SIAM J. IMAGING SCIENCES
Vol. 3, No. 2, pp. 153–175

Application of the Topological Gradient Method to Color Image Restoration*

Didier Auroux†, Lamia Jaafar Belaid‡, and Badreddine Rjaibi‡

Abstract. The aim of this paper is to generalize to color images the topological gradient method for the restoration problem. First, we consider a simple extension from grey-level to color images, by considering a channel by channel method. Then, as there should be some coupling between the different channels, we propose considering a tensorial approach: the Di Zenzo gradient. We derive an asymptotic expansion of this gradient with respect to the insertion of a small crack in the image, and we propose a new algorithm based on the topological gradient for the color image restoration problem. Finally, we illustrate these two approaches with numerical results.

Key words. topological gradient, topological asymptotic expansion, image restoration, Di Zenzo gradient, color vision

AMS subject classifications. 94A08, 94A40, 49Q10, 49Q12

DOI. 10.1137/080721017

1. Introduction.

The goal of this paper is to propose a new method for color image restoration. This method is based on the topological gradient approach, which has been introduced for topological optimization purposes [1, 3, 16, 17, 18, 20, 24, 26, 28, 33]. The idea of topological optimization is to create a partition of a given domain (in our case, an image), as one usually looks for an optimal design and its complement. A common way to consider the restoration problem is to identify the edges of the image in order to preserve them in the restoration process. Then, the image is smoothed elsewhere. For this reason, topological shape optimization and image processing problems have common mathematical methods, such as level set approaches, material properties optimization, or variational methods [7, 13, 14, 25, 27, 29, 31, 37, 39, 40].

The basic idea is to adapt to color images the topological gradient approach applied to diffusive image restoration [21] and the grey-level classification problem [8]. In these cases, an optimal material distribution is obtained at the first iteration by considering topological optimization tools for the detection of edges. More precisely, if we consider an open bounded domain $\Omega \subset \mathbb{R}^2$ and $j(\Omega) = J(u_\Omega)$ as a cost function to be minimized, where u_Ω is the solution to a given partial differential equation (PDE) problem defined in Ω, the topological gradient provides the derivative of the cost function with respect to the insertion of a small crack in the domain. We then define an edge set by thresholding the topological gradient, and we then consider only two values of the conductivity coefficient in the diffusion process: 0 (or a very small value) on the edge set, and a constant ($\simeq 1$) outside.

*Received by the editors April 11, 2008; accepted for publication (in revised form) January 11, 2010; published electronically April 7, 2010.
http://www.siam.org/journals/siims/3-2/72101.html
†Laboratoire J. A. Dieudonné, Université de Nice Sophia Antipolis, 06108 Nice cedex 2, France (auroux@unice.fr).
‡ENIT-LAMSIN, Campus Universitaire, BP.37, 1002 Le Belvédère, Tunis, Tunisia (lamia.belaid@esstt.rnu.tn, badreddine.rjaibi@lamsin.rnu.tn).

Several variational methods have been proposed for defining an edge indicator function, mainly based on the Mumford–Shah functional or on some approximations [25, 2, 38, 15]. Indeed, finding an edge set is equivalent to identifying its characteristic function. This problem does not seem differentiable, and standard ways of dealing with this issue are either to consider level-set functions or to relax the problem by considering an approximation of the characteristic function. In both cases, the conductivity (or diffusion) coefficient can take all the values of an interval (for instance, $[0; 1]$), which can be seen as a relaxation of our topological gradient approach. But, by enlarging the set of admissible solutions, relaxation increases the instability of the restoration process, and this could explain why our method usually needs only one iteration, whereas the other variational approaches need several iterations to achieve the convergence of the edge indicator function.

If we now consider color or multispectral images, they can be represented or modeled in various ways [19], and many approaches to studying classical problems in color image processing, such as restoration, segmentation, and classification problems, can be found in the literature [4, 5, 10, 12, 30, 34, 35, 36]. In this paper, we shall focus on the red-green-blue (RGB) model in which a color image I is usually represented by a three-dimensional (3D) vector and the topological gradient approach. For each pixel $x = (x_1, x_2)$ of the image, the vector $I(x) \in \mathbb{R}^3$ represents the intensity of the three colors red, green, and blue. Since each component can be treated as a grey-level image, the topological gradient approach can easily be used, provided that the cost function to be minimized has no coupling terms between the several channels. For the restoration problem, one usually sets the cost function to the square norm of the gradient of the image. If we consider the most simple case, i.e., the standard L^2 norm of the gradient, then the cost function can be easily written as a sum of three independent terms, one for each channel, and then the color image problem can be decomposed into three channel by channel independent problems.

On the contrary, the Di Zenzo gradient consists of coupling the three components of the standard gradient by considering a tensor gradient of the image [41]. This tensor defines a higher Di Zenzo gradient when all three components of the image vary in the same direction, whereas independent variations of the channels lead to a smaller gradient. In this case, by considering the square of this gradient as the cost function to be minimized, the three channels are highly coupled, and it is no longer possible to consider three problems separately.

Some other coupling techniques have been defined in order to process multichannel images. We can cite, for instance, the Beltrami diffusion-type process, in which images are considered as manifolds. A nonlinear structure tensor is defined and locally adjusted according to a gradient-type measure, allowing one to control the diffusion process [32, 11, 22]. As explained further (see section 3.3), the Beltrami and Di Zenzo structure tensors are nearly equivalent. But within the topological gradient framework, it seems more natural to first study the Di Zenzo approach for the channel coupling.

This paper is organized as follows. We first consider in section 2 the standard gradient norm, for which it is possible to decouple the vectorial minimization problem into three scalar problems. For each of them, we recall the topological asymptotic expansion of the cost function, and we give the corresponding expansion for the color problem. For comparison purposes, we also briefly recall the Ambrosio–Tortorelli variational approach (minimization of an approximation of the Mumford–Shah functional). Then section 3 is devoted to the Di

Zenzo gradient. First, we define a vectorial minimization problem and then give the main result of this article: the topological asymptotic expansion of the square norm of the Di Zenzo gradient. Next, we propose an algorithm for the color image restoration based on this result. We briefly compare the Di Zenzo approach and the Beltrami framework. We report the results of many numerical experiments in section 4, in order to validate the theoretical results and also to compare several approaches. Finally, some concluding remarks are given in section 5.

2. Application of the topological asymptotic expansion to color image restoration using a standard gradient norm. This section is an extension of the topological gradient method applied to grey-level restoration problems [21] to color images, by considering a standard gradient norm. First, we recall the principle of the topological asymptotic expansion [24], within the restoration problem.

2.1. Problem formulation. Let Ω be an open bounded domain of $\mathbb{R}^2$, and let v be a given noisy color image, defined as a vectorial function

$$v : x = (x_1, x_2) \in \Omega \subset \mathbb{R}^2 \longmapsto v(x) = (v^1(x), v^2(x), v^3(x)) \in \mathbb{R}^3, \tag{2.1}$$

where v^k is the value of the pixel $x = (x_1, x_2)$ in each color subspace. In image processing, color has been represented or modeled in various ways [19]. In this paper, we have chosen the RGB space in which images are usually stored and displayed. In the RGB representation, at each pixel $x = (x_1, x_2)$, the vector value $v(x)$ represents the intensity of the three primary colors separately: red, green, and blue. Each monochromatic component v^k is called one channel.

The standard variational approach for the restoration problem consists of minimizing the following cost function:

$$\mathcal{J}(u) = \frac{1}{2}\int_\Omega |u - v|^2 \, dx \; + \frac{1}{2}\int_\Omega c|\nabla u|^2 \, dx, \tag{2.2}$$

where $|.|$ represents the standard norm in $\mathbb{R}^2$. The first term of the cost function measures the distance to the data, and the second is a Tikhonov regularization term. The regularization makes the inverse problem of finding a smooth image $u \in H^1(\Omega)$ close to a nonsmooth image $v \in L^2(\Omega)$ well-posed.

The derivation of the Euler–Lagrange equation associated to the minimization of $\mathcal{J}$ leads to the following restoration equation:

$$\begin{cases} -\mathrm{div}\,(c\nabla u) + u = v & \text{in} \quad \Omega, \\ \partial_n u = 0 & \text{on} \quad \partial\Omega, \end{cases} \tag{2.3}$$

where n denotes the outward unit normal to $\partial\Omega$ and c is a positive conductivity coefficient [21]. In the topological gradient approach, c takes only two values: $c_0 \sim 1$ in the smooth part of the image and a small value $\varepsilon > 0$ in its complement.

Let $j(\Omega) = J(u_\Omega)$ be a cost function to be minimized, where u_Ω is the solution to a given PDE problem defined in Ω. For a small $\rho \geq 0$, let $\Omega_\rho = \Omega \backslash \sigma_\rho$ be the perturbed domain by the insertion of a crack $\sigma_\rho = x_0 + \rho\sigma(n)$, where $x_0 \in \Omega$, $\sigma(n)$ is a straight crack, and n is a

unit vector normal to the crack. The topological sensitivity theory provides an asymptotic expansion of j when ρ tends to zero. It takes the general form

$$j(\Omega_\rho) - j(\Omega) = f(\rho)g(x_0) + \circ(f(\rho)), \tag{2.4}$$

where $f(\rho)$ is an explicit positive function going to zero with ρ and $g(x_0)$ is called the topological gradient at point x_0.

