interested in texture modeling with functional analysis spaces. We focus on the case of color image processing, and in particular color image decomposition. The problem of image decomposition consists in splitting an original image f into two components u and v . u should contain the geometric information of the original image, while v should be made of the oscillating patterns of f , such as textures. We propose here a scheme based on a projected gradient algorithm to compute the solution of various decomposition models for color images or vector-valued images. We provide a direct convergence proof of the scheme, and we give some analysis on color texture modeling.

Key-words: Texture modeling, color image decomposition, projected gradient algorithm, color texture modeling.

1 Introduction

Since the seminal work by Rudin et al [59], total variation based image restoration and decomposition has drawn a lot of attention (see [26, 7, 57, 5] and references therein for instance). We are interested in minimizing energies of the following type where f is the original image:

$$\int |Du| + \mu \|f - u\|_T^k. \quad (1)$$

$\int |Du|$ stands for the total variation; in the case when u is smooth, then $\int |Du| = \int |\nabla u| dx$. $\|\cdot\|_T$ stands for a norm which captures the noise and/or the textures of the original image f (in the sense that it is not too large for such features) and k is a positive exponent.

The most basic choice for $\|\cdot\|_T$ is the L^2 norm, and $k = 2$. From a Bayesian point of view, this is also the norm which appears naturally when assuming that the image f has been corrupted by some additive Gaussian white noise (see e.g. [26]). However, since the book by Y. Meyer [53], other spaces have been considered for modeling oscillating patterns such as textures or noise. The problem of image decomposition has been a very active field of research during the last past five years. [53], was the

inspiration source of many works, e.g. [65, 58, 9, 6, 64, 10, 15, 24, 35, 69, 70, 71, 44, 11, 43, 48, 50]. Image decomposition consists in splitting an original image f into two components, u and $v = f - u$. u is supposed to contain the geometrical component of the original image (it can be seen as some kind of sketch of the original image), while v is made of the oscillatory component (the texture component in the case when the original image f is noise free).

In this paper, we focus on color image processing. While some authors deal with color images using a Riemannian framework, like G. Sapiro and D. L. Ringach [60] or N. Sochen, R. Kimmel and R. Malladi [62], others combine a functional analysis viewpoint with the Chromaticity-Brightness representation [14]. The model we use is more basic: it is the same as the one used in [19] (and related with [18]). Its advantage is to have a rich functional analysis interpretation. Note that in [66], the authors also propose a cartoon + texture color decomposition and denoising model inspired from Y. Meyer, using the vectorial versions of total variation and approximations of the space $G(\Omega)$ for textures (to be defined later); unlike the work presented here, they use Euler-Lagrange equations and gradient descent scheme for the minimization, which should be slower than by projection methods.

Here, we give some insight into the definition of a texture space for color images. In [12], a TV-Hilbert model was proposed for image restoration and/ or decomposition:

$$\int |Du| + \mu \|f - u\|_{\mathcal{H}}^2 \quad (2)$$

where $\|\cdot\|_{\mathcal{H}}$ stands for the norm of some Hilbert space \mathcal{H} . This is a particular case of problem (1). Due to the Hilbert structure of \mathcal{H} , there exist many different methods to minimize (2), such as a projection algorithm [12]. We extend (2) to the case of color images.

From a numerical point of view, (1) is not straightforward to minimize. Depending on the choice for $\|\cdot\|_T$, the minimization of (1) can be quite challenging. Nevertheless, even in the simplest case when $\|\cdot\|_T$ is the L^2 norm, handling the total variation term $\int |Du|$ needs to be done with care. The most classical approach consists in writing the associated Euler-Lagrange equation for problem (1). In [59], a fixed step gradient descent scheme is used to compute the solution. This method has on the one hand the advantage of being very easy to implement, and on the other hand the disadvantage of being quite slow. To improve the convergence speed, quasi-Newton methods have been proposed [23, 67, 36, 1, 28, 55, 56]. Iterative methods have proved successful [17, 34, 15]. A projected-subgradient method can be found in [30].

Duality based schemes have also drawn a lot of attention to solve (1): first by Chan and Golub in [25], later by A. Chambolle in [20] with a projection algorithm. A multiscale version of this algorithm has just been introduced in [22]. This projection algorithm has recently been extended to the case of color images in [19]. Chambolle's projection algorithm [20] has grown very popular, since it is the first algorithm solving exactly the total variation regularization problem and not an approximation, with a complete proof of convergence. In [73], a very interesting combination of the primal and dual problems has been introduced. Second order cone programming ideas and interior point methods have proved interesting approaches [45, 42]. Recently, it has been shown that graph cuts based algorithms could also be used [21, 33]. Finally, let us notice that it is shown in [68, 8] that Nesterov's scheme [54] provides fast algorithms for minimizing (1).

Another variant of Chambolle's projection algorithm [20] is to use a projected gradient algorithm [21, 8, 72]. Here we have decided to use this approach which has both advantages of being easy to implement and of being quite efficient.

Notice that all the schemes based on duality proposed in the literature can be seen as particular

instance of proximal algorithms [31, 51, 37, 41, 61].

The plan of the paper is the following. In Section 2, we define and provide some analysis about the spaces we consider in the paper. In Section 3, we extend the TV-Hilbert model originally introduced in [12] to the case of color images. In Section 4, we present a projected gradient algorithm to compute a minimizer of problem (2). This projected gradient algorithm has first been proposed by A. Chambolle in [21] for total variation regularization. A proof of convergence was given in [8] relying on optimization results by Bermudez and Moreno [16]. A proof based on results for proximal algorithms was proposed in [72]. We derive here a simple and direct proof of convergence. In Section 5, we apply this scheme to solve various classical denoising and decomposition problems. We illustrate our approach with many numerical examples. We discuss qualitative properties of image decomposition models in Section 6.

2 Definitions and properties of the considered color spaces

In this section, we introduce several notations and we provide some analysis of the functional spaces we consider to model color textures.

2.1 Introduction

Let Ω be a Lipschitz convex bounded open set in \mathbb{R}^2 . We model color images as \mathbb{R}^M -valued functions defined on Ω . The inner product in $L^2(\Omega, \mathbb{R}^M)$ is denoted as:

$$\langle \mathbf{u}, \mathbf{v} \rangle_{L^2(\Omega, \mathbb{R}^M)} = \int_{\Omega} \sum_{i=1}^M u_i v_i.$$

For a vector $\xi \in \mathbb{R}^M$, we define the norms:

$$|\xi|_1 = \sum_{i=1}^M |\xi_i|, \quad |\xi|_2 = \sqrt{\sum_{i=1}^M \xi_i^2}, \quad |\xi|_{\infty} = \max_{i=1 \dots M} |\xi_i|. \quad (3)$$

We will sometimes refer to the space of zero-mean functions in $L^2(\Omega, \mathbb{R}^M)$ by V_0 :

$$V_0 = \{ \mathbf{f} \in L^2(\Omega, \mathbb{R}^M), \int_{\Omega} \mathbf{f} = 0 \}.$$

We say that a function $\mathbf{f} \in L^1(\Omega, \mathbb{R}^M)$ has bounded variation if the following quantity is finite:

$$\begin{aligned} |\mathbf{f}|_{TV} &= \sup \left\{ \int_{\Omega} \sum_{i=1}^M f_i \operatorname{div} \vec{\varphi}_i, \vec{\varphi} \in C_c^1(\Omega, \mathcal{B}) \right\} \\ &= \sup_{\vec{\varphi} \in C_c^1(\Omega, \mathcal{B})} \langle \mathbf{f}, \operatorname{div} \vec{\varphi} \rangle_{L^2(\Omega, \mathbb{R}^M)} \end{aligned}$$

where \mathcal{B} is a centrally symmetric convex body of $\mathbb{R}^{2 \times M}$. This quantity is called the total variation. For more information on its properties, we refer the reader to [4]. The set of functions with bounded variation is a vector space classically denoted by $BV(\Omega, \mathbb{R}^M)$.

In this paper, we will consider the following set of test-functions, which is the classical choice ([4], [7]) :

$$\mathcal{B} = \{\vec{w} \in \mathbb{R}^{2 \times M} / |\vec{w}|_2 \leq 1\}.$$

Then, for \mathbf{f} smooth enough, the total variation of \mathbf{f} is:

$$|\mathbf{f}|_{TV} = \int_{\Omega} \sqrt{\sum_{i=1}^M |\nabla f_i|^2} dx.$$

As X. Bresson and T. Chan notice in [19], the choice of the set \mathcal{B} is crucial. If one chooses:

$$\mathcal{B} = \{\vec{w} \in \mathbb{R}^{2 \times M} / |\vec{w}|_{\infty} \leq 1\},$$

then one has:

$$|\mathbf{f}|_{TV} = \sum_{i=1}^M \int_{\Omega} |\nabla f_i| dx = \sum_{i=1}^M |f_i|_{TV}.$$

These two choices are mathematically equivalent and define the same BV space, but in practice the latter induces no coupling between the channels. As a consequence, if the data fidelity term does not introduce a coupling either, the minimization of an energy of type (1) amounts to a series of independent scalar TV minimization problems, which gives visual artifacts in image processing (see [18, 27]).

2.2 The color $G(\Omega)$ space

The $G(\mathbb{R}^2)$ space was introduced by Y. Meyer in [53] to model textures in grayscale images. It was defined as $\text{div}(L^{\infty}(\mathbb{R}^2))$, but one could show that this space was equal to $W^{-1,\infty}(\mathbb{R}^2)$ (the dual of $W^{1,1}(\mathbb{R}^2)$). For the generalization to color images, we will adopt the framework of [6] (the color space $G(\Omega)$ is also used in [66], as a generalization of [65] to color image decomposition and color image denoising). Let us insist on the fact that the results of this section (notably Proposition 2.1) are specific to the case of 2-dimensional images (i.e. $\Omega \subset \mathbb{R}^2$).

Definition 2.1. The space $G(\Omega)$ is defined by:

$$G(\Omega) = \{v \in L^2(\Omega, \mathbb{R}^M) / \exists \vec{\xi} \in L^{\infty}(\Omega, (\mathbb{R}^2)^M), \forall i = 1, \dots, M, v_i = \text{div } \vec{\xi}_i \text{ and } \vec{\xi}_i \cdot \vec{N} = 0 \text{ on } \partial\Omega\}$$

(where $\vec{\xi}_i \cdot \vec{N}$ refers to the normal trace of $\vec{\xi}_i$ over $\partial\Omega$). One can endow it with the norm:

$$\|v\|_G = \inf\{\|\vec{\xi}\|_{\infty}, \forall i = 1, \dots, M, v_i = \text{div } \xi_i, \vec{\xi}_i \cdot \vec{N} = 0 \text{ on } \partial\Omega\}$$

with $\|\vec{\xi}\|_{\infty} = \sup \text{ess} \sqrt{\sum_{i=1}^M |\xi_i|^2}$.

The following result, proved in [6] for grayscale images, can be easily extended working component by component: it characterizes $G(\Omega)$.

Proposition 2.1.

$$G(\Omega) = \left\{ v \in L^2(\Omega, \mathbb{R}^M) / \int_{\Omega} v = 0 \right\}.$$

Following the framework of [49], we define $\|\cdot\|_*$ below. Since the total variation does not change by the addition of constants, we also have:

Lemma 2.1. *For $\mathbf{f} \in L^2(\Omega, \mathbb{R}^M)$, let us consider the semi-norm:*

$$\|\mathbf{f}\|_* = \sup_{\mathbf{u} \in BV(\Omega, \mathbb{R}^M), |\mathbf{u}|_{TV} \neq 0} \frac{\langle \mathbf{f}, \mathbf{u} \rangle_{L^2(\Omega, \mathbb{R}^M)}}{|\mathbf{u}|_{TV}} = \sup_{\mathbf{u} \in BV(\Omega, \mathbb{R}^M), |\mathbf{u}|_{TV} \neq 0} \frac{\sum_{i=1}^M \int_{\Omega} f_i u_i}{|\mathbf{u}|_{TV}}.$$

If $\|\mathbf{f}\|_* < +\infty$, then $\int_{\Omega} \mathbf{f} = 0$.

