FAST DUAL MINIMIZATION OF THE VECTORIAL TOTAL VARIATION NORM AND APPLICATIONS TO COLOR IMAGE PROCESSING

Xavier Bresson and Tony F. Chan

Department of Mathematics, University of California
Los Angeles, CA 90095-1555, USA

(Communicated by Margaret Cheney)

Abstract. We propose a regularization algorithm for color/vectorial images which is fast, easy to code and mathematically well-posed. More precisely, the regularization model is based on the dual formulation of the vectorial Total Variation (VTV) norm and it may be regarded as the vectorial extension of the dual approach defined by Chambolle in [13] for gray-scale/scalar images. The proposed model offers several advantages. First, it minimizes the exact VTV norm whereas standard approaches use a regularized norm. Then, the numerical scheme of minimization is straightforward to implement and finally, the number of iterations to reach the solution is low, which gives a fast regularization algorithm. Finally, and maybe more importantly, the proposed VTV minimization scheme can be easily extended to many standard applications. We apply this L^1 vectorial regularization algorithm to the following problems: color inverse scale space, color denoising with the chromaticity-brightness color representation, color image inpainting, color wavelet shrinkage, color image decomposition, color image deblurring, and color denoising on manifolds. Generally speaking, this VTV minimization scheme can be used in problems that required vector field (color, other feature vector) regularization while preserving discontinuities.

1. **Introduction.** This paper is devoted to the regularization of multidimensional/ vectorial signals, such as color images or other vector fields (vector of features, normal of level sets, etc.), based on the vectorial TV norm. The regularization of vectorial images in the context of continuous models have already been proposed in the literature. In [33], Sapiro and Ringach proposed an anisotropic diffusion model for vectorial images based on the computation of the eigenvalues of the M-D image defined as a 2-D manifold. The eigenvectors determine the directions of maximal and minimal change of the vectorial image, and their associated eigenvalues give their rate of change. In [8], Blomgren and Chan defined a definition of the VTV norm which satisfies several good properties. Sochen, Kimmel and Malladi in [35] defined a vectorial image denoising models based on a 2-D manifold embedded in a $(M + 2)$-D space and the Polyakov energy. We propose in this paper a vectorial regularization algorithm that is as efficient as [35] and improve the regularization algorithms [33, 8].

©2008 American Institute of Mathematical Sciences

The proposed VTV minimization model is based on the dual formulation of the vectorial TV norm. The idea of dualizing the (scalar) TV norm was introduced by Chan, Golub and Mulet (CGM) in [15] to solve the denoising Rudin-Osher-Fatemi (ROF) model [32] for gray-scale images in a faster way than the previous models. CGM introduced in the TV minimization model an extra maximization problem w.r.t. a new (dual) variable. The new minimization-maximization problem can be solved using the Sion's theorem [34]. Sion's theorem is a generalization of John Von Neumann's minimax theorem [40] (see also [2, 22]) introduced in game theory for minimizing the maximum possible loss.

Based on the CGM model, Chambolle in [13] developed an efficient dual approach to minimize the scalar ROF model. Chambolle's algorithm is faster than CGM even if the convergence of Chambolle's scheme is linear and the CGM's scheme is quadratic. Chambolle's algorithm is faster because the cost per iteration to use CGM is higher (CGM needs to solve a linear system at each iteration). A recent fast minimization algorithm for the scalar ROF model was proposed by Darbon and Sigelle (DS) in [20] based on graph cuts. Although Chambolle's algorithm is not as fast as the model of DS to solve the variational scalar ROF model, it is still fast and presents some advantages compared with CGM and DS. First, Chambolle's model use the exact scalar TV norm whereas CGM model regularizes it to minimize it. Then, the numerical scheme of [13] is straightforward to implement unlike the CGM and DS algorithms. Besides, the TV norm of DS is anisotropic whereas the TV norm of Chambolle is isotropic. Finally, we will see that the Chambolle's model extends nicely to color/vector images whereas the question of extension is open for the CGM model and the generalization of DS model to color images is not as efficient as in the scalar case (see [19]).

This paper proposes these two following contributions:
• Extension of Chambolle's model [13] to multidimensional/vectorial images. We will see that the extended algorithm presents the same nice properties than the scalar algorithm since the VTV minimization algorithm is fast, easy to code and well-posed. Unlike [8, 4, 16], the proposed vectorial scheme does not regularize the VTV to minimize it. Finally, the numerical solution converges to the continuous minimizing solution in the vectorial BV space.
• Extension of the proposed VTV minimization scheme to several standard applications such as deblurring, inpainting, decomposition, denoising on manifolds. In fact, this vectorial regularization scheme can be applied to any problems that require a L^1 regularization process for vectorial components.

The paper is organized in two parts. The first part is the introduction of the VTV minimization algorithm. We introduce some notations and the dual formulation of the VTV. We then introduce the VTV minimization algorithm and prove its convergence to the continuous minimizing solution. Then, we compare our model to related denoising models. The second part of this paper aims at showing that this vectorial regularization process can be easily extended to any problems that require L^1 vector field regularization. We will focus on color image processing problems such as color inverse scale space, color denoising with the chromaticity-brightness color representation, color inpainting, color wavelet shrinkage, color decomposition, color deblurring or color denoising on manifold. But we can also extend this algorithm to

other problems such as the vector field regularization of diffusion tensors in medical imaging.

2. **Part I: Dual vectorial TV minimization algorithm.**

2.1. **Notation and definition.** In this paper, we are interested in the standard dual definition of the VTV in order to extend the Chambolle's scalar algorithm [13] to vectorial images such as color images or any vector fields. This section introduces some notations and definitions regarding the vector-valued TV norm used throughout the paper. Let us consider a vectorial (or M-dimensional or multichannel) function $\mathbf{u}$, such as a color image or a vector field, defined on a bounded open domain $\Omega \subset \mathbb{R}^N$, as follows:

$$\begin{aligned} \mathbf{u} \; : \quad & \Omega \quad \rightarrow \quad \mathbb{R}^M, \\ & \mathbf{x} \quad \rightarrow \quad \mathbf{u}(\mathbf{x}) := (u_1(\mathbf{x}), ..., u_M(\mathbf{x})), \end{aligned}$$

where each scalar function $u_i : \Omega \to \mathbb{R}, \; 1 \le i \le M$ is one of the M components (or channels) of the vector valued function $\mathbf{u}$.

Definition 2.1. For a given vector valued function $\mathbf{u} : \Omega \to \mathbb{R}^M$, the vectorial TV norm is denoted the finite positive measure:

$$(1) \qquad \int_\Omega |D\mathbf{u}| := \sup_{\mathbf{p} \in P} \left\{ \int_\Omega <\mathbf{u}, \nabla \cdot \mathbf{p}> d\mathbf{x} \right\}$$

where $\mathbf{p} := (\mathbf{p}_1, ..., \mathbf{p}_M) : \Omega \to \mathbb{R}^{M \times N}$, $\mathbf{p}_i := (p_i^{x_1}, ..., p_i^{x_N}) : \Omega \to \mathbb{R}^N$, $\forall i \in [1, M]$, $\nabla \cdot$ is the divergence operator such that $\nabla \cdot \mathbf{q} := (\nabla \cdot \mathbf{q}_1, ..., \nabla \cdot \mathbf{q}_N) : \Omega \to \mathbb{R}^M$, $\forall \mathbf{q} : \Omega \to \mathbb{R}^{M \times N}$, $\nabla \cdot \mathbf{q}_i := \sum_{j=1}^N \partial_{x_j} q_i^{x_j} : \Omega \to \mathbb{R}$, $\forall i \in [1, M]$, the product $<., >$ is the Euclidean scalar product defined as $<\mathbf{v}, \mathbf{w}> := \sum_{i=1}^M <v_i, w_i>$, $\forall(\mathbf{v}, \mathbf{w}) \in (\mathbb{R}^M)^2$, which implies that $<\mathbf{u}, \nabla \cdot \mathbf{p}> = \sum_{i=1}^M <u_i, \nabla \cdot \mathbf{p}_i>$, and the L^2/Euclidean norm $|.|$ is naturally defined by $|\mathbf{v}| := \sqrt{<\mathbf{v}, \mathbf{v}>} = \sqrt{\sum_{i=1}^L v_i^2}$, $\forall \mathbf{v} \in \mathbb{R}^L$.

Thus, the VTV norm (1) can be defined of different ways, depending on the set P of functions of the dual variable $\mathbf{p}$. Let us consider two sets:

$$\begin{aligned} P_1 \quad &:= \quad \{\mathbf{p} \in C_c^1(\Omega; \mathbb{R}^{M \times N}) \; : \; |\mathbf{p}|_\infty \le 1\} \\ P_2 \quad &:= \quad \{\mathbf{p} \in C_c^1(\Omega; \mathbb{R}^{M \times N}) \; : \; |\mathbf{p}| \le 1\} \end{aligned}$$

where $|.|_\infty$ in P_1 is the infinity norm such that $|\mathbf{p}|_\infty = \max\limits_{i=1,...,M} |\mathbf{p}_i|$ and $|.|$ in P_2 is the L^2 norm such that $|\mathbf{p}| = \sqrt{\sum_{i=1}^M <\mathbf{p}_i, \mathbf{p}_i>} = \sqrt{\sum_{i=1}^M \sum_{j=1}^N (p_i^{x_j})^2}$. If we assume that $\mathbf{u} \in C^1(\Omega; \mathbb{R}^M)$, i.e. $\mathbf{u}$ are smooth functions, and $\mathbf{p} \in C_c^1(\Omega; \mathbb{R}^{M \times N})$ are with compact support, then integration by parts gives:

$$\begin{aligned} \int_\Omega |D\mathbf{u}| \quad &= \quad \sup_{\mathbf{p} \in P} \left\{ \int_\Omega <\mathbf{u}, \nabla \cdot \mathbf{p}> d\mathbf{x} \right\} \\ &= \quad \sup_{\mathbf{p} \in P} \left\{ \int_\Omega \sum_{i=1}^M <\nabla u_i, \mathbf{p}_i> d\mathbf{x} \right\}. \end{aligned}$$

Then if $\mathbf{p} \in P_1$ then the supremum of (1) is obtained for:

$$\mathbf{p}_i = \left\{ \begin{array}{ll} \frac{\nabla u_i}{|\nabla u_i|} & \text{if } \nabla u_i \neq 0 \\ \text{any} & \text{if } \nabla u_i = 0 \end{array} \right. , \ 1 \leq i \leq M,$$

which defines the vectorial TV norm for smooth functions $\mathbf{u}$ as:

$$(2) \qquad \int_\Omega |D\mathbf{u}| = \sum_{i=1}^M \int_\Omega |\nabla u_i| d\mathbf{x} = \sum_{i=1}^M TV(u_i),$$

i.e. the sum of the TV of each channel. Then if $\mathbf{p} \in P_2$ then the supremum is obtained for:

$$\mathbf{p}_i = \left\{ \begin{array}{ll} \frac{\nabla u_i}{|\nabla \mathbf{u}|} & \text{if } \nabla \mathbf{u} \neq 0 \\ \text{any} & \text{if } \nabla \mathbf{u} = 0 \end{array} \right. , \ 1 \leq i \leq M,,$$

where the vectorial gradient of $\mathbf{u}$ is defined as $\nabla \mathbf{u} := (\nabla u_1, ..., \nabla u_N) : \Omega \to \mathbb{R}^{M \times N}$ and $|\nabla \mathbf{u}| = \sqrt{\sum_{i=1}^M < \nabla u_i, \nabla u_i >}$. Thus, the vectorial TV norm given by P_2 for smooth functions $\mathbf{u}$ is:

$$(3) \int_\Omega |D\mathbf{u}| = \int_\Omega \frac{\sum_{i=1}^M < \nabla u_i, \nabla u_i >}{|\nabla \mathbf{u}|} d\mathbf{x} = \int_\Omega \sqrt{\sum_{i=1}^N |\nabla u_i|^2} d\mathbf{x} = \int_\Omega |\nabla \mathbf{u}| d\mathbf{x}.$$

Hence, the VTV defined in (1) have different expressions for smooth functions $\mathbf{u}$ for different sets P_1 or P_2. The difference between sets P_1 or P_2 is important in the denoising context. Indeed, the vectorial TV based on P_1 is defined as the sum of the TV of each channel. This means that channels are considered as independent in the denoising process, which is not true on real-world images. The vectorial TV based on P_2 introduces a coupling between channels. Each channel use information coming from other channels to improve the denoising model as we will see in Section 2.5.

Finally, we want to emphasize that the vector valued TV norm defined by the set P_2, i.e. the L^2-norm, is the most standard definition of the VTV norm as introduced in the book of Ambrosio, Fusco and Pallara [1]. From this point, we will use the L^2-norm for the VTV norm throughout this paper.

We now present standard results regarding the VTV. These results can be found e.g. in [1, 27, 24].

Definition 2.2. Definition of Vector Valued Space $BV(\Omega; \mathbb{R}^M)$, $M \geq 1$.

We define the space $BV(\Omega; \mathbb{R}^M)$ of vector valued functions as the set of functions $\mathbf{u} \in L^1(\Omega; \mathbb{R}^M)$ such that $\int_\Omega |D\mathbf{u}| < \infty$, where the vectorial TV norm $\int_\Omega |D\mathbf{u}|$ is defined in Definition 2.1 with the set of functions $\mathbf{p} \in P_2$. The space $BV(\Omega; \mathbb{R}^M)$ endowed with the following norm:

$$\|\mathbf{u}\|_{L^1(\Omega; \mathbb{R}^M)} + \|\mathbf{u}\|_{BV(\Omega; \mathbb{R}^M)},$$

where $\|\mathbf{u}\|_{BV(\Omega; \mathbb{R}^M)} := \int_\Omega |D\mathbf{u}|$, is a Banach space.

Theorem 2.3. *Lower Semicontinuity Property.*

Let $\{\mathbf{u}_n\}$ a sequence of vector valued functions in $BV(\Omega; \mathbb{R}^M)$ which converge in $L^1(\Omega; \mathbb{R}^M)$ to a function $\mathbf{u}$. Then

$$\|\mathbf{u}\|_{BV(\Omega; \mathbb{R}^M)} \leq \liminf_{n \to \infty} \|\mathbf{u}_n\|_{BV(\Omega; \mathbb{R}^M)}.$$

Theorem 2.4. *Compactness Property.*
    Every sequence $\{\mathbf{u}_n\} \in BV(\Omega; \mathbb{R}^M)$ *such that* $\|\mathbf{u}_n\|_{L^\infty(\Omega; \mathbb{R}^M)} \leq M$, $\forall n \in \mathbb{N}$, *admits a subsequence* $\{\mathbf{u}_{n_k}\}$ *converging in* $L^1(\Omega; \mathbb{R}^M)$ *to a function* $\mathbf{u} \in BV(\Omega; \mathbb{R}^M)$.