For v a given function in $L^2(\Omega)$, we consider the following problem: find $u_\rho \in H^1(\Omega_\rho)$ such that

$$\begin{cases} -\text{div}\,(c\nabla u_\rho) + u_\rho = v & \text{in} \quad \Omega_\rho, \\ \partial_n u_\rho = 0 & \text{on} \quad \partial\Omega_\rho, \end{cases} \tag{2.5}$$

where n denotes the outward unit normal to $\partial\Omega_\rho$ and c is a constant function. Edge detection is equivalent to looking for a subdomain of Ω where the energy is small. So our goal is to minimize the energy norm outside edges

$$j(\rho) = J(u_\rho) = \int_{\Omega_\rho} |\nabla u_\rho|^2 \, dx, \tag{2.6}$$

where $|\,.\,|$ represents the standard norm in $\mathbb{R}^2$.

The sign of the topological sensitivity $g(x_0)$ called the topological gradient provides information concerning the insertion of a small crack at point x_0, as it will make the cost function increase or decrease (asymptotically). Hence the minimization of the cost function will be performed by the insertion of small cracks at the points where the topological gradient is the most negative.

2.2. Decomposition into three separate problems. The noisy color image v is defined as a vectorial function and is given by (2.1). Each function v^k is a scalar function from Ω to $\mathbb{R}$. In a similar way, the solution u of (2.3) can be written as $u = (u^1, u^2, u^3)$, where u^k are solutions of the following problems:

$$\begin{cases} -\text{div}\left(c\nabla u^k\right) + u^k = v^k & \text{in} \quad \Omega, \\ \partial_n u^k = 0 & \text{on} \quad \partial\Omega \end{cases} \tag{2.7}$$

for $k = 1, 2, 3$.

The perturbed problems corresponding to the insertion of a small crack in the domain are the following: for each k, find $u_\rho^k \in H^1(\Omega_\rho)$ such that

$$\begin{cases} -\text{div}\left(c\nabla u_\rho^k\right) + u_\rho^k = v^k & \text{in} \quad \Omega_\rho, \\ \partial_n u_\rho^k = 0 & \text{on} \quad \partial\Omega_\rho. \end{cases} \tag{2.8}$$

The corresponding variational formulation is given by

$$\begin{cases} \text{find } u_\rho^k \in H^1(\Omega_\rho) \text{ such that} \\ a_\rho^k(u_\rho^k, w) = l_\rho^k(w) \quad \forall w \in H^1(\Omega_\rho), \end{cases} \tag{2.9}$$

where a_ρ^k is the bilinear form, defined on $H^1(\Omega_\rho) \times H^1(\Omega_\rho)$ by

$$a_\rho^k(u^k, w) = \int_{\Omega_\rho} \left(c\nabla u^k \nabla w + u^k w \right) \, dx, \tag{2.10}$$

and l_ρ^k is the linear form, defined on $L^2(\Omega_\rho)$ by

$$l_\rho^k(w) = \int_{\Omega_\rho} v^k w \, dx. \tag{2.11}$$

The cost function j can also be decomposed into three terms, one for each channel, as follows:

$$j(\rho) = \sum_{k=1}^{3} j^k(\rho), \tag{2.12}$$

with

$$j^k(\rho) = \int_{\Omega_\rho} |\nabla u_\rho^k|^2 \, dx. \tag{2.13}$$

We refer the reader to [3] for the proof of the following result.

Theorem 2.1. *For each k, there exist a function $f_k : \mathbb{R}^+ \to \mathbb{R}^+$ going to zero with ρ, a linear form $L_\rho^k : H^1(\Omega_\rho) \to \mathbb{R}$, and four real numbers δJ_1^k, δJ_2^k, δa^k, and δl^k such that*

$$J_\rho^k(u_\rho^k) - J_\rho^k(u_0^k) = L_\rho^k(u_\rho^k - u_0^k) + f_k(\rho)\delta J_1^k + o(f_k(\rho)), \tag{2.14}$$

$$J_\rho^k(u_0^k) - J_0^k(u_0^k) = f_k(\rho)\delta J_2^k + o(f_k(\rho)), \tag{2.15}$$

$$(a_\rho^k - a_0^k)(u_0^k, v_\rho^k) = f_k(\rho)\delta a^k + o(f_k(\rho)), \tag{2.16}$$

$$(l_\rho^k - l_0^k)(v_\rho^k) = f_k(\rho)\delta l^k + o(f_k(\rho)), \tag{2.17}$$

where v_ρ^k is the solution to the adjoint problem

$$a_\rho^k(w, v_\rho^k) = -L_\rho^k(w) \quad \forall w \in H^1(\Omega_\rho), \tag{2.18}$$

L_ρ^k is the first derivative of J_ρ^k (see (2.14)) given by

$$L_\rho^k(w) = 2\int_{\Omega_\rho} \nabla u_0^k \nabla w \, dx, \tag{2.19}$$

and u_ρ^k is the solution to the direct problem

$$a_\rho^k(u_\rho^k, w) = l_\rho^k(w) \quad \forall w \in H^1(\Omega_\rho). \tag{2.20}$$

Moreover, for $f_k(\rho) = \rho^2$ we have the following results (see [3] for computational details):

$$\delta a^k = -\pi c(\nabla u_0^k(x_0).n)(\nabla v_0^k(x_0).n), \tag{2.21}$$

$$\delta l^k = 0, \tag{2.22}$$

$$\delta J_1^k = -\pi |\nabla u_0^k(x_0).n|^2, \tag{2.23}$$

$$\delta J_2^k = 0. \tag{2.24}$$

Equalities (2.21) and (2.23) take into account the variation of the gradient of u_0^k when the crack is created. Since the crack is of null measure, the variation of an integral over Ω is equal to zero, and this is why δl^k and δJ_2^k are equal to zero.

Remark. The solution v_ρ^k is considered only for mathematical analysis. In practice, for numerical calculations, only the adjoint v_0^k, calculated for $\rho = 0$, is considered. Hence, in our case, the direct and adjoint problems read, respectively, as

$$\begin{cases} -\mathrm{div}(c\nabla u_0^k) + u_0^k = v^k & \text{in} \quad \Omega, \\ \partial_n u_0^k = 0 & \text{on} \quad \partial\Omega \end{cases} \tag{2.25}$$

and

$$\begin{cases} -\mathrm{div}(c\nabla v_0^k) + v_0^k = -\partial_{u^k} J(u_0) & \text{in} \quad \Omega, \\ \partial_n v_0^k = 0 & \text{on} \quad \partial\Omega. \end{cases} \tag{2.26}$$

These two equations can be seen as an optimality system for the minimization of the cost function under the direct problem constraint. More precisely, if we define the Lagrangian operator

$$\mathcal{L}_\rho^k(u, v) = J_\rho^k(u) + a_\rho^k(u, v) - l_\rho^k(v), \tag{2.27}$$

then, for $\rho = 0$, $\partial_u(\mathcal{L}_\rho^k) = 0$ leads to (2.26) and $\partial_v(\mathcal{L}_\rho^k) = 0$ leads to (2.25).

From a vectorial point of view, the adjoint state v_0 is solution to the following adjoint problem:

$$\begin{cases} -\mathrm{div}(c\nabla v_0) + v_0 = -\partial_u J(u_0) & \text{in} \quad \Omega, \\ \partial_n v_0 = 0 & \text{on} \quad \partial\Omega. \end{cases} \tag{2.28}$$

For each channel $k = 1, 2, 3$, the asymptotic expansion of $j^k(\rho)$ is given by the following proposition.

Proposition 2.2. *The function j^k has the asymptotic expansion*

$$j^k(\rho) = j^k(0) + f_k(\rho)(\delta a^k - \delta l^k + \delta J_1^k + \delta J_2^k) + o(f_k(\rho)). \tag{2.29}$$

Proof. By considering (2.14)–(2.17), we have

$$\begin{aligned} j^k(\rho) - j^k(0) &= J_\rho^k\left(u_\rho^k\right) - J_0^k\left(u_0^k\right) \\ &= J_\rho^k\left(u_\rho^k\right) - J_\rho^k\left(u_0^k\right) + J_\rho^k\left(u_0^k\right) - J_0^k\left(u_0^k\right) \\ &= L_\rho^k\left(u_\rho^k - u_0^k\right) + f_k(\rho)\left(\delta J_1^k + \delta J_2^k\right) + o\left(f_k(\rho)\right) \\ &= a_\rho^k\left(u_0^k, v_\rho^k\right) - a_\rho^k\left(u_\rho^k, v_\rho^k\right) + f_k(\rho)\left(\delta J_1^k + \delta J_2^k\right) + o\left(f_k(\rho)\right) \\ &= \left(a_\rho^k - a_0^k\right)\left(u_0^k, v_\rho^k\right) - \left(l_\rho^k - l_0^k\right)\left(v_\rho^k\right) + f_k(\rho)\left(\delta J_1^k + \delta J_2^k\right) + o\left(f_k(\rho)\right) \\ &= f_k(\rho)(\delta a^k - \delta l^k + \delta J_1^k + \delta J_2^k) + o(f_k(\rho)). \end{aligned}$$

■

2.3. Asymptotic expansion and corresponding algorithm. Let J be the following cost function:

$$J(u) = \|\nabla u\|^2_{L^2(\Omega,\mathbb{R}^3)} = \int_\Omega \left(|\nabla u^1(x)|^2 + |\nabla u^2(x)|^2 + |\nabla u^3(x)|^2\right) dx, \tag{2.30}$$

where $|.|$ represents the standard norm in $\mathbb{R}^2$; then using the previous results and according to [21], we deduce that the cost function j has an asymptotic expansion given by the following theorem.