Comparing this property to Proposition 2.1, we can deduce that any function \mathbf{f} such that $\|\mathbf{f}\|_* < \infty$ belongs to $G(\Omega)$. The converse is also true:

Lemma 2.2. *Let $\mathbf{f} \in G(\Omega)$. Then $\|\mathbf{f}\|_* < \infty$.*

Proof: Let $\mathbf{f} \in G(\Omega)$. Since $\int_{\Omega} \mathbf{f} = 0$, the quantity $\frac{\langle \mathbf{f}, \mathbf{u} \rangle_{L^2(\Omega, \mathbb{R}^M)}}{|\mathbf{u}|_{TV}}$ does not change with the addition of constants to \mathbf{u} . Thus

$$\begin{aligned} \sup_{\mathbf{u} \in BV, |\mathbf{u}|_{TV} \neq 0} \frac{\langle \mathbf{f}, \mathbf{u} \rangle_{L^2(\Omega, \mathbb{R}^M)}}{|\mathbf{u}|_{TV}} &= \sup_{\mathbf{u} \in BV, \int \mathbf{u} = 0} \frac{\langle \mathbf{f}, \mathbf{u} \rangle_{L^2(\Omega, \mathbb{R}^M)}}{|\mathbf{u}|_{TV}} \\ &\leq \sup_{\mathbf{u} \in BV, \int \mathbf{u} = 0} \frac{\|\mathbf{f}\|_{L^2(\Omega, \mathbb{R}^M)} \|\mathbf{u}\|_{L^2(\Omega, \mathbb{R}^M)}}{|\mathbf{u}|_{TV}} \\ &\leq C \|\mathbf{f}\|_{L^2(\Omega)} < \infty, \end{aligned}$$

where we used Poincaré inequality : $\|\mathbf{u} - \mathbf{u}_{\Omega}\|_{L^2(\Omega, \mathbb{R}^M)} \leq C|\mathbf{u}|_{TV}$. ■

The following theorem completes the above two lemmas:

Theorem 2.2. *The following equality holds :*

$$G(\Omega) = \{\mathbf{f} \in L^2(\Omega) / \sup_{\mathbf{u} \in BV(\Omega, \mathbb{R}^M), |\mathbf{u}|_{BV} \neq 0} \frac{\langle \mathbf{f}, \mathbf{u} \rangle_{L^2(\Omega, \mathbb{R}^M)}}{|\mathbf{u}|_{TV}} < +\infty\}$$

and for all function $\mathbf{f} \in L^2(\Omega, \mathbb{R}^M)$, $\|\mathbf{f}\|_* = \|\mathbf{f}\|_G$.

Moreover, the infimum in the definition of $\|\cdot\|_G$ is reached.

Proof:

- (i) Let \mathbf{f} be a function in the set on the right hand-side. Thanks to Lemma 2.1, we know that $\forall i \in \{1, \dots, M\}, \int_{\Omega} f_i = 0$. By Proposition 2.1, $\mathbf{f} \in G(\Omega)$.

Now let $\mathbf{u} \in BV(\Omega, \mathbb{R}^M)$ such that $|\mathbf{u}|_{TV} \neq 0$. By Meyers-Serrin's theorem, one can find a sequence $\mathbf{u}_n \in C^{\infty}(\Omega, \mathbb{R}^M) \cap W^{1,1}(\Omega; \mathbb{R}^M)$ such that $\|\mathbf{u} - \mathbf{u}_n\|_2 \rightarrow 0$ and $|\mathbf{u}_n|_{TV} \rightarrow |\mathbf{u}|_{TV}$.

Then, for all \vec{g} such that $\mathbf{f} = \operatorname{div} \vec{g}$ and $\vec{g} \cdot \vec{N} = 0$ on $\partial\Omega$:

$$\begin{aligned} \langle \mathbf{f}, \mathbf{u}_n \rangle_{L^2(\Omega, \mathbb{R}^M)} &= \sum_{i=1}^M \int_{\Omega} \operatorname{div} \vec{g} \, \mathbf{u}_n \\ &= - \int_{\Omega} \left(\sum_{i=1}^M \vec{g}_i \cdot \nabla u_{i,n} \right) \\ &\leq \int_{\Omega} |\vec{g}| |\nabla \mathbf{u}_n| \\ &\leq \|\vec{g}\|_{\infty} |\mathbf{u}_n|_{TV}. \end{aligned}$$

Since $\mathbf{f} \in L^2(\Omega)$, we can pass to the limit in both sides of the inequality, and by construction of \mathbf{u}_n , we get:

$$\|\mathbf{f}\|_* \leq \|\mathbf{f}\|_G.$$

(ii) For the converse inequality, the proof is standard (see e.g. [46], [2]).

Let $\mathbf{f} \in L^2(\Omega, \mathbb{R}^M)$ such that $\sup_{\mathbf{u} \in BV(\Omega, \mathbb{R}^M), |\mathbf{u}|_{TV} \neq 0} \frac{\langle \mathbf{f}, \mathbf{u} \rangle_{L^2(\Omega, \mathbb{R}^M)}}{|\mathbf{u}|_{TV}} < +\infty$. Let us define :

$$T : \begin{cases} \mathcal{D}(\bar{\Omega}, \mathbb{R}^M) & \rightarrow & L^1(\Omega, \mathbb{R}^{2M}) \\ \varphi & \mapsto & \left(\frac{\partial \varphi^1}{\partial x_1}, \frac{\partial \varphi^1}{\partial x_2}, \dots, \frac{\partial \varphi^M}{\partial x_1}, \frac{\partial \varphi^M}{\partial x_2} \right) \end{cases}$$

To each vector $\left(\frac{\partial \varphi^1}{\partial x_1}, \frac{\partial \varphi^1}{\partial x_2}, \dots, \frac{\partial \varphi^M}{\partial x_1}, \frac{\partial \varphi^M}{\partial x_2} \right) \in T(\mathcal{D}(\bar{\Omega}, \mathbb{R}^M))$, we can associate $\int_{\Omega} \sum_{i=1}^M f^i \varphi^i dx$ (without ambiguity since f_i has zero-mean, and if two functions have the same gradient over Ω they only differ by a constant on the convex domain Ω). We have $\int_{\Omega} \sum_{i=1}^M f^i \varphi^i dx \leq \|\mathbf{f}\|_* |\varphi|_{BV} = \|\mathbf{f}\|_* \left\| \left(\frac{\partial \varphi^1}{\partial x_1}, \dots, \frac{\partial \varphi^M}{\partial x_2} \right) \right\|_1$ thus we have defined a bounded linear form on $T(\mathcal{D}(\bar{\Omega}, \mathbb{R}^M))$. Using Hahn-Banach's theorem, we can extend it to $L^1(\bar{\Omega}, \mathbb{R}^{2M})$ with the same norm $\|\mathbf{f}\|_*$. Since $L^\infty(\Omega, \mathbb{R}^{2M})$ is identifiable with the dual of $L^1(\Omega, \mathbb{R}^{2M})$, there exists $g \in L^\infty(\Omega, \mathbb{R}^{2M})$ with $\|g\|_{L^\infty(\Omega, \mathbb{R}^{2M})} = \|\mathbf{f}\|_*$, such that:

$$\forall \varphi \in \mathcal{D}(\bar{\Omega}, \mathbb{R}^M), \int_{\Omega} \sum_{i=1}^M f_i \varphi_i = - \int_{\Omega} \sum_{i=1}^M \sum_{j=1}^2 \frac{\partial \varphi_i}{\partial x_j} g_{i,j} = - \sum_{i=1}^M \int_{\Omega} \vec{g}_i \cdot \nabla \varphi_i. \quad (4)$$

This is true in particular for $\varphi \in \mathcal{D}(\Omega, \mathbb{R}^M)$, thus $\mathbf{f} = \operatorname{div} \vec{g}$ in the distributional sense, and since the functions are in $L^2(\Omega, \mathbb{R}^M)$ there is equality in $L^2(\Omega, \mathbb{R}^M)$. Since $\operatorname{div} \vec{g} \in L^2(\Omega, \mathbb{R}^M)$, we can then consider the normal trace of \vec{g} .

If $\varphi \in \mathcal{D}(\bar{\Omega}, \mathbb{R}^M)$, we have by (4):

$$\sum_{i=1}^M \int_{\Omega} f_i \varphi_i = - \sum_{i=1}^M \int_{\Omega} g_i \cdot \nabla \varphi_i.$$

But on the other hand, using integration by parts:

$$\sum_{i=1}^M \int_{\Omega} \operatorname{div} \vec{g}_i \varphi_i = - \sum_{i=1}^M \int_{\Omega} \vec{g}_i \cdot \nabla \varphi_i + \sum_{i=1}^M \int_{\partial\Omega} \varphi_i \vec{g}_i \cdot \vec{N}.$$

The equality $\mathbf{f} = \operatorname{div} \vec{\mathbf{g}}$ in $L^2(\Omega, \mathbb{R}^M)$ shows that the boundary contribution vanishes for $\varphi \in \mathcal{D}(\bar{\Omega}, \mathbb{R}^M)$. Thus $\vec{g}_i \cdot \vec{N} = 0$ over $\partial\Omega$.

Incidentally, we notice that the infimum in the G-norm is reached. ■

Remark 2.3. From the proof of Lemma 2.2, we can deduce that the topology induced by the G -norm on $G(\Omega)$ is coarser than the one induced by the L^2 norm. More generally, there exists a constant $C > 0$ (depending only on Ω), such that for all $\mathbf{f} \in G(\Omega)$, $\|\mathbf{f}\|_* \leq C\|\mathbf{f}\|_{L^2(\Omega, \mathbb{R}^M)}$.

In fact the G norm is strictly coarser than the L^2 norm. Here is an example for $M = 2$ color channels. Let us consider the family of functions $(\mathbf{f}^{(m_1, m_2)})_{m_1, m_2 \in \mathbb{N}^*}$ defined on $(-\pi, \pi)^2$ by $f_1^{(m_1, m_2)}(x, y) = \cos m_1 x + \cos m_1 y$ and $f_2^{(m_1, m_2)}(x, y) = \cos m_2 x + \cos m_2 y$. The vector field $\vec{\xi}^{(m_1, m_2)}$ defined by $\xi_k^{(m_1, m_2)} = \left(\frac{1}{m_k} \sin(m_k x), \frac{1}{m_k} \sin(m_k y)\right)$ for $k \in \{1, 2\}$ satisfies the boundary condition, and its divergence is equal to $\mathbf{f}^{(m_1, m_2)}$. As a consequence $\|\mathbf{f}^{(m_1, m_2)}\|_G \leq \sqrt{2 \left(\frac{1}{m_1^2} + \frac{1}{m_2^2}\right)}$ and $\lim_{m_1, m_2 \rightarrow +\infty} \|\mathbf{f}^{(m_1, m_2)}\|_G = 0$.

Yet,

$$\|\mathbf{f}^{(m_1, m_2)}\|_{L^2(\Omega, \mathbb{R}^M)}^2 = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} (\cos m_1 x + \cos m_1 y)^2 + (\cos m_2 x + \cos m_2 y)^2 dx dy = 8\pi^2.$$

The sequence \mathbf{f}_m converges to 0 for the topology induced by the G -norm, but not for the one induced by the L^2 norm.

The following result proposed by Y. Meyer in [53] specifies this idea.

Proposition 2.2. *Let $\mathbf{f}_n, n \geq 1$ be a sequence of functions of $L^q(\Omega, \mathbb{R}^M) \cap G(\Omega)$ with the following properties :*

1. *There exists $q > 2$ and $C > 0$ such that $\|\mathbf{f}_n\|_{L^q(\Omega, \mathbb{R}^M)} \leq C$.*
2. *The sequence \mathbf{f}_n converges to 0 in the distributional sense.*

Then $\|\mathbf{f}_n\|_G$ converges to 0 when n goes to infinity.

It means that oscillating patterns with zero mean have a small G-norm. Incidentally, notice that in the above example, the frequencies of all color channels had to go to infinity in order to have convergence to zero for the G-norm. Otherwise, assumption 2) in the above proposition fails.

3 Color TV-Hilbert model: presentation and mathematical analysis

The G-norm detailed in Section 2 is the main tool to study the TV-Hilbert problem on which all the algorithms described in this paper rely.

3.1 Presentation

The TV-Hilbert framework was introduced for grayscale images by J.-F. Aujol and G. Gilboa in [12] as a way to approximate the BV-G model. They prove that one can extend Chambolle's algorithm to this model. In this section we show that this is still true for color images. We are interested in solving the following problem:

$$\inf_{\mathbf{u}} |\mathbf{u}|_{TV} + \frac{1}{2\lambda} \|\mathbf{f} - \mathbf{u}\|_{\mathcal{H}}^2 \quad (5)$$

where $\mathcal{H} = V_0$ (the space of zero-mean functions of $L^2(\Omega, \mathbb{R}^M)$) is a Hilbert space endowed with the following norm:

$$\|\mathbf{v}\|_{\mathcal{H}}^2 = \langle \mathbf{v}, K\mathbf{v} \rangle_{L^2(\Omega, \mathbb{R}^M)}$$

and where $K : \mathcal{H} \rightarrow L^2(\Omega, \mathbb{R}^M)$

- is a bounded linear operator (for the topology induced by the $L^2(\Omega, \mathbb{R}^M)$ norm on \mathcal{H})
- is symmetric positive definite

and K^{-1} is bounded on $Im(K)$.