Theorem 2.5. *Decomposition of* $\|\mathbf{u}\|_{BV(\Omega; \mathbb{R}^M)} = \int_\Omega |D\mathbf{u}|$.
    Functional $\int_\Omega |D\mathbf{u}|$ *can be decomposed as follows:*

$$\int_\Omega |D\mathbf{u}| = \int_\Omega |\nabla \mathbf{u}| d\mathbf{x} + |D^c \mathbf{u}|(\Omega) + \int_{J_\mathbf{u} \cap \Omega} |\mathbf{u}^+ - \mathbf{u}^-| dH^{N-1},$$

where $\nabla \mathbf{u}$ *is the standard gradient of* $\mathbf{u}$ *with* $\int_\Omega |\nabla \mathbf{u}| d\mathbf{x} = \int_\Omega \sqrt{\sum_{i=1}^N |\nabla u_i|^2} d\mathbf{x}$ *for* $\mathbf{u} \in C^1(\Omega; \mathbb{R}^M)$, $D^c \mathbf{u}$ *is the Cantor part of* $D\mathbf{u}$, $J_\mathbf{u}$ *is the set of all jump points of* $\mathbf{u}$, $\mathbf{u}^+$, $\mathbf{u}^-$ *are the jump functions and* H^{N-1} *denotes the* $(N-1)$*-dimensional Hausdorff measure in* $\mathbb{R}^N$.

2.2. **Vectorial Rudin-Osher-Fatemi (VROF) model.** The Rudin-Osher-Fatemi model [32] is one of the most influential variational and PDE-based image denoising models in image processing. This denoising model removes noise in gray-scale images while preserving main features such as edges. In the case of vector-valued images, the ROF model is naturally extended as follows:

$$(4) \qquad \inf_{\mathbf{u}} \left\{ F(\mathbf{u}) := \|\mathbf{u}\|_{BV(\Omega; \mathbb{R}^M)} + \frac{1}{2\lambda} \|\mathbf{f} - \mathbf{u}\|_{L^2(\Omega; \mathbb{R}^M)}^2 \right\},$$

where $\mathbf{f} \in L^\infty(\Omega; \mathbb{R}^M)$, $\mathbf{f} := (f_1, ..., f_M)$ is the given (noisy) vector valued image and $\|\mathbf{f} - \mathbf{u}\|_{L^2(\Omega; \mathbb{R}^M)}^2 = \int_\Omega |\mathbf{f} - \mathbf{u}|^2 \, d\mathbf{x}$ is the L^2 fidelity norm.

Existence of a minimizer for the variational vectorial ROF model (4) can be proved by standard arguments of the vectorial BV space. We include the existence proof in this section for the sake of completeness. First, let us introduce truncated vectorial functions:

Theorem 2.6. *Truncated Functions.*
    Let $\mathbf{f} \in L^\infty(\Omega; \mathbb{R}^M)$ *and* $\mathbf{u} : \Omega \to \mathbb{R}^M$. *We define the truncated function* $\hat{\mathbf{u}} := (\|f_i\|_{L^\infty(\Omega; \mathbb{R})} \wedge u_i \vee -\|f_i\|_{L^\infty(\Omega; \mathbb{R})})_{i=1}^M$.
    For any $\mathbf{u} \in BV(\Omega; \mathbb{R}^M)$, *we have* $\hat{\mathbf{u}} \in BV(\Omega; \mathbb{R}^M)$ *and*
    (j) $\int_\Omega |\mathbf{f} - \hat{\mathbf{u}}|^2 \, d\mathbf{x} \leq \int_\Omega |\mathbf{f} - \mathbf{u}|^2 \, d\mathbf{x}$,
    (jj) $\|\hat{\mathbf{u}}\|_{BV(\Omega; \mathbb{R}^M)} \leq \|\mathbf{u}\|_{BV(\Omega; \mathbb{R}^M)}$.

Proof. The proof is, for example, based on [23]. The definition of the truncated function implies that $|\mathbf{f} - \hat{\mathbf{u}}| \leq |\mathbf{f} - \mathbf{u}|$ from which follows (j).

The truncated function is given by $\hat{\mathbf{u}}_i = h_i \circ \mathbf{u}_i$ where $h_i : \mathbb{R} \to \mathbb{R}$ is a Lipschitz function such that:

$$h_i(t) = \begin{cases} t & \text{if } -\|f_i\|_{L^\infty} \leq t \leq \|f_i\|_{L^\infty} \\ \|f_i\|_{L^\infty} & \text{if } t > \|f_i\|_{L^\infty} \\ -\|f_i\|_{L^\infty} & \text{if } t < -\|f_i\|_{L^\infty} \end{cases}.$$

Using the chain rule, we have $\hat{\mathbf{u}} \in BV(\Omega; \mathbb{R}^M)$.

Following [23], we have $|\nabla \hat{\mathbf{u}}_i(\mathbf{x})| \leq |\nabla \mathbf{u}_i(\mathbf{x})|$ then $\int_\Omega |\nabla \hat{\mathbf{u}}| d\mathbf{x} < \int_\Omega |\nabla \mathbf{u}| d\mathbf{x}$. Since $\mathbf{u}^+(\mathbf{x}) \geq \mathbf{u}^-(\mathbf{x})$ by construction and by definition of the function h_i we have $J_{\hat{\mathbf{u}}_i} \subseteq J_{\mathbf{u}_i}$ and $\hat{\mathbf{u}}^+(\mathbf{x}) - \hat{\mathbf{u}}^-(\mathbf{x}) \leq \mathbf{u}^+(\mathbf{x}) - \mathbf{u}^-(\mathbf{x}), \forall \mathbf{x} \in J_{\hat{\mathbf{u}}_i}$. Then it follows that $\int_{J_{\hat{\mathbf{u}}} \cap \Omega} |\hat{\mathbf{u}}^+ - \hat{\mathbf{u}}^-| dH^{N-1} \leq \int_{J_\mathbf{u} \cap \Omega} |\mathbf{u}^+ - \mathbf{u}^-| dH^{N-1}$. We also have $|D^c \hat{\mathbf{u}}_i| \leq |D^c \mathbf{u}_i|$, i.e. $|D^c \hat{\mathbf{u}}| \leq |D^c \mathbf{u}|$. The terms are summed over and we obtain (jj). $\quad\square$

Then, it is easy to prove the existence of a minimizer for the VROF.

Theorem 2.7. *Existence of a Solution for the Vectorial ROF Model ([4]).*
The vectorial ROF model

$$\inf_\mathbf{u} \left\{ F(\mathbf{u}) := \|\mathbf{u}\|_{BV(\Omega; \mathbb{R}^M)} + \frac{1}{2\lambda} \|\mathbf{f} - \mathbf{u}\|_{L^2(\Omega; \mathbb{R}^M)}^2 \right\},$$

where $\mathbf{f} \in L^\infty(\Omega; \mathbb{R}^M)$, has a solution in $BV(\Omega; \mathbb{R}^M)$.

Proof. Let $\{\mathbf{u}_n\}$ be a minimizing sequence in $BV(\Omega; \mathbb{R}^M)$. By definition of the total variation norm, $\|\mathbf{u}_n\|_{BV(\Omega; \mathbb{R}^M)}$ is uniformly bounded by a strict positive constant M such that $\|\mathbf{u}_n\|_{BV(\Omega; \mathbb{R}^M)} \leq M, \ \forall n \in \mathbb{N}$.

By Theorem 2.6, we can modify the minimizing sequence by truncation, obtaining a new minimizing sequence $\hat{\mathbf{u}}_n$ that is uniformly bounded in $BV(\Omega; \mathbb{R}^M)$, i.e.

$$\|\hat{\mathbf{u}}_n\|_{L^\infty(\Omega; \mathbb{R}^M)} \leq K, \quad \|\hat{\mathbf{u}}_n\|_{BV(\Omega; \mathbb{R}^M)} \leq M, \quad \forall n \in \mathbb{N},$$

where $K = \max_{i=1,\dots,M} \|f_i\|_{L^\infty(\Omega; \mathbb{R})}$.

Therefore, according to Theorem 2.4, there exists a subsequence $\{\hat{\mathbf{u}}_{n_k}\}$ converging to a function $\mathbf{u} \in BV(\Omega; \mathbb{R}^M)$. Since TV is lower semicontinuous in $BV(\Omega; \mathbb{R}^M)$ by definition, we have:

$$\|\mathbf{u}\|_{BV(\Omega; \mathbb{R}^M)} \leq \liminf_{k \to \infty} \|\hat{\mathbf{u}}_{n_k}\|_{BV(\Omega; \mathbb{R}^M)} \leq \liminf_{k \to \infty} \|\mathbf{u}_{n_k}\|_{BV(\Omega; \mathbb{R}^M)},$$

and from Fatou's Lemma:

$$\|\mathbf{f} - \mathbf{u}\|_{L^2(\Omega; \mathbb{R}^M)}^2 \leq \liminf_{k \to \infty} \|\mathbf{f} - \hat{\mathbf{u}}_{n_k}\|_{L^2(\Omega; \mathbb{R}^M)}^2 \leq \liminf_{k \to \infty} \|\mathbf{f} - \mathbf{u}_{n_k}\|_{L^2(\Omega; \mathbb{R}^M)}^2,$$

which implies that

$$F(\mathbf{u}) \leq \liminf_{k \to \infty} F(\mathbf{u}_{n_k}).$$

It follows that $\mathbf{u} \in BV(\Omega; \mathbb{R}^M)$ is a minimizer of F and the vectorial ROF model ([4]). $\quad\square$

Finally, the VROF model ([4]) is convex, which guaranties the uniqueness of the minimizer.

2.3. **Proposed algorithm for VROF minimization.** In this section, we present the solution minimizing the vectorial ROF model (4). Using the dual definition of the vectorial TV norm given in (1), the variational model (4) can be written as

$$(5) \quad \inf_{\mathbf{u}} \sup_{|\mathbf{p}| \leq 1} \left\{ < \mathbf{u}, \nabla \cdot \mathbf{p} >_{L^2(\Omega;\mathbb{R}^M)} + \frac{1}{2\lambda} \|\mathbf{f} - \mathbf{u}\|^2_{L^2(\Omega;\mathbb{R}^M)} \right\}$$

Using the minimax theorem [2, 22], inf and sup can be swapped since (5) is convex in $\mathbf{u}$ and concave in $\mathbf{p}$ and the set $\{|\mathbf{p}| \leq 1\}$ is bounded and convex. Thus (5) is first minimized w.r.t. $\mathbf{u}$ using the Euler-Lagrange's technique, which gives the minimizing solution of the VROF model:

$$(6) \quad \mathbf{u} = \mathbf{f} - \lambda \nabla \cdot \mathbf{p},$$

and for each channel:

$$u_i = f_i - \lambda \nabla \cdot \mathbf{p}_i, \quad 1 \leq i \leq M.$$

Substituting solution (6) in (5), we get the constrained maximization problem:

$$\sup_{|\mathbf{p}| \leq 1} \left\{ < \mathbf{f}, \nabla \cdot \mathbf{p} >_{L^2(\Omega;\mathbb{R}^M)} - \frac{\lambda}{2} \|\nabla \cdot \mathbf{p}\|^2_{L^2(\Omega;\mathbb{R}^M)} \right\},$$

which is equivalent to this minimization problem:

$$(7) \quad \inf_{|\mathbf{p}| \leq 1} \|\nabla \cdot \mathbf{p} - \mathbf{f}/\lambda\|^2_{L^2(\Omega;\mathbb{R}^M)}.$$

The Euler-Lagrange's technique gives the necessary condition for optimality at each $\mathbf{x}$:

$$(8) \quad -\nabla \big(\lambda \nabla \cdot \mathbf{p} - \mathbf{f}\big) + \alpha \mathbf{p} = 0,$$

where $\alpha(\mathbf{x})$ is the Lagrange multiplier associates with the constraint $|\mathbf{p}| \leq 1$. As noticed by Chambolle in [13] for the scalar case, it can also be eliminated in the vectorial case as follows. If $|\mathbf{p}| < 1$, then the Lagrange multiplier is not active, i.e. $\alpha = 0$ and $\nabla\big(\lambda\nabla \cdot \mathbf{p} - \mathbf{f}\big) = 0$ and if $|\mathbf{p}| = 1$, then the Lagrange multiplier becomes active and $\alpha = |\nabla\big(\lambda\nabla \cdot \mathbf{p} - \mathbf{f}\big)| > 0$. In any case, the Lagrange multiplier is equal to:

$$(9) \quad \alpha = |\nabla\big(\lambda\nabla \cdot \mathbf{p} - \mathbf{f}\big)|.$$

Introducing (9) in (8), we get:

$$\nabla\big(\lambda\nabla \cdot \mathbf{p} - \mathbf{f}\big) - |\nabla\big(\lambda\nabla \cdot \mathbf{p} - \mathbf{f}\big)| \, \mathbf{p} = 0,$$

which can be solved as in [13] with a semi-implicit gradient descent scheme as follows: $\mathbf{p}^{n=0} = 0$,

$$(10) \quad \nabla\big(\nabla \cdot \mathbf{p}^n - \mathbf{f}/\lambda\big) - |\nabla\big(\nabla \cdot \mathbf{p}^n - \mathbf{f}/\lambda\big)| \, \mathbf{p}^{n+1} = (\mathbf{p}^{n+1} - \mathbf{p}^n)/\delta t,$$

such that:

$$(11) \quad \mathbf{p}^{n+1} = \frac{\mathbf{p}^n + \delta t \nabla\big(\nabla \cdot \mathbf{p}^n - \mathbf{f}/\lambda\big)}{1 + \delta t |\nabla\big(\nabla \cdot \mathbf{p}^n - \mathbf{f}/\lambda\big)|},$$

which means for each channel:

$$(12) \quad \mathbf{p}_i^{n+1} = \frac{\mathbf{p}_i^n + \delta t \nabla\big(\nabla \cdot \mathbf{p}_i^n - f_i/\lambda\big)}{1 + \delta t \sqrt{\sum_{j=1}^M |\nabla\big(\nabla \cdot \mathbf{p}_j^n - f_j/\lambda\big)|^2}}, \quad 1 \leq i \leq M.$$

As in [13], the convergence of the iterative scheme (11) also holds since we have the following theorem.

Theorem 2.8. *Let* $\mathbf{f} \in L^\infty(\Omega; \mathbb{R}^M)$, *the iterative scheme (11) converges to the minimizing solution (7) for* $\delta t \le 1/8$.