Theorem 2.3. *The cost function j has the following asymptotic expansion:*

$$j(\rho) - j(0) = \rho^2 g(x_0) + o(\rho^2), \tag{2.31}$$

with $g(x_0) = \sum_{k=1}^3 g^k(x_0)$ *and*

$$g^k(x_0) = -\pi c(\nabla u_0^k(x_0).n)(\nabla v_0^k(x_0).n) - \pi|\nabla u_0^k(x_0).n|^2.$$

The topological gradient $g(x)$ can be rewritten in the following way:

$$g(x) = \langle M(x)n, n\rangle, \tag{2.32}$$

where $M(x)$ is the symmetric 2×2 matrix given by

$$M(x) = \sum_{k=1}^3 \left[-\pi c\frac{\nabla u_0^k(x)\nabla v_0^k(x)^T + \nabla v_0^k(x)\nabla u_0^k(x)^T}{2} - \pi\nabla u_0^k(x)\nabla u_0^k(x)^T\right]. \tag{2.33}$$

For a given point x, $g(x)$ takes its minimal value when n is the eigenvector associated to the lowest eigenvalue λ_{min} of M. This value will be considered as the topological gradient associated to the optimal orientation of the crack σ_ρ. We then consider the following algorithm.

Algorithm 1.

- Initialization: $c = c_0$;
- Calculation of $u_0 = (u_0^1, u_0^2, u_0^3)$ and $v_0 = (v_0^1, v_0^2, v_0^3)$, solutions of the direct (2.25) and adjoint (2.28) problems;
- Computation of the 2×2 matrix $M(x)$ and its lowest eigenvalue $\lambda_{min}(x)$ at each point of the domain;
- Set

$$c_1 = \begin{cases} \varepsilon & \text{if} \quad x\in\Omega \quad \text{such} \quad \text{that} \quad \lambda_{min}(x) < \alpha < 0, \quad \varepsilon > 0, \\ c_0 & \text{elsewhere;} \end{cases} \tag{2.34}$$

- Compute u_1, solution to the problem (2.3) with $c = c_1$.

In this algorithm, $\varepsilon > 0$ is assumed to be small, and α is a negative (small) threshold.

The justification of (2.34) is the following: when λ_{min} (or in a similar way, the topological gradient g) takes large negative values, the asymptotic expansion (2.31) of the energy tells us that we are looking at very energetic zones, i.e., image edges. It is then important to preserve these pixels in the restoration process by considering a small conductivity coefficient. And in the less energetic zones of the image, one can consider a larger conductivity in order to smooth and denoise the image.

2.4. Remarks. In this subsection, we compare the topological gradient approach with the Mumford–Shah models. The formulation of Mumford and Shah [25] is based on a functional minimization, in which a piecewise smooth approximation of the noisy image and an edge set are simultaneously recovered. Using the previous notation, the Mumford–Shah functional is

$$E_{MS}(u,\sigma) = \int_{\Omega\backslash\sigma} \left(\alpha|\nabla u|^2 + \beta(u-v)^2\right)\, dx + \mathcal{H}^1(\sigma), \tag{2.35}$$

where $\mathcal{H}^1$ represents the Hausdorff one-dimensional measure in $\mathbb{R}^2$. For a given edge set σ, the minimizer u of E_{MS} satisfies

$$\begin{cases} -\mathrm{div}\,(\alpha\nabla u) + \beta u = \beta v & \text{in} \quad \Omega\backslash\sigma, \\ \partial_n u = 0 & \text{on} \quad \partial(\Omega\backslash\sigma). \end{cases} \tag{2.36}$$

Equation (2.36) is then equivalent to (2.5) with a constant parameter c.

The Ambrosio–Tortorelli approximation of E_{MS} is

$$E^{\rho}_{AT}(u,\varphi) = \int_{\Omega} \left(\rho|\nabla\varphi|^2 + \alpha\left(\varphi^2|\nabla u|^2 + \frac{(\varphi-1)^2}{4\alpha\rho}\right) + \beta(u-v)^2\right) dx, \tag{2.37}$$

where $\rho > 0$ is assumed to be small. The idea of this model is to introduce a smooth edge indicator function φ, which is an approximation of $1-\chi_\sigma$ [2, 38].

This model has been extended to color images by replacing the vector norm of the gradient $|\nabla u|$ in (2.37) by its Frobenius norm, denoted by $\|\,.\,\|_{L^2(\Omega,\mathbb{R}^3)}$ in (2.30) [11]. It is also possible to consider the geometric model of the Beltrami flow for generalizing the regularization norm [32, 11, 9]:

$$\|\nabla u\|_{BEL} = \int_{\Omega} \sqrt{\det G}\, dx, \tag{2.38}$$

where the metric tensor G is defined by

$$G = \begin{pmatrix} 1 + \sum_{k=1}^{3} \left(\frac{\partial u^k}{\partial x_1}\right)^2 & \sum_{k=1}^{3} \frac{\partial u^k}{\partial x_1}\frac{\partial u^k}{\partial x_2} \\ \sum_{k=1}^{3} \frac{\partial u^k}{\partial x_1}\frac{\partial u^k}{\partial x_2} & 1 + \sum_{k=1}^{3} \left(\frac{\partial u^k}{\partial x_2}\right)^2 \end{pmatrix}. \tag{2.39}$$

This determinant measures not only the image smoothness, but also the color channel alignment. Note that this tensor is not very different from the Di Zenzo gradient that we study in section 3.

The optimal solution is usually deduced from an alternate minimization of (2.37). If we assume that v is given, the minimization of (2.37) with respect to u leads to

$$\begin{cases} -\mathrm{div}\left(\varphi^2\frac{\alpha}{\beta}\nabla u\right) + u = v & \text{in} \quad \Omega, \\ \partial_n u = 0 & \text{on} \quad \partial\Omega, \end{cases} \tag{2.40}$$

which is the same equation as in our approach, with a conductivity given by $(\frac{\alpha}{\beta})\varphi^2$.

Note that in our approach, the idea is to use the topological gradient as an edge indicator function and to threshold it in order to get the edge set. This approach has several advantages. First, the resolution of the PDE model is much simpler when considering a piecewise constant edge indicator function (see, e.g., [21]). Second, an edge indicator function taking all values of an interval (e.g., $[0;1]$) leads to an increasing instability of the minimization process, by enlarging the set of admissible solutions. For these reasons, the variational approaches based on the Mumford–Shah functional could be seen as a relaxation of our method. We refer to section 4 for numerical comparisons between the Ambrosio–Tortorelli approach and our topological gradient algorithm.

A straightforward way to smooth a multichannel image is to diffuse each channel independently. However, different edge locations will be created in different channels. To overcome this inconsistency, we diffuse each channel with a common edge indicator function. Note that in this case, in both the topological gradient (current section) and Mumford–Shah (see, e.g., [11]) approaches, the global edge indicator is exactly the sum on all channels of the edge indicators.

3. Vectorial approach with the Di Zenzo gradient.

In general, when processing of vector-valued images is considered, there should be some coupling between the different channels; however, it is not clear what the correct coupling is. Several approaches have been considered for the extension of the Mumford–Shah functional to color images. For instance, one can consider a geometrical approach, in which the image is seen as a manifold equipped with a metric and a structure tensor is defined from the induced metric (Beltrami framework; see, e.g., [32, 11, 22]). We give some remarks about this approach in section 3.3.

We propose in this part to consider the Di Zenzo gradient for this purpose, with the aim of extending the topological gradient method to this coupled approach [41]. In order to identify the local variations of the image, Di Zenzo defines a multispectral tensor associated to the image vector field. The largest eigenvalue of the tensor will correspond to the norm of the gradient called the Di Zenzo gradient.

Throughout this section, we assume that $\nabla u \neq (0,0,0)$; otherwise the Di Zenzo approach has no interest, and one can consider the standard vectorial approach defined in the previous section.