Examples:

- The Rudin-Osher Fatemi model

It was proposed in [59] by L. Rudin, S. Osher, and E. Fatemi for grayscale images. It was then extended to color images using different methods, for instance by G. Sapiro and D.L. Ringach [60], or Blomgren and T. Chan [18]. In [19], X. Bresson and T. Chan use another kind of color total variation, which is the one we use in this paper. The idea is to minimize the functional:

$$|\mathbf{u}|_{TV} + \frac{1}{2\lambda} \|\mathbf{f} - \mathbf{u}\|_{L^2(\Omega, \mathbb{R}^M)}^2. \quad (6)$$

It is clear that the problem commutes with the addition of constants. If the (unique) solution associated to \mathbf{f} is \mathbf{u} , then the solution associated to $\mathbf{f} + C$ is $(\mathbf{u} + C)$. As a consequence we can always assume that \mathbf{f} has zero mean. Then this model becomes a particular case of the TV-Hilbert model with $K = Id$.

- The OSV model

In [58], S. Osher, A. Solé and L. Vese propose to model textures by the H^{-1} space. In order to generalize this model, we must be cautious about the meaning of our notations but it is natural to introduce the following functional :

$$\inf_{\mathbf{u}} |\mathbf{u}|_{TV} + \frac{1}{2\lambda} \int_{\Omega} |\nabla \Delta^{-1}(\mathbf{f} - \mathbf{u})|^2 \quad (7)$$

where $\Delta^{-1}\mathbf{v} = \begin{pmatrix} \Delta^{-1}v_1 \\ \vdots \\ \Delta^{-1}v_M \end{pmatrix}$, $\nabla \boldsymbol{\rho} = \begin{pmatrix} \nabla \rho_1 \\ \vdots \\ \nabla \rho_M \end{pmatrix}$, $|\nabla \boldsymbol{\rho}|^2 = |\nabla \rho_1|^2 + |\nabla \rho_2|^2 + \dots + |\nabla \rho_M|^2$ and

$$\begin{aligned} \int_{\Omega} |\nabla \Delta^{-1}(\mathbf{f} - \mathbf{u})|^2 &= \int_{\Omega} \sum_{i=1}^M |\nabla \Delta^{-1}(f^i - u^i)|^2 = - \int_{\Omega} \sum_{i=1}^M (f^i - u^i) \Delta^{-1}(f^i - u^i) \\ &= \langle \mathbf{f} - \mathbf{u}, -\Delta^{-1}(\mathbf{f} - \mathbf{u}) \rangle_{L^2(\Omega, \mathbb{R}^M)}. \end{aligned}$$

The inversion of the Laplacian is defined component by component. For one component, it is defined by the following consequence of the Lax-Milgram theorem:

Proposition 3.1. *Let $X_0 = \{P \in H^1(\Omega, \mathbb{R}) : \int_{\Omega} P = 0\}$. If $v \in L^2(\Omega)$, with $\int_{\Omega} v = 0$, then the problem:*

$$-\Delta P = v, \quad \frac{\partial P}{\partial n} \Big|_{\partial\Omega} = 0$$

admits a unique solution in X_0 .

For $K = -\Delta^{-1}$, the Osher-Solé-Vese problem is a particular case of the TV-Hilbert framework.

3.2 Mathematical study

In this subsection, we study the existence and uniqueness of minimizers of Problem (5). A characterization of these minimizers is then given in Theorem 3.1. Notice that this theorem and its reformulation could be easily obtained using convex analysis results ([40]), but for a pedagogical purpose, we give an elementary proof inspired from [53].

Proposition 3.2 (Existence and Uniqueness). *Let $\mathbf{f} \in L^2(\Omega, \mathbb{R}^M)$. The minimization problem:*

$$\inf \left\{ |\mathbf{u}|_{TV} + \frac{1}{2\lambda} \langle \mathbf{f} - \mathbf{u}, K(\mathbf{f} - \mathbf{u}) \rangle_{L^2(\Omega, \mathbb{R}^M)}, \mathbf{u} \in BV(\Omega, \mathbb{R}^M), (\mathbf{f} - \mathbf{u}) \in V_0 \right\}$$

has a unique solution $\mathbf{u} \in BV(\Omega, \mathbb{R}^M)$.

Proof: Let $E(\mathbf{u})$ denote the functional defined on $L^2(\Omega, \mathbb{R}^M)$ (with $E(\mathbf{u}) = +\infty$ if $\mathbf{u} \notin BV(\Omega, \mathbb{R}^M)$ or $(\mathbf{f} - \mathbf{u}) \notin V_0$).

Let us notice that $E \neq +\infty$ since $\mathbf{f}_{\Omega} = \frac{1}{|\Omega|} \int_{\Omega} \mathbf{f}$ belongs to $BV(\Omega, \mathbb{R}^M)$ and $(\mathbf{f} - \mathbf{f}_{\Omega}) \in V_0$.

The functional E is convex. Since K is bounded, we deduce that E is lower semi-continuous for the weak $L^2(\Omega)$ topology. E is coercive: by Poincaré inequality,

$$\exists C > 0, \quad \|\mathbf{u} - \mathbf{u}_{\Omega}\|_2 \leq C |\mathbf{u}|_{TV}$$

with $\mathbf{u}_{\Omega} = \frac{1}{|\Omega|} \int_{\Omega} \mathbf{u} = \frac{1}{|\Omega|} \int_{\Omega} \mathbf{f}$ for $E(\mathbf{u}) < +\infty$. Thus E has a minimizer. Since the functional is strictly convex, this minimizer is unique. ■

We introduce the notation $\mathbf{v} = \mathbf{f} - \mathbf{u}$, when \mathbf{u} is the unique minimizer of the TV-Hilbert problem.

Theorem 3.1 (Characterization of minimizers). *Let $\mathbf{f} \in L^2(\Omega, \mathbb{R}^M)$.*

(i) *If $\|K\mathbf{f}\|_* \leq \lambda$ then the solution of the TV-Hilbert problem is given by $(\mathbf{u}, \mathbf{v}) = (0, \mathbf{f})$.*

(ii) *If $\|K\mathbf{f}\|_* > \lambda$ then the solution (\mathbf{u}, \mathbf{v}) is characterized by:*

$$\|K\mathbf{v}\|_* = \lambda \text{ and } \langle \mathbf{u}, K\mathbf{v} \rangle_{L^2(\Omega, \mathbb{R}^M)} = \lambda |\mathbf{u}|_{TV}.$$

Proof:

(i) $(0, \mathbf{f})$ is a minimizer iff

$$\begin{aligned} \forall \mathbf{h} \in BV(\Omega, \mathbb{R}^M), \forall \epsilon \in \mathbb{R}, |\epsilon \mathbf{h}|_{TV} + \frac{1}{2\lambda} \|\mathbf{f} - \epsilon \mathbf{h}\|_{\mathcal{H}}^2 &\geq \frac{1}{2\lambda} \|\mathbf{f}\|_{\mathcal{H}}^2, \\ \text{which is equivalent to } |\epsilon| |\mathbf{h}|_{TV} + \frac{1}{2\lambda} \epsilon^2 \|\mathbf{h}\|_{\mathcal{H}}^2 - \frac{1}{\lambda} \epsilon \langle \mathbf{f}, \mathbf{h} \rangle_{\mathcal{H}} &\geq 0. \end{aligned}$$

We can divide by $|\epsilon| \rightarrow 0$, and depending on the sign of ϵ we get :

$$\pm \langle \mathbf{f}, \mathbf{h} \rangle_{\mathcal{H}} \leq \lambda |\mathbf{h}|_{TV}.$$

If $(0, \mathbf{f})$ is a minimizer, we can take the supremum for $\mathbf{h} \in BV(\Omega, \mathbb{R}^M)$. By definition of the *-norm, we have : $\|K\mathbf{f}\|_* \leq \lambda$.

Conversely, if $\mathbf{f} \in L^2(\Omega, \mathbb{R}^M)$ is such that $\|K\mathbf{f}\|_* \leq \lambda$, the second inequality is true , thus $(0, \mathbf{f})$ is a minimizer.

(ii) As before, let us characterize the extremum: (\mathbf{u}, \mathbf{v}) is a minimizer iff

$$\begin{aligned} \forall \mathbf{h} \in BV(\Omega, \mathbb{R}^M), \forall \epsilon \in \mathbb{R}, |\mathbf{u} + \epsilon \mathbf{h}|_{TV} + \frac{1}{2\lambda} \|\mathbf{v} - \epsilon \mathbf{h}\|_{\mathcal{H}}^2 &\geq |\mathbf{u}|_{TV} + \frac{1}{2\lambda} \|\mathbf{v}\|_{\mathcal{H}}^2, \\ \text{or } |\mathbf{u} + \epsilon \mathbf{h}|_{TV} + \frac{1}{2\lambda} \epsilon^2 \|\mathbf{h}\|_{\mathcal{H}}^2 - \frac{1}{\lambda} \epsilon \langle \mathbf{v}, \mathbf{h} \rangle_{\mathcal{H}} &\geq |\mathbf{u}|_{TV}. \end{aligned}$$

By triangle inequality, we have:

$$\begin{aligned} |\mathbf{u}|_{TV} + |\epsilon| |\mathbf{h}|_{TV} + \frac{1}{2\lambda} \epsilon^2 \|\mathbf{h}\|_{\mathcal{H}}^2 - \frac{1}{\lambda} \epsilon \langle \mathbf{v}, \mathbf{h} \rangle_{\mathcal{H}} &\geq |\mathbf{u}|_{TV} \\ |\mathbf{h}|_{BV} &\geq \frac{1}{\lambda} \langle \mathbf{v}, \mathbf{h} \rangle_{\mathcal{H}} \end{aligned}$$

Taking the supremum, we obtain $\|K\mathbf{v}\|_* \leq \lambda$.

Moreover, choosing $\mathbf{h} = \mathbf{u}$, $\epsilon \in]-1, 1[$:

$$\begin{aligned} (1 + \epsilon) |\mathbf{u}|_{BV} &\geq \frac{1}{\lambda} \epsilon \langle \mathbf{v}, \mathbf{u} \rangle_{\mathcal{H}} + |\mathbf{u}|_{TV} \\ \text{For } \epsilon > 0 : |\mathbf{u}|_{TV} &\geq \frac{1}{\lambda} \langle \mathbf{v}, \mathbf{u} \rangle_{\mathcal{H}} \\ \text{For } \epsilon < 0 : |\mathbf{u}|_{TV} &\leq \frac{1}{\lambda} \langle \mathbf{v}, \mathbf{u} \rangle_{\mathcal{H}} \end{aligned}$$

We deduce that $\|K\mathbf{v}\|_*|\mathbf{u}|_{TV} \geq \langle \mathbf{v}, \mathbf{u} \rangle_{\mathcal{H}} = \lambda|\mathbf{u}|_{TV}$, and by the first upper-bound inequality, we have $\|K\mathbf{v}\|_* = \lambda$.

Conversely, let us assume these equalities hold. Then:

$$\begin{aligned} |\mathbf{u} + \epsilon\mathbf{h}|_{TV} + \frac{1}{2\lambda}\|\mathbf{v} - \epsilon\mathbf{h}\|_{\mathcal{H}}^2 &\geq \frac{1}{\lambda}\langle (\mathbf{u} + \epsilon\mathbf{h}), K\mathbf{v} \rangle_{L^2(\Omega, \mathbb{R}^M)} + \frac{1}{2\lambda}\|\mathbf{v}\|_{\mathcal{H}}^2 + \frac{1}{2\lambda}\|\mathbf{h}\|_{\mathcal{H}}^2\epsilon^2 - \frac{1}{\lambda}\epsilon\langle \mathbf{h}, \mathbf{v} \rangle_{\mathcal{H}} \\ &\geq |\mathbf{u}|_{TV} + \frac{1}{2\lambda}\|\mathbf{v}\|_{\mathcal{H}}^2. \end{aligned}$$

■

The mapping $v \mapsto \sup_{|\mathbf{u}|_{TV} \neq 0} \frac{\langle \mathbf{u}, K\mathbf{v} \rangle_{L^2(\Omega, \mathbb{R}^M)}}{|\mathbf{u}|_{TV}}$ is convex, lower semi-continuous for the \mathcal{H} strong topology as a supremum of convex lower semi-continuous functions. As a consequence, for $\lambda > 0$ the set

$$G_\lambda = \{\mathbf{v} \in L^2(\Omega, \mathbb{R}^M), \|\mathbf{v}\|_* \leq \lambda\}$$

is a closed convex set, as well as $K^{-1}G_\lambda$. The orthogonal projection of this set is well-defined and we can notice that Theorem 3.1 reformulates :

$$\begin{cases} \mathbf{v} &= P_{K^{-1}G_\lambda}^{\mathcal{H}}(\mathbf{f}) \\ \mathbf{u} &= \mathbf{f} - \mathbf{v}. \end{cases}$$

Indeed, if (\mathbf{u}, \mathbf{v}) is a minimizer of the TV-Hilbert problem, with $\mathbf{f} = \mathbf{u} + \mathbf{v}$, we have $\mathbf{v} \in K^{-1}G_\lambda$ and for any $\tilde{\mathbf{v}} \in K^{-1}G_\lambda$,

$$\langle \mathbf{f} - \mathbf{v}, \tilde{\mathbf{v}} - \mathbf{v} \rangle_{\mathcal{H}} = \langle \mathbf{u}, K\tilde{\mathbf{v}} \rangle_{L^2(\Omega, \mathbb{R}^M)} - \langle \mathbf{u}, K\mathbf{v} \rangle_{L^2(\Omega, \mathbb{R}^M)} \leq \|K\tilde{\mathbf{v}}\|_*|\mathbf{u}|_{TV} - \lambda|\mathbf{u}|_{TV} \leq 0,$$

thus by the equivalent definition of the projection on a closed convex set (see Ekeland-Témam [40]), we obtain the desired result.

