nD variational restoration of curvilinear structures with prior-based directional regularization

Odyssée Merveille, Benoît Naegel, Hugues Talbot, Nicolas Passat

nD variational restoration of curvilinear structures with prior-based directional regularization

Odyssée Merveille, Benoît Naegel, Hugues Talbot, Nicolas Passat

***Abstract*—Curvilinear structure restoration in image processing procedures is a difficult task, which can be compounded when these structures are thin, i.e. when their smallest dimension is close to the resolution of the sensor. Many recent restoration methods involve considering a local gradient-based regularization term as prior, assuming gradient sparsity. An isotropic gradient operator is typically not suitable for thin curvilinear structures, since gradients are not sparse for these. In this article, we propose a mixed gradient operator that combines a standard gradient in the isotropic image regions, and a directional gradient in the regions where specific orientations are likely. In particular, such information can be provided by curvilinear structure detectors (e.g. RORPO or Frangi filters). Our proposed mixed gradient operator, that can be viewed as a companion tool of such detectors, is proposed in a discrete framework and its formulation / computation holds in any dimension; in other words, it is valid in $\mathbb{Z}^n$, $n \geq 1$. We show how this mixed gradient can be used to construct image priors that take edge orientation as well as intensity into account, and then involved in various image processing tasks while preserving curvilinear structures. Experiments carried out on 2D, 3D, real and synthetic images illustrate the relevance of the proposed gradient and its use in variational frameworks for both denoising and segmentation tasks.**

***Index Terms*—directional gradient, proximity operator, nD images, variational framework, filtering, segmentation**

I. Introduction

CURVILINEAR structures are among the most difficult to preserve when carrying out image processing tasks. By curvilinear, we mean objects with a spatial dimension d that can be considered strictly lower than the dimension n of the space in which they are embedded. In this article, we are interested in objects of dimension $d = 1$ within spaces of dimension $n \geq 2$. For $n = 2$, such objects can be, for instance, roads in remote sensing images, vessels in eye fundus imaging; for $n = 3$, they can be fibres in composite materials, complex vascular networks in medical images; for $n = 4$, they can be tracked trajectories of moving objects in 3D+time images; etc.

The preservation of curvilinear structures is challenging for various reasons. They are generally sparsely distributed within images, due to their low dimension. They are also often thin structures, with a thickness similar to the image resolution. In addition, they can have complex geometry and topology, with high curvatures, tortuosity, junctions and bifurcations, etc. All these properties make observing curvilinear structures difficult. In particular, they are easily corrupted by noise, and they often suffer from partial volume effect, due to the above mentioned properties. In this context, many existing image processing methods cannot efficiently discriminate them from noise and artifacts.

For tackling these issues, a relevant solution consists of guiding the process by information related to the location and geometric properties (size, orientation) of curvilinear structures. The robustness of the framework then depends on (1) the reliability of this prior information, and (2) the ability to efficiently embed and use this information for improving the restoration or segmentation process.

Regarding point (1), several contributions have been specifically devoted to compute information of curvilinear structures from nD images (often with $n = 2$ or 3). Two representative examples of such contributions are RORPO [1] and Frangi vesselness [2], that both allow to get access to orientation information on putative curvilinear structures. A complete state of the art is beyond the scope of this article; a brief survey is proposed in Section II-A.

In this article, our purpose is related to point (2). Indeed, we aim at developing a framework that explicitly models and uses information about curvilinear structures for improving their efficient processing. This framework is indeed a companion tool of the curvilinear structure estimators of point (1). Our approach relies on a variational paradigm, which is versatile in terms of applications and image dimensions.

Many variational formulations stem from a Maximum a Posteriori Bayesian interpretation that is expressed in a sum of two terms, namely a data-fidelity term and a regularization term. The regularization term corresponds to an image model. Among these, Total-Variation-like approaches aim at regularizing the result versus image noise, by minimizing the overall gradient of the segmentation result. Resulting from several decades of research, this classical regularization term usually provides good results except on curvilinear structures. The reasons of this failure are discussed in Section III-A.

In this context, we define a mixed gradient, which merges a standard gradient and a directional gradient that derives from the directional information provided by any curvilinear structure estimator, for instance [2], [1]. This mixed gradient is then embedded in a variational formulation to form a directional regularization term adapted to curvilinear structures. In addition, we propose an algebraic formulation of this directional regularization, providing a unified definition, independent from the image dimension.

This article is an extended and improved version of the conference paper [3]. It is organized as follows. In Section II,

Odyssée Merveille and Benoît Naegel are with Université de Strasbourg, CNRS, ICube, Strasbourg, France ({merveille,b.naegel}@unistra.fr).

Hugues Talbot is with Centralesupélec, CVN, Paris, France (hugues.talbot@centralesupelec.fr).

Nicolas Passat is with Université de Reims Champagne-Ardenne, CReSTIC, Reims, France (nicolas.passat@univ-reims.fr).

we propose a brief survey of curvilinear structure processing. In Section III, we describe the drawbacks of standard gradient-based regularization, and we summarize our strategy for coping with the identified issues. In Section IV, we propose our gradient definitions. In Section V, we show how both standard and directional gradients can be expressed in a matrix formalism that remains homogeneous in any dimensions. In Section VI, experiments involving our gradient-based regularization are proposed for various kinds of 2D / 3D and synthetic / real images. Concluding remarks are provided in Section VII.

II. Curvilinear structure processing

The state of the art of curvilinear structure processing and analysis is, amusingly, both thin and wide. It is thin because it is not a mainstream topic in image processing, and it is wide because applications cover many different fields. In particular vascular imaging —i.e., 2D eye fundus imaging and 3D, 3D+time angiographic imaging— is certainly the object of the most intensive activity in this context. The bibliography published over the 25 last years took advantage of most paradigms classically proposed in image or signal processing and, more recently, in machine learning.

The short bibliographic discussion proposed hereafter is necessarily incomplete and focuses on the topics that are the most related to our proposed contributions. The readers interested in a more general discussion will find a global survey in [4, Chapters 1, 2]; a medical-oriented survey may also be found in [5], [6].

A. Curvilinear structure detection / description

Before actually processing / analysing a curvilinear structure, it is often relevant to make it more easily detectable in the image. This can be viewed as a filtering task, by improving the signal-to-noise and / or signal-to-background ratio. It can also be viewed as a more semantic task, by determining higher level information, for instance, local size and / or orientation of the curvilinear structure. In this context, two main families of approaches have been developed.

The first relies on linear operators, based on local, differential analysis of images. In particular, the analysis of second-order derivatives of 3D images were proposed in [7], [8]. In these pioneering works, the eigenvectors of multiscale Hessian matrices and their associated eigenvalues are analysed to characterise blobs (3D), planar (2D) and curvilinear (1D) structures as well as their scale and orientation. This led to the proposal of measures combining differential information into heuristic formulations. The measure proposed in [2] is often considered the current gold-standard. Several variants have been proposed since then, for instance in [9], [10]. Alternatively, steerable filters [11] can be expressed, for similar purposes, in terms of a linear combination of basis filters. As such, they are often used to detect oriented features such as curvilinear structures. A framework for 3D steerable filters was, in particular, proposed in [12], using a n^{th} Gaussian derivative basis filter.

Anisotropic diffusion methods [13], [14] were also proposed to filter curvilinear structures. Of particular interest is the Coherence Enhancing Diffusion (CED) filter [15], designed to perform actual anisotropic diffusion by adapting the diffusion direction using the structure tensor, i.e the tensor product of the gradient. More recently, the Hybrid Diffusion using Continuous Switch (HDCS) [16] proposed to combine CED and a diffusion filter with actual anisotropic effects, namely Edge-Enhancing Diffusion (EED) [14]. This combination allows for a strong denoising effect both in the background and inside curvilinear structures, while still preserving their edges. The direction estimation used by both methods depends on the structure tensor which is a reliable estimator for large scale curvilinear structures, but is less robust at small scales. In [17], authors consider local analytical and semi-analytical schemes that can be used for implementing true anisotropic diffusion filters such as EED. They also make an interesting link with Haar wavelets shrinkage, allowing for a multiscale approach.

The second family relies on nonlinear approaches. In particular, notions of optimal path detection in graph and mathematical morphology were involved in the development of these approaches. At the frontier between these two domains, geodesic paths [18] were introduced to consider long-range, non-local interactions while still coping with the constraints of thin objects, in particular noise. A curvilinear object detector was also proposed in [19] using geodesic voting, similar to path density. In [20], a notion of local optimal path was pioneered. Its purpose is to restrict the research to a given distance, and in a given cone of orientations. This paradigm led to the development of a notion of path opening [21], enabling a higher flexibility in geometry and size, while preserving a 1D semantics. Algorithmic efforts were conducted to make such approach computationally efficient [22] and robust to noise [23], [24], leading to a notion of robust path opening. In [1], [25], the notion of RORPO (ranking the orientation responses of path operators), finally built upon these notions, in order to provide a semi-global, nonlinear alternative to the Hessian based approaches in the 2D and 3D cases.

B. Variational paradigm

In this article, we consider the classical variational image restoration problem expressed as the minimization of a two-term energy defined as follows:

$$\hat{f} = \arg\min_{f} \; E_{\text{data}}(f) + \lambda E_{\text{reg}}(f) \tag{1}$$

where $\hat{f}$ is the restored image, E_{data} is the data fidelity term and E_{reg} is the regularization term. This framework was popularized in the 1990s with active contour, level sets and image restoration models.