Proof. We follow the proof given by Chambolle in [13] in the case of gray-scale images. Let us denote $|.|_{X,Y}, < .,. >_{X,Y}$ the Euclidean norm and inner product defined in $X := L^2(\Omega; \mathbb{R}^M)$ and $Y := L^2(\Omega; \mathbb{R}^{M \times N})$ with $|\mathbf{f}|_X := \sqrt{\sum_{i=1}^M f_i^2}$ and $|\mathbf{p}|_Y := \sqrt{\sum_{i=1}^M \sum_{j=1}^N (p_i^{x_j})^2}$. We call $\boldsymbol{\eta} := (\mathbf{p}^{n+1} - \mathbf{p}^n)/\delta t$ and we have:

$$
\begin{aligned}
|\nabla \cdot \mathbf{p}^{n+1} - \mathbf{f}/\lambda|_X^2 &= |\nabla \cdot \mathbf{p}^n - \mathbf{f}/\lambda|_X^2 + \\
&\quad\; 2\delta t < \nabla \cdot \boldsymbol{\eta}, \nabla \cdot \mathbf{p}^n - \mathbf{f}/\lambda >_X + \delta t^2 |\nabla \cdot \boldsymbol{\eta}|_X^2, \\
&\le |\nabla \cdot \mathbf{p}^n - \mathbf{f}/\lambda|_X^2 - \\
&\quad\; \delta t \big(2 < \boldsymbol{\eta}, \nabla(\nabla \cdot \mathbf{p}^n - \mathbf{f}/\lambda) >_Y - \delta t \tau^2 |\boldsymbol{\eta}|_Y^2 \big),
\end{aligned}
$$

where $\tau := |\nabla \cdot|_X$ is the norm of the divergence operator $\nabla \cdot : Y \to X$. From (10), we denote $\boldsymbol{\rho} := |\nabla(\nabla \cdot \mathbf{p}^n - \mathbf{f}/\lambda)|_Y \mathbf{p}^{n+1} = \nabla(\nabla \cdot \mathbf{p}^n - \mathbf{f}/\lambda) - \boldsymbol{\eta}$ such that we have:

$$
\begin{aligned}
2 < \boldsymbol{\eta}, \nabla(\nabla \cdot \mathbf{p}^n - \mathbf{f}/\lambda) >_Y - \delta t \tau^2 |\boldsymbol{\eta}|_Y^2 &= \\
(1 - \delta t \tau^2)|\boldsymbol{\eta}|_Y^2 + (|\nabla(\nabla \cdot \mathbf{p}^n - \mathbf{f}/\lambda)|_Y^2 - |\boldsymbol{\rho}|_Y^2). &
\end{aligned}
$$

We have $|\nabla(\nabla \cdot \mathbf{p}^n - \mathbf{f}/\lambda)|_Y^2 - |\boldsymbol{\rho}|_Y^2 = |\nabla(\nabla \cdot \mathbf{p}^n - \mathbf{f}/\lambda)|_Y^2 (1 - |\mathbf{p}^{n+1}|_Y^2) \ge 0$ since $|\mathbf{p}^n|_Y \le 1$ for $n \ge 0$. Hence, if $\delta t \le 1/\tau^2$, we see that $|\nabla \cdot \mathbf{p}^{n+1} - \mathbf{f}/\lambda|^2$ is decreasing with n, unless $\boldsymbol{\eta} = 0$, which means that $\mathbf{p}^{n+1} = \mathbf{p}^n$. It is clear for $\delta t < 1/\tau^2$. For the case $\delta t = 1/\tau^2$, i.e. $|\nabla \cdot \mathbf{p}^{n+1} - \mathbf{f}/\lambda|^2 = |\nabla \cdot \mathbf{p}^n - \mathbf{f}/\lambda|^2$, we have $|\boldsymbol{\rho}|_Y = |\nabla(\nabla \cdot \mathbf{p}^n - \mathbf{f}/\lambda)|_Y$ such that either $|\nabla(\nabla \cdot \mathbf{p}^n - \mathbf{f}/\lambda)|_Y = 0$ or $|\mathbf{p}^{n+1}|_Y = 1$, which means that $\mathbf{p}^{n+1} = \mathbf{p}^n$ in both cases.

Let $m := \lim_{n \to \infty} |\nabla \cdot \mathbf{p}^n - \mathbf{f}/\lambda|_X^2$ and $\bar{\mathbf{p}}$ be the limit of a converging subsequence $\{\mathbf{p}^{n_k}\}$ of $\{\mathbf{p}^n\}$. If $\bar{\mathbf{p}}'$ is the limit of $\{\mathbf{p}^{n_k+1}\}$ then we have:

$$
\bar{\mathbf{p}}' = \frac{\bar{\mathbf{p}} + \delta t \nabla(\nabla \cdot \bar{\mathbf{p}} - \mathbf{f}/\lambda)}{1 + \delta t |\nabla(\nabla \cdot \bar{\mathbf{p}} - \mathbf{f}/\lambda)|_Y},
$$

and if the previous calculation is repeated then we see that $m = |\nabla \cdot \bar{\mathbf{p}} - \mathbf{f}/\lambda|_X = |\nabla \cdot \bar{\mathbf{p}}' - \mathbf{f}/\lambda|_X$, which implies that $\bar{\boldsymbol{\eta}} = (\bar{\mathbf{p}}' - \bar{\mathbf{p}})/\delta t = 0$ and $\bar{\mathbf{p}}' = \bar{\mathbf{p}}$. Hence,

$$
\nabla(\nabla \cdot \bar{\mathbf{p}} - \mathbf{f}/\lambda) - |\nabla(\nabla \cdot \bar{\mathbf{p}} - \mathbf{f}/\lambda)|_Y \, \bar{\mathbf{p}} = 0,
$$

which is the Euler-Lagrange equation of (7). Hence $\bar{\mathbf{p}}$ solves (7). Finally, since (7) is convex, then all sequences $\lambda \nabla \cdot \mathbf{p}^n$ converge to $\lambda \nabla \cdot \bar{\mathbf{p}}$.

It remains to show that $\delta t \le 1/8$. Let us define the discrete divergence operator applied to a vector $\mathbf{p}_i = (p_i^x, p_i^y)$, $1 \le i \le M$, M being the number of channels, at the grid point (i_x, i_y) (following [13]):

$$
(\nabla \cdot \mathbf{p}_i)_{i_x, i_y} =
\begin{cases}
(p_i^x)_{i_x, i_y} - (p_i^x)_{i_x-1, i_y} & \text{if} \quad 1 < i_x < N_x, \\
(p_i^x)_{i_x, i_y} & \text{if} \quad i_x = 1, \\
-(p_i^x)_{i_x-1, i_y} & \text{if} \quad i_x = N_x,
\end{cases}
$$

$$
+
\begin{cases}
(p_i^y)_{i_x, i_y} - (p_i^y)_{i_x, i_y-1} & \text{if} \quad 1 < i_y < N_y, \\
(p_i^y)_{i_x, i_y} & \text{if} \quad i_y = 1, \\
-(p_i^y)_{i_x, i_y-1} & \text{if} \quad i_y = N_y,
\end{cases}
$$

Then, the norm of the divergence operator is defined by $\tau^2 := \sup_{|\mathbf{p}| \leq 1} |\nabla \cdot \mathbf{p}|^2$. The discrete square norm of $\mathbf{p}$ is

$$|\mathbf{p}|^2 = \sum_{1 \leq i \leq M} \sum_{1 \leq i_x \leq Nx} \sum_{1 \leq i_y \leq Ny} (p_i^x)_{i_x,i_y}^2 + (p_i^y)_{i_x,i_y}^2.$$

We have:

$$|\nabla \cdot \mathbf{p}|^2$$
$$= \sum_{1 \leq i \leq M} \sum_{1 \leq i_x \leq Nx} \sum_{1 \leq i_y \leq Ny} \left((p_i^x)_{i_x,i_y} - (p_i^x)_{i_x-1,i_y} + (p_i^y)_{i_x,i_y} - (p_i^y)_{i_x,i_y-1} \right)^2$$
$$\leq 4 \sum_{1 \leq i \leq N} \sum_{1 \leq i_x \leq Nx} \sum_{1 \leq i_y \leq Ny} (p_i^x)_{i_x,i_y}^2 + (p_i^x)_{i_x-1,i_y}^2 + (p_i^y)_{i_x,i_y}^2 + (p_i^y)_{i_x,i_y-1}^2$$
$$\leq 8 \, |\mathbf{p}|^2,$$

using the zero boundary condition $\mathbf{p}_{0,i_y} = \mathbf{p}_{N_x,i_y} = \mathbf{p}_{i_x,0} = \mathbf{p}_{i_x,N_y} = 0$, $\forall (i_x, i_y)$, which implies that $\tau^2 = 8$ and so $\delta t = 1/8$. We notice that the temporal bound does not depend on the number of channels. $\square$

2.4. **Convex analysis and projection point of view.** The VROF minimization algorithm introduced in the previous section can also be deduced from convex analysis. Indeed, the projection algorithm of Chambolle [13] for gray-scale images can be easily extended to the vectorial case.

Theorem 2.9. *(VROF Minimization Through Projection) Let* $\mathbf{f} \in L^\infty(\Omega; \mathbb{R}^M)$, $\mathbf{u} : \Omega \to \mathbb{R}^M$, $\lambda > 0$, *then the minimizer of*

$$(13) \qquad ||\mathbf{u}||_{BV(\Omega;\mathbb{R}^M)} + \frac{1}{2\lambda}||\mathbf{f} - \mathbf{u}||_{L^2(\Omega;\mathbb{R}^M)}^2,$$

is given by

$$(14) \qquad \mathbf{u} = (1 - \Pi_{\lambda K_{BV(\Omega;\mathbb{R}^M)}})\mathbf{f},$$

where $\Pi_{\lambda K_{BV(\Omega;\mathbb{R}^M)}}$ *is the orthogonal projection operator onto the closed convex set associated to* $||.||_{BV(\Omega;\mathbb{R}^M)}$ *and defined as follows:*

$$K_{BV(\Omega;\mathbb{R}^M)} = \left\{ \nabla \cdot \mathbf{p} \in L^2(\Omega; \mathbb{R}^M), \ \forall \mathbf{p} \in L^2(\Omega; \mathbb{R}^{M \times N}) : |\mathbf{p}| \leq 1 \right\}.$$

Proof. Let us call $J_{BV}(\mathbf{u}) = ||\mathbf{u}||_{BV(\Omega;\mathbb{R}^M)}$. The Euler-Lagrange of (13) is equal to:

$$(15) \qquad \partial J_{BV}(\mathbf{u}) + \frac{\mathbf{u} - \mathbf{f}}{\lambda} \ni 0,$$

where ∂J_{BV} is called the sub-differential of J_{BV}. The sub-differential $\partial J(\mathbf{u})$ of a functional J at $\mathbf{u} \in X$ is defined by $\mathbf{z} \in \partial J(\mathbf{u}) \Leftrightarrow J(\mathbf{v}) \geq J(\mathbf{u}) + <\mathbf{z}, \mathbf{v} - \mathbf{u}>_X$, $\forall \mathbf{v} \in X$. If $J(\mathbf{u}) < \infty$ and J is differentiable at $\mathbf{u}$, then $\partial J(\mathbf{u}) = \{\nabla J(\mathbf{u})\}$. (15) can be re-written as $\frac{\mathbf{f}-\mathbf{u}}{\lambda} \in \partial J_{BV}(\mathbf{u})$, which, according to convex analysis [22], is equivalent to

$$(16) \qquad \mathbf{u} \in \partial J_{BV}^\star(\frac{\mathbf{f} - \mathbf{u}}{\lambda}),$$

where $J_{BV}^\star$ is the Legendre-Fenchel (LF) transform (or conjugate function) of J_{BV}. In convex analysis [22], the LF of a function $J \neq \infty$ convex, one-homogeneous

$(J(\lambda \mathbf{u}) = \lambda J(\mathbf{u}),\ \lambda > 0,\ \forall \mathbf{u} \in X)$ and lower semi-continuous, defined as $J^\star(\mathbf{v}) = \sup\limits_{u} < \mathbf{u}, \mathbf{v} >_X -J(\mathbf{u})$, is the characteristic function of a closed convex set K_J:

(17)
$$J^\star(\mathbf{v}) = \chi_{K_J}(\mathbf{v}) = \begin{cases} 0 & \text{if } \mathbf{v} \in K_J, \\ +\infty & \text{otherwise} \end{cases}$$

where K_J is defined as:

(18)
$$K_J = \big\{ \mathbf{s} \in X |\ \forall \mathbf{x} \in X,\ < \mathbf{s}, \mathbf{x} >_X \leq J(\mathbf{x}) \big\}.$$

Writing (16) as $\frac{\mathbf{f}}{\lambda} \in \mathbf{z} + \frac{1}{\lambda}\partial J^\star_{BV}(\mathbf{z})$, defining $\mathbf{z} = \frac{\mathbf{f}-\mathbf{u}}{\lambda}$. We get that $\mathbf{z}$ is the minimizer of $\frac{1}{2}||\mathbf{z} - \frac{\mathbf{f}}{\lambda}||^2 + \frac{1}{\lambda}J^\star_{BV}(\mathbf{z})$. Since $J^\star_{BV}$ is given by (17), we deduce that $\mathbf{z}$ is the orthogonal projection of $\mathbf{f}/\lambda$ on K_{BV}. $\Pi_{K_{BV}}$ is the projection operator onto K_{BV}. We have $\mathbf{z} = \Pi_{\lambda K_{BV}}(\frac{\mathbf{f}}{\lambda}) = \frac{\mathbf{f}-\mathbf{u}}{\lambda}$ and finally $\mathbf{u} = (1 - \Pi_{\lambda K_{BV}})\mathbf{f}$.

The closed convex set K_{BV} can be defined according to (18):

$$K_{BV} = \big\{ \mathbf{s} \in L^2(\Omega; \mathbb{R}^M) |\ \forall \mathbf{x} \in L^2(\Omega; \mathbb{R}^M),\ < \mathbf{s}, \mathbf{x} >_{L^2(\Omega;\mathbb{R}^M)} \leq ||\mathbf{x}||_{BV(\Omega;\mathbb{R}^M)} =$$
$$\sup\limits_{|\mathbf{p}|\leq 1} < \mathbf{p}, \nabla \mathbf{x} >_{L^2(\Omega;\mathbb{R}^M)} = \sup\limits_{|\mathbf{p}|\leq 1} < \nabla \cdot \mathbf{p}, \mathbf{x} >_{L^2(\Omega;\mathbb{R}^M)} \big\},$$

which naturally gives:

$$K_{BV} = \big\{ \nabla \cdot \mathbf{p} \in L^2(\Omega;\mathbb{R}^M),\ \forall \mathbf{p} \in L^2(\Omega;\mathbb{R}^M) \times L^2(\Omega;\mathbb{R}^M) : |\mathbf{p}| \leq 1 \big\}.$$

Finally, since (6) gives $\mathbf{u} = \mathbf{f} - \lambda \nabla \cdot \mathbf{p}$, then the projection operator in (14) is determined by the iterative scheme $\lambda \nabla \cdot \mathbf{p}_n \to \lambda \nabla \cdot \bar{\mathbf{p}} = \Pi_{\lambda K_{BV}}(\mathbf{f})$, introduced in the previous section.   □

2.5. **Code and results.** The iterative scheme (12) is straightforward to implement[1]. We already defined in the previous section the discrete divergence operator. It remains to define the discrete gradient operator applied to a function v at the point (i_x, i_y) (following [13]):

$$(\nabla v)_{i_x, i_y} = \big((\nabla v)^x_{i_x, i_y}, (\nabla v)^y_{i_x, i_y} \big)$$

with

$$(\nabla v)^x_{i_x, i_y} = \begin{cases} v_{i_x+1, i_y} - v_{i_x, i_y} & \text{if } i_x < N_x, \\ 0 & \text{if } i_x = N_x, \end{cases}$$

$$(\nabla v)^y_{i_x, i_y} = \begin{cases} v_{i_x, i_y+1} - v_{i_x, i_y} & \text{if } i_y < N_y, \\ 0 & \text{if } i_y = N_y. \end{cases}$$

The iterative process (12) is stopped if $\max(|\mathbf{u}^{n+1} - \mathbf{u}^n|) \leq r$, where r is a given residue and we recall that $\mathbf{u}^{n+1} = \mathbf{f} - \lambda \nabla \cdot \mathbf{p}^{n+1}$. In all experiments, initial values are chosen to be $\mathbf{p} = 0$, the residual $r = 10^{-4}$.