3.1. Di Zenzo gradient.

More precisely, given an image $u = (u^1, u^2, u^3)$ defined in an open set Ω of $\mathbb{R}^2$ into $\mathbb{R}^3$, the structure tensor is given by

$$T = \begin{pmatrix} w_{11} & w_{12} \\ w_{12} & w_{22} \end{pmatrix}, \tag{3.1}$$

where

$$w_{ij} = \sum_{k=1}^{3} \frac{\partial u^k}{\partial x_i} \frac{\partial u^k}{\partial x_j}, \quad 1 \leq i, j \leq 2. \tag{3.2}$$

This tensor then describes the bidimensional first-order differential structure at a given point (x_1, x_2) in the image, and the eigenvalue analysis of this tensor gives a biggest eigenvalue

defined by

$$\lambda_{max} = \frac{w_{11} + w_{22} + \sqrt{(w_{11} - w_{22})^2 + 4w_{12}^2}}{2}. \tag{3.3}$$

We note here that, according to Di Zenzo's approach, the vector image gradient is considered as a random vector and the tensor T defines its covariance matrix. The biggest eigenvalue λ_{max} of T provides the maximal variation of the image gradient, along the direction given by the corresponding eigenvector. It is straightforward to see that the maximum variation direction is given by

$$n = \begin{pmatrix} 2w_{12} \\ w_{22} - w_{11} + \sqrt{(w_{11} - w_{22})^2 + 4w_{12}^2} \end{pmatrix}, \tag{3.4}$$

and the second eigenvector τ gives the direction of minimal variation of the image gradient. The eigenvector n corresponding to the largest eigenvalue then provides the normal to the edge. The Di Zenzo gradient is finally given by the square root of the largest eigenvalue of the structure tensor as follows:

$$\|\nabla u\|_{DZ} = \frac{1}{\sqrt{2}} \left[w_{11} + w_{22} + \sqrt{(w_{11} - w_{22})^2 + 4w_{12}^2} \right]^{\frac{1}{2}}. \tag{3.5}$$

One should note that if $w_{12}^2 = w_{11} w_{22}$, which happens, for instance, when the three components of the image are proportional, then $\|\nabla u\|_{DZ} = \sqrt{w_{11} + w_{22}} = \sqrt{|\nabla u^1|^2 + |\nabla u^2|^2 + |\nabla u^3|^2}$ $= |\nabla u|$, the standard norm of the gradient of u in $\mathbb{R}^3$.

In the following, we will denote by $|\,.\,|$ the standard norm of $\mathbb{R}^3$ and by $\|\,.\,\|$ the standard norm of $L^2(\Omega, \mathbb{R}^3)$.

In order to simplify further calculations, we will consider an equivalent formulation of the Di Zenzo gradient, given by the following lemma (see Appendix A for computational details).

Lemma 3.1. *For $u = (u^1, u^2, u^3)$ in $L^2\left(\Omega, \mathbb{R}^3\right)$, the Di Zenzo gradient is given by*

$$\|\nabla u\|_{DZ} = \sqrt{(|\nabla u^1|^2 + |\nabla u^2|^2 + |\nabla u^3|^2) \frac{1 + \sqrt{1 - 4f(\nabla u)}}{2}}, \tag{3.6}$$

where f is the function defined by

$$f(\nabla u) = \frac{det^2(\nabla u^1, \nabla u^2) + det^2(\nabla u^1, \nabla u^3) + det^2(\nabla u^2, \nabla u^3)}{(|\nabla u^1|^2 + |\nabla u^2|^2 + |\nabla u^3|^2)^2}, \tag{3.7}$$

with

$$det^2(\nabla u^s, \nabla u^t) = \left(\frac{\partial u^s}{\partial x_1} \frac{\partial u^t}{\partial x_2} - \frac{\partial u^t}{\partial x_1} \frac{\partial u^s}{\partial x_2} \right)^2 \qquad for \;\; s, t = 1, 2, 3, \tag{3.8}$$

and $|.|$ is the standard norm in $\mathbb{R}^2$.

3.2. Asymptotic expansion. The idea is the same as in the previous section: minimize the energy norm outside the edges of the image, and the cost function we now consider is the following:

$$j(\rho) = J(u_\rho) = \int_{\Omega_\rho} \|\nabla u_\rho\|^2_{DZ}\, dx, \tag{3.9}$$

where u_ρ is the solution of problem (2.8). We refer to section 2 for a precise description of the topological asymptotic analysis and for the notation. From the previous work, we can easily deduce that the cost function j has the following asymptotic expansion:

$$j(\rho) - j(0) = \rho^2 \left(\sum_{k=1}^{3} \delta a^k - \sum_{k=1}^{3} \delta l^k + \delta J_1 + \delta J_2 \right) + o(\rho^2). \tag{3.10}$$

As the bilinear a^k and linear l^k forms are kept unchanged, their variations with respect to the insertion of a small crack at point $x_0 \in \Omega$ are still given by (2.21) and (2.22). Since (2.21) takes into account the variation of the gradient of u_0^k when the crack is inserted and since the crack is of null measure, the variation of l^k is equal to zero. As the cost function changed in comparison with section 2, we have to compute δJ_1 and δJ_2. These variations are given by the following proposition.

Proposition 3.2. *We have the following results:*

$$\delta J_1 = -\pi H(\nabla u_0(x_0)) \sum_{k=1}^{3} |\nabla u_0^k(x_0).n|^2, \tag{3.11}$$

$$\delta J_2 = 0, \tag{3.12}$$

where H is given by

$$H(\nabla u_0) = \frac{1 + \sqrt{1 - 4f(\nabla u_0)}}{2} \tag{3.13}$$

and f is defined by (3.7).

The proof of Proposition 3.2 is given in Appendix B. Finally, from all the previous results, we obtain the following theorem.

Theorem 3.3. *The cost function j defined by*

$$j(\rho) = J_\rho(u_\rho) = \int_{\Omega_\rho} \|\nabla u_\rho\|^2_{DZ}\, dx, \tag{3.14}$$

where u_ρ^k are solutions of the perturbed problems

$$\begin{cases} -\mathrm{div}(c\nabla u_\rho^k) + u_\rho^k = v^k & in \quad \Omega_\rho, \\ \partial_n u_\rho^k = 0 & on \quad \partial\Omega_\rho, \end{cases} \tag{3.15}$$

has the asymptotic expansion

$$j(\rho) - j(0) = \rho^2 G(x_0, n) + o(\rho^2), \tag{3.16}$$

with

$$G(x_0,n)=\sum_{k=1}^{3}\left[-\pi c(\nabla u_0^k(x_0).n)(\nabla v_0^k(x_0).n)-\pi H(\nabla u_0(x_0))\left|\nabla u_0^k(x_0).n\right|^2\right], \tag{3.17}$$

where u_0^k *are solutions of the direct problems*

$$\begin{cases} -\mathrm{div}(c\nabla u_0^k)+u_0^k=v^k & \text{in} \quad \Omega,\\ \partial_n u_0^k=0 & \text{on} \quad \partial\Omega \end{cases} \tag{3.18}$$

and v_0^k *are solutions of the adjoint problems*

$$\begin{cases} -\mathrm{div}(c\nabla v_0^k)+v_0^k=-\partial_{u^k}J(u_0) & \text{in} \quad \Omega,\\ \partial_n v_0^k=0 & \text{on} \quad \partial\Omega. \end{cases} \tag{3.19}$$

As in the previous section, the topological gradient can be written as

$$G(x,n)=\langle M(x)n,n\rangle, \tag{3.20}$$

where $M(x)$ is the 2×2 symmetric matrix given by

$$M(x)=\sum_{k=1}^{3}\left[-\pi c\frac{\nabla u_0^k(x)\nabla v_0^k(x)^T+\nabla v_0^k(x)\nabla u_0^k(x)^T}{2}-\pi H(\nabla u_0(x))\nabla u_0^k(x)\nabla u_0^k(x)^T\right]. \tag{3.21}$$

The restoration algorithm is then the following.

Algorithm 2.

- Initialization: $c=c_0$;
- Calculation of u_0^k and v_0^k, solutions of the direct (3.18) and adjoint (3.19) problems for each k;
- Computation of the 2×2 matrix $M(x)$ defined by (3.21) and its lowest eigenvalue $\lambda_{min}(x)$ at each point of the domain;
- Set

$$c_1=\begin{cases} \varepsilon & \text{if} \quad x\in\Omega \quad \text{such} \quad \text{that} \quad \lambda_{min}(x)<\alpha<0, \quad \varepsilon>0,\\ c_0 & \text{elsewhere;} \end{cases} \tag{3.22}$$

- Compute u_1^k, solutions to the problems (3.18) with $c=c_1$ for each k;
- Set $u_1=(u_1^1,u_1^2,u_1^3)$.

3.3. Remarks. In this section, we briefly present the Beltrami framework for processing multichannel images and compare this approach with ours. The idea of the Beltrami framework is to consider a distance function represented by the following 2×2 structure tensor g:

$$g=\begin{pmatrix} g_{11} & g_{12}\\ g_{21} & g_{22}\end{pmatrix}, \tag{3.23}$$

where

$$g_{11} = 1 + \left(\frac{\partial u^1}{\partial x_1}\right)^2 + \left(\frac{\partial u^2}{\partial x_1}\right)^2 + \left(\frac{\partial u^3}{\partial x_1}\right)^2, \tag{3.24}$$

$$g_{12} = g_{21} = \frac{\partial u^1}{\partial x_1}\frac{\partial u^1}{\partial x_2} + \frac{\partial u^2}{\partial x_1}\frac{\partial u^2}{\partial x_2} + \frac{\partial u^3}{\partial x_1}\frac{\partial u^3}{\partial x_2}, \tag{3.25}$$

$$g_{22} = 1 + \left(\frac{\partial u^1}{\partial x_2}\right)^2 + \left(\frac{\partial u^2}{\partial x_2}\right)^2 + \left(\frac{\partial u^3}{\partial x_2}\right)^2. \tag{3.26}$$

Note that $g = T + Id$, where T is the structure tensor in the Di Zenzo approach (see (3.1) and (3.2)) and Id is the 2×2 identity matrix. Then, in the geometric approach, the image evolves via the Beltrami flow

$$\partial_t u = \Delta_g u, \tag{3.27}$$

where Δ_g is the Beltrami operator [23].