Mathematically speaking, image restoration is an ill-posed inverse problem. For solving it, it is necessary to impose some regularity on the solution. In [26], a quadratic regularization term is used. This is highly efficient, however, this can generate blurring effects. This quadratic regularization can be replaced by a ℓ_1 gradient norm, called total variation (TV), which better preserves edges [27].

This framework is very flexible and can be used with sparse and/or a blurring operator in the data fidelity for non-blind deconvolution, for instance in the context of alpha-matting [28]

and super-resolution [29], [30]. With missing image values in the input, it can be used for image inpainting [31], [32].

By constraining the output image to take a restricted set of values, this framework can be used for segmentation. In particular, the Chan-Vese model [33] divides the image into two regions of piecewise constant intensities. If these two constant values are known, it results in a convex problem that can be solved exactly [34].

In the context of curvilinear structure segmentation, extensions of the Chan-Vese model were proposed by adding curvilinear priors, for instance, superellipsoids [35], B-splines framelet [36], adaptive dictionaries [37] and elastical regularization [38]. In [39], we also proposed a variational approach for tubular structure restoration. By considering the Frangi measure [2], we designed an adaptive regularization parameter to avoid intensity loss in curvilinear structures, which is an intrinsic problem of classical variational frameworks. However, this approach can prevent regularization within the curvilinear structures, thus leading to various problems, such as potential disconnections.

In the current manuscript, we propose a framework for defining and embedding a mixed gradient operator coupling the directional and standard gradient. This allows us to more robustly take advantage of prior knowledge related to curvilinear structure analysis. Indeed, regularization can be carried out everywhere in the image while remaining adapted to the geometric context. In addition our proposed variational framework remains fully n-dimensional, thus encompassing any images considered in potential applications.

III. Curvilinear structure restoration using directional regularization

In this section, we first explain the limitations of classical regularization terms in the context of curvilinear structure restoration. Then, we give the intuition behind our directional regularization and why it is a better suited term for curvilinear structure restoration. More formal explanations on the directional regularization and its implementation details are exposed in Section IV.

A. Regularization principle and limitations

The regularization term in a variational restoration problem can be interpreted as choosing a solution with desirable properties within a solutions space. A property often desired is to minimize image noise while retaining image content. This property is translated into a regularization term that minimizes the norm of the image gradient, under the assumption that the image content has sparse contours. When this norm is the ℓ_1 norm, the regularization term is called *total variation* (TV):

$$E_{\text{reg TV}}(f) = \|\nabla f\|_1 \tag{2}$$

where ∇f is the discrete gradient of f and $\|\cdot\|_1$ is the ℓ_1 norm.

Interpreting this regularization term as a statistical model, it says that gradient intensities follow an exponential distribution. This is indeed observed to be the case for natural images [40]. Interpreting it as an image model, it says that natural images should have a sparse gradient, meaning that most variations in images are small, except near a few contours where they can be large. This corresponds to an intuitive piecewise smooth image model.

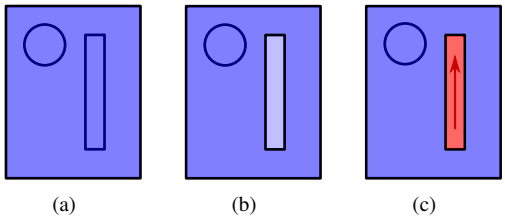


Fig. 1. Regularization principle. The dark (resp. bright) blue areas represent high (resp. low) isotropic (i.e. with no privileged direction) regularization, whereas the red area represents a directional regularization. (a) The classical regularization is performed in all directions with the same intensity everywhere. (b) The regularization proposed by Miraucourt *et al.* [39] acts with a low intensity inside curvilinear structures and a high intensity elsewhere, but always in all directions. (c) Our directional regularization acts with the same intensity inside and outside curvilinear structures; however, inside curvilinear structures, we only regularize along the curvilinear structure direction.

This regularization term is effective at decreasing the image noise in natural images; however it may also suppress image contours when they are not highly contrasted. Moreover, structures with a high perimeter over surface ratio (for $n = 2$) or a high surface over volume ratio (for $n = 3$) are highly penalized. For these reasons, classical regularization terms are not a good model for curvilinear structures, which tend to disappear in the resulting restored images.

To cope with this problem, Miraucourt et al. [39] proposed to include a curvilinear structure position prior in order to regularize more strongly outside the curvilinear structures than inside (see Fig. 1(b)). This strategy effectively prevents curvilinear structures from disappearing, but is not effective at decreasing noise inside the curvilinear structures, where only a weak regularization is applied.

B. Directional regularization motivation

The motivation behind [39] was to keep the good behavior of the total variation outside curvilinear structures and decrease its effects within them. However, nothing specific was done to regularize and restore the curvilinear structures specifically. We propose to extend the total variation formulation to better preserve curvilinear structures, while keeping its good properties on other structures. To this end, we consider not only a positional prior, as in [39], but also an orientation prior. Instead of a weaker regularization inside curvilinear structures, we apply a strong regularization but solely along the local curvilinear structure direction. In other words, we propose an intensity and directional spatially variant regularization term (see Fig. 1(c)).

The total variation is an isotropic regularization term as the discrete gradient lacks a privileged direction. Indeed, the discrete gradient is computed by finite differences in all the directions of the Cartesian basis. Our goal is to change the total variation behavior inside curvilinear structures to better preserve them. We define a directional gradient $\nabla_{\mathbf{d}} f$ which has a privileged direction, namely the local direction of the curvilinear structure itself.

Based on both isotropic (∇f) and anisotropic ($\nabla_{\mathbf{d}} f$) gradients, we propose to adapt the total variation formulation to better preserve curvilinear structures, while keeping its good properties on other structures. This new regularization term, called directional regularization is defined by:

$$E_{\text{reg}}(f(\mathbf{x})) = \|\nabla_m f(\mathbf{x})\|_1 \tag{3}$$

where:

$$\nabla_m f(\mathbf{x}) = \begin{cases} \nabla_{\mathbf{d}} f(\mathbf{x}) & \text{if } \mathbf{x} \in \text{curvilinear structure} \\ \nabla f(\mathbf{x}) & \text{otherwise} \end{cases} \tag{4}$$

C. Implementation

In practice, we aim at solving the following problem (see Eq. (1)):

$$\begin{aligned} \hat{f}(\mathbf{x}) &= \arg\min_{f \in C} E_{\text{data}}(f(\mathbf{x})) + \lambda E_{\text{reg}}(f(\mathbf{x})) \\ &= \arg\min_{f \in C} E_{\text{data}}(f(\mathbf{x})) + \lambda \|\nabla_m f(\mathbf{x})\|_1 \end{aligned} \tag{5}$$

where C is a closed convex subset of the image domain, $E_{\text{data}}(f)$ is a convex, differentiable function and $\|\nabla_m f\|_1$ is a convex, but non-differentiable function that forbids using a gradient descent approach. For many inverse problems including segmentation, the convex C is a centered unit ball which ensures that output values are consistent with the input image data.

It can be shown from [41] that Eq. (5) admits a solution given by the proximal point splitting algorithm:

$$u_{n+1} = \text{prox}_{\gamma h}\big(u_n - \gamma \nabla E_{\text{data}}(u_n)\big) \tag{6}$$

where $\gamma \in (0, +\infty)$ is a step-size parameter and $h = \lambda \|\nabla_m f(u_n)\|_1$.

The reader may refer to [42] for a review on proximal splitting point algorithm. Here, we only recall the definition of the proximity operator of a function ϕ:

$$\text{prox}_{\phi}(y) = \arg\min_x \big(\phi(x) + \frac{1}{2}\|x - y\|_2^2\big) \tag{7}$$

where $\|\cdot\|_2$ is the ℓ_2 norm. There is no generic strategy in the literature, that allows one to compute the proximity operator of any function ϕ. Nonetheless, several algorithms were proposed to compute the proximity operator of specific functions.

To solve our restoration problem (Eq. (5)), we use the iterative method of Eq. (6), that requires to compute the proximity operator of γh. In particular, we choose the Beck and Teboulle algorithm [43], called Fast Gradient Projection (FGP), that was designed to solve this specific proximity operator (see Algorithm 1).

In this algorithm, $\mathbb{D}$ is the domain of the image as described below in Eq. (8), P_{ℓ_2} is the projection on the ℓ_2 unit ball, P_C is the projection on the convex set C, γ is set to the Lipschitz constant of the gradient of E_{data}. ∇ and div are the gradient and divergence operators, respectively.

Algorithm 1: FGP algorithm

Data: $y \in \mathbb{R}^{\mathbb{D}}, x_0 \in \mathbb{R}^{n\mathbb{D}}, z_1 = x_0$ and $t_0 = 1$

for $n \geq 1$ **do**

$$\begin{aligned} y_n &= z_n + \frac{1}{\gamma\lambda}\nabla\Big(P_C\big[y - \lambda\,\text{div}(z_n)\big]\Big) \\ x_n &= P_{\ell_2}[y_n] \\ t_{n+1} &= \frac{1 + \sqrt{4t_n^2 + 1}}{2} \\ \lambda_n &= \frac{t_n - 1}{t_{n+1}} \\ z_{n+1} &= x_n + \lambda_n(x_n - x_{n-1}) \end{aligned}$$

return $\text{prox}_{\gamma h}(y) = P_C\big(y - \lambda\,\text{div}(z_N)\big)$

IV. Gradients formulations

In this section, we define formally the operators presented in the previous section: the gradient, the directional gradient and the mixed gradient.