Before testing the vectorial denoising ROF model, we come back to (12) where it is interesting to notice that a coupling term between the M channels appears, precisely $\sqrt{\sum_{j=1}^N |\nabla(\nabla \cdot \mathbf{p}^n_j - f_j/\lambda)|^2}$. Blomgren and Chan have already noticed in [8] the existence of a coupling term in the minimization equation of the vectorial ROF model. They observed that the coupling improves the denoising process compared with a direct application of the scalar ROF model channel-by-channel. Figure 1

----

[1]A Matlab version is available at http://www.math.ucla.edu/~xbresson

gives an illustrative example of comparison between the vectorial ROF model and the scalar ROF model channel-by-channel. Globally speaking the coupling term helps to better restore parts in images where intensities are weak.

As a summary of the first part of this paper, the standard dual definition of the vectorial TV norm, introduced in (1), will be our basic tool to develop different color image processing models such as image denoising, vector field denoising, image decomposition, image deblurring and image inpainting. This definition of the vectorial TV norm uses many advantages. First of all, there is no need to introduce a regularization parameter of the TV norm as it was done in [8, 4, 16] and many other papers to carry out the minimization process. Indeed, the vectorial TV norm is usually not directly minimized but the regularized version $\int_\Omega \sqrt{|\nabla \mathbf{u}|^2 + \epsilon} \, d\mathbf{x}$, ϵ being a very small parameter to be faithful to the original norm and useful to avoid numerical instabilities in the minimization flow. Secondly, the minimization process based on the dual approach is much faster than in a standard regularized approach [8] (explicit gradient descent flow). Indeed, the direct consequence of the regularization parameter ϵ is the obligation to use a very small temporal step to ensure a correct and stable minimization process. Thus the minimization process is slow because a large number of iterations is necessary to reach the steady state minimizing solution. Figure 3 presents a comparison between our dual-based ROF model and the standard ϵ-based ROF model used in [8]. Our model converges fast to the denoising result whereas the standard model converges slowly because of the regularization parameter. Besides, the quality of the denoising process is slightly better with our model, still because of the ϵ parameter. Finally, another advantage of our approach is to provide a numerical minimization scheme easy to code.

2.6. **Other vectorial image Denoising models.**

2.6.1. *Sapiro-Ringach and Blomgren-Chan's Models.* In [33], Sapiro and Ringach propose an anisotropic diffusion model for vectorial images based on Riemannian geometry. From a Riemannian point of view, the vectorial image $\mathbf{u} : (x, y) \rightarrow (u_1(x, y), ..., u_M(x, y))$ can be seen as a parametric 2-D manifold embedded in a M-D Euclidean space. The (square) distance between two close points on the 2-D Riemannian manifold is given by $ds^2 := \sum_{\mu,\nu} g_{\mu\nu} d\mu d\nu$, $\mu, \nu = \{x, y\}$, and $g_{\mu\nu} := < \partial_\mu \mathbf{u}, \partial_\nu \mathbf{u} >$, $\partial_\mu \mathbf{u} := \partial \mathbf{u} / \partial \mu$, is the first fundamental form or metric tensor [28]. The metric tensor $g_{\mu\nu}$ is considered as a "vectorial edge" indicator function, i.e. a detector of discontinuities for vectorial images. More precisely, the eigenvectors of the image metric $g_{\mu\nu}$ determine the directions of maximal and minimal change at a give point on the manifold-image, and their associated eigenvalues give their rate of change. Eigenvalues of $g_{\mu\nu}$ are provided by:

$$\lambda_\pm = \frac{1}{2}\Big(g_{xx} + g_{yy} \pm \sqrt{(g_{xx} - g_{yy})^2 + 4g_{xy}^2} \Big)$$

and the eigenvectors are $(\cos\theta_\pm, \sin\theta_\pm)$, where $\theta_+ = \frac{1}{2}\arctan\frac{2g_{xy}}{g_{xx}-g_{yy}}$ and $\theta_- = \theta_+ + \frac{\pi}{2}$.

Based on the previous analysis of the image metric $g_{\mu\nu}$, Sapiro and Ringach propose to define the anisotropic diffusion for vectorial images by smoothing in the direction of minimal change, θ_-, i.e. parallel to the "vectorial" edge, using the flow: $\frac{\partial \mathbf{u}}{\partial t} = g(\lambda_+, \lambda_-)\frac{\partial^2 \mathbf{u}}{\partial \theta_-^2}$. This approach is consistent with the scalar case presented in

[12]. Indeed, in the case of gray-scale images, i.e. $M = 1$, it is easy to show that $(\cos\theta_-, \sin\theta_-) = \frac{\nabla u}{|\nabla u|}$ which implies that $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial \theta_-^2} = \nabla \cdot (\frac{\nabla u}{|\nabla u|})$ and the TV norm is equal to $TV(u) = \int_\Omega \sqrt{\lambda_+} d\mathbf{x}$ since $\lambda_+ = |\nabla u|^2$ and $\lambda_- = 0$. From the latter scalar approach, Sapiro and Ringach propose in [33] the following vectorial TV norm:

$$(19) \qquad TV_{SR}(\mathbf{u}) = \int_\Omega f(\lambda_+, \lambda_-) \, d\mathbf{x}.$$

However, they do not recommend a specific f. One natural choice, as proposed by Blomgren and Chan in [8], is $f := \sqrt{\lambda_+ + \lambda_-}$ because it is equivalent to the TV norm for the scalar case and it is a non-decreasing function of $\lambda_\pm$.

We notice that (19) while considering $f = \sqrt{\lambda_+ + \lambda_-}$ is strictly equal to the definition of the vectorial TV norm used in this paper and presented in Equation (3) for smooth functions $\mathbf{u} \in C^1(\Omega; \mathbb{R}^M)$. Indeed, $\sqrt{\lambda_+ + \lambda_-} = \sqrt{g_{xx} + g_{yy}}$, $g_{xx} = \sum_{i=1}^M (\partial_x u_i)^2$ and $g_{yy} = \sum_{i=1}^M (\partial_y u_i)^2$ implies that

$$(20) \quad TV_{SR}(\mathbf{u}) = \int_\Omega \sqrt{\lambda_+ + \lambda_-} d\mathbf{x} = \int_\Omega \sqrt{\sum_{i=1}^M |\nabla u_i|^2} d\mathbf{x} = \|\mathbf{u}\|_{BV(\Omega; \mathbb{R}^M)}.$$

Finally, Blomgren and Chan noticed that the norm (20) has the tendency to produce a color smearing contrary to their definition of vectorial TV given by $TV_{BC}(\mathbf{u}) = \sqrt{\sum_{i=1}^M \left(\int_\Omega |\nabla u_i| d\mathbf{x} \right)^2}$. Since our model corresponds to the norm (20) for smooth functions $\mathbf{u}$, our algorithm still gets the color smearing. However, despite of this drawback, as we have already said, the norm (20) offers many important advantages. We remind that no regularization of the TV norm is done. Besides, the number of iterations to converge to the denoised vectorial image is much less important in our approach than in [33, 8] as shown on Figure 3 and our numerical scheme is straightforward to implement. Finally, a dual formulation of the vectorial TV norm provides the basic element to prove the existence of a denoising solution in Section 2.2, which is difficult to prove with TV_{BC} in [8].

2.6.2. *Sochen-Kimmel-Malladi's Model.* In [35], Sochen, Kimmel and Malladi introduce another Riemannian framework to smooth vectorial images such as color or texture images. In their approach, they represent the given vectorial image as a parametric 2-D manifold embedded in a $(M+2)$-D space, M being the dimension of the vectorial image. The core of their framework is the Polyakov functional, which measures the weight of a mapping $X : \Sigma \to \mathcal{M}$ between an embedded manifold (the image manifold) Σ and an embedding manifold $\mathcal{M}$ as follows:

$$(21) \qquad P(X, \Sigma, \mathcal{M}) = \int_\Sigma d\Sigma \, \sqrt{g} g^{\mu\nu} \partial_\mu X^i \partial_\nu X^j h_{ij},$$

where $g_{\mu\nu}$ is the first fundamental form of the manifold Σ, $d\Sigma$ is the integration element w.r.t. the local coordinates on Σ, h_{ij} is the metric tensor of the embedding space $\mathcal{M}$, $g^{\mu\nu}$ is the inverse metric of $g_{\mu\nu}$, g is the determinant of $g_{\mu\nu}$, $\mu, \nu = x, y$ and $i, j = 1, ..., M$ in our paper and $\partial_\mu X^i := \partial X^i / \partial \mu$. Finally, the Einstein summation convention is used in (21), which means that when identical indices appear one up and one down, they are summed over. The Polyakov functional is

related with harmonic maps which are geodesics or minimal surfaces for curves and surfaces. Indeed, if the metric tensor of the embedded manifold Σ is chosen to be the induced metric tensor, $g_{\mu\nu} := \partial_\mu X^i \partial_\nu X^j h_{ij}$, by pullback procedure, then X which minimizes the Polyakov action are called harmonic maps and the Polyakov functional is reduced to the Euler functional:

$$S = \int_\Sigma d\Sigma \ \sqrt{g},$$

which describes the length/(hyper-)area of a curve/(hyper-)surface Σ and $\sqrt{g}$ being the invariant-area element on Σ.

In the case of vectorial images, we have $X : (x, y) \to (x, y, u_1(x, y), ..., u_M(x, y))$ and the components of the metric tensor are thus: $g_{xx} = 1 + \sum_{i=1}^{M} (\partial_x u_i)^2$, $g_{yy} = 1 + \sum_{i=1}^{M} (\partial_y u_i)^2$ and $g_{xy} = \sum_{i=1}^{M} \partial_x u_i \partial_y u_i$. This implies the denoising functional for vectorial images [35]:

$$(22) \qquad S = \int_\Omega \sqrt{1 + \sum_{i=1}^{M} |\nabla u_i|^2 + \frac{1}{2} \sum_{i,j=1}^{M} < \nabla u_i, \nabla u_j >^2} \ d\mathbf{x},$$

where the term $\sum_{i,j=1}^{M} < \nabla u_i, \nabla u_j >^2$ is a cross-correlation term of orientation between different channels that directs different channels to align together. Minimizing (22) w.r.t. u_i gives the Beltrami flow: $\frac{\partial u_i}{\partial t} = \frac{1}{\sqrt{g}} \partial_\mu(\sqrt{g} g^{\mu\nu} \partial_\nu u_i) =: \Delta_g u_i$, where Δ_g is the Laplace-Beltrami operator.

There exist some connections between the previous model and the dual TV norm. If we define a new vectorial image as $\mathbf{u} : (x, y) \to (\frac{1}{\sqrt{2}}x, \frac{1}{\sqrt{2}}y, u_1(x, y), ..., u_M(x, y))$, then the dual vector TV norm is equal to:

$$(23) \qquad \|\mathbf{u}\|_{BV(\Omega;\mathbb{R}^M)} = \int_\Omega \sqrt{1 + \sum_{i=1}^{M} |\nabla u_i|^2} \ d\mathbf{x}.$$

Functional (23) is equal to (22) when the cross-correlation term is null. The solution of the vectorial ROF model $\inf_{\mathbf{u}} \|\mathbf{u}\|_{BV(\Omega;\mathbb{R}^M)} + \frac{1}{2\lambda}\|\mathbf{f} - \mathbf{u}\|^2_{L^2(\Omega;\mathbb{R}^M)}$, where $\mathbf{f} = (\frac{1}{\sqrt{2}}x, \frac{1}{\sqrt{2}}y, f_1, ..., f_M)$ as the given image, is given by:

$$(24) \qquad u_i = f_i - \lambda \nabla \cdot \mathbf{p}_i, \quad 1 \leq i \leq M,$$

and $\mathbf{p}_i$ is the steady state solution of the following flow:

$$(25) \qquad \mathbf{p}_i^{n+1} = \frac{\mathbf{p}_i^n + \delta t \nabla (\lambda \nabla \cdot \mathbf{p}_i^n - f_i)}{1 + \delta t \sqrt{1 + \sum_{j=1}^{M} |\nabla(\lambda \nabla \cdot \mathbf{p}_j^n - f_j)|^2}}, \quad 1 \leq i \leq M.$$

An interesting application of (24) and (25) is for gray-scale images. Indeed, Functionals (22) and (23) are equal for gray-scale images, $S(u) = \|u\|_{BV(\Omega)} = \int_\Omega \sqrt{1 + |\nabla u|^2} d\mathbf{x}$, since the cross-correlation term is null. This means that the Beltrami flow can be "dualized" in the case of gray-scale images as proposed by Chambolle in [13]. In other words, the minimizing solution of

$$(26) \qquad \int_\Omega \sqrt{1 + |\nabla u|^2} d\mathbf{x} + \frac{1}{2\lambda}\|f - u\|^2_{L^2(\Omega)},$$

where $u : \Omega \to \mathbb{R}$, is given either in the standard Euler-Lagrange approach by:

$$(27) \qquad \frac{\partial u}{\partial t} = \underbrace{\frac{(1+u_y^2)u_{xx} - 2u_x u_y u_{xy} + (1+u_x^2)u_{yy}}{(1+u_x^2+u_y^2)^2}}_{\Delta_g u} + \frac{1}{2\lambda}(u-f).$$

or in the dual approach by:

$$(28) \qquad\qquad\qquad u = f - \lambda \nabla \cdot \mathbf{p},$$

where $\mathbf{p}$ is the steady state of:

$$\mathbf{p}^{n+1} = \frac{\mathbf{p}^n + \delta t \nabla(\lambda \nabla \cdot \mathbf{p}^n - f)}{1 + \delta t \sqrt{1 + |\nabla(\lambda \nabla \cdot \mathbf{p}^n - f)|^2}}.$$

We have noticed that the dual approach (28) is faster than the Euler-Lagrange (EL) approach (27). We mean that the computation of the minimum of (26) is obtained faster with solution (28) than solution (27), even if we consider an explicit steepest gradient descent scheme for both approaches. We think that the explicit computation of the Lagrange multiplier in the dual approach helps to speed up the minimization speed. We have already observed that the dual approach is faster than the EL approach in the ROF denoising model. Indeed, Figure 3 presents a comparison between the dual and the EL approaches, where the dual is clearly faster. Finally, we have also noticed that the semi-implicit scheme does not really speed up the minimization in the dual approach. Similar minimization speeds are obtained with the implementation of the explicit gradient descent scheme.