Consequently, \mathbf{v} is the orthogonal projection of \mathbf{f} on the set $K^{-1}G_\lambda$ where $G_\lambda = \{\lambda \operatorname{div} \vec{\mathbf{p}}, |\vec{\mathbf{p}}| \leq 1\}$ (see Theorem 2.2), and the problem is equivalent to its dual formulation:

$$\inf_{|\vec{\mathbf{p}}| \leq 1} \|\lambda K^{-1} \operatorname{div} \vec{\mathbf{p}} - \mathbf{f}\|_{\mathcal{H}}^2. \quad (8)$$

4 Projected gradient algorithm

In the present section, we are interested in solving the dual formulation (8), which can be done using fast algorithms. For grayscale images, the famous projection algorithm by A. Chambolle [20] was the inspiration for all the following algorithms. We present here a projection algorithm, and we provide a complete proof of convergence of this scheme. We note that an independent work has just been reported in [52], where the authors M. Zhu, S.J. Wright, and T.F. Chan have also applied the projected gradient method for solving the dual formulation of total variation minimization for image denoising. They have a general framework also, although applied only to scalar image denoising and not related to image decompositions.

As explained in the introduction, the scheme we propose here can be seen as a proximal point algorithm [31]. The proof of convergence could then be derived from results in [51] (the Lions-Mercier algorithm being itself a particular instance of proximal algorithms [39]). However, we have decided to present here an elementary proof of the convergence for didactic reasons.

4.1 Discrete setting

From now on, we will work in the discrete case, using the following conventions. A grayscale image is a matrix of size $N \times N$. We write $X = \mathbb{R}^{N \times N}$ the space of grayscale images. Their gradients belong to the space $Y = X \times X$. The discrete L^2 inner product is $\langle u, v \rangle_X = \sum_{1 \leq i, j \leq N} u_{i,j} v_{i,j}$.

The gradient operator is defined by $(\nabla u)_{i,j} = ((\nabla u)_{i,j}^x, (\nabla u)_{i,j}^y)$ with:

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

The divergence operator is defined as the opposite of the adjoint operator of ∇ :

$$\forall \vec{p} \in Y, \langle -\operatorname{div} \vec{p}, u \rangle_X = \langle \vec{p}, \nabla u \rangle_Y$$

$$(\operatorname{div} \vec{p})_{i,j} = \begin{cases} p_{i+1,j}^x - p_{i,j}^x & \text{if } 1 < i < N \\ p_{i,j}^x & \text{if } i = 1 \\ -p_{i-1,j}^x & \text{if } i = N \end{cases} + \begin{cases} p_{i,j+1}^y - p_{i,j}^y & \text{if } 1 < j < N \\ p_{i,j}^y & \text{if } j = 1 \\ -p_{i,j-1}^y & \text{if } j = N. \end{cases}$$

A color image is an element of X^M . The gradient and the divergence are defined component by component, and the discrete L^2 inner product is given by :

$$\forall \mathbf{u}, \mathbf{v} \in X^M, \langle \mathbf{u}, \mathbf{v} \rangle_{X^M} = \sum_{k=1}^M \langle u^{(k)}, v^{(k)} \rangle_X$$

$$\forall \vec{p}, \vec{q} \in Y^M, \langle \vec{p}, \vec{q} \rangle_{Y^M} = \sum_{k=1}^M \langle p^{(k)}, q^{(k)} \rangle_Y$$

so that the color divergence is still the opposite of the adjoint of the color gradient.

4.2 Bresson-Chan algorithm

Problem (8) for grayscale images was tackled in [20] in the case $K = Id$, then in [12] for a general K . For color images, X. Bresson and T. Chan [19] showed that Chambolle's projection algorithm could still be used when $K = Id$. It is actually easy to check that it can be used with a general K for color images as well.

Following the steps of [12], one can notice that, provided $\tau \leq \frac{1}{8\|K^{-1}\|_{L^2}}$, the fixed point iteration:

$$\vec{p}^{n+1} = \frac{\vec{p}^n + \tau(\nabla(K^{-1}\operatorname{div} \vec{p}^n - \mathbf{f}/\lambda))}{1 + \tau|\nabla(K^{-1}\operatorname{div} \vec{p}^n - \mathbf{f}/\lambda)|} \quad (9)$$

gives a sequence $(\vec{p}^n)_{n \in \mathbb{N}}$ such that $\lambda K^{-1}\operatorname{div} \vec{p}^{n+1} \rightarrow v_\lambda$, and $\mathbf{f} - \lambda K^{-1}\operatorname{div} \vec{p}^{n+1} \rightarrow u_\lambda$ when $n \rightarrow +\infty$.

Notice that the upper bound on τ is the same as for grayscale images.

4.3 Projected gradient

It was recently noticed ([21], [8]) that problem (8) for grayscale images could be solved using a projected gradient descent. This is the algorithm we decided to extend to the case of color images.

Let \mathcal{B} be the discrete version of our set of test-functions :

$$\mathcal{B} = \{\mathbf{v} \in Y^M, \forall 1 \leq i, j \leq N, |v_{i,j}|_2 \leq 1\}.$$

The orthogonal projection on \mathcal{B} is easily computed:

$$P_{\mathcal{B}}(x) = \left(\frac{x_1}{\max\{1, |x|_2\}}, \frac{x_2}{\max\{1, |x|_2\}} \right).$$

The projected gradient descent scheme is defined by :

$$\bar{\mathbf{p}}^{m+1} = P_{\mathcal{B}} \left(\bar{\mathbf{p}}^{m+1} + \tau \nabla (K^{-1} \operatorname{div} \bar{\mathbf{p}}^m - \mathbf{f}/\lambda) \right) \quad (10)$$

which amounts to:

$$\mathbf{p}_{i,j}^{m+1} = \frac{\mathbf{p}_{i,j}^m + \tau \nabla (K^{-1} \operatorname{div} \mathbf{p}^m - \frac{\mathbf{f}}{\lambda})_{i,j}}{\max \left(1, |\mathbf{p}_{i,j}^m + \tau \nabla (K^{-1} \operatorname{div} \mathbf{p}^m - \frac{\mathbf{f}}{\lambda})_{i,j}|_2 \right)} \quad (11)$$

where $\tau = \frac{\lambda}{\mu}$.

The convergence result for the projected gradient descent in the case of elliptic functions is standard (see [29], Theorem 8.6-2). Yet in our case the functional is not elliptic, and the proof needs a little more work.

Proposition 4.1. *If $0 < \tau < \frac{1}{4\|K^{-1}\|}$, then algorithm (11) converges. More precisely, there exists $\vec{\mathbf{p}} \in \mathcal{B}$ such that :*

$$\lim_{m \rightarrow \infty} (K^{-1} \operatorname{div} \bar{\mathbf{p}}^m) = K^{-1} \operatorname{div} \vec{\mathbf{p}}$$

and

$$\|\lambda K^{-1} \operatorname{div} \vec{\mathbf{p}} - \mathbf{f}\|_{\mathcal{H}}^2 = \inf_{\vec{\mathbf{p}} \in \mathcal{B}} \|\lambda K^{-1} \operatorname{div} \vec{\mathbf{p}} - \mathbf{f}\|_{\mathcal{H}}^2$$

Proof: Writing

$$\|K^{-1} \operatorname{div} \vec{\mathbf{q}} - \mathbf{f}/\lambda\|_{\mathcal{H}}^2 = \|K^{-1} \operatorname{div} \vec{\mathbf{p}} - \mathbf{f}/\lambda\|_{\mathcal{H}}^2 + \langle K K^{-1} \operatorname{div} (\vec{\mathbf{q}} - \vec{\mathbf{p}}), K^{-1} \operatorname{div} \vec{\mathbf{p}} - \mathbf{f}/\lambda \rangle_{L^2} + O(\|\vec{\mathbf{q}} - \vec{\mathbf{p}}\|^2),$$

we begin by noticing that $\vec{\mathbf{p}}$ is a minimizer if and only if,

$$\vec{\mathbf{p}} \in \mathcal{B} \text{ and } \forall \vec{\mathbf{q}} \in \mathcal{B}, \forall \tau > 0, \langle \vec{\mathbf{q}} - \vec{\mathbf{p}}, \vec{\mathbf{p}} - (\vec{\mathbf{p}} + \tau \nabla (K^{-1} \operatorname{div} \vec{\mathbf{p}} - \mathbf{f}/\lambda)) \rangle_{L^2} \geq 0.$$

Or equivalently:

$$\vec{\mathbf{p}} = P_{\mathcal{B}} \left(\vec{\mathbf{p}} + \tau (\nabla (K^{-1} \operatorname{div} \vec{\mathbf{p}} - \mathbf{f}/\lambda)) \right)$$

where $P_{\mathcal{B}}$ is the orthogonal projection on \mathcal{B} with respect to the L^2 inner product.

We know that such a minimizer exists. Let us denote it by $\vec{\mathbf{p}}$.

- Now let us consider a sequence defined by (10), and write $A = -\nabla K^{-1} \operatorname{div} \cdot$. We have:

$$\begin{aligned} \|\vec{p}^{k+1} - \vec{p}\|^2 &= \|P_{\mathcal{B}}(\vec{p}^k + \tau \nabla(K^{-1} \operatorname{div} \vec{p}^k - \mathbf{f}/\lambda)) - P_{\mathcal{B}}(\vec{p} + \tau \nabla(K^{-1} \operatorname{div} \vec{p} - \mathbf{f}/\lambda))\|^2 \\ &\leq \|\vec{p} - \vec{p}^k + \tau \nabla K^{-1} \operatorname{div}(\vec{p} - \vec{p}^k)\|^2 \text{ since } P_{\mathcal{B}} \text{ is 1-Lipschitz [29]} \\ &\leq \|(I - \tau A)(\vec{p} - \vec{p}^k)\|^2 \end{aligned}$$

Provided $\|I - \tau A\| \leq 1$, we can deduce:

$$\|\vec{p}^{k+1} - \vec{p}\| \leq \|\vec{p}^k - \vec{p}\| \quad (12)$$

and the sequence $(\|\vec{p}^k - \vec{p}\|)$ is convergent.

- A is a symmetric positive semi-definite operator. By writing $E = \ker A$ and $F = \operatorname{Im} A$, we have:

$$Y^M = E \oplus F$$

and we can decompose any $\vec{q} \in Y^M$ as the sum of two orthogonal components $\vec{q}_E \in E$ and $\vec{q}_F \in F$. Notice that by injectivity of K^{-1} , E is actually equal to the kernel of the divergence operator.

Let $\mu_1 = 0 < \mu_2 \leq \dots \leq \mu_a$ be the ordered eigenvalues of A .

$$\begin{aligned} \|I - \tau A\| &= \max(|1 - \tau \mu_1|, |1 - \tau \mu_a|) \\ &= \max(1, |1 - \tau \mu_a|) \\ &= 1 \text{ for } 0 \leq \tau \leq \frac{2}{\mu_a} \end{aligned}$$

We can restrict $I - \tau A$ to F and then define:

$$\begin{aligned} g(\tau) = \|(I - \tau A)|_F\| &= \max(|1 - \tau \mu_2|, |1 - \tau \mu_a|) \\ &< 1 \text{ for } 0 < \tau < \frac{2}{\mu_a} \end{aligned}$$

- Now we assume that $0 < \tau < \frac{2}{\mu_a}$. Therefore, inequality (12) is true and the sequence (\vec{p}^k) is bounded, and so is the sequence $(K^{-1} \operatorname{div} \vec{p}^k)$. We are going to prove that the sequence $(K^{-1} \operatorname{div} \vec{p}^k)$ has a unique cluster point. Let $(K^{-1} \operatorname{div} \vec{p}^{\varphi(k)})$ be a convergent subsequence. By extraction, one can assume that $\vec{p}^{\varphi(k)}$ is convergent too, and denote by \vec{p} its limit.