In the following we define a nD image f ($n \in \mathbb{N}$, $n \geq 1$) as follows:

$$\left| \begin{array}{ccccc} f & : & \mathbb{D} & \to & \mathbb{R} \\ & & \mathbf{x} & \mapsto & f(\mathbf{x}) \end{array} \right. \tag{8}$$

where $\mathbb{D} = \prod_{i=1}^{n} [\![0, d_i - 1]\!] \subset \mathbb{Z}^n$ and $(d_i)_{i \in [\![1,n]\!]}$ are the dimensions of the image.

A. Standard gradient

The discrete gradient of an image can be expressed, via finite differences such that:

$$\nabla f(\mathbf{x}) = \big(f(\mathbf{x} + \mathbf{e}_i) - f(\mathbf{x})\big)_{i=1}^{n} \tag{9}$$

where $(\mathbf{e}_i)_{i=1}^n$ is the canonical basis of $\mathbb{R}^n$ and $\mathbb{Z}^n$, namely $\mathbf{e}_i = (\delta_{i,j})_{j=1}^n$ for any $i \in [\![1, n]\!]$, (with $\delta_{i,j}$ the Kronecker delta).

B. Directional gradient

When a curvilinear structure is identified at a point $\mathbf{x} \in \mathbb{D}$, one may wish to define a gradient operator locally oriented in the direction of this structure, in order to allow for noise removal without altering this structure.

Let $\mathbf{d}(\mathbf{x})$ be the unit vector lying in the direction of the curvilinear structure observed at $\mathbf{x}$. We assume that $\mathbf{d}(\mathbf{x})$ is oriented in the half-space of $\mathbb{R}^n$ such that $d_n \geq 0$ (see discussion in Appendix A). We define the directional gradient $\nabla_{\mathbf{d}} f : \mathbb{D} \to \mathbb{R}^n$ via finite differences such that:

$$\nabla_{\mathbf{d}} f(\mathbf{x}) = \big(f(\mathbf{x} + \mathbf{d}(\mathbf{x})) - f(\mathbf{x})\big).\mathbf{d}(\mathbf{x}) \tag{10}$$

However, $f(\mathbf{x} + \mathbf{d}(\mathbf{x}))$ is generally undefined, since $\mathbf{d}(\mathbf{x}) \in \mathbb{R}^n$ while f is a function on $\mathbb{D} \subset \mathbb{Z}^n$. It is then necessary to consider an interpolation of f on the part of $\mathbb{R}^n$ associated to $\mathbb{D}$. In particular, the standard n-linear interpolation can be considered, namely:

$$f(\mathbf{x} + \mathbf{d}(\mathbf{x})) = \sum_{\mathbf{u} \in [\![-1,1]\!]^{n-1} \times [\![0,1]\!]} \lambda_{\mathbf{d}(\mathbf{x}),\mathbf{u}} f(\mathbf{x} + \mathbf{u}) \tag{11}$$

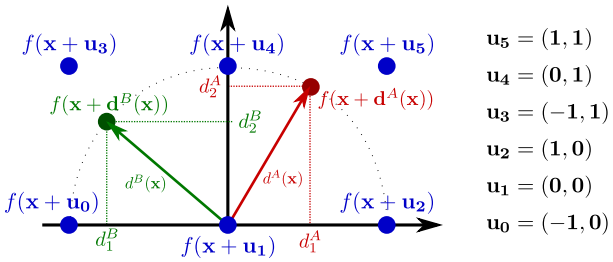


Fig. 2. Illustration of two interpolations (in red and green) in the case $n=2$. $\mathbf{d}^A$ (resp. $\mathbf{d}^B$) is the local orientation of a curvilinear structure at point $\mathbf{x}$ ($\mathbf{x}=\mathbf{x}+\mathbf{u_1}$). To compute the directional gradient, the value at point $f(\mathbf{x}+\mathbf{d}^A(\mathbf{x}))$ (resp. $f(\mathbf{x}+\mathbf{d}^B(\mathbf{x}))$) is required.

where:

$$\lambda_{\mathbf{d}(\mathbf{x}),\mathbf{u}} = \prod_{i=1}^{n} w(d_i(\mathbf{x}),u_i) \tag{12}$$

$$w(d_i(\mathbf{x}),u_i) = \begin{cases} (1-|u_i|)(1-|d_i(\mathbf{x})|) \\ +|u_i|(u_i.d_i(\mathbf{x})+|d_i(\mathbf{x})|)/2 \end{cases} \tag{13}$$

with $\mathbf{x}=(x_i)_{i=1}^n$, $\mathbf{d}(\mathbf{x})=(d_i(\mathbf{x}))_{i=1}^n$, $\mathbf{u}=(u_i)_{i=1}^n$.

Fig. 2 shows two examples of interpolation in the case $n=2$.

C. Mixed gradient

It is possible to associate to the function $f:\mathbb{D}\to\mathbb{R}$, a directional vector field $\mathbf{d}:\mathbb{D}\to\mathbb{R}^n$ that yields a unit vector $\mathbf{d}(\mathbf{x})$ for each point $\mathbf{x}\in\mathbb{D}$ if a specific orientation is detected at $f(\mathbf{x})$, and the null vector $\mathbf{0}$ otherwise.

In particular, the definition of $\mathbf{d}(\mathbf{x})$ either as a unit vector providing a specific orientation, or a null vector expressing the absence of such orientation, is conditioned by the curvilinear structure estimator used beforehand. For instance, this can be a thresholding of the RORPO filter or of the Frangi vesselness output. In case of a positive answer, both RORPO and Frangi vesselness provide orientation information.

This vector field provides the information required to compute the directional gradient $\nabla_{\mathbf{d}}f$, see Section IV-B. It is also useful for computing and modeling in a matrix fashion a mixed directional-standard gradient of f.

In particular, let us consider the function $\alpha:\mathbb{D}\to\{0,1\}$ defined, for any $\mathbf{x}\in\mathbb{D}$ by $\alpha(\mathbf{x})=\|\mathbf{d}(\mathbf{x})\|_2$, namely the norm of $\mathbf{d}(\mathbf{x})$. This function takes its values in the binary set $\{0,1\}$. It is equal to 1 when a unit vector $\mathbf{d}(\mathbf{x})$ is defined at $\mathbf{x}$, and 0 when $\mathbf{d}(\mathbf{x})$ is the null vector $\mathbf{0}$, i.e. when no specific orientation is defined at $\mathbf{x}$.

Then, according to the binary values of α, we can build a mixed gradient $\nabla_m f$ from the standard and directional gradients ∇f and $\nabla_{\mathbf{d}}f$ as:

$$\nabla_m f(\mathbf{x}) = \begin{cases} \nabla f(\mathbf{x}) & \text{if } \alpha(\mathbf{x})=0 \\ \nabla_{\mathbf{d}} f(\mathbf{x}) & \text{if } \alpha(\mathbf{x})=1 \end{cases} \tag{14}$$

that is:

$$\nabla_m f(\mathbf{x}) = \alpha(\mathbf{x})\nabla_{\mathbf{d}}f(\mathbf{x}) + (1-\alpha(\mathbf{x}))\nabla f(\mathbf{x}) \tag{15}$$

V. Matrix formulation of the directional regularization

Even though the vector expression of the mixed gradient presented above (see Eq. (15)) is simple, its implementation and manipulation in nD ($n>2$) becomes somewhat complex. Indeed, Algorithm 1 requires the definition of our mixed gradient operator and its associated divergence given by the following adjoint relation:

$$-\langle \operatorname{div} p, u\rangle_S = \langle p, \nabla u\rangle_{S^n} \tag{16}$$

with $S=\mathbb{R}^{\mathbb{D}}$. The higher the dimension of the image, the more complex the divergence definition. This is especially true in terms of limit cases due to the discrete nature of an image. To tackle these issues, we propose to consider the discrete calculus framework, by proposing a matrix formulation M^{∇_m} of the mixed gradient operator. Thus, the adjoint divergence of this matrix operator, M^{∇_m}, will be simply obtained as the transpose matrix $(M^{\nabla_m})^T$.

A. Vector formulation of functions

A function $h:X\to Y$ can be modeled as a vector $V^h\in Y^{|X|}$ of $|X|$ elements, representing the image of each element of X by h, that is:

$$V^h=[h(x)]_{x\in X}=\begin{bmatrix} h(a) \\ \vdots \\ h(z)\end{bmatrix} \tag{17}$$

In particular, the function f defined in Eq. (8) is modeled by the vector:

$$V^f=[f(\mathbf{x})]_{x\in\mathbb{D}}=\begin{bmatrix} f((0,\ldots,0)) \\ \vdots \\ f((d_1-1,\ldots,d_n-1))\end{bmatrix}=\begin{bmatrix} v_1^f \\ \vdots \\ v_D^f\end{bmatrix} \tag{18}$$

where $D=|\mathbb{D}|$.

In particular, for modeling $f:\mathbb{D}\to\mathbb{R}$ as a vector $V^f\in\mathbb{R}^D$, we need a transfer function σ between $\mathbb{D}\subset\mathbb{Z}^n$ and $[\![1,D]\!]\subset\mathbb{Z}$. Such a function can be defined as:

$$\left|\begin{array}{ccccc} \sigma & : & \mathbb{D} & \to & [\![1,D]\!] \\ & & \mathbf{x}=(x_i)_{i=1}^n & \mapsto & 1+\sum_{i=1}^n(\prod_{j=1}^{i-1}d_j)x_i \end{array}\right. \tag{19}$$

The function σ is bijective. In the sequel, we note $\mathbf{x}_j=\sigma^{-1}(j)\in\mathbb{D}$.