3. **Part II: Extension to standard color image processing.** In the first part of this paper, we have introduced a regularization algorithm for color/vectorial images based on the VTV norm. In the second part of this paper, we extend this L^1 vectorial regularization algorithm to standard color image processing problems. We will see that the extension works well for many color problems such as the color inverse scale space, the color denoising with the chromaticity-brightness color representation, the color wavelet shrinkage, the color inpainting, the color decomposition, the color deblurring and finally the color denoising on manifolds. Generally speaking, the proposed VTV minimization algorithm can be used in problems that required vector field (color, other vector features) L^1 regularization while preserving discontinuities.

3.1. **Color inverse scale space.** In this section, we improve the VROF model, introduced in Section 2.2, with the iterative regularization process proposed by Osher, Burger, Goldfarb, Xu and Yin in [31]. This process, also called inverse scale space, enhances the denoising task defined by the ROF model by better preserving the contrast. The proposed iterative regularization process is based on geometry. The basic idea is to match not only the intensity values between the denoising image and the noisy one but also the normals of level sets in each image while applying the TV norm which preserves the location of edges while smoothing out the noise. In this section, we improve the VROF model by extending the iterative regularization scheme [31] to color images.

We consider $\mathbf{u} \in C^1(\Omega; \mathbb{R}^M)$ to explain the model. Given the noisy vectorial image $\mathbf{f} \in L^\infty(\Omega, \mathbb{R}^M)$, an "approximation of normals of level sets" can be computed by carrying out the minimization model:

$$\mathbf{u}_1 = \arg\inf_{\mathbf{u}} \|\mathbf{u}\|_{BV(\Omega;\mathbb{R}^M)} + \frac{1}{2\lambda} \|\mathbf{f} - \mathbf{u}\|^2_{L^2(\Omega;\mathbb{R}^M)},$$

where $\mathbf{n}_1 := \frac{\nabla \mathbf{u}_1}{|\nabla \mathbf{u}_1|}$ is called the approximation of "normals of level sets". The Euler-Lagrange equation of the previous equation is equal to:

$$(29) \qquad \sum_{i=1}^{M} \nabla \cdot \frac{\nabla u_{1i}}{|\nabla \mathbf{u}_1|} + \frac{1}{\lambda} v_i = 0,$$

where $v_i := f_i - u_i, \forall i$. We now introduce $\mathbf{n}_1$ in the following minimization model:

$$\mathbf{u}_2 = \arg\inf_{\mathbf{u}} \|\mathbf{u}\|_{BV(\Omega;\mathbb{R}^M)} - < \mathbf{n}_1, \nabla \mathbf{u} >_{L^2(\Omega;\mathbb{R}^M)} + \frac{1}{2\lambda} \|\mathbf{f} - \mathbf{u}\|^2_{L^2(\Omega;\mathbb{R}^M)}.$$

In the previous equation, we have:

$$- < \mathbf{n}_1, \nabla \mathbf{u} >_{L^2(\Omega;\mathbb{R}^M)} = \int_{\Omega} - \sum_{i=1}^{M} < n_{1i}, \nabla u_i >$$

$$= \int_{\Omega} \sum_{i=1}^{M} u_i \nabla \cdot n_{1i} = \int_{\Omega} \sum_{i=1}^{M} u_i \nabla \cdot \frac{\nabla u_{1i}}{|\nabla \mathbf{u}_1|}.$$

Using (29), we obtain:

$$\int_{\Omega} \sum_{i=1}^{M} u_i \nabla \cdot \frac{\nabla u_{1i}}{|\nabla \mathbf{u}_1|} = \int_{\Omega} \sum_{i=1}^{M} u_i \left(-\frac{1}{\lambda} v_{1i} \right),$$

which gives:

$$\mathbf{u}_2 = \arg\inf_{\mathbf{u}} \quad \|\mathbf{u}\|_{BV(\Omega;\mathbb{R}^M)} + \frac{1}{2\lambda} \int_{\Omega} \sum_{i=1}^{M} \left(-2 u_i v_{1i} + (f_i - u_i)^2 \right)$$

$$\arg\inf_{\mathbf{u}} \quad \|\mathbf{u}\|_{BV(\Omega;\mathbb{R}^M)} + \frac{1}{2\lambda} \int_{\Omega} \sum_{i=1}^{M} \left((f_i + v_{1i} - u_i)^2 + \eta \right),$$

where η is independent of $\mathbf{u}$ and therefore useless in the minimization process. Finally, we end up with the modified vectorial ROF model:

$$\mathbf{u}_2 = \arg\inf_{\mathbf{u}} \|\mathbf{u}\|_{BV(\Omega;\mathbb{R}^M)} + \frac{1}{2\lambda} \|\mathbf{f} + \mathbf{v}_1 - \mathbf{u}\|^2_{L^2(\Omega;\mathbb{R}^M)},$$

where $\mathbf{v}_1 := \mathbf{f} - \mathbf{u}_1$. Hence, the implementing iterative procedure for the vectorial image case is as follows:

1. Initialize: $\mathbf{u}_0 = 0$ and $\mathbf{v}_0 = 0$.
2. For $k = 0, 1, 2, \ldots$: compute $\mathbf{u}_{k+1}$ as a minimizer of the modified vectorial ROF model:

$$\mathbf{u}_{k+1} = \arg\min_{\mathbf{u}} \|\mathbf{u}\|_{BV(\Omega;\mathbb{R}^M)} + \frac{1}{2\lambda} \|\mathbf{f} + \mathbf{v}_k - \mathbf{u}\|^2_{L^2(\Omega;\mathbb{R}^M)},$$

and update $\mathbf{v}_{k+1} := \mathbf{v}_k + \mathbf{f} - \mathbf{u}_{k+1}$.

Figure 4 presents an application of the proposed iterative regularization process, which better preserve the color contrast.

3.2. **Color Denoising with the Chromaticity-Brightness representation.**
In Section 2.2, the color denoising model has been formulated as a vectorial model
in the RGB representation of color images. In [16], authors showed that the
Chromaticity-Brightness model or the Hue-Saturation-Value (HSV) enable to bet-
ter restore color images because these models are closer to human perception. In
this section, we propose to perform the color image denoising in the CB repre-
sentation of color images. In the RGB color model, the given image is given by
$\mathbf{f} := (f_1, ..., f_M) : \Omega \to \mathbb{R}^M$, $f_i \geq 0$, $1 \leq i \leq M$. In the CB color model, the image
$\mathbf{f}$ is separated into the brigthness component $B_{\mathbf{f}} := |\mathbf{f}| : \Omega \to \mathbb{R}$ and the chromatic-
ity component $\mathbf{C}_{\mathbf{f}} := \mathbf{f}/|\mathbf{f}| = \mathbf{f}/B_{\mathbf{f}} : \Omega \to \mathbb{S}^2$. Thus, the brightness component
is a gray-scale image and the chromaticity component stores the color information
which takes values on the unit sphere. Following [27], the previous constraint on
the sphere is introduced as a penalization term in the variational model:

$$(30)\inf_{\mathbf{C}} \left\{ \|\mathbf{C}\|_{BV(\Omega;\mathbb{R}^M)} + \frac{1}{2\lambda}\|\mathbf{C} - \mathbf{C_f}\|_{L^2(\Omega;\mathbb{R}^M)}^2 + \frac{1}{2\alpha}\||\mathbf{C}|^2 - 1\|_{L^2(\Omega;\mathbb{R}^M)}^2 \right\},$$

which denoises the chromaticity component. (30) is regularized using the Ginzburg-
Landau functional [11] as follows:

$$\inf_{\mathbf{C},\mathbf{D}} \left\{ \|\mathbf{C}\|_{BV(\Omega;\mathbb{R}^M)} + \frac{1}{2\lambda}\|\mathbf{C} - \mathbf{C_f}\|_{L^2(\Omega;\mathbb{R}^M)}^2 + \frac{1}{2\alpha}\||\mathbf{D}|^2 - 1\|_{L^2(\Omega;\mathbb{R}^M)}^2 + \right.$$
$$\left. \frac{1}{2\beta}\|\mathbf{D} - \mathbf{C}\|_{L^2(\Omega;\mathbb{R}^M)}^2 \right\},$$

The Ginzburg-Landau functional ($\||\mathbf{D}|^2 - 1\|_2^2$) is introduced in order to use the fast
minimization algorithm proposed in Section 2.3. Indeed, we have two minimization
problems. The first one is the VROF:

$$\inf_{\mathbf{C}} \|\mathbf{C}\|_{BV(\Omega;\mathbb{R}^M)} + \frac{\beta + \lambda}{2\lambda\beta}\|\mathbf{C} - (\frac{\beta}{\lambda + \beta}\mathbf{C_f} + \frac{\lambda}{\lambda + \beta}\mathbf{D})\|_{L^2(\Omega;\mathbb{R}^M)}^2,$$

and the second one is defined by:

$$(31) \qquad \inf_{\mathbf{D}} \frac{1}{2\alpha}\||\mathbf{D}|^2 - 1\|_{L^2(\Omega;\mathbb{R}^M)}^2 + \frac{1}{2\beta}\|\mathbf{D} - \mathbf{C}\|_{L^2(\Omega;\mathbb{R}^M)}^2,$$

whose the Euler-Lagrange equation is given by $\mathbf{D}(\frac{2\beta}{\alpha}(|\mathbf{D}|^2 - 1) + 1) + \mathbf{C} = 0$. We
use the following semi-implicit iterative scheme to minimize (31):

$$\mathbf{D}^{n+1} = \frac{\mathbf{D}^n + \delta t \mathbf{C}}{1 + \delta t(\frac{2\beta}{\alpha}(|\mathbf{D}^n|^2 - 1) + 1)}.$$

The brightness component is a gray-scale image that can be denoised with the
model [13]. Finally, the denoised image is given by $\mathbf{u} = \mathbf{C}/B$, where $\mathbf{C}$ and B
are respectively the denoised chromaticity component and the denoised brightness
component.

Figure 5 presents the denoised color image with the CB color representation.
The SNR is 14.57 with the RGB color representation and 14.72 with the CB repre-
sentation.

3.3. **Color image inpainting.** Image inpainting consists in recovering/interpolating an image in regions where the original information is missing, see [6, 17]. Given the image domain Ω, a color image $\mathbf{f} \in L^\infty(\Omega, \mathbb{R}^M)$, a region $\Omega_I \subset \Omega$ where the image will be inpainted, the variational color image inpainting model is as follows:

$$\begin{cases} \inf_{\mathbf{u}} \|\mathbf{u}\|_{BV(\Omega_I;\mathbb{R}^M)} \\ \mathbf{u}_{|\partial \Omega_I} = \mathbf{f} \end{cases}.$$

We propose to solve the previous minimization problem with boundary condition with the following convex variational model:

$$\begin{cases} \inf_{\mathbf{u},\mathbf{v}} \|\mathbf{u}\|_{BV(\Omega;\mathbb{R}^M)} + \frac{1}{2\lambda}\|\mathbf{u} - \mathbf{v}\|^2_{L^2(\Omega;\mathbb{R}^M)} \\ \mathbf{v}_{|\Omega\setminus\Omega_I} = \mathbf{f} \end{cases}.$$

The minimization problem w.r.t. $\mathbf{u}$ is the vectorial ROF model and the minimization w.r.t. $\mathbf{v}$ gives $\mathbf{v}_{|\Omega_I} = \mathbf{u}$ and $\mathbf{v}_{|\Omega\setminus\Omega_I} = \mathbf{f}$.

Figure 8 presents a color image inpainting result.

3.4. **Color wavelet shrinkage.** In this paper, we have considered for vectorial images the space of regularization as the BV space. However, other spaces of smoothness can obviously be used. A popular regularization space is $L^2(\Omega;\mathbb{R}^M)$. In this case, the minimizer (not really useful for denoising) of $\|\mathbf{u}\|^2_{L^2(\Omega;\mathbb{R}^M)} + \frac{1}{2\lambda}\|\mathbf{f} - \mathbf{u}\|^2_{L^2(\Omega;\mathbb{R}^M)}$ is given by the Tikhonov method [38]. In this section, we consider the Besov space for color images, which will relate variational model to wavelet shrinkage through a denoising process as shown in [14].