Passing to the limit in (10), the sequence $(\vec{p}^{\varphi(k)+1})$ is convergent towards:

$$\vec{p} = P_{\mathcal{B}} \left(\vec{p} + \tau \nabla(K^{-1} \operatorname{div} \vec{p} - \mathbf{f}/\lambda) \right)$$

Using (12), we also notice that:

$$\|\vec{p} - \vec{p}\| = \|\vec{p} - \vec{p}\|$$

As a consequence:

$$\begin{aligned} \|\vec{p} - \vec{p}\|^2 &= \|\vec{p} - \vec{p}\|^2 \\ &= \|P_{\mathcal{B}} \left(\vec{p} + \tau \nabla(K^{-1} \operatorname{div} \vec{p} - \mathbf{f}/\lambda) \right) - P_{\mathcal{B}}(\vec{p} + \tau \nabla(K^{-1} \operatorname{div} \vec{p} - \mathbf{f}/\lambda))\|^2 \\ &\leq \|(I - \tau A)(\vec{p} - \vec{p})\|^2 \\ &\leq \|(\vec{p} - \vec{p})_E\|^2 + g(\tau)^2 \|(\vec{p} - \vec{p})_F\|^2 \\ &< \|\vec{p} - \vec{p}\|^2 \text{ if } (\vec{p} - \vec{p})_F \neq 0 \end{aligned}$$

Of course, this last inequality cannot hold, which means that $\|(\vec{\tilde{p}} - \vec{p})_F\| = 0$. Hence $(\vec{\tilde{p}} - \vec{p}) \in E = \ker A$ and $K^{-1} \operatorname{div} \vec{\tilde{p}} = K^{-1} \operatorname{div} \vec{p}$: the sequence $(K^{-1} \operatorname{div} \vec{p}^k)$ is convergent.

- The last remark consists in evaluating μ_a . We have:

$$\mu_a = \|\nabla K^{-1} \operatorname{div}\| \leq \|\nabla\| \|K^{-1}\| \|\operatorname{div}\|$$

Since $\|\operatorname{div}\|^2 = \|\nabla\|^2 = 8$ (see [20], the result is still true for color images), we deduce that

$$\mu_a \leq 8\|K^{-1}\|.$$

■

Since we are only interested in $\mathbf{v} = \lambda K^{-1} \operatorname{div} \vec{p}$, Proposition (4.1) justifies the validity of algorithm (10). Yet it is possible to prove that \vec{p} itself converges.

Lemma 4.1. *Let P be the orthogonal projection on a nonempty closed convex set K . Let $Q = Id - P$. Then we have:*

$$\|Q(v_1) - Q(v_2)\|^2 + \|P(v_1) - P(v_2)\|^2 \leq \|v_1 - v_2\|^2. \quad (13)$$

Proof:

$$\begin{aligned} & \|Q(v_1) - Q(v_2)\|^2 + \|P(v_1) - P(v_2)\|^2 \\ = & \|v_1 - v_2 + P(v_1) - P(v_2)\|^2 + \|P(v_1) - P(v_2)\|^2 \\ = & \|v_1 - v_2\|^2 + 2\|P(v_1) - P(v_2)\|^2 - 2\langle P(v_1) - P(v_2), v_1 - v_2 \rangle \\ = & \|v_1 - v_2\|^2 + 2\langle P(v_1) - P(v_2), P(v_1) - P(v_2) - v_1 + v_2 \rangle \\ = & \|v_1 - v_2\|^2 + 2\underbrace{\langle P(v_1) - P(v_2), P(v_1) - v_1 \rangle}_{\leq 0} + 2\underbrace{\langle P(v_2) - P(v_1), P(v_2) - v_2 \rangle}_{\leq 0} \end{aligned}$$

(using the characterization of the projection on a nonempty closed convex set [29]).

■

Remark: We have:

$$\vec{p} = P_{\mathcal{B}}(\vec{p} - \tau(A\vec{p} + \nabla \mathbf{f}/\lambda)) = \vec{p} - \tau(A\vec{p} + \nabla \mathbf{f}/\lambda) - Q_{\mathcal{B}}(\vec{p} - \tau(A\vec{p} + \nabla \mathbf{f}/\lambda)),$$

thus

$$-\tau(A\vec{p} + \nabla \mathbf{f}/\lambda) = Q_{\mathcal{B}}(\vec{p} - \tau(A\vec{p} + \nabla \mathbf{f}/\lambda)). \quad (14)$$

Corollary 4.1. *The sequence \vec{p}^k defined by (10) converges to \vec{p} unique solution of problem (8).*

Proof: Using the above lemma, we have:

$$\begin{aligned}
& \|Q_{\mathcal{B}}(\vec{\mathbf{p}} - \tau(A\vec{\mathbf{p}} + \nabla \mathbf{f}/\lambda)) - Q_{\mathcal{B}}(\vec{\mathbf{p}}^k - \tau(A\vec{\mathbf{p}}^k + \nabla \mathbf{f}/\lambda))\|^2 \\
& + \|P_{\mathcal{B}}(\vec{\mathbf{p}} - \tau(A\vec{\mathbf{p}} + \nabla \mathbf{f}/\lambda)) - P_{\mathcal{B}}(\vec{\mathbf{p}}^k - \tau(A\vec{\mathbf{p}}^k + \nabla \mathbf{f}/\lambda))\|^2 \\
\leq & \|\vec{\mathbf{p}} - \tau(A\vec{\mathbf{p}} + \nabla \mathbf{f}/\lambda) - \vec{\mathbf{p}}^k + \tau(A\vec{\mathbf{p}}^k + \nabla \mathbf{f}/\lambda)\|^2 \\
\leq & \|\vec{\mathbf{p}} - \vec{\mathbf{p}}^k\|^2 + \tau^2 \|A(\vec{\mathbf{p}} - \vec{\mathbf{p}}^k)\|^2 - 2\tau \langle \vec{\mathbf{p}} - \vec{\mathbf{p}}^k, A(\vec{\mathbf{p}} - \vec{\mathbf{p}}^k) \rangle
\end{aligned}$$

Hence:

$$\begin{aligned}
\|Q_{\mathcal{B}}(\vec{\mathbf{p}} - \tau(A\vec{\mathbf{p}} + \nabla \mathbf{f}/\lambda)) - Q_{\mathcal{B}}(\vec{\mathbf{p}}^k - \tau(A\vec{\mathbf{p}}^k + \nabla \mathbf{f}/\lambda))\|^2 + \|\vec{\mathbf{p}} - \vec{\mathbf{p}}^{k+1}\|^2 \\
\leq \|\vec{\mathbf{p}} - \vec{\mathbf{p}}^k\|^2 + \tau^2 \|A(\vec{\mathbf{p}} - \vec{\mathbf{p}}^k)\|^2 - 2\tau \langle \vec{\mathbf{p}} - \vec{\mathbf{p}}^k, A(\vec{\mathbf{p}} - \vec{\mathbf{p}}^k) \rangle.
\end{aligned}$$

But we have already shown that $\|\vec{\mathbf{p}} - \vec{\mathbf{p}}^k\|$ converges, and that $K^{-1} \operatorname{div}(\vec{\mathbf{p}} - \vec{\mathbf{p}}^k) \rightarrow 0$ (therefore $A(\vec{\mathbf{p}} - \vec{\mathbf{p}}^k) \rightarrow 0$). Hence, passing to the limit in the above equation, we get:

$$Q(\vec{\mathbf{p}}^k - \tau(A\vec{\mathbf{p}}^k + \nabla \mathbf{f}/\lambda)) \rightarrow Q_{\mathcal{B}}(\vec{\mathbf{p}} - \tau(A\vec{\mathbf{p}} + \nabla \mathbf{f}/\lambda)).$$

We thus deduce from (14) that

$$Q_{\mathcal{B}}(\vec{\mathbf{p}}^k - \tau(A\vec{\mathbf{p}}^k + \nabla \mathbf{f}/\lambda)) \rightarrow -\tau(A\vec{\mathbf{p}} + \nabla \mathbf{f}/\lambda).$$

Remembering that $P_{\mathcal{B}} + Q_{\mathcal{B}} = Id$, we get:

$$\vec{\mathbf{p}}^{k+1} = P_{\mathcal{B}}(\vec{\mathbf{p}}^k - \tau(A\vec{\mathbf{p}}^k + \nabla \mathbf{f}/\lambda)) = (Id - Q_{\mathcal{B}})(\vec{\mathbf{p}}^k - \tau(A\vec{\mathbf{p}}^k + \nabla \mathbf{f}/\lambda)).$$

Hence,

$$\vec{\mathbf{p}}^{k+1} - \vec{\mathbf{p}}^k = -\tau(A\vec{\mathbf{p}}^k + \nabla \mathbf{f}/\lambda) - Q_{\mathcal{B}}(\vec{\mathbf{p}}^k - \tau(A\vec{\mathbf{p}}^k + \nabla \mathbf{f}/\lambda)).$$

Passing to the limit, we obtain

$$\vec{\mathbf{p}}^{k+1} - \vec{\mathbf{p}}^k \rightarrow 0.$$

We can now pass to the limit in (10) and get that $\vec{\mathbf{p}}^k \rightarrow \vec{\mathbf{p}}$. ■

5 Applications to color image denoising and decomposition

The projected gradient algorithm may now be applied to solve various color image problems.

5.1 TV-Hilbert model

The color ROF model

As an application of (11) with $K = Id$, we use the following scheme for the ROF model (6):

$$\mathbf{p}_{i,j}^{m+1} = \frac{\mathbf{p}_{i,j}^m + \tau \nabla(\operatorname{div} \mathbf{p}^m - \frac{\mathbf{f}}{\lambda})_{i,j}}{\max\left(1, |\mathbf{p}_{i,j}^m + \tau \nabla(\operatorname{div} \mathbf{p}^m - \frac{\mathbf{f}}{\lambda})_{i,j}|_2\right)}. \quad (15)$$

The color OSV model

As for the OSV model (7), $K = -\Delta^{-1}$, we use:

$$\mathbf{p}_{i,j}^{m+1} = \frac{\mathbf{p}_{i,j}^m - \tau \nabla(\Delta \operatorname{div} \mathbf{p}^m + \frac{\mathbf{f}}{\lambda})_{i,j}}{\max\left(1, |\mathbf{p}_{i,j}^m - \tau \nabla(\Delta \operatorname{div} \mathbf{p}^m + \frac{\mathbf{f}}{\lambda})_{i,j}|_2\right)}. \quad (16)$$

5.2 The color A2BC algorithm

Following Y. Meyer [53], one can use the $G(\Omega)$ space to model textures, and try to solve the problem:

$$\inf_{\mathbf{u}} |\mathbf{u}|_{TV} + \alpha \|\mathbf{f} - \mathbf{u}\|_G \quad (17)$$

In [9], J.-F. Aujol, G. Aubert, L. Blanc-Féraud and A. Chambolle approximate this problem by minimizing the following functional:

$$F_{\mu,\lambda}(\mathbf{u}, \mathbf{v}) = \begin{cases} |\mathbf{u}|_{TV} + \frac{1}{2\lambda} \|\mathbf{f} - \mathbf{u} - \mathbf{v}\|_{L^2(\Omega)}^2 & \text{if } (\mathbf{u}, \mathbf{v}) \in BV(\Omega, \mathbb{R}) \times G_\mu \\ +\infty & \text{otherwise} \end{cases} \quad (18)$$

or equivalently,

$$F_{\mu,\lambda}(\mathbf{u}, \mathbf{v}) = |\mathbf{u}|_{TV} + \frac{1}{2\lambda} \|\mathbf{f} - \mathbf{u} - \mathbf{v}\|_{L^2(\Omega)}^2 + \chi_{G_\mu}(\mathbf{v}_n)$$

$$\text{with } \chi_{G_\mu}(\mathbf{v}) = \begin{cases} 0 & \text{if } \mathbf{v} \in G_\mu \\ +\infty & \text{otherwise} \end{cases}.$$

The generalization to color images was done by J.-F. Aujol and S. H. Kang in [14] using a chromaticity-brightness model. The authors used gradient descent in order to compute the projections, which is rather slow, and requires to regularize the total variation.

In [19], X. Bresson and T. Chan used the following scheme (but relying on Chambolle's algorithm) for color images in order to compute the projections. As in the grayscale case, the minimization is done using an alternate scheme (but in the present paper we use the projected gradient descent scheme described before to compute the projections):

- Initialization:

$$\mathbf{u}_0 = \mathbf{v}_0 = 0$$

- Iterations:

$$\mathbf{v}_{n+1} = P_{G_\mu}(\mathbf{f} - \mathbf{u}_n) \quad (19)$$

$$\mathbf{u}_{n+1} = \mathbf{f} - \mathbf{v}_{n+1} - P_{G_\lambda}(\mathbf{f} - \mathbf{v}_{n+1}) \quad (20)$$

- Stop if the following condition is true:

$$\max(|\mathbf{u}_{n+1} - \mathbf{u}_n|, |\mathbf{v}_{n+1} - \mathbf{v}_n|) \leq \epsilon$$

In [6], it is shown that under reasonable assumptions, the solutions of problem (18) converge when $\lambda \rightarrow 0$ to a solution of problem (17) for some α .