Following the notations of Eq. (18), we then have:

$$f(\mathbf{x})=v^f_{\sigma(\mathbf{x})} \tag{20}$$

and we can then establish a one-to-one correspondence between the terms of f and V^f

B. Matrix formulation of the gradient operator

We are now ready to establish the matrix formulation of the standard gradient operator ∇f of f. We have $\nabla f:\mathbb{D}\to\mathbb{R}^n$. Then, it can be defined as a vector $V^{\nabla f}\in(\mathbb{R}^n)^{\mathbb{D}}$ of D elements or, equivalently, as a vector $V^{\nabla f}\in\mathbb{R}^{n.\mathbb{D}}$ of $n.D$ elements. Practically, this vector can be split into n vectors $V_i^{\nabla f}$ of D elements, each of them giving the coordinates of

the gradient in the subspace $\mathbb{R}$ of $\mathbb{R}^n$ induced by a basis vector $\mathbf{e}_i$.

In other words, we have:

$$V^{\nabla f} = \begin{bmatrix} V_1^{\nabla f} \\ \vdots \\ V_n^{\nabla f} \end{bmatrix} \tag{21}$$

with, for all $i \in [\![1, n]\!]$ (see Eq. (9)):

$$V_i^{\nabla f} = \begin{bmatrix} \frac{\partial f}{\partial x_i}(\mathbf{x}_1) \\ \vdots \\ \frac{\partial f}{\partial x_i}(\mathbf{x}_j) \\ \vdots \\ \frac{\partial f}{\partial x_i}(\mathbf{x}_D) \end{bmatrix} \tag{22}$$

$$= \begin{bmatrix} f(\mathbf{x}_1 + \mathbf{e}_i) - f(\mathbf{x}_1) \\ \vdots \\ f(\mathbf{x}_j + \mathbf{e}_i) - f(\mathbf{x}_j) \\ \vdots \\ f(\mathbf{x}_D + \mathbf{e}_i) - f(\mathbf{x}_D) \end{bmatrix} \tag{23}$$

Consequently, $V_i^{\nabla f}$ can be expressed as:

$$V_i^{\nabla f} = (G_i - I_D) \cdot V^f \tag{24}$$

where I_D is the $D \times D$ identity matrix and $G_i = (g_i(j,k))_{(j,k)\in[\![1,D]\!]^2}$ is the $D \times D$ matrix defined by:

$$g_i(j,k) = \delta_{k,\sigma(\mathbf{x}_j+\mathbf{e}_i)} \tag{25}$$

For the correct handling of boundary effects on the frontier of the domain $\mathbb{D}$, whenever we have $\mathbf{x}_j + \mathbf{e}_i \notin \mathbb{D}$ we set:

$$g_i(j,k) = \delta_{k,\sigma(\mathbf{x}_j)} = \delta_{k,j} \tag{26}$$

In other words, in such a case, we consider a Neumann condition of nullity of the gradient coordinate. (Note that this strategy will be also involved in anisotropy effects correction; see discussion in Appendix A.)

Finally, we can then define the vectorial expression of ∇f by combining its n coordinate vectors, as:

$$V^{\nabla f} = M^{\nabla} \cdot V^f = \left(\begin{pmatrix} G_1 \\ \vdots \\ G_i \\ \vdots \\ G_n \end{pmatrix} - \begin{pmatrix} I_D \\ \vdots \\ I_D \\ \vdots \\ I_D \end{pmatrix} \right) \cdot V^f \tag{27}$$

where M^{∇} is a $n.D \times D$ matrix.

C. Matrix formulation of directional gradient

In a similar way, it is also possible to express a matrix formulation of the directional gradient. This formulation is, however, slightly different from that of the standard gradient. Indeed, for any $\mathbf{x} \in \mathbb{D}$, $\nabla f(\mathbf{x})$ is expressed by its n coordinates with respect to the canonical basis $(\mathbf{e}_i)_{i=1}^n$. In contrast, $\nabla_{\mathbf{d}} f(\mathbf{x})$ has only one coordinate with respect to a unit vector $\mathbf{d}(\mathbf{x})$ on the 1D line locally oriented in the specific direction observed at $\mathbf{x}$. In particular, this vector $\mathbf{d}(\mathbf{x})$ depends on $\mathbf{x}$, and the coordinate system is then spatially variant.

As a consequence, we have to compute only one coordinate of the directional gradient for each $\mathbf{x} \in \mathbb{D}$, but this computation is more complex than for standard gradient, as it theoretically involves 3^n points of $\mathbb{D}$ around $\mathbf{x}$ (in practice, only 2^n actually contribute to the result); see Eqs. (11–13).

The vectorial representation $V^{\nabla_{\mathbf{d}} f}$ of $\nabla_{\mathbf{d}} f$ is then expressed as:

$$V^{\nabla_{\mathbf{d}} f} = \begin{bmatrix} \frac{\partial f}{\partial d(\mathbf{x}_1)}(\mathbf{x}_1) \\ \vdots \\ \frac{\partial f}{\partial d(\mathbf{x}_j)}(\mathbf{x}_j) \\ \vdots \\ \frac{\partial f}{\partial d(\mathbf{x}_D)}(\mathbf{x}_D) \end{bmatrix} \tag{28}$$

$$= \begin{bmatrix} f(\mathbf{x}_1 + \mathbf{d}(\mathbf{x}_1)) - f(\mathbf{x}_1) \\ \vdots \\ f(\mathbf{x}_j + \mathbf{d}(\mathbf{x}_j)) - f(\mathbf{x}_j) \\ \vdots \\ f(\mathbf{x}_D + \mathbf{d}(\mathbf{x}_D)) - f(\mathbf{x}_D) \end{bmatrix} \tag{29}$$

Consequently, $V^{\nabla_{\mathbf{d}} f}$ can be written as:

$$V^{\nabla_{\mathbf{d}} f} = M^{\nabla_{\mathbf{d}}} \cdot V^f = E_{\mathbb{D}} \cdot (G_{\mathbf{d}} - I_D) \cdot V^f \tag{30}$$

where I_D is the $D \times D$ identity matrix, $G_{\mathbf{d}} = (g_{\mathbf{d}}(j,k))_{(j,k)\in[\![1,n]\!]^2}$ is the $D \times D$ matrix defined by:

$$g_{\mathbf{d}}(j,k) = \sum_{\mathbf{u}\in[\![-1,1]\!]^n} \delta_{k,\sigma(\mathbf{x}_j+\mathbf{u})} \cdot \lambda_{\mathbf{d}(\mathbf{x}_j),\mathbf{u}} \tag{31}$$

and $E_{\mathbb{D}} = [\varepsilon_j]_{j=1}^D \in \{0,1\}^D$ is defined by:

$$\varepsilon_j = 1 - \delta_{0,\prod_{\mathbf{u}\in[\![-1,1]\!]^n} \sigma(\mathbf{x}_j+\mathbf{u})} \tag{32}$$

For the correct handling of boundary effects on the frontier of the domain $\mathbb{D}$, we still extend —by convention— σ out of $\mathbb{D}$ by setting $\sigma(\mathbf{x}) = 0$ for any $\mathbf{x} \in \mathbb{Z}^n \setminus \mathbb{D}$. With this formulation, a useful feature is that the directional gradient $\nabla_{\mathbf{d}} f$ vanishes exactly on the border of the domain $\mathbb{D}$, i.e. for points $\mathbf{x} = (x_i)_{i=1}^n$ with at least one coordinate $x_i = 0$ or $d_i - 1$.

D. Matrix formulation of the mixed gradient

A matrix representation M^{∇_m} of the mixed gradient ∇_m can be formulated from the above matrix representations of $\nabla_{\mathbf{d}}$ and ∇, see Eqs. (27) and (30). Indeed, we have:

$$M^{\nabla_m} = \begin{pmatrix} (\mathbf{1}_{\mathbb{D}} - V^{\alpha}) \cdot (G_1 - I_D) \\ \vdots \\ (\mathbf{1}_{\mathbb{D}} - V^{\alpha}) \cdot (G_i - I_D) \\ \vdots \\ (\mathbf{1}_{\mathbb{D}} - V^{\alpha}) \cdot (G_n - I_D) \\ V^{\alpha} \cdot E_{\mathbb{D}} \cdot (G_{\mathbf{d}} - I_D) \end{pmatrix} \tag{33}$$

where $\mathbf{1}_{\mathbb{D}}$ is the vector representation of the constant function $\mathbf{x} \mapsto 1$ on $\mathbb{D}$, and V^{α} is the vector representation of the function α. Note that M^{∇_m} is a $(n+1).D \times D$ matrix.

VI. Experiments

In this section, we experimentally assess the relevance of the proposed restoration framework. To emphasize the versatility of our approach in terms of dimensionality and application, we first present a segmentation application on 2D images; then we show denoising results on 3D images. For both experiments, the prior information on estimated orientations of putative curvilinear structures was obtained by the RORPO operator [1].

A. Segmentation

The DRIVE database [44] is a well known dataset of 40 eye fundus images associated with their blood vessels ground-truth. In this section, we apply our restoration framework to the blood vessel segmentation problem. In this case, we use the data fidelity term proposed by Chan *et al.* [34], $E_{\text{data}} = \langle c_f, u \rangle_F$, where:

- $c_f \in \mathbb{R}^{\mathbb{D}}$ is the function $\mathbf{x} \mapsto (c_1 - f(\mathbf{x}))^2 - (c_2 - f(\mathbf{x}))^2$ (with $c_\star \in \mathbb{R}$ some constant values assumed to correspond to the background and foreground of the image);
- $\langle \cdot, \cdot \rangle_F$ the Frobenius product.