Proposition 1. *(Color Wavelet Shrinkage) Let $\mathbf{f} \in L^\infty(\Omega;\mathbb{R}^M)$, $\mathbf{u} : \Omega \to \mathbb{R}^M$, $\lambda > 0$, then the minimizer of*

$$(32) \qquad \|\mathbf{u}\|^p_{B^\alpha_p(L_p(\Omega;\mathbb{R}^M))} + \frac{1}{2\lambda}\|\mathbf{f} - \mathbf{u}\|^2_{L^2(\Omega;\mathbb{R}^M)},$$

is given by

$$(33) \qquad \mathbf{u} = (1 - \Pi_{\lambda K_{B^\alpha_p(L_p(\Omega;\mathbb{R}^M))}})\mathbf{f},$$

where $B^\alpha_p(L_p(\Omega;\mathbb{R}^M))$ is a Besov space. In the case of $p = 1, \alpha = 1$, the closed convex set $K_{B^\alpha_p}$ is given by:

$$K_{B^1_1(L_1(\Omega;\mathbb{R}^M))} = \left\{ \mathbf{s} \in L^2(\Omega;\mathbb{R}^M) | \; \forall i,j,k,\psi, \; |c^{s_i}_{j,k,\psi}| \leq 1 \right\}.$$

Proof. The proof is the straightforward extension of the paper [14]. In this paper, Chambolle, DeVore, Lee and Lucier showed relations between wavelet-based algorithms and variational models in the case of the Besov space and gray-scale images. Let us express the vectorial image $\mathbf{u} : \Omega \to \mathbb{R}^M$ in an orthogonal wavelet basis for $L^2(\Omega)$ that we denote $\psi_{j,k}$, then for $\mathbf{f} \in L^2(\Omega;\mathbb{R}^M)$, we have:

$$\mathbf{f} = \sum_{j,k,\psi} c^{\mathbf{f}}_{j,k,\psi} \psi_{j,k},$$

where $c^{\mathbf{f}}_{j,k,\psi} : \Omega \to \mathbb{R}^M$. The norm of $\mathbf{f}$ in $L^2(\Omega;\mathbb{R}^M)$ is:

$$||\mathbf{f}||_{L^2(\Omega;\mathbb{R}^M)} = \sum_{i,j,k,\psi} |c^{f_i}_{j,k,\psi}|^2,$$

where $i \in [1, M]$ and the norm of $\mathbf{f}$ in $B^\alpha_p(L_p(\Omega;\mathbb{R}^M))$ is ([14]):

$$||\mathbf{f}||_{B^\alpha_p(L_p(\Omega;\mathbb{R}^M))} = \Big(\sum_{i,j,k,\psi} |c^{f_i}_{j,k,\psi}|^p \Big)^{1/p}.$$

It can be shown that minimizing (32) is equivalent to the wavelet shrinkage algorithm of Donoho and Johnstone defined in [21]. If we consider $\mathbf{f} = \sum_{j,k,\psi} c^{\mathbf{f}}_{j,k,\psi}\psi_{j,k}$ and $\mathbf{u} = \sum_{j,k,\psi} c^{\mathbf{u}}_{j,k,\psi}\psi_{j,k}$, then (32) is equal to:

$$\sum_{i,j,k,\psi} |c^{u_i}_{j,k,\psi}|^p + \frac{1}{2\lambda} \sum_{i,j,k,\psi} |c^{f_i}_{j,k,\psi} - c^{u_i}_{j,k,\psi}|^2,$$

which implies the minimization of

$$(34) \qquad |c^{u_i}_{j,k,\psi}|^p + \frac{1}{2\lambda}|c^{f_i}_{j,k,\psi} - c^{u_i}_{j,k,\psi}|^2,$$

for each i, j, k, ψ. (34) is equivalent to minimize $E(s) = |s|^p + \frac{1}{2\lambda}|s - t|^2$. In the special case of $p = 1$, the exact minimizer is given by $s = \frac{t}{|t|}\max(|t|-\lambda, 0)$. Thus the wavelet coefficients $c^{u_i}_{j,k,\psi}$ are shrinked toward zero by an amount of λ to obtain the exact minimizer. This corresponds to the wavelet shrinkage algorithm of Donoho and Johnstone. Thus, wavelet shrinkage can be obtained with the variational model (32). And the minimizer of the minimization problem (32) can be obtained by a projection algorithm, whose the closed convex set $K_{B^1_1(L_1(\Omega;\mathbb{R}^M))}$ (with $p = 1, \alpha = 1$) is defined as follows:

$$K_{B^1_1(L_1(\Omega;\mathbb{R}^M))} = \big\{ \mathbf{s} \in L^2(\Omega;\mathbb{R}^M)| \ \forall \mathbf{x} \in L^2(\Omega;\mathbb{R}^M), \ < \mathbf{s}, \mathbf{x} >_{L^2(\Omega;\mathbb{R}^M)} \leq$$
$$||\mathbf{x}||_{B^1_1(L_1(\Omega;\mathbb{R}^M))} = \sum_{i,j,k,\psi} |c^{s_i}_{j,k,\psi}| \big\},$$

and so:

$$K_{B^1_1(L_1(\Omega;\mathbb{R}^M))} = \big\{ \mathbf{s} \in L^2(\Omega;\mathbb{R}^M)| \ \forall i, j, k, \psi, \ |c^{s_i}_{j,k,\psi}| \leq 1 \big\}.$$

$\square$

3.5. Color image decomposition.

In this section, we consider the color image decompsition task introduced by Aujol and Kang introduced in [4] (another relevant work is [39]). Image decomposition aims at splitting an image into a smooth/geometrical component and a textural/oscillating component. This decomposition can be used to enhance image processing tasks such as the denoising [3], segmentation [10] or inpainting [7]. The model of Aujol and Kang considers a regularized version of the vectorial TV norm, $\int_\Omega \sqrt{|\nabla \mathbf{u}|^2 + \epsilon} \, d\mathbf{x}, \ \epsilon > 0$, which induces a slow decomposition process as we explained in Section 2.1. We will use the vectorial dual approach introduced in this paper to speed up the color image decomposition process.

Proposition 2. *(Color Image Decomposition) Let* $\mathbf{f} \in L^\infty(\Omega; \mathbb{R}^M)$, $\mathbf{u}, \mathbf{v} : \Omega \to \mathbb{R}^M$, $\lambda > 0$, *then the minimizer of*

$$(35) \qquad J_{\mathbf{u}}(\mathbf{u}) + J_{\mathbf{v}}(\mathbf{v}) + \frac{1}{2\lambda}||\mathbf{f} - (\mathbf{u}+\mathbf{v})||^2_{L^2(\Omega; \mathbb{R}^M)},$$

is given by

$$(36) \qquad \left\{ \begin{array}{rcl} \mathbf{u} & = & (1 - \Pi_{\lambda K_{\mathbf{u}}})(\mathbf{f} - \mathbf{v}), \\ \mathbf{v} & = & (1 - \Pi_{\lambda K_{\mathbf{v}}})(\mathbf{f} - \mathbf{u}), \end{array} \right.$$

where $J_{\mathbf{u}}, J_{\mathbf{v}}$ *are the functionals whose describe the smooth and textural parts of* $\mathbf{f}$*. Since we are interested in the image decomposition algorithm proposed by Aujol and Kang, then* $J_{\mathbf{u}} = ||\mathbf{u}||_{BV(\Omega; \mathbb{R}^M)}$ *and* $J_{\mathbf{v}} = ||\mathbf{v}||_{BV^\star(\Omega; \mathbb{R}^M)}$ *such that:*

$$(37) \qquad \left\{ \begin{array}{rcl} \mathbf{u} & = & (1 - \Pi_{\lambda K_{BV(\Omega;\mathbb{R}^M)}})(\mathbf{f} - \mathbf{v}), \\ \mathbf{v} & = & \Pi_{\lambda K_{BV(\Omega;\mathbb{R}^M)}}(\mathbf{f} - \mathbf{u}), \end{array} \right.$$

We observe that the norm $||.||_{BV^\star(\Omega;\mathbb{R}^M)}$ *was introduced by Aujol and Kang in* [4] *and it corresponds to the multidimensional version of the G-norm introduced by Meyer in* [30] *to capture oscillating patterns in gray-scale images.*

Proof. It is easy to prove (37) using the following result. The minimizer of $J_{BV}(\mathbf{u}) + \frac{1}{2}||\mathbf{u}-\mathbf{g}||^2$ is also the minimizer of $J_{BV^\star}(\mathbf{z}_1) + \frac{1}{2}||\mathbf{z}_1-\mathbf{g}||^2$ with $\mathbf{z}_1 = \Pi_{K_{BV}}(\mathbf{g})$. Hence, the minimizer of $J_{BV^\star}(\mathbf{u}) + \frac{1}{2\lambda}||\mathbf{u}-\mathbf{g}||^2$ is the minimizer of $J_{BV^{\star\star}}(\mathbf{z}_2) + \frac{1}{2}||\mathbf{z}_2 - \mathbf{g}||^2$ with $\mathbf{z}_2 = \Pi_{\lambda K_{BV^\star}}(\mathbf{g})$. Since we have $J_{BV^{\star\star}} = J_{BV}$, this implies that we also have $\mathbf{z}_2 = (1 - \Pi_{\lambda K_{BV}})(\mathbf{g})$. This means that $\Pi_{\lambda K_{BV^\star}} = 1 - \Pi_{\lambda K_{BV}}$, which gives (37). $\square$

Figure 6 and 7 present two color image decomposition results.

3.6. **Color image deblurring.** We consider the fidelity term like $||\mathbf{f} - A\mathbf{u}||^2$, where $A : X \to X$ is a linear self-adjoint and positive operator satisfying $||A|| < 1$. This new fidelity term allows to deal with the image deblurring problem. We remind that image deblurring aims at compensating the motion of the camera during the exposure, see [18, 25] for gray-scale TV deblurring. Given a color blurred and noisy image $\mathbf{f}$, we propose the following variational color image deblurring and denoising model:

Proposition 3. *(Color Image Deblurring) Let* $\mathbf{f} \in L^\infty(\Omega; \mathbb{R}^M)$, $\mathbf{u}, \mathbf{v} : \Omega \to \mathbb{R}^M$, $A : \mathbb{R}^M \to \mathbb{R}^M$, $||A|| < 1$, $A\mathbf{u} := (A_1 u_1, ..., A_N u_N) : \Omega \to \mathbb{R}^M$, $\lambda, \eta > 0$, *then the minimizer of*

$$(38) \qquad ||\mathbf{u}||_{BV(\Omega;\mathbb{R}^M)} + \frac{1}{2\lambda}||\mathbf{f} - A\mathbf{u}||^2_{L^2(\Omega;\mathbb{R}^M)},$$

is computed via the following regularized energy:

$$(39) \qquad ||\mathbf{u}||_{BV(\Omega;\mathbb{R}^M)} + \frac{1}{2\lambda}||\mathbf{f} - A\mathbf{v}||^2_{L^2(\Omega;\mathbb{R}^M)} + \frac{1}{2\eta}||\mathbf{u} - \mathbf{v}||^2_{L^2(\Omega;\mathbb{R}^M)},$$

Since (38) is convex, the minimizer is given by:

$$\left\{ \begin{array}{rcl} \mathbf{u} & = & (1 - \Pi_{\eta K_{BV(\Omega;\mathbb{R}^M)}})(\mathbf{v}), \\ \mathbf{v} & = & (1 + \frac{\lambda}{\eta}A^\star A)^{-1}(\mathbf{u} + \frac{\lambda}{\eta}A^\star \mathbf{f}), \end{array} \right.$$

where $A^\star$ is the L^2-conjugate of A such that $A^\star\mathbf{f} := (A_1^\star f_1, ..., A_N^\star f_N) : \Omega \to \mathbb{R}^M$.

Proof. The proof is immediate using the VROF model and computing the Euler-Lagrange of (39) w.r.t. $\mathbf{v}$. □

Figure 9 presents a color image deblurring and denoising result. The deblurring process is a *non-blind* deblurring process since the kernel A is assumed to be known.

3.7. **Generalized projection algorithm of chambolle.** Based on the work of Bect, Blanc-Feraud, Aubert and Chambolle [5], the regularization operator can be generalized to any linear operator for color images.

Proposition 4. *(Generalized Projection Algorithm of Chambolle) In the case of a regularization term defined as $J(\mathbf{u}) = ||Q\mathbf{u}||_1$, where Q is a linear operator $Q : X \to Y$, the projection $\Pi_{\lambda K}$ onto the convex set closed set λK can be numerically computed by a fixed point method. We consider vectors $\mathbf{p} \in \{\mathbf{p} \in Y : |\mathbf{p}| \le 1\}$ such that:*

$$J(\mathbf{u}) = ||Q\mathbf{u}||_1 = \sup_{|\mathbf{p}|\le 1} \int < Q\mathbf{u}, \mathbf{p} > .$$

The projection $\Pi_{\lambda K}$ in the case of $||Q\mathbf{u}||_1 + \frac{1}{2\lambda}||\mathbf{f} - \mathbf{u}||^2$ is given by the iterative sequence:

$$(40) \qquad \mathbf{p}_{n+1} = \frac{\mathbf{p}_n + \tau Q(Q^\star\mathbf{p}_n - \mathbf{f}/\lambda)}{1 + \tau|Q(Q^\star\mathbf{p}_n - \mathbf{f}/\lambda)|}.$$

The previous iterative algorithm converges assuming that the time step $\tau \le \frac{1}{||Q^\star||^2}$ and we have $\lambda Q^\star\mathbf{p}_n \to \lambda Q^\star\bar{\mathbf{p}} = \Pi_{\lambda K_J}(\mathbf{f})$.

Proof. The proof is similar to the proof of Theorem 2.3.1. We have:

$$\inf_{\mathbf{u}} ||Q\mathbf{u}||_1 + \frac{1}{2\lambda}||\mathbf{f} - \mathbf{u}||_{L^2}^2,$$

$$\inf_{\mathbf{u}} \sup_{|\mathbf{p}|\le 1} < Q\mathbf{u}, \mathbf{p} >_{L^2} + \frac{1}{2\lambda}||\mathbf{f} - \mathbf{u}||_{L^2}^2,$$

$$(41) \qquad \inf_{\mathbf{u}} \sup_{|\mathbf{p}|\le 1} < \mathbf{u}, Q^\star\mathbf{p} >_{L^2} + \frac{1}{2\lambda}||\mathbf{f} - \mathbf{u}||_{L^2}^2,$$

inf and sup can be swapped using the minimax theorem. The Euler-Lagrange equation w.r.t. $\mathbf{u}$ is $\mathbf{u} = \mathbf{f} - \lambda Q^\star\mathbf{p}$, introduced in (41) gives $\sup_{|\mathbf{p}|\ge 1} < \mathbf{f}, Q^\star\mathbf{p} >_{L^2}$ $-\frac{\lambda}{2}||Q^\star\mathbf{p}||_{L^2}^2$. The constrained Euler-Lagrange equation is $-Q(Q^\star\mathbf{p}-\mathbf{f}/\lambda)+\alpha\mathbf{p} = 0$. Following Section 2.3, we have (40). Besides, Section 2.3 shows that the temporal step that guarantees the stability is $\delta t \ge 1/\tau^2$ with $\tau = ||Q^\star||$. □

Application 1: VTV. In this case, we have $Q = \nabla$, $Q^\star = -\nabla\cdot$.

Application 2: VTV on Manifold. Using Theorem 3.4, the VROF model can also be used to denoise color/vectorial images defined on a manifold Σ:

$$(42) \qquad ||\mathbf{u}||_{BV(\Sigma;\mathbb{R}^M)} + \frac{1}{2\lambda}||\mathbf{f} - A\mathbf{u}||_{L^2(\Sigma;\mathbb{R}^M)}^2,$$

where $\lambda > 0$, $\mathbf{f}$ is the given (noisy) vectorial image on Σ, $||\mathbf{u}||_{BV(\Sigma;\mathbb{R}^M)} = \int_\Sigma |\nabla_\Sigma \mathbf{u}|$, ∇_Σ is the gradient operator on Σ and $||\mathbf{f} - A\mathbf{u}||^2_{L^2(\Sigma;\mathbb{R}^M)} = \int_\Sigma |\mathbf{f} - \mathbf{u}|^2 \, d\Sigma$ where $d\Sigma$ is the invariant-area element. It follows from the Stokes' theorem (integration by parts on the manifold Σ) that $< \nabla_\Sigma u, \mathbf{p} > = - < u, \nabla_\Sigma \cdot \mathbf{p} >$, where $\nabla_\Sigma \cdot$ is the divergence operator on Σ, which is the adjoint of the gradient operator ∇_Σ. This implies that the solution of (42) is equal to:

$$
(43) \qquad\qquad \mathbf{u} = \mathbf{f} - \lambda \nabla_\Sigma \cdot \mathbf{p},
$$

where $\mathbf{p}$ is the steady state of the following flow:

$$
(44) \qquad\qquad \mathbf{p}^{n+1} = \frac{\mathbf{p}^n + \delta t \nabla_\Sigma \big(\lambda \nabla_\Sigma \cdot \mathbf{p}^n - \mathbf{f} \big)}{1 + \delta t |\nabla_\Sigma \big(\lambda \nabla_\Sigma \cdot \mathbf{p}^n - \mathbf{f} \big)|},
$$

given $Q = \nabla_\Sigma$ and $Q^\star = -\nabla_\Sigma \cdot$.