5.3 The color TV-L1 model

The TV-L1 model is very popular for grayscale images. It benefits from having both good theoretical properties (it is a morphological filter) and fast algorithms (see [32]). In order to extend it to color images, we consider the following problem:

$$\inf_{\mathbf{u}} |\mathbf{u}|_{TV} + \lambda \|\mathbf{f} - \mathbf{u}\|_1 \quad (21)$$

with the notation

$$\|\mathbf{u}\|_1 = \int_{\Omega} \sqrt{\sum_{l=1}^M |u_l|^2}$$

(as for the total variation, we have decided to have a coupling between channels).

Our approach is different from the one used by J. Darbon in [32], since it was using a channel by channel decomposition with the additional constraint that no new color is created. As for the A2BC algorithm, we are led to consider the approximation:

$$\inf_{\mathbf{u}, \mathbf{v}} |\mathbf{u}|_{TV} + \frac{1}{2\alpha} \|\mathbf{f} - \mathbf{u} - \mathbf{v}\|_2^2 + \lambda \|\mathbf{v}\|_1.$$

Once again, having a projection algorithm for color images allows us to generalize easily this problem. In order to generalize the TV-L1 algorithm proposed by J.-F. Aujol, G. Gilboa, T. Chan and S. Osher ([13]), we aim at solving the alternate minimization problem:

(i)

$$\inf_{\mathbf{u}} |\mathbf{u}|_{TV} + \frac{1}{2\alpha} \|\mathbf{f} - \mathbf{u} - \mathbf{v}\|_2^2$$

(ii)

$$\inf_{\mathbf{v}} \frac{1}{2\alpha} \|\mathbf{f} - \mathbf{u} - \mathbf{v}\|_2^2 + \lambda \|\mathbf{v}\|_1$$

The first problem is a Rudin-Osher-Fatemi problem. Scheme (11) with $K = Id$ is well adapted for solving it. For the second one, the following property shows that a "vectorial soft thresholding" gives the solution:

Proposition 5.1. *The solution of problem (ii) is given by:*

$$\mathbf{v}(x) = VT_{\alpha\lambda}(\mathbf{f}(x) - \mathbf{u}(x)) = \frac{\mathbf{f}(x) - \mathbf{u}(x)}{|\mathbf{f}(x) - \mathbf{u}(x)|_2} \max(|\mathbf{f}(x) - \mathbf{u}(x)|_2 - \alpha\lambda, 0) \text{ almost everywhere.}$$

The proof of this last result is a straightforward extension of Proposition 4 in [13].

Henceforth, we propose the following generalization of the TV-L1 algorithm:

- Initialization:

$$\mathbf{u}_0 = \mathbf{v}_0 = 0$$

- Iterations:

$$\mathbf{v}_{n+1} = VT_{\alpha\lambda}(\mathbf{f} - \mathbf{u}_n) \quad (22)$$

$$\mathbf{u}_{n+1} = \mathbf{f} - \mathbf{v}_{n+1} - P_{G_\alpha}(\mathbf{f} - \mathbf{v}_{n+1}) \quad (23)$$

- Stop if the following condition is satisfied:

$$\max(|\mathbf{u}_{n+1} - \mathbf{u}_n|, |\mathbf{v}_{n+1} - \mathbf{v}_n|) \leq \epsilon$$

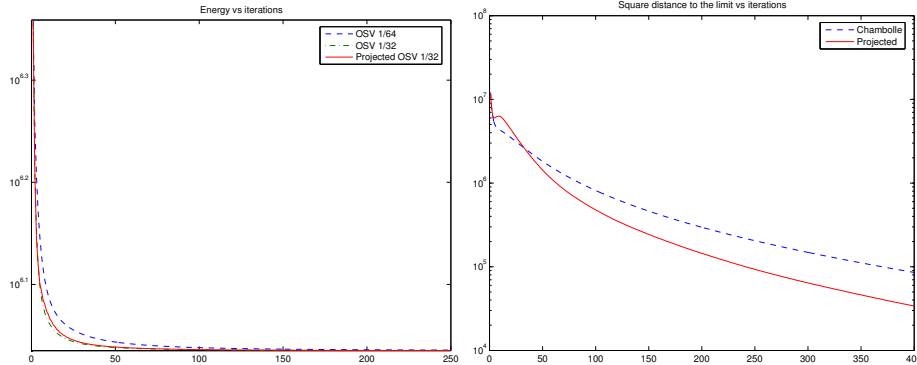


Figure 1: Left: Energy vs iterations of the Osher-Solé-Vese model with Chambolle’s projection algorithm (in green and blue - stepsize 1/64 and 0.031) and with the Projected gradient algorithm (in red - stepsize 0.031). Right: L^2 square distance (on a logarithmic scale) between the limit value (2000 iterations) vs the number of iterations, for OSV using 1/64 stepsize.

5.4 Numerical experiments

In this subsection, we present the results of the algorithms described above. Figure 1 shows the decrease of the energy and the convergence of the projected gradient algorithm for the OSV model (7). We compare scheme (11) with the projection algorithm of [10] (which is a straightforward modification of Chambolle’s projection algorithm [20]). Both algorithms behave similarly and it is hard to tell whether one converges faster than the other. Figures 2 and 3 depict denoising results using ROF (6) and OSV (7) models. The images look very similar but since the OSV model penalizes much more the highest frequencies than the ROF model [13], the denoised image still shows the lowest frequencies of the noise.

Figures 5 and 6 depict cartoon-texture decomposition experiments using different kinds of textures (the original images are shown on Figure 4), using the A2BC and TVL1 algorithms. Both results look good. However, one may observe, notably on Figures 6 and 7, a sort of halo in the texture part of the A2BC (by comparison, with TVL1 the texture of the caps looks flatter), and edges that, in our opinion, should remain in the cartoon part only. Both texture parts look ”gray”.



Figure 2: First row: original and noisy images (PSNR = 56.9 dB). Second row: denoised with color ROF ($\lambda = 15$, PSNR= 73.8 dB) and with color OSV ($\lambda = 15$, PSNR= 73.4 dB).

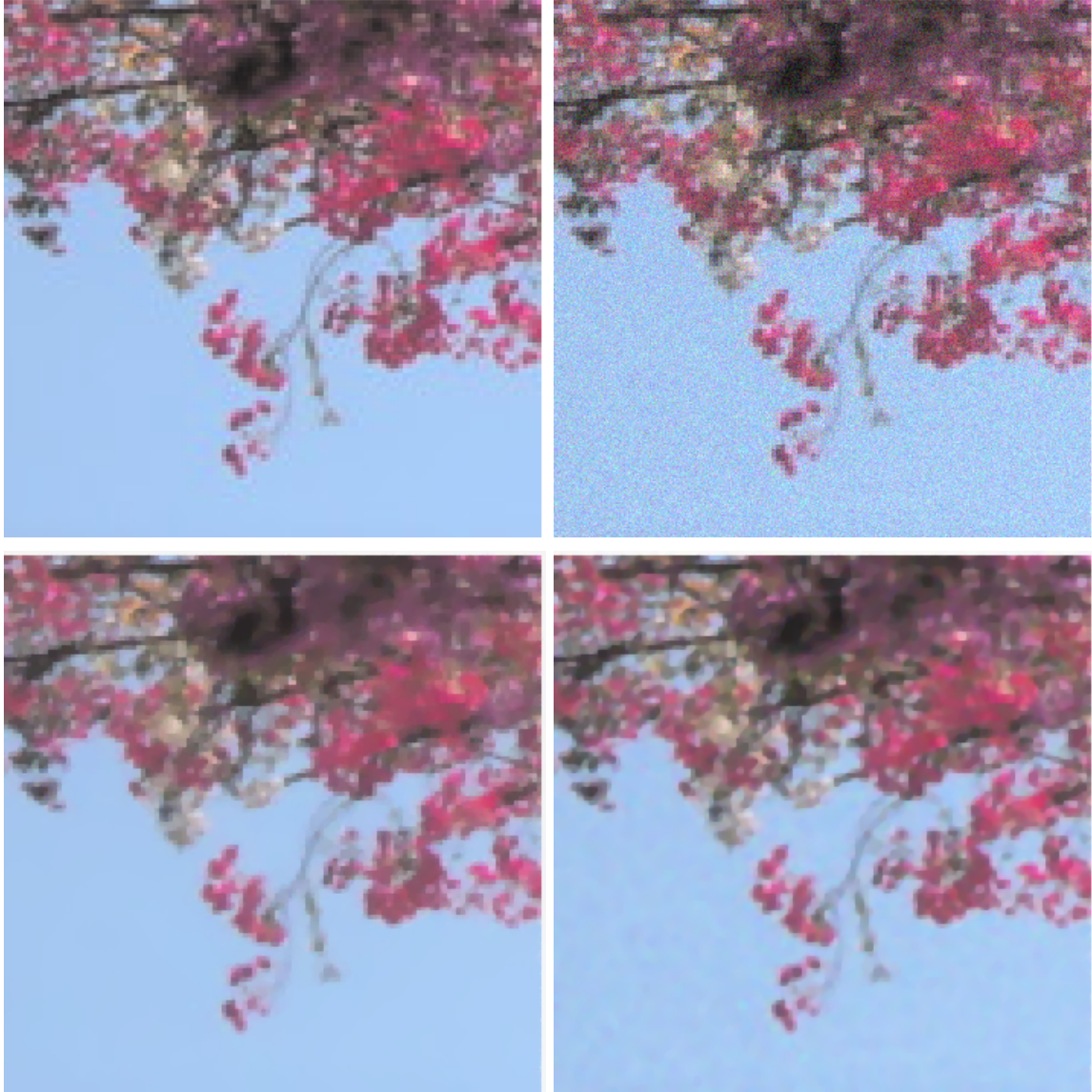


Figure 3: First row: original and noisy images (PSNR = 57.3 dB). Second row: denoised with color ROF ($\lambda = 25$, PSNR= 74.2 dB) and with color OSV ($\lambda = 25$, PSNR= 74.1 dB).



Figure 4: Original images.

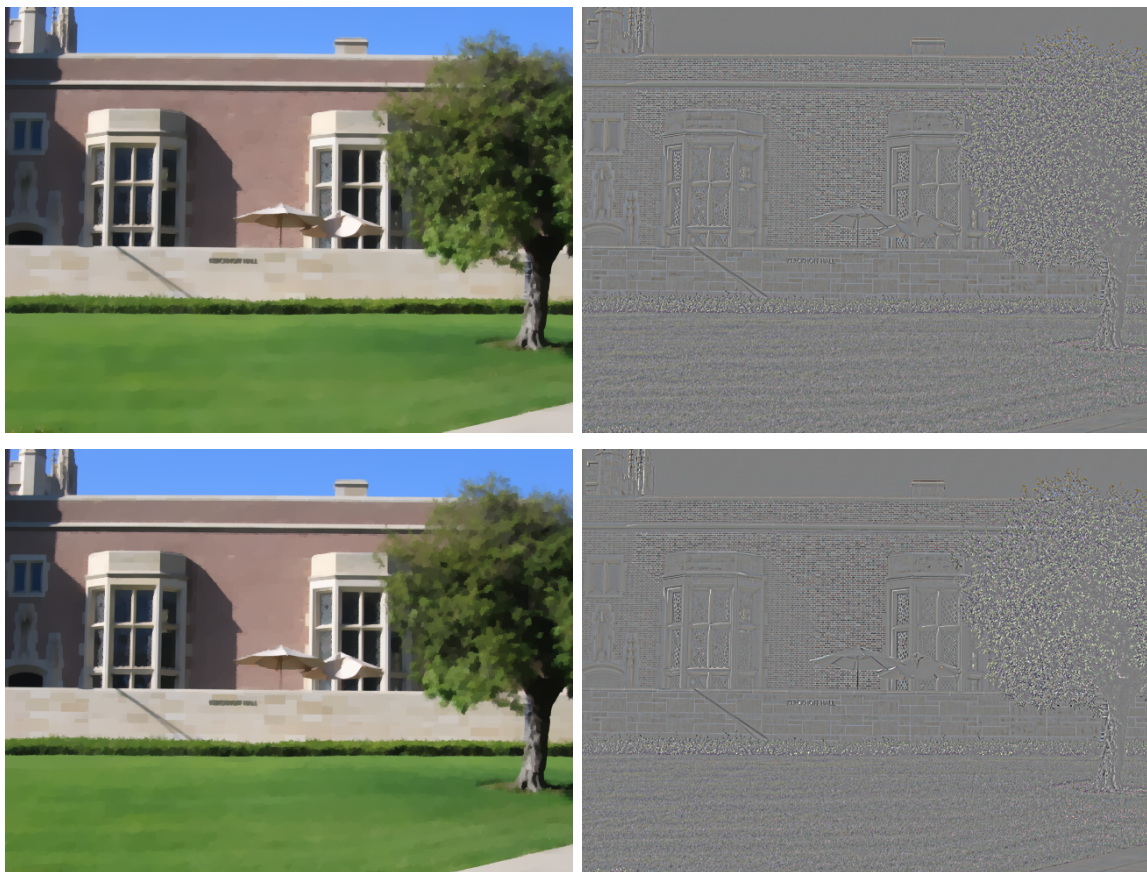


Figure 5: Cartoon-texture decomposition using color A2BC algorithm (upper row) and color TVL1 (lower row).