To assess the relevance of our regularization term in a segmentation framework, we compare the classic Chan *et al.* segmentation model [34] that uses the total variation:

$$\hat{u} = \arg \min_{u \in [0,1]^{\mathbb{D}}} \langle c_f, u \rangle_F + \lambda \|\nabla u\|_{2,1} \tag{34}$$

with our segmentation model that uses the proposed directional regularization:

$$\hat{u} = \arg \min_{u \in [0,1]^{\mathbb{D}}} \langle c_f, u \rangle_F + \lambda \|\nabla_m u\|_{2,1} \tag{35}$$

Both standard and directional versions of the method share the same parameters: c_1, c_2 and λ. We optimized these parameters based on the images of the DRIVE training set. The optimization scheme and parameters values are presented in Appendix B.

For our experiments, we consider the single-channel version of these images, i.e. the grey-level version of the data. We subtracted the median filter to each image in order to homogenize the image background, since the Chan *et al.* data fidelity assumes homogeneous background and foreground intensities. An example of such image is shown in Fig. 3(a).

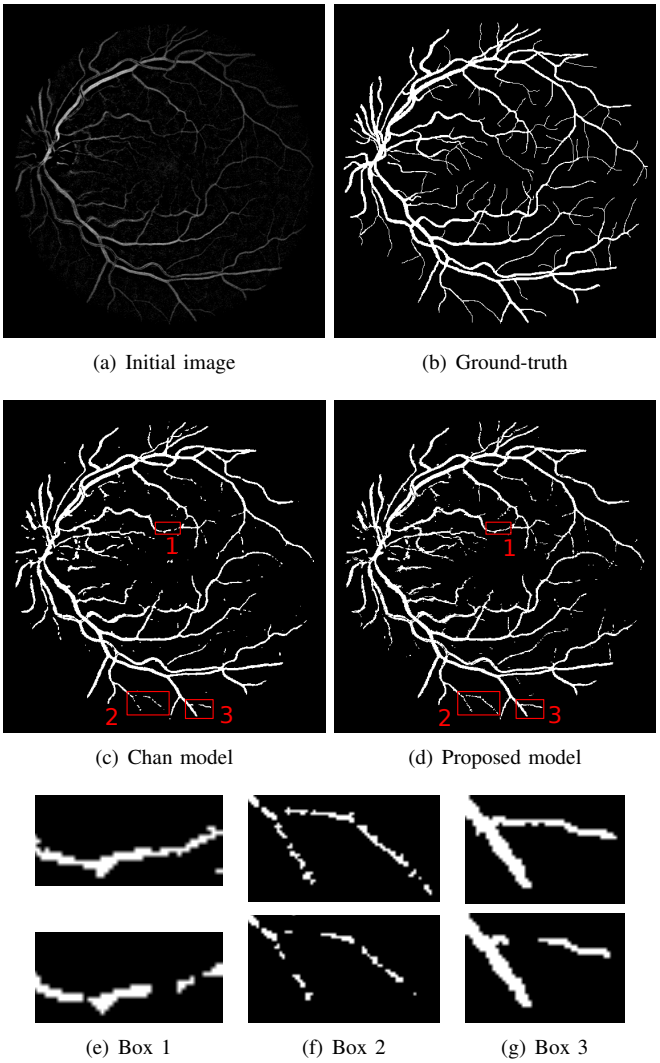
We computed the true positive rate / sensivity (TPR) and true negative rate / specificity (TNR), along with the accuracy[1] (Acc) in the field of view area provided by the DRIVE dataset of each segmentation:

$$\begin{aligned} \text{TPR} &= \frac{\text{TP}}{\text{TP} + \text{FN}} \\ \text{TNR} &= \frac{\text{TN}}{\text{TN} + \text{FP}} \\ \text{Acc} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \end{aligned} \tag{36}$$

[1]In Section VI-B, we will consider the Matthews correlation coefficient (MCC) [45] instead of accuracy, as it is better suited in the case of sparse structures. Here, the choice of accuracy was motivated by the fact that most of the methods benchmarked on the DRIVE dataset were compared using the accuracy measure.

TABLE I
Quantitative segmentation results on the DRIVE database [44]. TPR: true positive ratio (sensitivity). TNR: true negative ratio (specificity). Acc: Accuracy.

	TPR	TNR	Acc
Chan [34]	0.6615	0.9833	0.9423
Our approach	0.6820	0.9809	0.9427
Staal *et al.* [44]	–	–	0.9442
Lupascu *et al.* [46]	–	–	0.9597
Al-Rawi *et al.* [47]	–	–	0.9514
Human observer	–	–	0.9473



(a) Initial image (b) Ground-truth

(c) Chan model (d) Proposed model

(e) Box 1 (f) Box 2 (g) Box 3

Fig. 3. Comparison of segmentation results with the classical Chan *et al.* model (c) and our directional TV model (d) on the first DRIVE image (a) and its ground-truth (b). (e–g) Zoom on distal parts of blood vessels of the Chan model results (bottom) and our proposed model results (top).

where TP (resp. FP) is the number of true (resp. false) positives and TN (res. FN) is the number of true (resp. false) negatives.

The mean quantitative results are summarized in Table I and illustrated in Fig. 3. For comparison, we also provide some state of the art results.

Qualitatively, the segmentation improvements mostly concern the distal part of the blood vessels. In particular, our directional regularization successfully reconnects these extremities, as shown in Fig. 3(e–g). This leads to a better connected vessel

network, which is a highly desired feature in blood vessel related applications.

Quantitatively, the directional regularization improves the accuracy of the classical Chan *et al.* segmentation. As the reconnections represent only a few pixels within the image, the accuracy of our directional regularization is only slightly higher than the Chan *et al.* accuracy even if the improvement is real and significant as can be seen on the images.

Moreover, the accuracy of our method is close to state of the art methods, even though our purpose was not to propose a dedicated segmentation model for retinal images, but only a generic directional regularization term for curvilinear structures.

B. Denoising

In this section, we assess the performance of our method in the case of 3D image denoising. It is difficult to have access to reliable ground-truth for 3D images with curvilinear structures. Thus, we generated synthetic images using the VascuSynth software package[2] [48]. We generated 10 images containing curvilinear patterns, organized as a connected acyclic structure, constituting our ground-truth. We then added a Gaussian random field background to these images, in order to simulate undesired, non-homogeneous, smooth, blob-like features. Finally, we introduced a non-additive, spatially variant noise by subjecting each image with 7 different intensities of Poisson noise, determined by the gain g. This resulted in a database of 70 images. An example of such image is presented in Figures 5(b) and 6(b).

We compared our results with three methods: the exact ROF model [27], [49], which is a classical variational denoising filter, and is thus used as the baseline experiment showing how our directional regularization improves the curvilinear structures restoration; the hybrid diffuson using continuous switch (HDCS) [16], which is a state of the art anisotropic diffusion filter designed to preserve and enhance curvilinear and blob-like structures while filtering the noise; and BM4D [50], which is the 3D version of BM3D, considered as one of the gold standard denoising algorithms. These experiments were performed using the TubeTK implementation of HDCS[3] and the original implementation of BM3D in Matlab[4].

For each of these methods, we optimized the parameters on the first image of the dataset and for each level of noise. Then, we used this fixed set of parameters on the whole dataset. Parameter optimization was performed by sampling the parameter space for each method in reasonable ranges. The sets of parameters used for each method is provided in Appendix B.

For each method and level of noise, we computed the mean ROC curve over the 10 images as the true positive rate / sensivity (TPR) against the false positives rate (FPR). As the amount of negatives is much higher than the amount of positives in these images, we define the FPR as the number of false positives over the number of positives (P), namely FPR = FP / P. For instance, a FPR of 2 means that there is twice as many false positives as ground truth positives. We also computed the Matthews correlation coefficient (MCC) [45] of the best threshold of each result:

$$\mathrm{MCC} = \frac{\mathrm{TP}\cdot\mathrm{TN} - \mathrm{FP}\cdot\mathrm{FN}}{\sqrt{(\mathrm{TP}+\mathrm{FP})(\mathrm{TP}+\mathrm{FN})(\mathrm{TN}+\mathrm{FP})(\mathrm{TN}+\mathrm{FN})}} \tag{37}$$

The choice of MCC is motivated by the fact that it is well suited for sparse images. The closer to 1 the MCC, the more similar the result to the ground-truth. We chose to analyze these denoising results based on two criteria that are usually considered as segmentation metrics because the ground-truths of this denoising are the binary images of curvilinear structures used for the dataset generation. In this case, the best denoising result is the closest to these binary structures.

A summary of the results is shown in Table II, and the mean ROC curves for a representative noise level (gain $g = 15$) is shown in Figure 4. From a quantitative point of view, we observe that the MCC scores are significantly higher with our directional regularization, compared to the other three methods irrespective of the noise level. The ROC curves show that our directional regularization outperforms all the compared methods when the FPR is between 0 and 2. For a FPR between 2 and 4, our proposal still outperforms ROF and

TABLE II
QUANTITATIVE DENOISING RESULTS ON VASCUSYNTH IMAGES. THE MEAN MCC VALUE OVER THE 10 IMAGES FOR THE SAME LEVEL OF NOISE IS GIVEN FOR EACH METHOD.