In the Di Zenzo approach, the edge set is defined according to the square root of the largest eigenvalue of the structure tensor T, whereas in the Beltrami framework, the idea is to control the diffusion process along and across the gradient of the image via an anisotropic tensor. These two approaches have the same goal—to extend the gradient operator and diffusion process to multichannel images—and they can be seen as natural extensions to color images of both the classical topological gradient approach and the Ambrosio–Tortorelli version of the Mumford–Shah functional. A nice future application of the topological gradient would be to consider the gradient norm defined by the square root of the Beltrami operator and to try to derive a corresponding topological gradient. The topological derivation with respect to the Di Zenzo structure tensor could be in some sense a preliminary work for studying the Beltrami framework, which is a bit more complicated.

4. Numerical experiments and comparisons.

4.1. Standard multichannel approaches. We first consider the natural extension of the topological gradient method to multichannel images, using the standard gradient norm (see section 2), and we compare this approach with the Ambrosio–Tortorelli version of the Mumford–Shah functional (see section 2.4). For several images, we compare the restored images, the signal-to-noise ratios (SNRs), and the CPU time of the computations.

Figure 1 shows for several test cases the original image, the noisy image (with the corresponding SNR), and the two restored images using, respectively, the topological gradient algorithm (see section 2) and the Ambrosio–Tortorelli algorithm (based on an approximation of the Mumford–Shah functional; see section 2.4), with their corresponding SNRs and computing times (on the same computer, using the same solver functions).

We can see that these two algorithms provide almost the same kind of results: visually equivalent and with similar SNRs. If we look in more detail, we see that the Ambrosio–Tortorelli algorithm provides better results where there is a low level of noise (e.g., in the second test case) or when the original image is already a bit noisy (e.g., in the fourth case). When the level of noise is increased, the topological gradient algorithm provides slightly better results. But in all cases, the computing time of the topological gradient method is smaller, with a speed-up factor of nearly 25%.

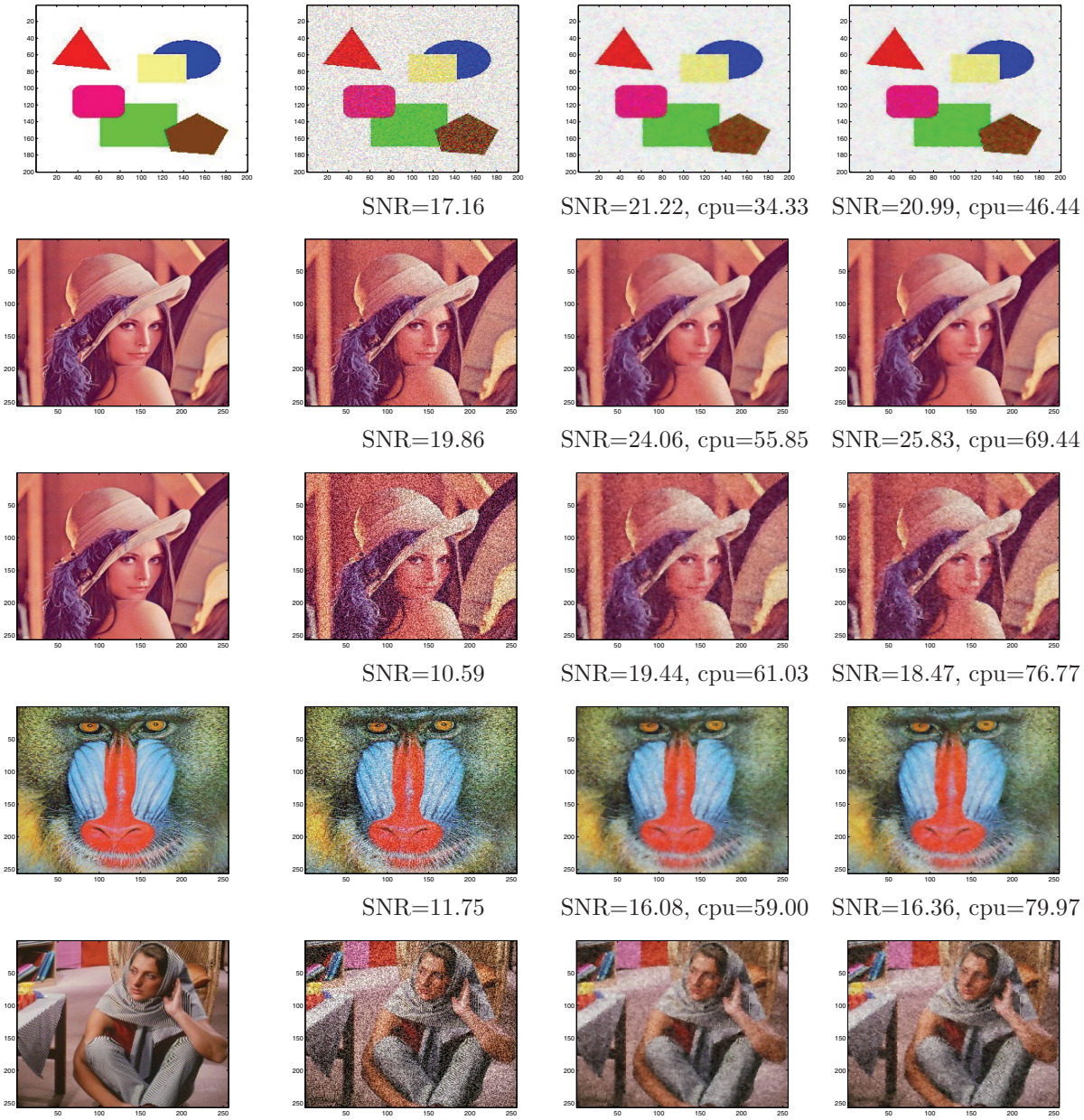
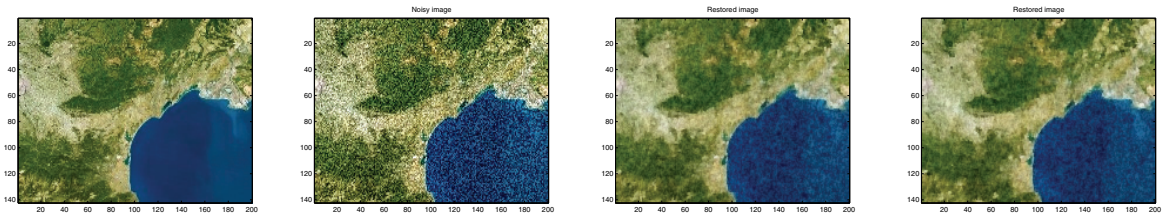


Figure 1. *From left to right: original image, noisy image, and restored images by the topological gradient and the Ambrosio–Tortorelli algorithm, respectively.*



SNR=10.34 SNR=16.29, cpu=25.67 SNR=16.57, cpu=26.64

Figure 2. *From left to right: original image, noisy image, and restored images by the topological gradient algorithm with the standard and the Di Zenzo gradient, respectively.*

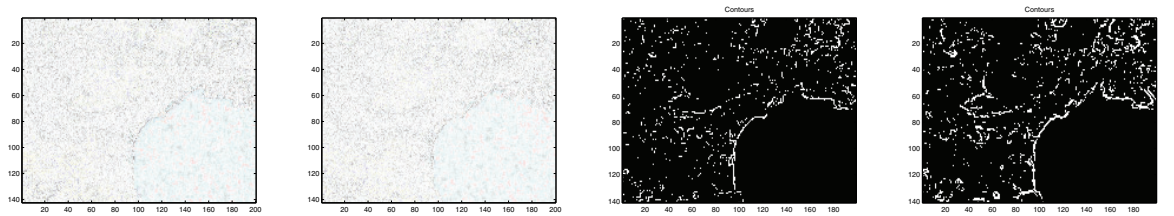


Figure 3. *From left to right: reconstruction error, using the standard and the Di Zenzo gradients in the topological gradient approach; corresponding identified edges.*

4.2. Improvement of our algorithm by using the Di Zenzo gradient. We now study the impact on our algorithm of considering the Di Zenzo gradient instead of the standard gradient for channel coupling.

Figure 2 shows the test image we have considered for numerical experiments. The original image is shown on the left, and the corresponding noisy image (shown on the right) has been perturbed with an additive Gaussian noise on each channel. We applied both topological gradient algorithms to this noisy image. Figure 2 then shows the two corresponding restored images: on the left, with the standard vectorial approach (Algorithm 1), and, on the right, with the Di Zenzo vectorial approach (Algorithm 2). In all the numerical experiments we have performed, the SNR of the restored image with the Di Zenzo vectorial approach was a little bit larger, of 0.1 to 0.3 dB. From a visual point of view, the restored images are very similar, but if we look carefully at the edges, for instance, between the blue and white/green parts, we can observe that they are quite blurred on the left, and a little bit more enhanced on the right. Note that the Di Zenzo approach does not degrade the computing time.

In order to quantify this remark, Figure 3 shows in the two left images the reconstruction error for both algorithms (standard and Di Zenzo gradients, respectively), i.e., the difference between the restored image and the original image. These differences have been multiplied by 3 in order to see something in this figure, but the scale is the same for the two images. We can see in this figure that there is some accumulation of the reconstruction error on the main edges of the image in the standard approach (Algorithm 1). On the contrary, with the tensor approach (Algorithm 2), the error is a little less located on the edges and is more uniformly distributed.