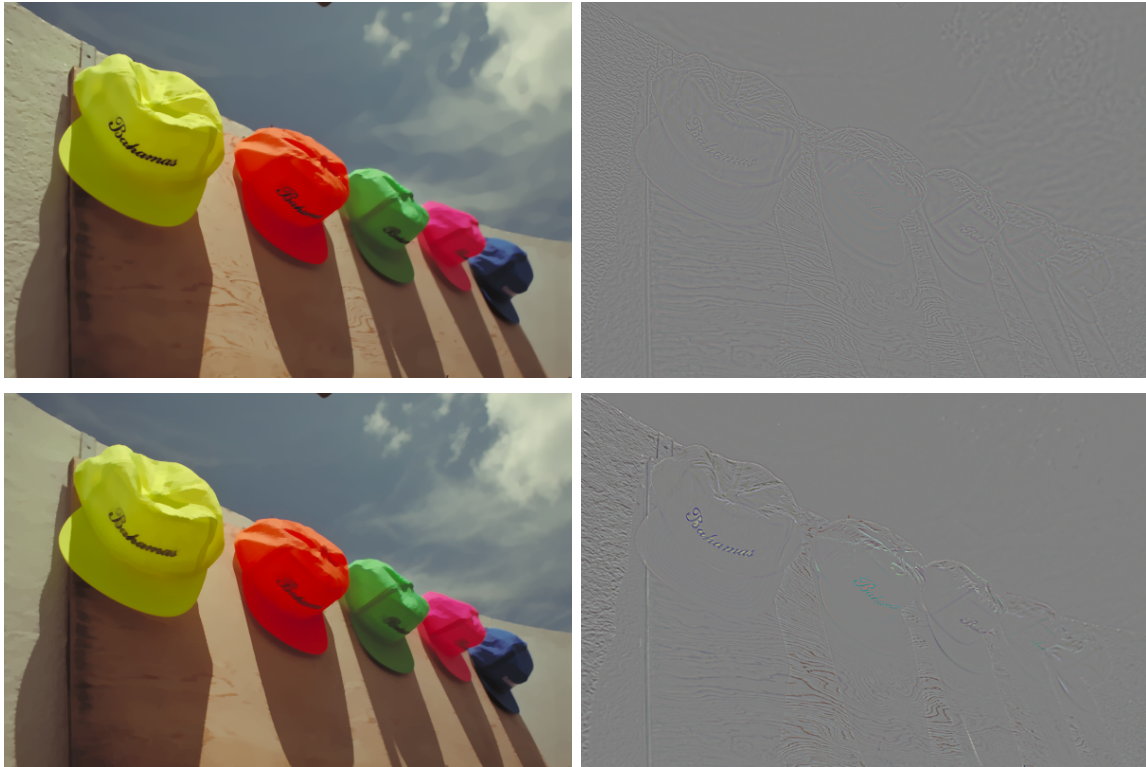


Figure 6: Top: cartoon-texture decomposition using the A2BC algorithm ($\lambda = 0.1$, $\mu = 10$). Bottom: cartoon-texture decomposition using the TVL1 algorithm ($\alpha = 0.1$, $\lambda = 0.7$). The parameters were chosen as a compromise so that the results are comparable. However, with the A2BC algorithm, if one wants to send the raised pattern on the left side of the wall to the texture part (as in the TVL1 model), one has to considerably degrade the clouds.

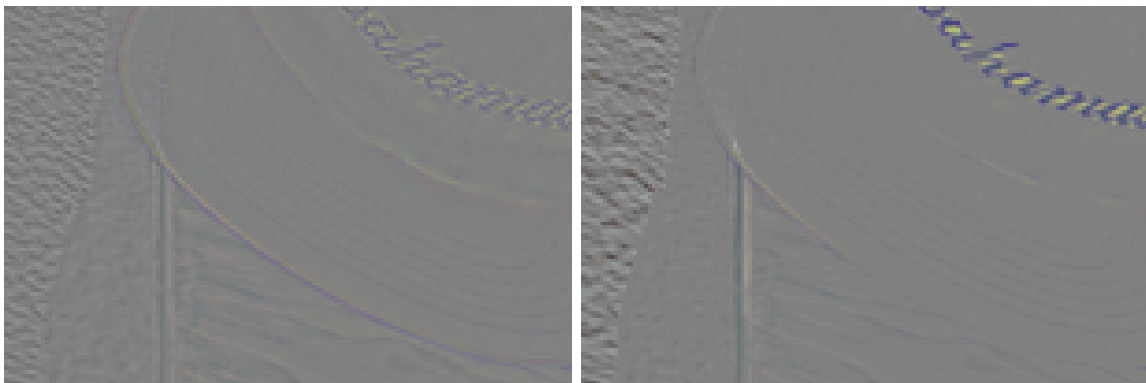


Figure 7: Detail of the texture parts of Figure 6 (left: A2BC, right TVL1). Notice how the edge of the cap appears in the A2BC model.

6 Qualitative analysis of decomposition models

In this section, we try to explain some of the visual results obtained in the last section.

6.1 Edges and halo

More precisely, we first focus on Figures 6 and 7 (cartoon-texture decomposition), where it seems that, with the A2BC algorithm, a sort of halo and edges appear in the texture part (at least more visibly than with the TVL1 algorithm). We believe that this phenomenon is not related to a wrong choice of parameters, but rather to an inherent limit of models that rely on the total variation plus a norm that favors oscillations for the texture part (like the BV-G model).

For the sake of simplicity, we will restrict our discussion to the simple case of one-dimensional single channel signals, but the core idea still applies to the case of color images. Suppose for instance, that one wants to decompose a signal f using the BV-G model, i.e. find a decomposition (u, v) that solves :

$$\min_{u+v=f} |u|_{TV} + \|v\|_G \quad (24)$$

where the functions are defined on $\Omega = (-1, 1)$. In dimension 1, the divergence operator reduces to the derivation, and the boundary condition on ξ implies that ξ is the antiderivative of v that cancels at -1 :

$$\|v\|_G = \sup_{t \in (-1, 1)} \left| \int_{-1}^t v(t) dt \right| \quad (25)$$

Now, consider a step edge, perturbed with some textures as in Figure 8 (a), for instance :

$$f(x) = \mathbb{1}_{(0,1)}(x) + \beta \sin(8p\pi x) \mathbb{1}_{\frac{1}{4} \leq |x| \leq \frac{3}{4}} \quad (26)$$

where $\beta > 0$, $p \in \mathbb{N}^*$. The ideal decomposition would be a perfect step $u(x) = \mathbb{1}_{(0,1)}(x)$ on the one hand, and the pure oscillation $v(x) = \beta \sin(8p\pi x) \mathbb{1}_{\frac{1}{4} \leq |x| \leq \frac{3}{4}}$ on the other hand (see Figure 8 (b) and (c)). The energy of the cartoon part is simply $|u|_{TV} = \lim_{t \rightarrow 1} u - \lim_{t \rightarrow -1} u = 1$, whereas the energy of the texture part is given by $\|v\|_G = \beta \int_{\frac{1}{4}}^{\frac{1}{4} + \frac{1}{8p}} \sin(8p\pi t) dt = \frac{\beta}{4p\pi}$. Yet, replacing u on $[-\frac{1}{4}, \frac{1}{4}]$ with any non-decreasing function \tilde{u} with the same limits at the boundary, say, a ramp $x \mapsto (\frac{1}{2} + \frac{1}{2\eta}x) \mathbb{1}_{[-\eta, \eta]}(x)$ as in Figure 8 (d)), one still gets the same cartoon energy $|\tilde{u}|_{TV} = 1$. As for the texture part, we should notice that one extra oscillation is added near the discontinuity of the original function f . But, precisely, the G-norm favors oscillations, so that this change in the texture part will not be penalized. Indeed:

$$\|\tilde{v}\|_G = \max \left(\frac{\beta}{4p\pi}, \int_{-\eta}^0 \left(\frac{1}{2} + \frac{1}{2\eta}t \right) dt \right) = \frac{1}{4p\pi} \quad (27)$$

for η small enough ($0 \leq \eta \leq \min(\frac{1}{4}, \frac{\beta}{p\pi})$).

To sum up, given any decomposition with sharp edges, there exists a decomposition with the same energy where shadows of edges appear in the texture part. So it is not surprising to see edges appear in the texture part of our experiments. This phenomenon is also a possible explanation of the

sort of halo that also appears. Replacing the sharp edge with a gradation, the same argument shows that one can add an oscillation to the texture part without changing the total variation of the cartoon part. This is why in Figure 6 the caps in the texture part look so "inflated" compared to those of TVL1.

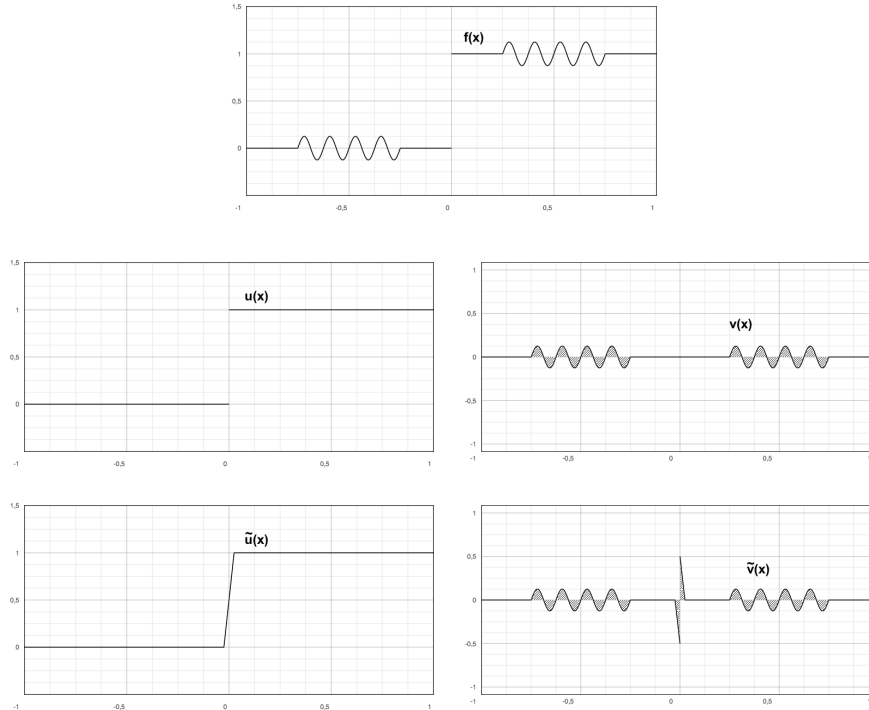


Figure 8: Decomposition of a 1-dimensional signal. Top: Original signal f . Middle row: ideal decomposition (left: cartoon part, right texture part). Bottom row: another decomposition. The total variation of the cartoon part is equal to 1 in both cases, and by (25), the G-norm of each texture part is equal to the maximum area of its "bumps". Therefore both decompositions have the same energy.

It is clear that a model like TVL1 does not suffer from such a drawback, since any extra oscillation in the texture part would be penalized. In fact, the TVL1 model does not favor oscillations at all, and it is rather surprising that this model performs well compared to oscillation-based models (see [33, 71], and Figures 5 and 6). In fact, the TVL1 model "works" because instead of considering textures as oscillations, one might as well regard them as objects of small scale. For grayscale images, since this filter is morphological, it is sufficient to study its effect on each level set of the data. The case of disks was studied in [24] and convex sets in [3, 38]. It can be shown that if the data is the characteristic function of a convex set, solutions of (21) are given by an opening of radius $\frac{1}{\lambda}$ followed by a test on the ratio perimeter/area. If that ratio is too large, then the solution is the emptyset; if it is small enough, the solution is the result of the opening. Therefore what we obtain when considering the family of TVL1 filters for $\lambda > 0$ is very similar to a granulometry, a tool that is used in mathematical morphology in the study of textures (see for instance [63]).

6.2 Color of textures

The second remark is that the extracted texture parts look like gray-level images. If one looks at them closely, one can see that there are colors, but they are spread on very small areas. In view of the definition of the G-norm and the example of Remark 2.3, it is clear that, on each channel and on every local neighborhood, the average value of the texture part should vanish (otherwise, the "antiderivative" ξ would take large values, so that the G-norm would be large). Therefore the overall impression is gray. The TVL1 model does not impose such a condition on the average value, as it can be seen on Figure 6. Notice the stripes of the wood and the holes inside the letters that are colored, contrary to the A2BC model. These were sent to the texture part because of their small scale, not because they were oscillations. However the overall impression is still gray too, since the colors of large areas, which matter for an impression of color, are kept in the cartoon part.