Gradient operator and divergence operator in (43) and (44) can be defined in a two different ways. If the manifold Σ is defined as a parametric manifold, then the gradient of a function $h : \Sigma \to \mathbb{R}$ and the divergence of a vector field $\mathbf{V} = (V_x, V_y) : \Sigma \to \mathbb{R}^2$ in local coordinates are given by [28]:

$$
\begin{cases} (\nabla_\Sigma h)_\mu &= \sum_\nu g^{\mu\nu} \partial_\nu h \\ \nabla_\Sigma \cdot \mathbf{V} &= \frac{1}{\sqrt{g}} \sum_\nu \partial_\nu (\sqrt{g} V_\nu) \end{cases} ,
$$

where $\mu, \nu = \{x, y\}$, $g^{\mu\nu}$ is the inverse of the first fundamental form $g_{\mu\nu}$ of the parametric manifold Σ and $\sqrt{g}$ is the square root of the determinant of $g_{\mu\nu}$.

If the manifold Σ is represented by an implicit level set function as described in [29], then the gradient and the divergence operator are given by:

$$
\begin{cases} \nabla_\Sigma h &= P_{\nabla\phi_\Sigma} \nabla h \\ \nabla_\Sigma \cdot \mathbf{V} &= \frac{1}{|\nabla\phi_\Sigma|} \nabla \cdot \left(\frac{P_{\nabla\phi_\Sigma} \mathbf{V}}{|P_{\nabla\phi_\Sigma} \mathbf{V}|} |\nabla\phi_\Sigma| \right) \end{cases} ,
$$

where ϕ_Σ is a level set function representing the surface $\Sigma := \{\mathbf{x} : \phi_\Sigma(\mathbf{x}) = 0\}$ and $P_{\nabla\phi_\Sigma} := I - \frac{\nabla\phi_\Sigma \otimes \nabla\phi_\Sigma}{|\nabla\phi_\Sigma|^2}$ is a projector operator onto the plane orthogonal to $\nabla\phi_\Sigma$, which is the normal to the surface Σ.

One interesting application of the previous denoising model is for omnidirectional images, which contain information of a scene in all directions simultaneously without rotating a camera. Figure 10 presents the denoising of a color omnidirectional image. The manifold considered is a paraboloid. Its metric, developed in [9], is:

$$
g_{\mu\nu} = \begin{pmatrix} 1 + 4x^2 & 4xy \\ 4xy & 1 + 4y^2 \end{pmatrix},
$$

which implies the following components of the gradient operator:

$$
\begin{cases} (\nabla_\Sigma h)_x &= g^{xx} \partial_x h + g^{xy} \partial_y h \\ (\nabla_\Sigma h)_y &= g^{xy} \partial_x h + g^{yy} \partial_y h \end{cases} ,
$$

and the divergence operator:

$$
\nabla_\Sigma \cdot \mathbf{V} = \frac{1}{\sqrt{g}} \Big(\partial_x (\sqrt{g} V_x) + \partial_y (\sqrt{g} V_y) \Big).
$$

4. **Conclusion.** In this paper, we have presented a regularization algorithm for color/vectorial images based on the dual definition of the vectorial TV norm. Unlike the CGM model [15] or the DS model [20], the dual approach of Chambolle [13] extends nicely to a variety of vector models, producing useful results which are better than standard methods in many cases.

Indeed, this regularization algorithm offers important advantages. Firstly, it minimizes the exact TV norm unlike usual approaches such as [8, 4, 16]. Secondly, the numerical scheme to carry out the minimization process is easy to code. Then, the number of iterations to reach the steady state solution is low, which offers a fast color denoising model compared with other related models such as [8, 33]. Finally, the dual formulation of the vectorial TV norm offers well-founded mathematical theorems to prove the existence of a minimizing solution.

We have compared our approach to related vector denoising models such as the model of Sapiro and Ringach in [33]. In our case, a fast and easy to implement denoising scheme is proposed. We have also proved the existence of a minimizer in the vector valued BV space. In [8], Blomgren and Chan propose a different version of the vectorial TV norm, which produces less color smearing than our model. However [8] does not minimize the exact TV norm and the minimization scheme is slow as shown in Figure 3. In [35], Sochen, Kimmel and Malladi propose a variational model which shares some connections with our model except the existence of a cross-correlation term in their energy which does not exist in ours. However, the cross-correlation term vanishes for gray-scale images, which induces the equivalence of our model with the Beltrami model. This also means that the Beltrami flow can be "dualized" for gray-scale images. Besides, it is difficult to study the existence of a denoising solution in the Beltrami framework unlike the TV framework.

We have also showed that the proposed fast vectorial regularization scheme can be easily applied to several standard color image processing such as the color inverse scale space, color denoising with the chromaticity-brightness color representation, color image inpainting, color wavelet shrinkage, color image decomposition, color image deblurring, and color denoising on manifolds.

Finally, and maybe more importantly, we would like to emphasize that this proposed L^1 vectorial regularization algorithm can be used in problems that required vector field regularization while preserving discontinuities. A potential interesting application of vector field denoising is in medical imaging where the proposed regularization process can be used to denoise Diffusion Tensor Images (DTI). Figure 11 presents the regularization of a synthetic vector field that has been damaged by noise. Note that the regularization is able to restore the four regions with same orientation although it is very difficult to see them in Figure 11(e). Other applications may be optical flow regularization [26], image colorization [27], image inpainting [37], multiscale hierarchical decomposition of color images [36], etc.

(a) Channel-By-Channel ROF: Channel #1.

(b) Channel-By-Channel ROF: Channel #2.

(c) Vectorial ROF: Channel #1.

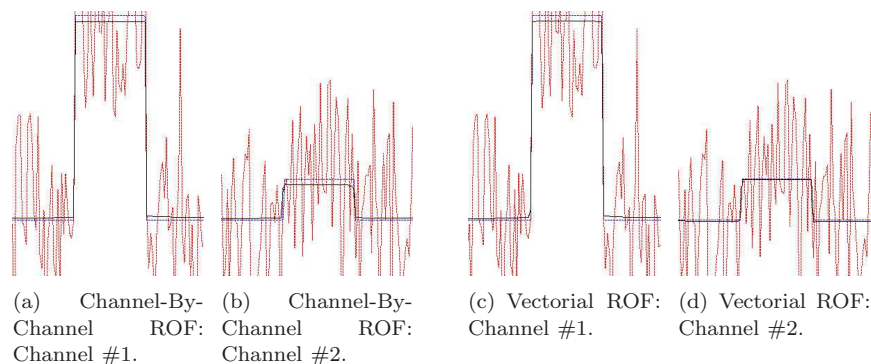(d) Vectorial ROF: Channel #2.

FIGURE 1. Comparison between the vectorial ROF model and the scalar ROF model channel-by-channel. The blue plot is the original signal, the red plot is the noisy signal and the black plot is the denoised signal. Figures (a-b) present the denoising result carried out by the scalar ROF model channel-by-channel. Figures (c-d) show the result obtained with the vectorial ROF model. The coupling term of channels in the vectorial ROF model allows to better restore Channel #2 on Figure (d) compared with Figure (b) where the direct scalar ROF model can not completely restore the weak signal.

(a) Original Color Image.

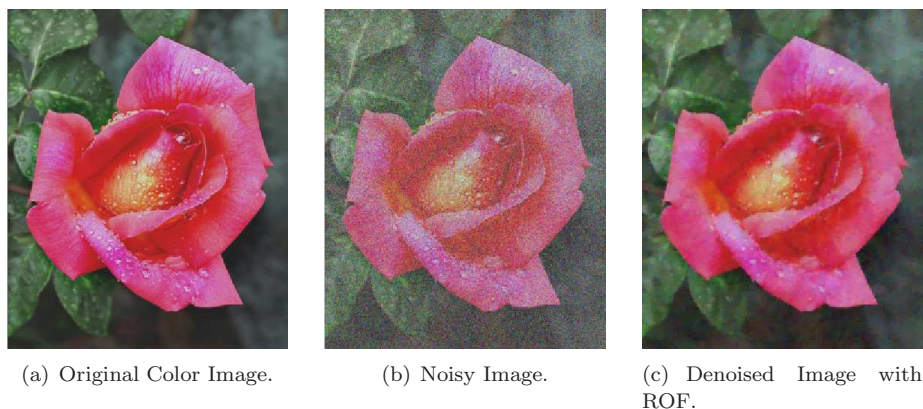(b) Noisy Image.

(c) Denoised Image with ROF.

FIGURE 2. Denoising of a color picture with the vectorial ROF model. The denoising process takes 1.8sec and 60 iterations. The picture size is 250x303. (Note: All pictures are in color.)

REFERENCES

[1] L. Ambrosio, N. Fusco and D. Pallara, "Functions of Bounded Variation and Free Discontinuity Problems," edition, Oxford University Press, 2000.

[2] J. P. Aubin, "Applied Functional Analysis," edition, John Wiley and Sons Inc., New York, 1979.

[3] J.-F. Aujol and A. Chambolle, *Dual norms and image decomposition models*, International Journal of Computer Vision, **63** (2005), 85–104.
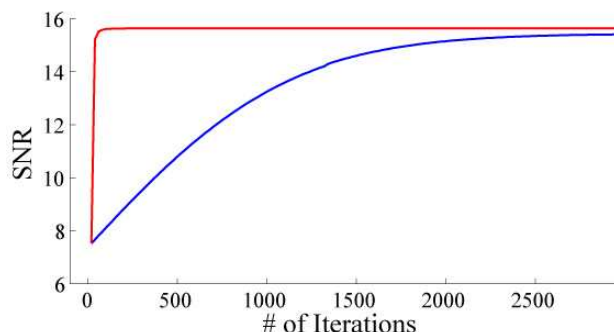
Figure 3. Comparison between our vectorial denoising ROF model (red plot) and the denoising ROF model (blue plot) proposed by Blomgren and Chan in [8]. Both models are applied on Figure 2. x-axis represents the number of iterations and y-axis the Signal-to-Noise Ratio (SNR). Our numerical scheme, based on the dual approach, converges much faster to the denoising solution than the original model [8]. Our model converges in 60 iterations (1.8sec) whereas the original model needs more than 2000 iterations (∼500sec) to reach the denoising solution. We also notice than the SNR of the original model is slightly smaller than the SNR obtained by the dual approach. This is due to the regularization of the vectorial TV norm in the standard approach [8], which does not exist in the dual approach.

(a) Noisy Color Image: $\mathbf{f}$. (b) 1st Iteration of ROF: $\mathbf{u}_1$. (c) 2nd Iteration of ROF: $\mathbf{u}_2$.

Figure 4. Application of the iterative regularization ROF model. The iterative denoising process takes 1.8sec and 100 iterations for each ROF iteration and the picture size is 250x303. The direct ROF model gives a SNR of 14.57 and the iterative ROF model produces a SNR equal to 15.03. There are less geometric residues with the iterative process because of the better contrast preservation. (Note: All pictures are in color.)

(a) Original Color Image.   (b) Noisy Image.   (c) Denoised Image with the CB Representation.
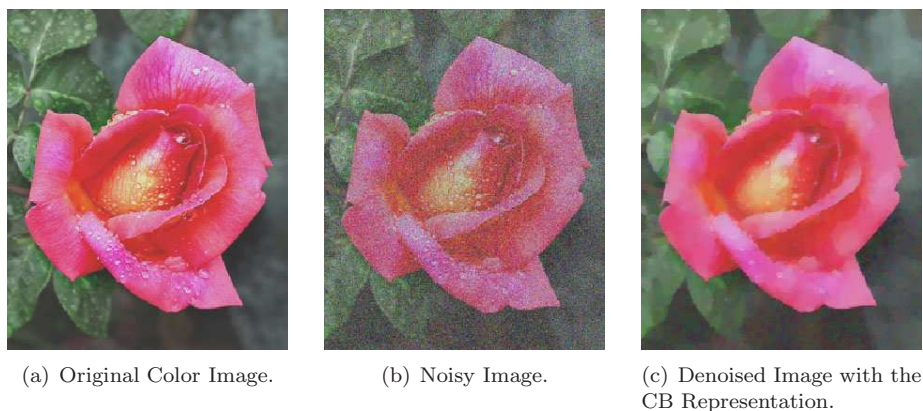
FIGURE 5. Application of the Chromaticity-Brightness (CB) color representation. Figure (c) presents the denoised color image with the CB representation. The SNR is 14.57 with the RGB color representation and 14.72 with the CB representation. (Note: All pictures are in color.)

(a) Original Color Image.   (b)   Smooth/Geometrical Part: **u**.   (c) Oscillating/Textural Part: **v**.
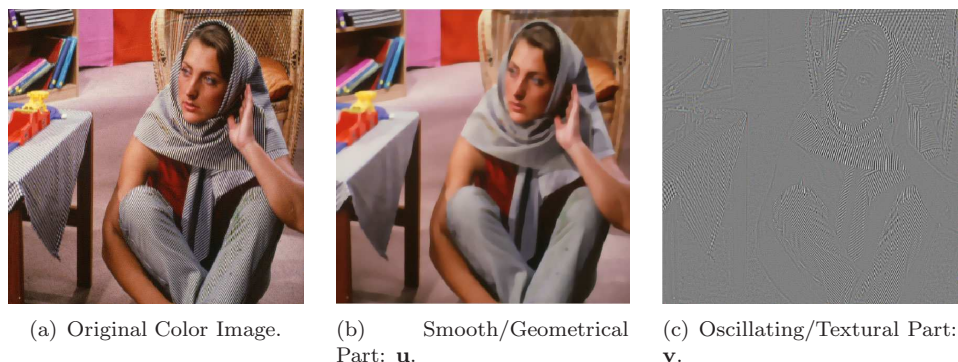
FIGURE 6. Color image decomposition process. Figure (b) presents the smooth/geometrical part of (a) and Figure (c) the oscillating/textural part of (a). The decomposition process takes 19sec and 60 iterations. The picture size is 512x512. (Note: All pictures are in color.)

[4] J. F. Aujol and S. H. Kang, *Color image decomposition and restoration*, Journal of Visual Communication and Image Representation, **17** (2006), 916–928.

[5] J. Bect, L. Blanc-Feraud, G. Aubert and A. Chambolle, *A l1-unified variational framework for image restoration*, in "European Conference on Computer Vision," (2004), 1–13.