The explanation of this phenomenon is given in the right side of Figure 3, where the

identified edges by the topological gradient are given for both approaches. It is now clear that the main edges are better identified by the tensorial approach (on the right) than by the standard approach (on the left). Consequently, these edges are more preserved in the tensorial restoration process, whereas they are partially blurred in the standard restoration process. This better detection of the main edges with the Di Zenzo gradient is well known, because it has been precisely defined in order to identify the points where all components of the gradient vary along similar directions.

5. Conclusion. In this paper, we have considered the topological gradient approach for color image restoration. For this purpose, we have considered two different approaches. The first is an extension to color images of the method introduced in [8, 21] for grey-level image processing. The second approach uses the Di Zenzo gradient, a tensor measuring the relative variations of the different components of the image, leading to a better identification of the main edges (for which all the components of the image vary along the same direction).

We have compared the topological gradient method with the Ambrosio–Tortorelli approximation of the Mumford–Shah model, and we have shown that our algorithm provides similar results in a shorter time. The topological gradient algorithm seems better when there is a high level of noise, whereas the Ambrosio–Tortorelli method usually provides better results when there is little noise.

We have shown that the tensorial approach can also be considered within the topological asymptotic analysis, and the corresponding asymptotic expansion has been derived. The numerical tests show that the tensorial approach slightly improves the restoration process, in comparison with the standard approach. The main differences are on the edges, as the tensorial approach provides a much better identification of the main image contours. As a consequence, the restored image has less blurred edges. But of course, in all other parts of the image, the Di Zenzo gradient is not very different from the standard image gradient, and hence the restoration process is almost similar. This explains the small difference in the reconstruction SNRs. Globally, the tensorial approach is more satisfactory, as it removes a similar level of noise and preserves more efficiently the global features of a color image. Finally, we mention that the complexity of both algorithms is the same, and, according to the work performed on the grey-level problems, the complexity is $\mathcal{O}(n.\log(n))$, where n is the size of the image.

Some extensions to 3D images (or movies) will be considered in a future work, as well as some other ways to couple the different channels of the image, e.g., by using a gradient based on the Beltrami operator. Also, inspired by recent works on multiplicative noise removal (see, e.g., [6]), we will study the application of the topological asymptotic analysis to variational models designed for multiplicative Gaussian noise.

Appendix A. Proof of Lemma 3.1.

Proof. For each $k = 1, 2, 3$, we consider the following notation:

$$u_{kj} = \frac{\partial u^k}{\partial x_j}, \quad 1 \leq j \leq 2.$$

Then, according to (3.2), we have

$$w_{11} + w_{22} = \left(u_{11}^2 + u_{12}^2\right) + \left(u_{21}^2 + u_{22}^2\right) + \left(u_{31}^2 + u_{32}^2\right) \tag{A.1}$$

and

$$
\begin{aligned}
(w_{11}-w_{22})^2+4w_{12}^2 &= \left(\left(u_{11}^2+u_{21}^2+u_{31}^2\right)-\left(u_{12}^2+u_{22}^2+u_{32}^2\right)\right)^2 \\
&\quad +4\left(u_{11}u_{12}+u_{21}u_{22}+u_{31}u_{32}\right)^2 \\
&= 2u_{11}^2u_{21}^2+2u_{11}^2u_{31}^2+2u_{11}^2u_{12}^2-2u_{11}^2u_{22}^2-2u_{11}^2u_{32}^2 \\
&\quad +2u_{21}^2u_{31}^2-2u_{21}^2u_{12}^2+2u_{21}^2u_{22}^2-2u_{21}^2u_{32}^2-2u_{31}^2u_{12}^2 \\
&\quad -2u_{31}^2u_{22}^2+2u_{31}^2u_{32}^2+2u_{12}^2u_{22}^2+2u_{12}^2u_{32}^2+2u_{22}^2u_{32}^2 \\
&\quad +u_{11}^4+u_{21}^4+u_{31}^4+u_{12}^4+u_{22}^4+u_{32}^4+8u_{11}u_{12}u_{21}u_{22} \\
&\quad +8u_{11}u_{12}u_{31}u_{32}+8u_{21}u_{22}u_{31}u_{32}.
\end{aligned} \tag{A.2}
$$

On the other hand, using (3.7) and (3.8), we obtain

$$
f\left(\nabla u\right)=\frac{\left(u_{11}u_{22}-u_{21}u_{12}\right)^2+\left(u_{11}u_{32}-u_{31}u_{12}\right)^2+\left(u_{21}u_{32}-u_{31}u_{22}\right)^2}{\left(\left(u_{11}^2+u_{12}^2\right)+\left(u_{21}^2+u_{22}^2\right)+\left(u_{31}^2+u_{32}^2\right)\right)^2},
$$

which permits us to determine the expression of the relation

$$
\left(1-4f\left(\nabla u\right)\right)\left(\left(u_{11}^2+u_{12}^2\right)+\left(u_{21}^2+u_{22}^2\right)+\left(u_{31}^2+u_{32}^2\right)\right)^2,
$$

given by the following computation:

$$
\begin{aligned}
&\left(u_{11}^2+u_{12}^2+u_{21}^2+u_{22}^2+u_{31}^2+u_{32}^2\right)^2-4\left(u_{11}u_{22}-u_{21}u_{12}\right)^2 \\
&-4\left(u_{11}u_{32}-u_{31}u_{12}\right)^2-4\left(u_{21}u_{32}-u_{31}u_{22}\right)^2 \\
&=2u_{11}^2u_{12}^2+2u_{11}^2u_{21}^2-2u_{11}^2u_{22}^2+2u_{11}^2u_{31}^2-2u_{11}^2u_{32}^2-2u_{12}^2u_{21}^2+2u_{12}^2u_{22}^2 \\
&-2u_{12}^2u_{31}^2+2u_{12}^2u_{32}^2+2u_{21}^2u_{22}^2+2u_{21}^2u_{31}^2-2u_{21}^2u_{32}^2-2u_{22}^2u_{31}^2+2u_{22}^2u_{32}^2+2u_{31}^2u_{32}^2 \\
&+u_{11}^4+u_{12}^4+u_{21}^4+u_{22}^4+u_{31}^4+u_{32}^4+8u_{11}u_{22}u_{21}u_{12}+8u_{11}u_{32}u_{31}u_{12}+8u_{21}u_{32}u_{31}u_{22}.
\end{aligned}
$$

From relation (A.2), we deduce that

$$
(w_{11}-w_{22})^2+4w_{12}^2=\left(1-4f\left(\nabla u\right)\right)\left(\left|\nabla u_1\right|^2+\left|\nabla u_2\right|^2+\left|\nabla u_3\right|^2\right)^2.
$$

Finally, using (A.1), we obtain the desired result. ■

Appendix B. Proof of Proposition 3.2. In order to prove (3.11), we need the following three lemmas, assuming $\nabla u \neq 0$. Otherwise, by extending the function H with the notation $H(0)=1$, the result clearly remains valid.

Lemma B.1. *For all $k=1,2,3$,*

$$
E_1=\int_{\sigma_\rho} H(\nabla u_0(x))\partial_n u_0^k[u_\rho^k]d\Gamma=\pi\rho^2 H(\nabla u_0(x_0))|\nabla u_0^k(x_0).n|^2+o(\rho^2), \tag{B.1}
$$

where $[u_\rho^k]=u_\rho^k|_{\sigma_\rho^+}-u_\rho^k|_{\sigma_\rho^-}$ represents the variation of u_ρ^k along the normal direction of the crack.

Proof. We have

$$E_1 = \int_{\sigma_\rho} \left(H(\nabla u_0(x)) - H(\nabla u_0(x_0))\right) \partial_n u_0^k[u_\rho^k] d\Gamma + H(\nabla u_0(x_0)) \int_{\sigma_\rho} \partial_n u_o^k[u_\rho^k] d\Gamma.$$

Or from [3], we have

$$\int_{\sigma_\rho} \partial_n u_0^k[u_\rho^k] d\Gamma = \pi\rho^2 |\nabla u_0^k(x_0).n|^2 + o(\rho^2); \tag{B.2}$$

then we deduce that

$$E_1 = \int_{\sigma_\rho} \left(H(\nabla u_0(x)) - H(\nabla u_0(x_0))\right) \partial_n u_0^k[u_\rho^k] d\Gamma + H(\nabla u_0(x_0)) \left(\pi\rho^2 |\nabla u_0^k(x_0).n|^2 + o(\rho^2)\right).$$

Now, we have to show that the first term is $o(\rho^2)$. For that we consider the change of variable $x = x_0 + \rho y$, which gives

$$\begin{aligned} &\int_{\sigma_\rho} \left(H(\nabla u_0(x)) - H(\nabla u_0(x_0))\right) \partial_n u_0^k[u_\rho^k] d\Gamma \\ &\quad = \rho \int_{\sigma} [H(\nabla u_0(\rho y)) - H(\nabla u_0(x_0))] \partial_n u_0^k(\rho y)[u_\rho^k(\rho y)] d\Gamma. \end{aligned}$$