6.3 Comments

The conclusion is that although the different algorithms studied in this paper produce apparently similar results, they actually have very different qualitative properties, and they rely on different definitions of textures. This is another illustration of the difficulty to define precisely the notion of texture. Even though we pointed out some qualitative differences in the visual results, the choice of the best decomposition is arguable. While it is clear that some textures should be dealt with working on their frequential content, some others require to work geometrically on their elementary patterns, in a way that is reminiscent of the texton theory [47].

References

- [1] R. Acar and C. Vogel. Analysis of total variation penalty methods for ill-posed problems. *Inverse Problems*, 10:1217–1229, 1994.
- [2] R.A. Adams. *Sobolev Spaces*, volume 65 of *Mathematics*. Academic Press, 1975.
- [3] W.K. Allard. Total variation regularization for image denoising II: Examples. *SIAM Journal on Imaging Sciences*, 2008. Accepted.
- [4] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford mathematical monographs. Oxford University Press, 2000.
- [5] F. Andreu-Vaillo, V. Caselles, and J. M. Mazón. *Parabolic quasilinear equations minimizing linear growth functionals*, volume 223 of *Progress in Mathematics*. Birkhauser, 2002.
- [6] G. Aubert and J.-F. Aujol. Modeling very oscillating signals. application to image processing. *Applied Mathematics and Optimization*, 51(2):163–182, mars 2005.
- [7] G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, volume 147. Springer Verlag, Applied Mathematical Sciences, 2001.
- [8] J.-F. Aujol. Some first-order algorithms for total variation based image restoration. *J. Math. Imaging Vis.*, 34(3):307–327, 2009.

- [9] J.-F. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle. Image decomposition into a bounded variation component and an oscillating component. *Journal of Mathematical Imaging and Vision*, 22(1):71–88, January 2005.
- [10] J.-F. Aujol and A. Chambolle. Dual norms and image decomposition models. *International Journal on Computer Vision*, 63(1):85–104, June 2005.
- [11] J.-F. Aujol and T. F. Chan. Combining geometrical and textured information to perform image classification. *Journal of Visual Communication and Image representation*, 17(5):1004–1023, October 2006.
- [12] J.-F. Aujol and G. Gilboa. Constrained and SNR-based solutions for TV-Hilbert space image denoising. *Journal of Mathematical Imaging and Vision*, 26(1-2):217–237, 2006.
- [13] J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher. Structure-texture image decomposition - modeling, algorithms, and parameter selection. *International Journal of Computer Vision*, 67(1):111–136, April 2006.
- [14] J.-F. Aujol and S. H. Kang. Color image decomposition and restoration. *Journal of Visual Communication and Image Representation*, 17(4):916–928, August 2006.
- [15] J. Bect, L. Blanc-Féraud, G. Aubert, and A. Chambolle. A 11-unified variational framework for image restoration. In *ECCV 04*, volume 3024 of *Lecture Notes in Computer Sciences*, pages 1–13, 2004.
- [16] A. Bermudez and C. Moreno. Duality methods for solving variational inequalities. *Comp. and Maths. with Appls.*, 7(1):43–58, July 1981.
- [17] J. Bioucas-Dias and M. Figueiredo. Thresholding algorithms for image restoration. *IEEE Transactions on Image processing*, 16(12):2980–2991, 2007.
- [18] P. Blomgren and T. Chan. Color TV: Total variation methods for restoration of vector valued images. *IEEE Transactions on Image Processing*, 7(3):304–309, 1998.
- [19] X. Bresson and T. Chan. Fast minimization of the vectorial total variation norm and applications to color image processing. *SIAM Journal on Imaging Sciences (SIIMS)*, Submitted, 2007.
- [20] A. Chambolle. An algorithm for total variation minimization and its applications. *Journal of Mathematical Imaging and Vision*, 20:89–97, 2004.
- [21] A. Chambolle. Total variation minimization and a class of binary MRF models. *EMMCVPR in Vol. 3757 of Lecture Notes in Computer Sciences*, pages 136–152, June 2005.
- [22] A. Chambolle, S.E. Levine, and B.J. Lucier. Some variations on total variation-based image smoothing. 2009.
- [23] A. Chambolle and P.L. Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76(3):167–188, 1997.
- [24] T. Chan and S. Esedoglu. Aspects of total variation regularized L^1 function approximation. *SIAM Journal of Applied Mathematics*, 65(5):1817–1837, 2005.

- [25] T. Chan, G. Golub, and P. Mulet. A nonlinear primal-dual method for total variation-based image restoration. *SIAM Journal on Scientific Computing*, 20(6):1964–1977, 1999.
- [26] T. Chan and J. Shen. *Image processing and analysis - Variational, PDE, wavelet, and stochastic methods*. SIAM Publisher, 2005.
- [27] Tony F. Chan, Sung Ha Kang, and Jianhong Shen. Total variation denoising and enhancement of color images based on the cb and hsv color models. *J. Visual Comm. Image Rep*, 12:2001, 2000.
- [28] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud. Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*, 6(2), 2007.
- [29] P. G Ciarlet. *Introduction à l'Analyse Numérique Matricielle et à l'Optimisation*. Dunod, 1998.
- [30] P.L. Combettes and J. Pesquet. Image restoration subject to a total variation constraint. *IEEE Transactions on Image Processing*, 13(9):1213–1222, 2004.
- [31] P.L. Combettes and V. Wajs. Signal recovery by proximal forward-backward splitting. *SIAM Journal on Multiscale Modeling and Simulation*, 4(4):1168–1200, 2005.
- [32] J. Darbon. Total variation minimization with l^1 data fidelity as a contrast invariant filter. *4th International Symposium on Image and Signal Processing and Analysis (ISPA 2005)*, pages 221–226, September 2005.
- [33] J. Darbon and M. Sigelle. Image restoration with discrete constrained total variation part I: Fast and exact optimization. *Journal of Mathematical Imaging and Vision*, 26(3):261–276, December 2006.
- [34] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57:1413–1457, 2004.
- [35] I. Daubechies and G. Teschke. Variational image restoration by means of wavelets: simultaneous decomposition, deblurring and denoising. *Applied and Computational Harmonic Analysis*, 19:1–16, 2005.
- [36] D. Dobson and C. Vogel. Convergence of an iterative method for total variation denoising. *SIAM Journal on Numerical Analysis*, 34:1779–1791, 1997.
- [37] S. Durand, J. Fadili, and M. Nikolova. Multiplicative noise removal using L^1 fidelity on frame coefficients. *CMLA Report*, 08-40, 2008.
- [38] V. Duval and Y. Gousseau J.-F. Aujol. The TVL1 model: a geometric point of view. *SIAM journal on Multiscale Modeling and Simulation*, April 2009. (submitted).
- [39] J. Eckstein. The Lions-Mercier algorithm and the alternating direction method are instances of the proximal point algorithm. 1988.
- [40] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems*, volume 28 of *Classics in Applied Mathematics*. SIAM, 1999.

- [41] J. Fadili and J.-L. Starck. Monotone operator splitting for fast sparse solutions of inverse problems., 2009. submitted.
- [42] H. Fu, M. Ng, M. Nikolova, and J. Barlow. Efficient minimization methods of mixed l1-l1 and l2-l1 norms for image restoration. *SIAM Journal on Scientific computing*, 27(6):1881–1902, 2006.
- [43] J. Garnett, P. Jones, T. Le, and L. Vese. Modeling oscillatory components with the homogeneous spaces $BMO^{-\alpha}$ and $W^{-\alpha,p}$. *UCLA CAM Report*, 07-21, July 2007.
- [44] J. Gilles. Noisy image decomposition: a new structure, textures and noise model based on local adaptivity. *Journal of Mathematical Imaging and Vision*, 28(3), July 2007.
- [45] D. Goldfarb and W. Yin. Second-order cone programming methods for total variation based image restoration. *SIAM Journal on Scientific Computing*, 27(2):622–645, 2005.
- [46] A. Haddad. *Méthodes variationnelles en traitement d’images*. PhD thesis, ENS Cachan, June 2005.
- [47] B. Julesz. Texton gradients: The texton theory revisited. *Biological Cybernetics*, 54:245–251, 1986.
- [48] T. Le and L. Vese. Image decomposition using total variation and $\text{div}(BMO)$. *Multiscale Modeling and Simulation, SIAM Interdisciplinary Journal*, 4(2):390–423, June 2005.
- [49] L. Lieu. *Contribution to Problems in Image Restoration, Decomposition, and Segmentation by Variational Methods and Partial Differential Equations*,. PhD thesis, UCLA, June 2006.
- [50] L.H. Lieu and L.A. Vese. Image restoration and decomposition via bounded total variation and negative Hilbert-Sobolev spaces. *Applied Mathematics & Optimization*, 58:167–193, 2009.
- [51] P.L. Lions and B. Mercier. Algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis*, 16(6):964–979, 1979.
- [52] S.J. Wright M. Zhu and T.F. Chan. Duality-based algorithms for total variation image restoration. Technical report, UCLA CAM Report 08-33, 2008.
- [53] Y. Meyer. *Oscillating patterns in image processing and nonlinear evolution equations*, volume 22 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2001. The fifteenth Dean Jacqueline B. Lewis memorial lectures.
- [54] Y. Nesterov. Smooth minimization of non-smooth functions. *Mathematical Programming (A)*, 103(1):127–152, 2005.
- [55] M.K. Ng, L. Qi, Y.F. Yang, and Y. Huang. On semismooth Newton methods for total variation minimization. *Journal of Mathematical Imaging and Vision*, 27:265–276, 2007.
- [56] M. Nikolova and R. Chan. The equivalence of half-quadratic minimization and the gradient linearization iteration. *IEEE Transactions on Image Processing*, 16(6):1623–1627, 2007.
- [57] S. Osher and N. Paragios. *Geometric Level Set Methods in Imaging, Vision, and Graphic*. Springer, 2003.

- [58] S. Osher, A. Sole, and L. Vese. 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.
- [59] L. Rudin, S. Osher, and E. Fatemi. Non linear total variation based noise removal algorithms. *Physica D*, 60:259–268, 2003.
- [60] G. Sapiro and D. L. Ringach. Anisotropic diffusion of multivalued images with applications to color filtering. *IEEE Transactions on Image Processing*, 5(11):1582–1586, 1996.
- [61] S. Setzer. Split Bregman Algorithm, Douglas-Rachford Splitting and frame shrinkage. *Preprint, University of Mannheim*, 2008.
- [62] N. Sochen, R. Kimmel, and R. Malladi. A general framework for low level vision. *IEEE Transactions on Image Processing*, 7(3):310–318, 1998.
- [63] P. Soille. *Morphological Image Analysis: Principles and Applications*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2003.
- [64] J.L. Starck, M. ELad, and D.L. Donoho. Image decomposition via the combination of sparse representations and a variational approach, 2005.
- [65] L.A. Vese and S. J. Osher. Modeling textures with total variation minimization and oscillating patterns in image processing. *Journal of Scientific Computing*, 19(1-3):553–572, December 2003.
- [66] L.A. Vese and S.J. Osher. Color texture modeling and color image decomposition in a variational-PDE approach . In *Proceedings of the Eighth International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC '06)*, pages 103 – 110. IEEE, 2006.
- [67] C.R. Vogel. *Computational Methods for Inverse Problems*, volume 23 of *Frontiers in Applied Mathematics*. SIAM, 2002.
- [68] P. Weiss, L. Blanc-Féraud, and G. Aubert. Efficient schemes for total variation minimization under constraints in image processing. *SIAM journal on Scientific Computing*, 31(3):2047–2080, 2009.
- [69] W. Yin and D. Goldfarb. Second-order cone programming methods for total variation based image restoration. *SIAM J. Scientific Computing*, 27(2):622–645, 2005.
- [70] W. Yin, D. Goldfarb, and S. Osher. Image cartoon-texture decomposition and feature selection using the total variation regularized l^1 . In *Variational, Geometric, and Level Set Methods in Computer Vision*, volume 3752 of *Lecture Notes in Computer Science*, pages 73–84. Springer, 2005.
- [71] W. Yin, D. Goldfarb, and S. Osher. A comparison of three total variation based texture extraction models. *Journal of Visual Communication and Image Representation*, 18(3):240–252, June 2007.
- [72] J. Yuan, C. Schnörr, and G. Steidl. Convex Hodge decomposition and regularization of image flows. *Journal of Mathematical Imaging and Vision*, 2009. in press.
- [73] M. Zhu and T.F. Chan. An efficient primal-dual hybrid gradient algorithm for total variation image restoration, May 2008. UCLA CAM Report 08-34.