Noise (g)	Ours	ROF [27]	HDCS [16]	BM3D [50]
2	**0.668**	0.659	0.593	0.564
5	**0.580**	0.560	0.528	0.460
10	**0.457**	0.441	0.438	0.375
15	**0.510**	0.442	0.464	0.400
20	**0.463**	0.407	0.432	0.379
30	**0.336**	0.275	0.321	0.264
50	**0.286**	0.227	0.265	0.212

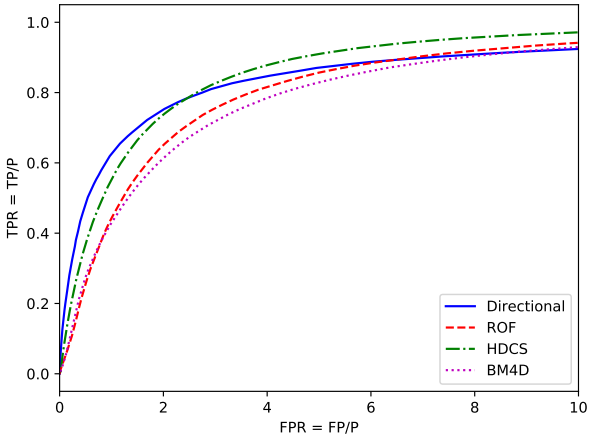
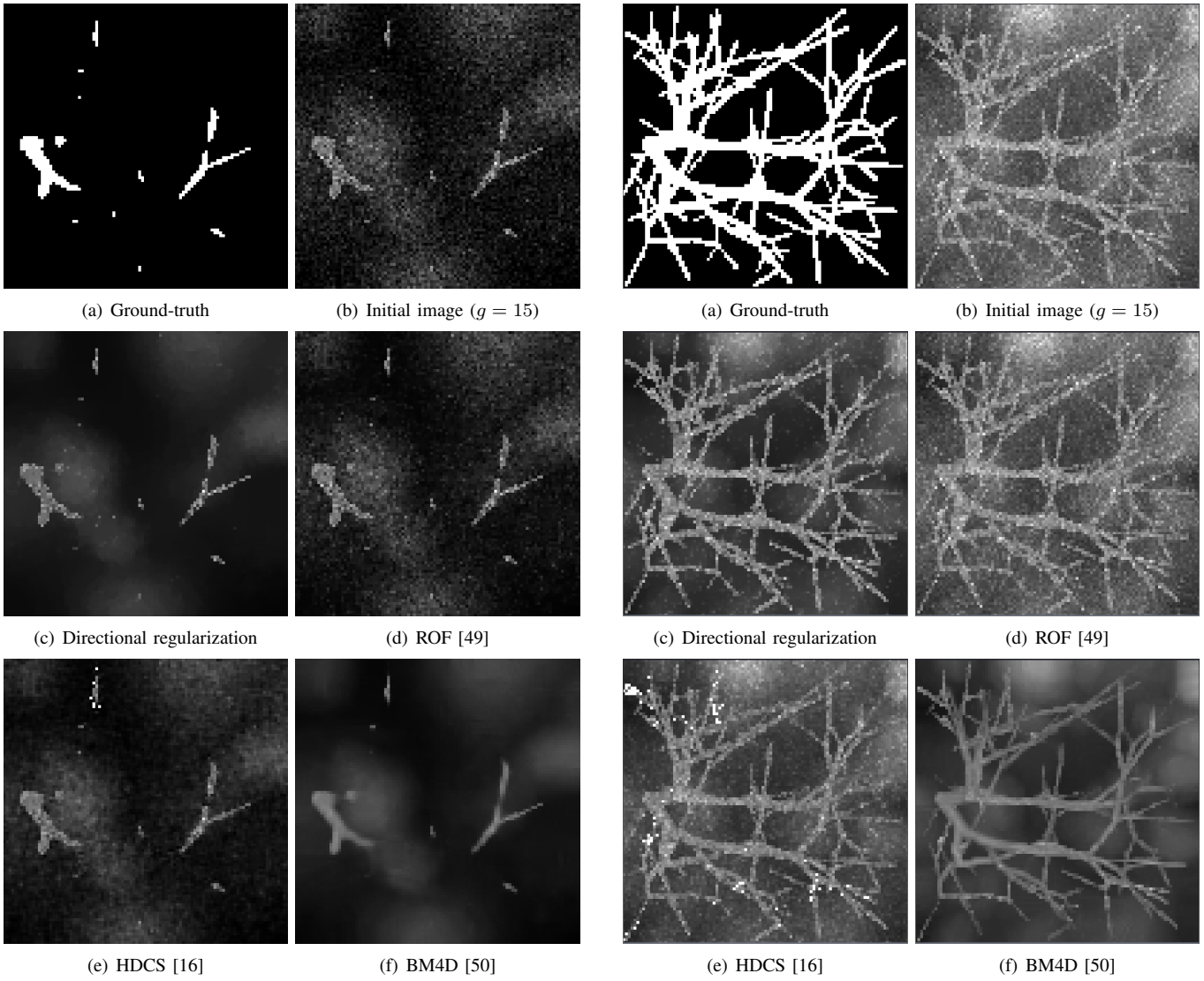


Fig. 4. Mean ROC curves over the 10 VascuSynth images with with a Poisson noise of gain $g = 15$, for FPR values between 0 and 10. (Note that due to the sparsity of the images, the maximum TPR is reached for all methods only for a FPR around 100.)

[2]http://vascusynth.cs.sfu.ca/Software.html

[3]https://public.kitware.com/Wiki/TubeTK/Anisotropic_Hybrid_Diffusion_with_Continuous_Switch

[4]http://www.cs.tut.fi/~foi/GCF-BM3D



(a) Ground-truth (b) Initial image ($g = 15$)

(c) Directional regularization (d) ROF [49]

(e) HDCS [16] (f) BM4D [50]

Fig. 5. Denoising results on a 3D synthetic image (2D slices). (a) Synthetic image generated with VascuSynth [48]. (b) Noisy image generated from (a) with Poisson noise (gain $g = 15$). (c) Denoising image with total variation (ROF model). (d) Denoising image with directional regularization.

(a) Ground-truth (b) Initial image ($g = 15$)

(c) Directional regularization (d) ROF [49]

(e) HDCS [16] (f) BM4D [50]

Fig. 6. Denoising results on a 3D synthetic image (maximum intensity projections). (a) Synthetic image generated with VascuSynth [48]. (b) Noisy image generated from (a) with Poisson noise (gain $g = 15$). (c) Denoising image with total variation (ROF model). (d) Denoising image with directional regularization.

BM4D, but in this situation HDCS provides the best results. In other words, our method is the best for preserving the 3D curvilinear structures in the most challenging and useful situations, corresponding to the cases were the TPR is high while the FPR must remains low.

This quantitative analysis is qualitatively confirmed by visual inspection of Figures 5 and 6. We observe that the reconstruction of the thinnest structures, with thickness close to the image resolution, is better with the directional gradient. This is indeed coherent since the directional regularization is carried out only in orientations that do not lead to alteration of the contours of the structures of interest. BM4D appears to have the best restoration result but it actually oversmooths.

VII. Conclusion

In this article, we have proposed a regularization term for curvilinear structure restoration based on the design of a mixed gradient operator. This spatially-variant operator can be computed from orientation measures provided by curvilinear structure detectors. This regularization term has a different behaviour depending on the location in the image. Within curvilinear structures, it regularizes only in the local main direction of the considered structure. Within isotropic structures, it behaves like an isotropic gradient. Using the formalism of discrete calculus, this operator has a consistent behaviour everywhere. In particular it handles border effects gracefully both on the border of the image and at the interface between oriented and isotropic areas. The adjoint divergence operator is immediately derived again thanks to the discrete calculus formulation.

The mixed gradient developed in this work was developed and formalized in a non-dimensional way, thus leading to restoration approaches that can proceed irrespective of the space dimension (2D, 3D, etc.).

We illustrated our approach on segmentation using the

classical Chan *et al.* data fidelity term, and image denoising using the ℓ_2 norm fidelity term. However, other formulations may be considered, as long as they include a regularization term expressed as a gradient measure. For instance, it would make sense to use the exact Poisson log-likelihood as data fidelity term for the Poisson noise case [51]. Our next work will augment these strategies with additional priors (for instance considering connectivity), as well as non-local oriented gradient approaches.

Our approach may also be compliant with deep-learning strategies recently investigated in the field of curvilinear structure analysis. Indeed, deep-learning is emerging in this domain, mainly focusing on 2D cases at this stage [52], with a specific focus, e.g. on retinal imaging [53]. However, very recent works also aim at considering more complex cases, in 3D (bio)medical imaging modalities [54], [55], [56]. In this context, our mixed gradient may be involved as an additional loss term for strengthening connectivity of thin, curvilinear structures in complex nD images.

ACKNOWLEDGEMENTS

This work was funded with French *Agence Nationale de la Recherche* grant agreements ANR-12-MONU-0010, ANR-15-CE40-0006, ANR-18-CE45-0014 and ANR-18-CE45-0018.