On the other hand, from [3] we have

$$\int_{\sigma} \left|\partial_n u_0^k(\rho y)[u_\rho^k(\rho y)]\right| d\Gamma = O(\rho).$$

As u_0 is of class $\mathcal{C}^2$ in a neighborhood of x_0, and as the gradient of u_0 is not equal to 0 in a neighborhood of x_0, the differentiability of H gives the existence of $c_1 > 0$ such that

$$|H(\nabla u_0(\rho y)) - H(\nabla u_0(x_0))| \leq c_1 \rho,$$

which implies that

$$\int_{\sigma} [H(\nabla u_0(\rho y)) - H(\nabla u_0(x_0))] \partial_n u_0^k(\rho y)[u_\rho^k(\rho y)] d\Gamma = O(\rho^2).$$

We deduce that

$$\int_{\sigma_\rho} \left(H(\nabla u_0(x)) - H(\nabla u_0(0))\right) \partial_n u_0^k[u_\rho^k] d\Gamma = o(\rho^2),$$

which completes the proof. ∎

Lemma B.2. *For all* $k = 1, 2, 3$,

$$E_2 = \int_{\Omega_\rho} \left(H(\nabla u_\rho(x)) - H(\nabla u_0(x))\right) \left|\nabla u_\rho^k\right|^2 dx = L_\rho^k(u_\rho - u_0) + o(\rho^2), \tag{B.3}$$

where L_ρ^k *is a linear form on* $L^2(\Omega_\rho)$.

Proof. By using a Taylor asymptotic expansion, we obtain

$$H(\nabla u_\rho(x)) - H(\nabla u_0(x)) = DH(\nabla u_0(x))(\nabla(u_\rho - u_0)(x)) + |\nabla(u_\rho - u_0)|\,\epsilon(\nabla(u_\rho - u_0)),$$

with $\epsilon(Z) \to 0$ when $Z \to 0$. Then we can write that

$$E_2 = A_1 + A_2,$$

with

$$A_1 = \int_{\Omega_\rho} DH(\nabla u_0(x))(\nabla(u_\rho - u_0)(x)) \left|\nabla u_\rho^k\right|^2 dx$$

and

$$A_2 = \int_{\Omega_\rho} \|\nabla(u_\rho - u_0)\|\,\epsilon(\nabla(u_\rho - u_0)) \left|\nabla u_\rho^k\right|^2 dx.$$

As

$$\|\nabla(u_\rho - u_0)\|_{L^2(\Omega_\rho)} = O(\rho^2), \tag{B.4}$$

it follows that

$$\begin{aligned} |A_2| &\le \int_{\Omega_\rho} \|\nabla(u_\rho - u_0)\|\,|\epsilon(\nabla(u_\rho - u_0))| \left|\nabla u_\rho^k\right|^2 dx \\ &\le O(\rho^2) \int_{\Omega_\rho} |\epsilon(\nabla(u_\rho - u_0))| \left|\nabla u_\rho^k\right|^2 dx. \end{aligned} \tag{B.5}$$

On the other hand, since $u_\rho^k \in H^2(\Omega_\rho)$, $|\nabla u_\rho^k|$ is L^2-bounded in Ω_ρ. Moreover, $\nabla(u_\rho - u_0) \to 0$ when $\rho \to 0$; then $\epsilon(\nabla(u_\rho - u_0)) \to 0$ when $\rho \to 0$. We deduce that

$$|A_2| \le O(\rho^2)\,\|\epsilon(\nabla(u_\rho - u_0))\|_{L^\infty(\Omega_\rho)} \left\|\nabla u_\rho^k\right\|_{L^2}^2 \le O(\rho^2)\ o(1)\ O(1). \tag{B.6}$$

From (B.6), we obtain

$$|A_2| = o(\rho^2).$$

Now we consider A_1:

$$\begin{aligned} A_1 = &\int_{\Omega_\rho} DH(\nabla u_0(x))(\nabla(u_\rho - u_0)(x)) \left|\nabla u_0^k\right|^2 dx \\ &+ \int_{\Omega_\rho} DH(\nabla u_0(x))(\nabla(u_\rho - u_0)(x)) \left(\left|\nabla u_\rho^k\right|^2 - \left|\nabla u_0^k\right|^2\right) dx. \end{aligned}$$

We denote these two terms by B_1 and B_2, respectively.

By continuity of $DH(\nabla u_0)$, we have

$$|B_2| \le M\,||\nabla(u_\rho - u_0)||_\infty \int_{\Omega_\rho} \left|\nabla u_\rho^k\right|^2 - \left|\nabla u_0^k\right|^2 dx.$$

Using (B.4), we deduce that

$$\int_{\Omega_\rho} \left| \left|\nabla u_\rho^k\right|^2 - \left|\nabla u_0^k\right|^2 \right| dx \leq \int_{\Omega_\rho} |\nabla(u_\rho^k - u_0^k)|\, |\nabla(u_\rho^k + u_0^k)|\, dx \leq O(\rho^2).$$

Moreover, as $\|\nabla(u_\rho - u_0)\|_\infty = o(1)$, we deduce that

$$|B_2| = o(\rho^2).$$

Finally, since B_1 is a linear form of $u_\rho - u_0$, we obtain

$$\int_{\Omega_\rho} DH(\nabla u_0(x))(\nabla(u_\rho - u_0)(x)) \left|\nabla u_0^k\right|^2 dx = L_\rho^k(u_\rho - u_0). \qquad \blacksquare$$

Lemma B.3. *For all $k = 1, 2, 3$,*

$$E_3 = \int_{\Omega_\rho} \operatorname{div}\left(H(\nabla u_0(x))\nabla(u_\rho^k - u_0^k)\right)(u_\rho^k - u_0^k)\, dx = o(\rho^2). \tag{B.7}$$

Proof. We have

$$\begin{aligned} E_3 &= \int_{\Omega_\rho} H(\nabla u_0(x))\Delta(u_\rho^k - u_0^k)(u_\rho^k - u_0^k)\, dx \\ &\qquad + \int_{\Omega_\rho} \nabla(H(\nabla u_0(x))).\nabla(u_\rho^k - u_0^k)\,(u_\rho^k - u_0^k)\, dx \\ &= \int_{\Omega_\rho} H(\nabla u_0(x))\,\frac{1}{c}\,(u_\rho^k - u_0^k)^2\, dx + \int_{\Omega_\rho} \nabla(H(\nabla u_0(x))).\nabla(u_\rho^k - u_0^k)\,(u_\rho^k - u_0^k)\, dx \end{aligned}$$

from (2.3) and (2.5) (recall that c is constant on Ω_ρ). As $H(\nabla u_0)$ is of class C^1 on Ω_ρ,

$$|E_3| \leq M_1 \|u_\rho^k - u_0^k\|^2_{L^2(\Omega_\rho)} + M_2 \|u_\rho^k - u_0^k\|_{L^2(\Omega_\rho)}\, \|\nabla(u_\rho^k - u_0^k)\|_{L^2(\Omega_\rho)}.$$

From (B.4) and Poincaré's inequality, we have $|E_3| = o(\rho^2)$. $\blacksquare$

We now consider the proof of Proposition 3.2.

Proof. We have

$$\begin{aligned} J_\rho(u_\rho) - J_\rho(u_0) &= \int_{\Omega_\rho} \|\nabla u_\rho\|^2_{DZ}\, dx - \int_{\Omega_\rho} \|\nabla u_0\|^2_{DZ}\, dx \\ &= \sum_{k=1}^{3} \left[\int_{\Omega_\rho} H(\nabla u_\rho)\, |\nabla u_\rho^k|^2\, dx - \int_{\Omega_\rho} H(\nabla u_0)\, |\nabla u_0^k|^2\, dx \right] \\ &= \sum_{k=1}^{3} \left[\int_{\Omega_\rho} (H(\nabla u_\rho) - H(\nabla u_0))\, |\nabla u_\rho^k|^2\, dx + \int_{\Omega_\rho} H(\nabla u_0) \left(|\nabla u_\rho^k|^2 - |\nabla u_0^k|^2\right) dx \right] \\ &= \sum_{k=1}^{3} \left[\int_{\Omega_\rho} H(\nabla u_0) \left(|\nabla u_\rho^k|^2 - |\nabla u_0^k|^2\right) dx + L_\rho^k(u_\rho - u_0) + o(\rho^2) \right] \end{aligned}$$

according to Lemma B.2. Moreover,

$$\int_{\Omega_\rho} H(\nabla u_0)\,(|\nabla u_\rho^k|^2 - |\nabla u_0^k|^2)\,dx$$
$$= 2\int_{\Omega_\rho} H(\nabla u_0)\,\nabla(u_\rho^k - u_0^k)\,\nabla u_0^k\,dx + \int_{\Omega_\rho} H(\nabla u_0)\,|\nabla u_\rho^k - \nabla u_0^k|^2\,dx.$$

The first term is clearly a linear function of $u_\rho^k - u_0^k$. Green's formula on the second term gives

$$\int_{\Omega_\rho} H(\nabla u_0)\,|\nabla u_\rho^k - \nabla u_0^k|^2\,dx$$
$$= -\int_{\Omega\rho} \operatorname{div}\left(H(\nabla u_0(x))\nabla(u_\rho^k - u_0^k)\right)\,(u_\rho^k - u_0^k)\,dx - \int_{\sigma_\rho} H(\nabla u_0(x))\partial_n u_o^k[u_\rho^k]\,d\Gamma$$
$$= -\pi\rho^2 H(\nabla u_0(x_0))|\nabla u_0^k(x_0).n|^2 + o(\rho^2)$$

according to Lemmas B.1 and B.3. Finally, we obtain

$$J_\rho(u_\rho) - J_\rho(u_0) = L_\rho(u_\rho - u_0) - \pi\rho^2 H(\nabla u_0(x_0))\sum_{k=1}^{3}|\nabla u_0^k(x_0).n|^2 + o(\rho^2),$$

where

$$L_\rho(u_\rho - u_0) = \sum_{k=1}^{3}\left[L_\rho^k(u_\rho - u_0) + 2\int_{\Omega_\rho} H(\nabla u_0)\,\nabla(u_\rho^k - u_0^k)\,\nabla u_0^k\,dx\right]$$

is the corresponding linear form. Thus, we deduce (3.11).