[6] M. Bertalmio, G. Sapiro, V. Caselles and C. Ballester, *Image inpainting*, in "International Conference on Computer Graphics and Interactive Techniques," (2000), 417–424.

[7] M. Bertalmio, L. Vese, G. Sapiro and S. Osher, *Simultaneous structure and texture image inpainting*, IEEE Transactions on Image Processing, **12** (2003), 882–889.

[8] P. Blomgren and T. F. Chan, *Color TV: Total variation methods for restoration of vector-valued images*, IEEE Transactions on Image Processing, **7** (1998), 304–309.

(a) Original Color Image.   (b)   Smooth/Geometrical   (c) Oscillating/Textural Part:
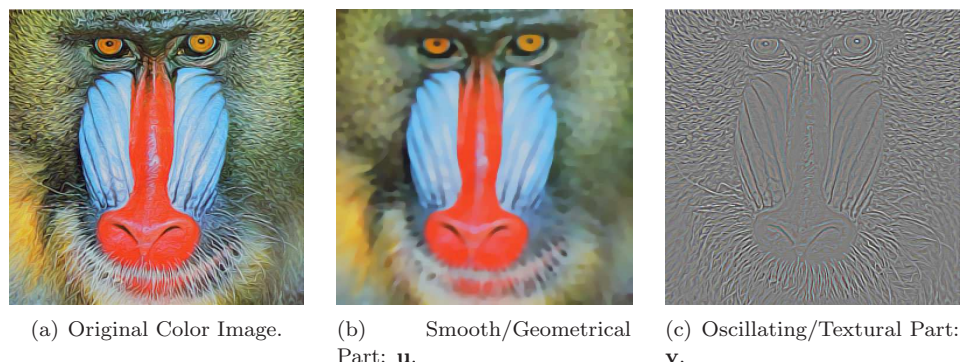   Part: **u**.   **v**.

Figure 7. Color image decomposition process. Figure (b) presents the smooth/geometrical part of (a) and Figure (c) the oscillating/textural part of (a). The decomposition process takes 27sec and 300 iterations. The picture size is 253x253. (Note: All pictures are in color.)

[9] I. Bogdanova, X. Bresson, J. Thiran and P. Vandergheynst, *Scale-Space analysis and active contours for omnidirectional images*, IEEE Transactions on Image Processing, **16** (2007), 1888–1901.

[10] X. Bresson and J.-P. Thiran, *Image segmentation model using active contour and image decomposition*, in "IEEE International Conference on Image Processing," (2006), 1657–1660.

[11] H. Brezis, F. Bethuel and F. Helein, *Singular limit for the minimization of Ginzburg-Landau functionals*, Comptes rendus de l'Academie des Sciences, Serie Mathematique, **314** (1992), 891-895.

[12] F. Catt, P. L. Lions, J. M. Morel and T. Coll, *Image selective smoothing and edge detection by nonlinear diffusion*, SIAM Journal on Numerical Analysis, **29** (1992), 182–193.

[13] A. Chambolle, *An algorithm for total variation minimization and applications*, Journal of Mathematical Imaging and Vision, **20** (2004), 89–97.

[14] A. Chambolle, R. DeVore, N. Lee and B. Lucier, *Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage*, IEEE Transactions on Image Processing, **7** (1998), 319–333.

[15] T. F. Chan, G. H. Golub and P. Mulet, *A nonlinear primal-dual method for total variation-based image restoration*, SIAM journal on scientific computing, **20** (1999), 1964–1977.

[16] T. F. Chan, S. H. Kang and J. Shen, *Total variation denoising and enhancement of color images based on the CB and HSV color models*, Journal of Visual Communication and Image Representation, **12** (2001), 422–435.

[17] T. F. Chan and J. Shen, *Variational image inpainting*, Communications on Pure and Applied Mathematics, **58** (2005), 579–619.

[18] T. F. Chan and C.-K. Wong, *Total variation blind deconvolution*, IEEE Transactions on Image Processing, **7** (1998), 370–375.

[19] J. Darbon and S. Peyronnet, *A vectorial self-dual morphological filter based on total variation minimization*, in "International Symposium on Visual Computing" (ISVC), **3804** (2005), 388–395.

[20] 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** (2006), 277–291.

[21] D. L. Donoho and I. M. Johnstone, *Ideal spatial adaptation by wavelet shrinkage*, Biometrika, **81** (1994), 425–455.

[22] I. Ekeland and R. Temam, "Convex Analysis and Variational Problems," edition, Society for Industrial and Applied Mathematics, 1976.

[23] M. Fornasier and R. March, *Restoration of color images by vector valued BV functions and variational calculus*, SIAM Journal on Applied Mathematics, **68** (2007), 437–460.
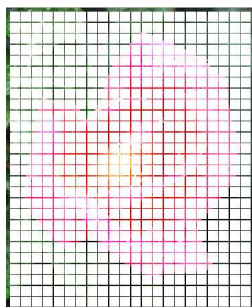
(a) Original Color Image.
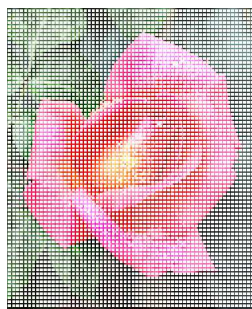
(b) Image which lost 56% of original information.

(c) Inpainting Result.

(d) Image which lost 79% of original information.

(e) Inpainting Result.

FIGURE 8. Color image inpainting process. Figures (b,c) present the image which lost 56% of original information and the inpainting result. The inpainting process takes 2.3sec and 60 iterations and the picture size is 250x303. Figures (d,e) show the image which lost 79% of original information and the inpainting result. The inpainting process takes 5.9sec and 150 iterations. (Note: All pictures are in color.)

[24] D. Graziani, "On the L^1-Lower Semicontinuity and Relaxation in $BV(\Omega, \mathbb{R}^M)$ for Integral Functionals with Discontinuous Integrand," Ph.D thesis, 2006.

[25] L. He, A. Marquina and S. Osher, *Blind deconvolution using TV regularization and bregman iteration*, International Journal of Imaging Systems and Technology, **15** (2003), 74–83.
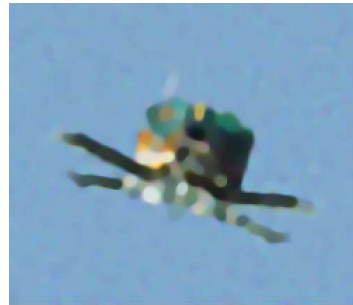
(a) Original Color Image.

(b) Blurred and Noisy Image #1.

(c) Deblurred and Denoised Result #1.

(d) Blurred and Noisy Image #2.

(e) Deblurred and Denoised Result #2.

FIGURE 9. Color image deblurring and denoising process. The simultaneous denoising and deblurring process takes 205sec and 600 iterations. (Note: All pictures are in color.)

[26] B. Horn and B. Schunck, *Determining optical flow*, Artificial Intelligence, **17** (1981), 185–203.
[27] S. H. Kang and R. March, *Variational models for image colorization via chromaticity and brightness decomposition*, IEEE Transactions on Image Processing, **16** (2007), 2251–2261.
[28] E. Kreyszig, "Differential Geometry," edition, Paperback, 1991.
[29] F. Memoli, M. Bertalmio, L. T. Cheng, G. Sapiro and S. Osher, *Variational problems and partial differential equations on implicit surfaces: Bye bye triangulated surfaces?*, in "Geometric Level Set Methods in Imaging, Vision, and Graphics," Springer, New York, (2003), 381–397.

(a) Noisy Omnidirectional Image on a Parabolic Manifold.

(b) Denoised Omnidirectional Image.

(c) Noisy Omnidirectional Image on the Parametric Plane.

(d) Zoom In the Noisy Image.
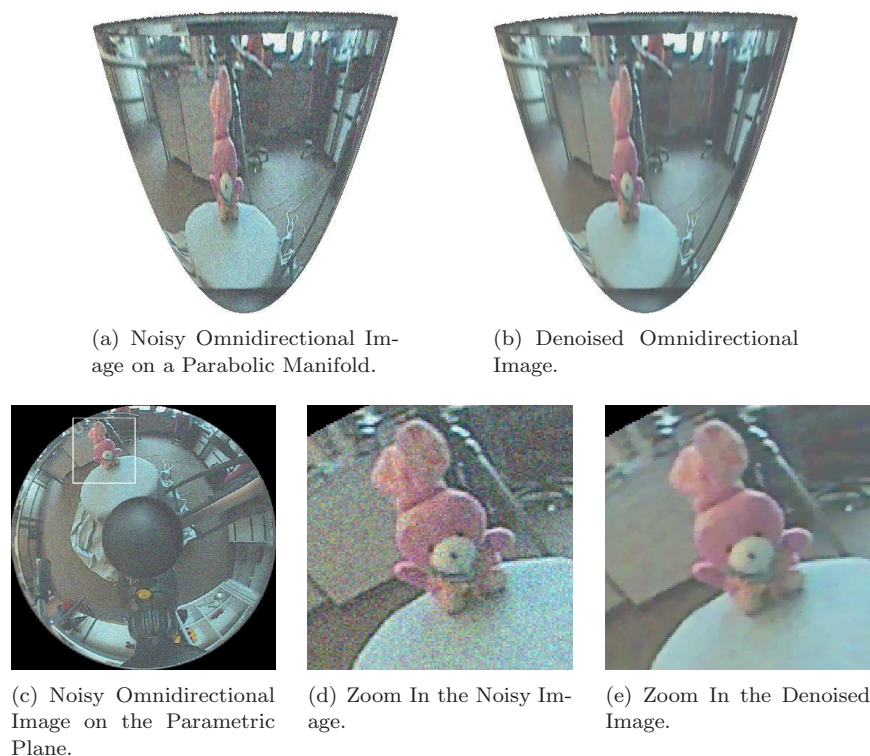
(e) Zoom In the Denoised Image.

FIGURE 10. Denoising of a color image defined on a manifold: an omnidirectional image painted on a parabolic manifold. The denoising process takes 19.7sec and 130 iterations and the picture size is 574x574. (Note: All pictures are in color.)

[30] Y. Meyer, "Oscillating Patterns in Image Processing and Nonlinear Evolution Equations," edition, American Mathematical Society, 22, 2001.

[31] S. Osher, M. Burger, D. Goldfarb, J. Xu and W. Yin, *An iterative regularization method for total variation-based image restoration*, SIAM Multiscale Modeling & Simulation, **4** (2005), 460–489.

[32] L. I. Rudin, S. Osher and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D, **60** (1992), 259–268.

[33] G. Sapiro and D. L. Ringach, *Anisotropic diffusion of multi-valued images with applications to color filtering*, IEEE Transactions on Image Processing, **5** (1996), 1582–1586.

[34] M. Sion, *On general minimax theorems*, Pacific Journal of Mathematics, **9** (1958), 171–176.

[35] N. Sochen, R. Kimmel and R. Malladi, *A general framework for low level vision*, IEEE Transactions on Image Processing, **7** (1998), 310–318.

[36] E. Tadmor, S. Nezzar and L. Vese, *Multiscale hierarchical decomposition of images with applications to deblurring, denoising and segmentation*, Communications on Mathematical Sciences, **6** (2008), 281–307.

[37] S. Osher, X.-C. Tai and R. Holm, *Image inpainting using a TV-stokes equation*, in "Image Processing Based on Partial Differential Equations," Tai, Lie, Chan and Osher eds, Springer, (2007), 3–22.

[38] A. N. Tikhonov and V. A. Arsenin, "Solution of Ill-posed Problems," edition, Winston & Sons, 1977.

[39] L. A. Vese and S. J. Osher, *Color texture modeling and color image decomposition in a variational-PDE approach*, in "Symbolic and Numeric Algorithms for Scientific Computing" (SYNASC), (2006), 103–110.

(a) Clean Vector Field.   (b) Noisy Vector Field #1.   (c) Denoised Vector Field #1 with ROF.

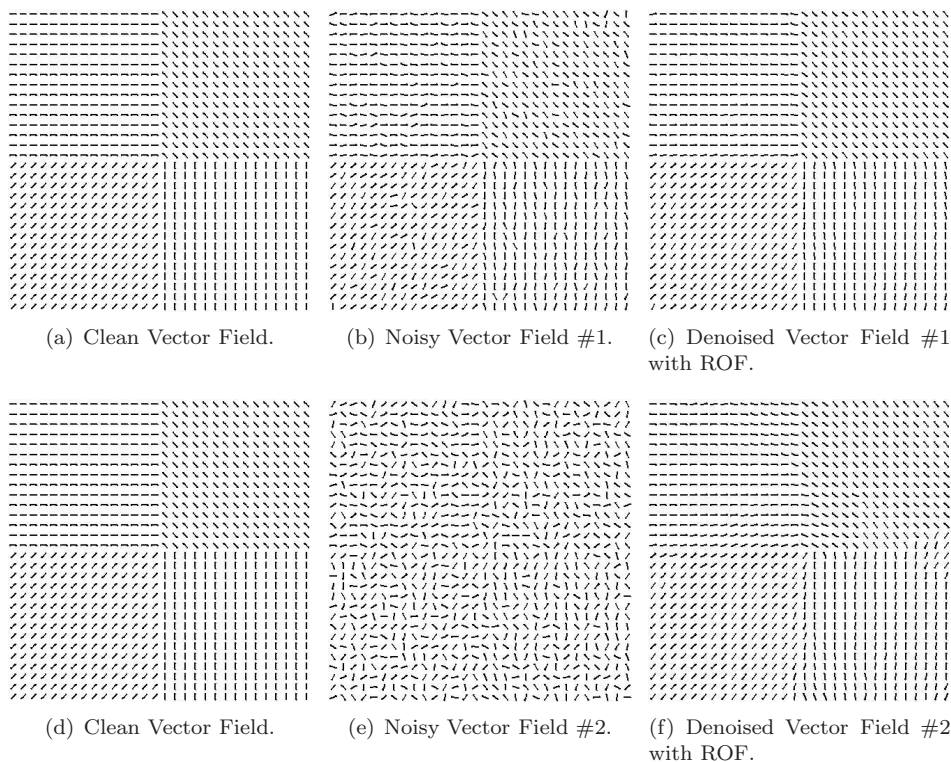(d) Clean Vector Field.   (e) Noisy Vector Field #2.   (f) Denoised Vector Field #2 with ROF.

FIGURE 11. Denoising of a vector field with the vectorial ROF model. The denoising process takes less than 0.1sec, 30 iterations for #1 and 100 iterations for #2. The picture size is 30x30. We can notice the good performance of our denoising model with the noisy vector field (e), which has lost the main directions (d). Our model is able to recover the main structure (f).

[40] J. Von Neumann, *Zur theorie der gesellschaftsspiele*, Math. Annalen., **100** (1928), 295–320.

E-mail address: xbresson@math.ucla.edu
E-mail address: chan@math.ucla.edu