APPENDIX

A. Gradient anisotropy correction

As most discrete gradient formulations, the one proposed in this article is anisotropic. Indeed, the standard gradient formulation described in Section IV-A is based on finite differences between values at $\mathbf{x}$ and at $\mathbf{x}+\mathbf{e}_i$ for the $\mathbf{e}_i$ vectors of the canonical basis of $\mathbb{Z}^n$ (see Eq. (9)). The choice of $\mathbf{x}+\mathbf{e}_i$ versus $\mathbf{x}-\mathbf{e}_i$ is arbitrary, and leads to compute the gradient in a cone that represents $1/2^n$ of the neighbourhood of $\mathbf{x}$ in $\mathbb{Z}^n$.

This anisotropy, together with the combined use of two kinds of gradients within a same image, may cause some slight boundary effects at the frontier between regions with / without specific directions. For instance, let us consider two neighbour points $\mathbf{x}$ and $\mathbf{y}$ such that $\alpha(\mathbf{x}) = 0$ and $\alpha(\mathbf{y}) = 1$ (see Eq. (14)), i.e. with a standard gradient at $\mathbf{x}$ and a directional gradient at $\mathbf{y}$.

At $\mathbf{y}$, the gradient formulation $\nabla_m f(\mathbf{y}) = \nabla_{\mathbf{d}} f(\mathbf{y})$ will take into account the difference of intensities between $I(\mathbf{y})$ and $I(\mathbf{x})$ only if the direction $\mathbf{d}(\mathbf{y})$ is mainly the same as $\mathbf{x}-\mathbf{y}$. Then, at $\mathbf{y}$, the global behaviour of $\nabla_m f$ is as expected.

By contrast, at $\mathbf{x}$, the gradient formulation $\nabla_m f(\mathbf{x}) = \nabla f(\mathbf{x})$ will take into account the difference of intensities between $I(\mathbf{y})$ and $I(\mathbf{x})$ whenever $\mathbf{y}-\mathbf{x}$ is one of the $\mathbf{e}_i$ vectors of the canonical basis. In such case, a gradient value will be computed in $\mathbf{x}$, while it may correspond to the (external) border of a curvilinear structure, where we expect to vanish the gradient. This phenomenon happens, by definition, on one side of the objects, oriented opposite to the canonical vector basis. This may cause, for instance, non-symmetric, blurring effects in segmentation results.

A simple way to get rid of such undesired effects consists of vanishing the i^{th} component of the standard gradient at $\mathbf{x}$ whenever $\mathbf{x}+\mathbf{e}_i$ is a point of $\mathbb{D}$ where a specific orientation is defined.

Practically, we only have to modify the conditions of validity of Eqs (25–26) as follows. We set:

$$\begin{aligned}\mathbb{D}_0 &= \{\mathbf{x} \in \mathbb{D} \mid \alpha(\mathbf{x}) = 0\} \\ \mathbb{D}_1 &= \{\mathbf{x} \in \mathbb{D} \mid \alpha(\mathbf{x}) = 1\}\end{aligned} \tag{38}$$

In other words, $\mathbb{D}_0$ (resp. $\mathbb{D}_1$) corresponds to the part of $\mathbb{D}$ where a standard (resp. directional) gradient is valid (see Eq. (14)). Then, we slightly modify the definition of the matrices G_i, $i \in [\![1, n]\!]$ (see Eq. (24)) constituting the gradient matrix of $V^{\nabla f}$ as follows:

$$g_i(j,k) = \begin{cases} \delta_{k,\sigma(\mathbf{x}_j+\mathbf{e}_i)} & \text{if } \mathbf{x}_j+\mathbf{e}_i \in \mathbb{D}_0 \\ \delta_{k,j} & \text{if } \mathbf{x}_j+\mathbf{e}_i \notin \mathbb{D}_0 \end{cases} \tag{39}$$

In other words, we now apply the Neumann condition of nullity of the gradient coordinates on the border of $\mathbb{D}_0$, i.e. on the frontier of the whole area of non-oriented points, and not only on the boundary of the image. The overall idea remains the same, that is avoiding to compute a classical gradient coordinate when the information carried by one of the two points is not known / not relevant for this kind of gradient.

B. Parameter optimization

In the following, the parameters of each method used in the experiments of Section VI are enumerated, and we present the parameter values used in these experiments.

1) Segmentation: Both compared methods share the same parameters: the background and foreground constants c_1 and c_2, and the regularization weight λ balancing between the data fidelity and the directional regularization. The constants c_1 and c_2 depend on the initial image. The images were preprocessed to homogenize the background, resulting in a mean background intensity of almost $c_1 = 0$. We then chose to set the c_2 constant to the Otsu threshold value of each image. Finally we performed a heuristic optimization of the λ parameter by looking for the best value in the range $[0.001, 0.5]$ for each training image for both methods; we chose for each method the most frequent λ value. This resulted in $\lambda = 0.10$ for the ROF method and $\lambda = 0.15$ for our method. One may note that our directional regularization allows for a higher regularization power without losing performance.

The directional regularization also requires a directional information provided here by the RORPO operator. RORPO has 3 parameters: the minimum scale s, the factor f and the number of scales n. They mostly depend on the dimension of the structure of interest, and the results are robust to parameter variations. We set them to: $s = 10$, $f = 2$, $n = 5$.

2) Denoising: The parameter optimization of this experiment was performed on the first image of the VascuSynth dataset. We sampled the parameter space of each method in reasonable ranges and we kept the set of parameters leading to the best MCC on this image.

- Directional regularization – The directional regularization requires a directional information provided in this work

TABLE III
ROF AND DIRECTIONAL REGULARIZATION PARAMETERS.

Noise	λ ROF	λ directional
2	0.001	0.010
5	0.010	0.050
10	0.010	0.010
15	0.050	0.050
20	0.010	0.050
30	0.050	0.050
50	0.050	0.050

TABLE IV
HDCS PARAMETERS.

Noise	λ_h	λ_c	λ_e	scale
2	20	40	20	0.5
5	20	40	20	0.5
10	40	20	20	0.5
15	30	30	10	1.0
20	20	40	10	1.0
30	40	30	10	1.0
50	40	20	10	1.0

by the RORPO operator. RORPO has 3 parameters: the minimum scale s, the factor f and the number of scales n. These parameters mostly depend on the size of the structures of interest, but the results are robust to parameter variations. We set them once for all noise levels: $s = 20$, $f = 1.7$ and $n = 4$. The directional regularization itself has one parameter, namely the regularization weight λ. The results of the parameter optimization are presented in Table III.

- ROF – The ROF model has one parameter, namely the regularization weight λ. The results of the parameter optimization are presented in Table III.
- HDCS – The HDCS model has 4 parameter: the hybrid contrast parameter λ_h, the CED contrast parameter λ_c, the EED contrast parameter λ_e and the scale parameter. The results of the parameters optimization are presented in Table IV.
- BM4D – The original BM4D implementation has a default option that estimates the best parameters for the given image. We used this option to compute all the BM4D denoising results.

REFERENCES

[1] O. Merveille, H. Talbot, L. Najman, and N. Passat, "Curvilinear structure analysis by ranking the orientation responses of path operators," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 2, pp. 304–317, 2018.

[2] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *MICCAI, Proceedings*, vol. 1496, 1998, pp. 130–137.

[3] O. Merveille, O. Miraucourt, S. Salmon, N. Passat, and H. Talbot, "A variational model for thin structure segmentation based on a directional regularization," in *ICIP, Proceedings*, 2016, pp. 4324–4328.

[4] H. Talbot, "Oriented patterns in image analysis: From thin objects to flow-based methods," Habilitation Thesis, Université Paris-Est, 2013, https://hal.archives-ouvertes.fr/tel-01099256.

[5] D. Lesage, E. Angelini, I. Bloch, and G. Funka-Lea, "A review of 3D vessel lumen segmentation techniques: Models, features and extraction schemes," *Medical Image Analysis*, vol. 13, no. 6, pp. 819–845, 2009.

[6] S. Moccia, E. De Momi, S. El Hadji, and L. S. Mattos, "Blood vessel segmentation algorithms — Review of methods, datasets and evaluation metrics," *Computer Methods and Programs in Biomedicine*, vol. 158, pp. 71–91, 2018.

[7] Y. Sato, S. Nakajima, H. Atsumi, T. Koller, G. Gerig, S. Yoshida, and R. Kikinis, "3D multi-scale line filter for segmentation and visualization of curvilinear structures in medical images," in *CVRMed-MRCAS, Proceedings*, vol. 1205, 1997, pp. 213–222.

[8] C. Lorenz, I.-C. Carlsen, T. M. Buzug, C. Fassnacht, and J. Weese, "Multi-scale line segmentation with automatic estimation of width, contrast and tangential direction in 2D and 3D medical images," in *CVRMed-MRCAS, Proceedings*, vol. 1205, 1997, pp. 233–242.

[9] K. Krissian, G. Malandain, N. Ayache, R. Vaillant, and Y. Trousset, "Model-based detection of tubular structures in 3D images," *Computer Vision and Image Understanding*, vol. 80, no. 2, pp. 130–171, 2000.

[10] P.-E. Danielsson, Q. Lin, and Q.-Z. Ye, "Efficient detection of second-degree variations in 2D and 3D images," *Journal of Visual Communication and Image Representation*, vol. 12, no. 3, pp. 255–305, 2001.

[11] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 9, pp. 891–906, 1991.

[12] K. G. Derpanis and J. M. Gryn, "Three-dimensional nth derivative of Gaussian separable steerable filters," in *ICIP, Proceedings*, vol. 3, 2005, pp. 553–556.

[13] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.

[14] J. Weickert, *Anisotropic diffusion in image processing*. Teubner Stuttgart, 1998, vol. 1.