From (2.12) and the proof of Theorem 2.1 in the channel by channel case, it is straightforward to deduce that δJ_2 is equal to zero. ∎

Acknowledgment. The authors would like to thank the referees for their useful comments and suggestions.

REFERENCES

[1] G. Allaire, F. Jouve, and A. M. Taoder, *A level set method for shape optimization*, C. R. Acad. Sci. Paris, 334 (2002), pp. 1125–1130.

[2] L. Ambrosio and V. Tortorelli, *Approximation of functionals depending on jumps by elliptic functionals via* Γ*-convergence*, Comm. Pure Appl. Math., 43 (1990), pp. 999–1036.

[3] S. Amstutz, I. Horchani, and M. Masmoudi, *Crack detection by the topological gradient method*, Control Cybernet., 34 (2005), pp. 119–138.

[4] J. Angulo and J. Serra, *Morphological color size distributions for classification and retrieval*, in Proceedings of the ACIVS 2002 (Advanced Concepts for Intelligent Vision Systems), Ghent, Belgium, 2002, pp. 46–53.

[5] J. Angulo and J. Serra, *Color segmentation by ordered mergings*, In Proceedings of the IEEE International Conference on Image Processing (ICIP'03), Vol. II, Barcelona, Spain, 2003, pp. 125–128.

[6] G. Aubert and J.-F. Aujol, *A variational approach to removing multiplicative noise*, SIAM J. Appl. Math., 68 (2008), pp. 925–946.

[7] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing*, Appl. Math. Sci. 147, Springer-Verlag, New York, 2002.

[8] D. Auroux, L. Jaafar Belaid, and M. Masmoudi, *A topological asymptotic analysis for the regularized grey-level image classification problem*, M2AN Math. Model. Numer. Anal., 41 (2007), pp. 607–625.
[9] L. Bar, A. Brook, N. Sochen, and N. Kiryati, *Deblurring of color images corrupted by impulsive noise*, IEEE Trans. Image Process, 16 (2007), pp. 1101–1111.
[10] P. Blomgren and T. Chan, *Color TV: Total variation methods for restoration of vector-valued images*, IEEE Trans. Image Process, 7 (1998), pp. 304–309.
[11] A. Brook, R. Kimmel, and N. Sochen, *Variational restoration and edge detection for color images*, J. Math. Imaging Vision, 18 (2003), pp. 247–268.
[12] T. Chan, S. H. Kang, and J. Shen, *Total variation denoising and enhancement of color images based on the CB and HSV color models*, J. Vis. Commun. Image Rep., 12 (2001), pp. 422–435.
[13] T. F. Chan and J. Shen, *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*, SIAM, Philadelphia, 2005, pp. 207–244.
[14] M. Elad and M. Aharon, *Image denoising via sparse and redundant representations over learned dictionaries*, IEEE Trans. Image Process., 15 (2006), pp. 3736–3745.
[15] E. Erdem and S. Tari, *Mumford-Shah regularizer with contextual feedback*, J. Math. Imaging Vision, 33 (2009), pp. 67–84.
[16] S. Garreau, P. Guillaume, and M. Masmoudi, *The topological asymptotic for PDE systems: The elasticity case*, SIAM J. Control Optim., 39 (2001), pp. 1756–1778.
[17] Ph. Guillaume and K. Sid Idris, *The topological asymptotic expansion for the Dirichlet problem*, SIAM J. Control Optim., 41 (2002), pp. 1042–1072.
[18] Ph. Guillaume and K. Sid Idris, *Topological sensitivity and shape optimization for the Stokes equations*, SIAM J. Control Optim., 43 (2004), pp. 1–31.
[19] R. Gonzalez and R. Wood, *Digital Image Processing*, 3rd ed., Prentice–Hall, Upper Saddle River, NJ, 2008.
[20] M. Hassine and M. Masmoudi, *The topological asymptotic expansion for the quasi-Stokes problem*, ESAIM Control Optim. Calc. Var., 10 (2004), pp. 478–504.
[21] L. Jaafar Belaid, M. Jaoua, M. Masmoudi, and L. Siala, *Image restoration and edge detection by topological asymptotic expansion*, C. R. Acad. Sci. Paris, 342 (2006), pp. 313–318.
[22] R. Kimmel, R. Malladi and N. Sochen, *Image processing via the Beltrami operator*, in Proceedings of the Third Asian Conference on Computer Vision, Hong Kong, 1998, pp. 574–581.
[23] E. Kreyszing, *Differential Geometry*, Dover, New York, 1991.
[24] M. Masmoudi, *The topological asymptotic*, in Computational Methods for Control Applications, GAKUTO Internat. Ser. Math. Sci. Appl. 16, R. Glowinski, H. Kawarada, and J. Périaux, eds., Gakkōtosho, Tokyo, Japan, 2001, pp. 53–72.
[25] D. Mumford and J. Shah, *Optimal approximations by piecewise smooth functions and associated variational problems*, Comm. Pure Appl. Math., 42 (1989), pp. 577–685.
[26] A. Novotny, R. Feijoo, C. Padra, and E. Taroco, *Derivada topologica via analise de sensibilidade a mudanca de forma na otimizacao topologica*, Rev. Internac. Métod. Numér. Cálc. Diseñ. Ingr., 18 (2002), pp. 499–519.
[27] N. Paragios and R. Deriche, *Geodesic active regions and level set methods for supervised texture segmentation*, Int. J. Comput. Vision, 46 (2002), pp. 223–247.
[28] B. Samet, S. Amstutz, and M. Masmoudi, *The topological asymptotic for the Helmholtz equation*, SIAM J. Control Optim., 42 (2003), pp. 1523–1544.
[29] C. Samson, L. Blanc-Féraud, G. Aubert, and J. Zerubia, *A variational model for image classification and restoration*, IEEE Trans. Pattern Anal. Mach. Intell., 22 (2000), pp. 460–472.
[30] G. Sapiro and P. L. Ringach, *Anisotropic diffusion of multivalued images with applications to color filtering*, IEEE Trans. Image Process., 5 (1996), pp. 1582–1585.
[31] J. A. Sethian, *Level Set Methods and Fast Marching Methods. Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Sciences*, Cambridge Monogr. Appl. Comput. Math. 3, Cambridge University Press, Cambridge, UK, 1999.
[32] N. A. Sochen, G. Gilboa, and Y. Y. Zeevi, *Color image enhancement by a forward-and-backward adaptive Beltrami flow*, in Proceedings of the Second International Workshop on Algebraic Frames for the Perception-Action Cycle (AFPAC 2000), Lecture Notes in Comput. Sci. 1888, G. Sommer and Y. Y. Zeevi, eds., Springer-Verlag, Berlin, 2000, pp. 319–328.

[33] J. Sokolowski and A. Zochowski, *Topological derivatives of shape functionals for elasticity systems*, Internat. Ser. Numer. Math., 139 (2002), pp. 231–244.
[34] B. Tang, G. Sapiro, and V. Caselles, *Color image enhancement via chromaticity diffusion*, IEEE Trans. Image Process., 10 (2001), pp. 701–707.
[35] D. Tschumperlé, *PDE's Based Regularization of Multivalued Images and Applications*, Ph.D. thesis, Université de Nice Sophia Antipolis, Nice, France, 2002.
[36] D. Tschumperlé and R. Deriche, *Diffusion PDE's on vector-valued images Local approach and geometric viewpoint*, IEEE Signal Processing Magazine, 19 (2002), pp. 16–25.
[37] D. Tschumperlé and R. Deriche, *Vector-valued image regularization with PDEs: A common framework for different applications*, IEEE Trans. Pattern Anal. Mach. Intell., 27 (2005), pp. 506–517.
[38] L. Vese and T. Chan, *Reduced Non-Convex Functional Approximations for Image Restoration and Segmentation*, CAM report 97-56, UCLA, Los Angeles, 1997.
[39] J. Weickert, *Anisotropic Diffusion in Image Processing*, Ph.D. thesis, University of Kaiserslautern, Kaiserslautern, Germany, 1996.
[40] J. Weickert, *Anisotropic Diffusion in Image Processing*, ECMI Series, Teubner, Stuttgart, Germany, 1998.
[41] S. Di Zenzo, *A note on the gradient of a multi-image*, Comput. Vision Graph. Image Process., 33 (1986), pp. 116–125.