[15] ——, "Coherence-enhancing diffusion filtering," *International Journal of Computer Vision*, vol. 31, no. 2, pp. 111–127, 1999.

[16] A. M. Mendrik, E.-J. Vonken, A. Rutten, M. A. Viergever, and B. van Ginneken, "Noise reduction in computed tomography scans using 3-D anisotropic hybrid diffusion with continuous switch," *IEEE Transactions on Medical Imaging*, vol. 28, no. 10, pp. 1585–1594, 2009.

[17] M. Welk, G. Steidl, and J. Weickert, "Locally analytic schemes: A link between diffusion filtering and wavelet shrinkage," *Applied and Computational Harmonic Analysis*, vol. 24, no. 2, pp. 195–224, 2008.

[18] L. Cohen and T. Deschamps, "Segmentation of 3D tubular objects with adaptive front propagation and minimal tree extraction for 3D medical imaging," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 10, no. 4, pp. 289–305, 2007.

[19] Y. Rouchdy and L. D. Cohen, "Geodesic voting for the automatic extraction of tree structures. Methods and applications," *Computer Vision and Image Understading*, vol. 117, no. 10, pp. 1453–1467, 2013.

[20] L. Vincent, "Minimal path algorithms for the robust detection of linear features in gray images," in *ISMM, Proceedings*, ser. Computational Imaging and Vision, vol. 12, 1998, pp. 331–338.

[21] H. J. A. M. Heijmans, M. Buckley, and H. Talbot, "Path openings and closings," *Journal of Mathematical Imaging and Vision*, vol. 22, no. 2–3, pp. 107–119, 2005.

[22] C. L. Luengo Hendriks, "Constrained and dimensionality-independent path openings," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1587–1595, 2010.

[23] H. Talbot and B. Appleton, "Efficient complete and incomplete path openings and closings," *Image and Vision Computing*, vol. 25, no. 4, pp. 416–425, 2007.

[24] F. Cokelaer, H. Talbot, and J. Chanussot, "Efficient robust d-dimensional path operators," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 7, pp. 830–839, 2012.

[25] O. Merveille, B. Naegel, H. Talbot, L. Najman, and N. Passat, "2D filtering of curvilinear structures by ranking the orientation responses of path operators (RORPO)," *Image Processing On Line*, vol. 7, pp. 246–261, 2017.

[26] A. N. Tikhonov, "Regularization of incorrectly posed problems," *Soviet Mathematics – Doklady*, vol. 4, no. 6, pp. 1624–1627, 1963.

[27] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1–4, pp. 259–268, 1992.

[28] C. Rhemann, C. Rother, P. Kohli, and M. Gelautz, "A spatially varying PSF-based prior for alpha matting," in *CVPR, Proceedings*. IEEE, 2010, pp. 2149–2156.

[29] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 1838–1857, 2011.

[30] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861–2873, 2010.

[31] T. F. Chan, A. M. Yip, and F. E. Park, "Simultaneous total variation image inpainting and blind deconvolution," *Journal of Imaging Systems and Technology*, vol. 15, pp. 92–102, 2005.

[32] M.-J. Fadili, J.-L. Starck, and F. Murtagh, "Inpainting and zooming using sparse representations," *The Computer Journal*, vol. 52, no. 1, pp. 64–79, 2009.

[33] T. Chan and L. Vese, "Active active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[34] T. F. Chan, S. Esedoglu, and M. Nikolova, "Algorithms for finding global minimizers of image segmentation and denoising models," *SIAM Journal on Applied Mathematics*, vol. 66, no. 5, pp. 1632–1648, 2006.

[35] J. A. Tyrrell, E. di Tomaso, D. Fuja, R. Tong, K. Kozak, R. K. Jain, and B. Roysam, "Robust 3-D modeling of vasculature imagery using superellipsoids," *IEEE Transactions on Medical Imaging*, vol. 26, no. 2, pp. 223–237, 2007.

[36] C. Tai, X. Zhang, and Z. Shen, "Wavelet frame based multi-phase image segmentation," *SIAM Journal on Imaging Sciences*, vol. 6, no. 4, pp. 2521–2546, 2013.

[37] R. Rigamonti and V. Lepetit, "Accurate and efficient linear structure segmentation by leveraging ad hoc features with learned filters," in *MICCAI, Proceedings*, vol. 7510, 2012, pp. 189–197.

[38] N. Y. El-Zehiry and L. Grady, "Vessel segmentation using 3D elastical regularization," in *ISBI, Proceedings*, 2012, pp. 1288–1291.

[39] O. Miraucourt, A. Jezierska, H. Talbot, S. Salmon, and N. Passat, "Variational method combined with Frangi vesselness for tubular object segmentation," in *CMBE, Proceedings*, 2015, pp. 485–488.

[40] D. L. Ruderman, "The statistics of natural images," *Network: Computation in Neural Systems*, vol. 5, no. 4, pp. 517–548, 1994.

[41] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Modeling & Simulation*, vol. 4, no. 4, pp. 1168–1200, 2005.

[42] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, 2011, pp. 185–212.

[43] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2419–2434, 2009.

[44] J. Staal, M. Abramoff, M. Niemeijer, M. Viergever, and B. van Ginneken, "Ridge based vessel segmentation in color images of the retina," *IEEE Transactions on Medical Imaging*, vol. 23, no. 4, pp. 501–509, 2004.

[45] B. Matthews, "Comparison of the predicted and observed secondary structure of t4 phage lysozyme," *Biochimica et Biophysica Acta (BBA) – Protein Structure*, vol. 405, no. 2, pp. 442 – 451, 1975.

[46] C. A. Lupaşcu, D. Tegolo, and E. Trucco, "FABC: Retinal vessel segmentation using AdaBoost," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 5, pp. 1267–1274, 2010.

[47] M. Al-Rawi, M. Qutaishat, and M. Arrar, "An improved matched filter for blood vessel detection of digital retinal images," *Computers in Biology and Medicine*, vol. 37, no. 2, pp. 262–267, 2007.

[48] G. Hamarneh and P. Jassi, "VascuSynth: Simulating vascular trees for generating volumetric image data with ground-truth segmentation and tree analysis," *Computerized Medical Imaging and Graphics*, vol. 34, no. 8, pp. 605–616, 2010.

[49] A. Chambolle, "An algorithm for total variation minimization and applications," vol. 20, no. 1, pp. 89–97, 2004.

[50] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi, "Nonlocal transform-domain filter for volumetric data denoising and reconstruction," *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 119–133, 2013.

[51] E. Chouzenoux, A. Jezierska, J.-C. Pesquet, and H. Talbot, "A convex approach for image restoration with exact poisson–gaussian likelihood," *SIAM Journal on Imaging Sciences*, vol. 8, no. 4, pp. 2662–2682, 2015.

[52] G. J. S. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.

[53] P. Liskowski and K. Krawiec, "Segmenting retinal blood vessels with deep neural networks," *IEEE Transactions on Medical Imaging*, vol. 35, no. 11, pp. 2369–2380, 2016.

[54] G. Tetteh, V. Efremov, N. D. Forkert, M. Schneider, J. Kirschke, B. Weber, C. Zimmer, M. Piraud, and B. H. Menze, "DeepVesselNet: Vessel segmentation, centerline prediction, and bifurcation detection in 3-D angiographic volumes," *CoRR*, vol. abs/1803.09340, 2018.

[55] M. Haft-Javaherian, L. Fang, V. Muse, C. B. Schaffer, N. Nishimura, and M. R. Sabuncu, "Deep convolutional neural networks for segmenting 3D in vivo multiphoton images of vasculature in Alzheimer disease mouse models," *CoRR*, vol. abs/1801.00880, 2018.

[56] P. Sanches, C. Meyer, V. Vigon, and B. Naegel, "Cerebrovascular network segmentation of MRA images with deep learning," in *ISBI, Proceedings*, 2019.

Odyssée Merveille received the engineering degree from the Institut Supérieur de BioSciences de Paris (Université Paris-Est Créteil / ESIEE Engineering) in 2013. She obtained the PhD from Université Paris-Est in 2016. She is currently post-doctoral fellow at Université de Strasbourg. Her scientific interests include mathematical morphology, machine learning and (bio)medical imaging.

Benoît Naegel studied Computer Science at the Universite Strasbourg 1 (M.Sc., 2000; Ph.D., 2004; Habilitation, 2018), specializing in image analysis. In 2005, he worked at the University of Applied Sciences Western Switzerland, Geneva, as a Research Assistant. Since 2007, he has been Assistant Professor of Computer Science at the Université de Nancy. Since 2010, he is Assistant Professor of Computer Science at the Université de Strasbourg. His scientific interests include mathematical morphology, document processing and medical imaging.

Hugues Talbot received the Habilitation from Université Paris-Est in 2013, the PhD from École des Mines de Paris in 1993 and the engineering degree from École Centrale de Paris in 1989. He was principal research scientist at CSIRO, Sydney, from 1994 to 2004. He is now a professor at CentraleSupelec, Université Paris-Saclay. He is the co-author of 6 books and over 200 articles in the area of mathematical morphology, discrete geometry, combinatorial and continuous optimization.

Nicolas Passat obtained the MSc and PhD from Université Strasbourg 1 in 2002 and 2005, and Habilitation from Université de Strasbourg in 2011. He was an assistant professor at Université de Strasbourg, between 2006 and 2012. He is now a professor at Université de Reims Champagne-Ardenne. His scientific interests include mathematical morphology, discrete geometry, medical imaging and remote sensing.