CrossMark

Variational Methods for Normal Integration

Yvain Quéau[1] · Jean-Denis Durou[2] · Jean-François Aujol[3,4]

Abstract
The need for an efficient method of integration of a dense normal field is inspired by several computer vision tasks, such as shape-from-shading, photometric stereo, deflectometry. Inspired by edge-preserving methods from image processing, we study in this paper several variational approaches for normal integration, with a focus on non-rectangular domains, free boundary and depth discontinuities. We first introduce a new discretization for quadratic integration, which is designed to ensure both fast recovery and the ability to handle non-rectangular domains with a free boundary. Yet, with this solver, discontinuous surfaces can be handled only if the scene is first segmented into pieces without discontinuity. Hence, we then discuss several discontinuity-preserving strategies. Those inspired, respectively, by the Mumford–Shah segmentation method and by anisotropic diffusion, are shown to be the most effective for recovering discontinuities.

Keywords 3D-reconstruction · Integration · Normal field · Gradient field · Variational methods · Photometric stereo · Shape-from-shading

1 Introduction

In this paper, we study several methods for numerical integration of a gradient field over a 2D-grid. Our aim is to estimate the values of a function $z: \mathbb{R}^2 \to \mathbb{R}$, over a set $\Omega \subset \mathbb{R}^2$ (reconstruction domain) where an estimate $\mathbf{g} = [p, q]^\top: \Omega \to \mathbb{R}^2$ of its gradient ∇z is available. Formally, we want to solve the following equation in the unknown *depth map z*:

$$\nabla z(u, v) = \underbrace{[p(u, v), q(u, v)]^\top}_{\mathbf{g}(u,v)}, \quad \forall (u, v) \in \Omega \tag{1}$$

In a companion survey paper [48], we have shown that an ideal numerical tool for solving Eq. (1) should satisfy the following properties, apart accuracy:

- $\mathcal{P}_{\text{Fast}}$: be as *fast* as possible;
- $\mathcal{P}_{\text{Robust}}$: be *robust* to a noisy gradient field;

✉ Yvain Quéau
yvain.queau@tum.de

[1] Technical University Munich, Garching, Germany

[2] IRIT, Université de Toulouse, Toulouse, France

[3] IMB, Université de Bordeaux, Talence, France

[4] Institut Universitaire de France, Paris, France

- $\mathcal{P}_{\text{FreeB}}$: be able to handle a *free boundary*;
- $\mathcal{P}_{\text{Disc}}$: preserve the *depth discontinuities*;
- $\mathcal{P}_{\text{NoRect}}$: be able to work on a *non-rectangular* domain Ω;
- $\mathcal{P}_{\text{NoPar}}$: have *no critical parameter* to tune.

Contributions This paper builds upon the previous conference papers [19,20,47] to clarify the building blocks of variational approaches to the integration problem, with a view to meeting the largest subset of these requirements. As discussed in Sect. 2, the variational framework is well adapted to this task, thanks to its flexibility. However, these properties are difficult, if not impossible, to satisfy simultaneously. In particular, $\mathcal{P}_{\text{Disc}}$ seems hardly compatible with $\mathcal{P}_{\text{Fast}}$ and $\mathcal{P}_{\text{NoPar}}$.

Therefore, we first focus in Sect. 3 on the properties $\mathcal{P}_{\text{FreeB}}$ and $\mathcal{P}_{\text{NoRect}}$. A new discretization strategy for normal integration is presented, which is independent from the shape of the domain and assumes no particular boundary condition. When used within a quadratic variational approach, this discretization strategy allows to ensure all the desired properties except $\mathcal{P}_{\text{Disc}}$. In particular, the numerical solution comes down to solving a symmetric, diagonally dominant linear system, which can be achieved very efficiently using preconditioning techniques. In comparison with our previous work [20] which considered only forward finite differences

and standard Jacobi iterations, the properties $\mathcal{P}_{\text{Robust}}$ and $\mathcal{P}_{\text{Fast}}$ are better satisfied.

In Sect. 4, we focus more specifically on the integration problem in the presence of discontinuities. Several variations of well-known models from image processing are empirically compared, while suggesting for each of them the appropriate state-of-the-art minimization method. Besides the approaches based on total variation and non-convex regularization, which we already presented, respectively, in [19,47], two new methods inspired by the Mumford–Shah segmentation method and by anisotropic diffusion are introduced. They are shown to be particularly effective for handling $\mathcal{P}_{\text{Disc}}$, although $\mathcal{P}_{\text{Fast}}$ and $\mathcal{P}_{\text{NoPar}}$ are lost.

These variational methods for normal integration are based on the same variational framework, which is detailed in the next section.

2 From Variational Image Restoration to Variational Normal Integration

In view of the $\mathcal{P}_{\text{Robust}}$ property, variational methods, which aim at estimating the surface by minimization of a well-chosen criterion, are particularly suited for the integration problem. Hence, we choose the variational framework as basis for the design of new methods. This choice is also motivated by the fact that the property which is the most difficult to ensure is probably $\mathcal{P}_{\text{Disc}}$. Numerous variational methods have been designed for edge-preserving image processing: such methods may thus be a natural source of inspiration for designing discontinuity-preserving integration methods.

2.1 Variational Methods in Image Processing

For a comprehensive introduction to this literature, we refer the reader to [4] and to pioneering papers such as [11,16,34, 38]. Basically, the idea in edge-preserving image restoration is that edges need to be processed in a particular way. This is usually achieved by choosing an appropriate energy to minimize, formulating the inverse problem as the recovery of a restored image $z \colon \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ minimizing the energy:

$$\mathcal{E}(z) = \mathcal{F}(z) + \mathcal{R}(z) \tag{2}$$

where

- $\mathcal{F}(z)$ is a *fidelity* term penalizing the difference between a corrupted image z^0 and the restored image:

$$\mathcal{F}(z) = \iint\limits_{(u,v) \in \Omega} \Phi\left(z(u,v) - z^0(u,v)\right) \mathrm{d}u\, \mathrm{d}v \tag{3}$$

- $\mathcal{R}(z)$ is a *regularization* term, which usually penalizes the gradient of the restored image:

$$\mathcal{R}(z) = \iint\limits_{(u,v) \in \Omega} \lambda(u,v)\, \Psi\left(\|\nabla z(u,v)\|\right) \mathrm{d}u\, \mathrm{d}v \tag{4}$$

In (3), Φ is chosen accordingly to the type of corruption the original image z^0 is affected by. For instance, $\Phi_{L_2}(s) = s^2$ is the natural choice in the presence of additive, zero-mean, Gaussian noise, while $\Phi_{L_1}(s) = |s|$ can be used in the presence of bi-exponential (Laplacian) noise, which is a rather good model when outliers come into play (e.g., "salt and pepper" noise).

In (4), $\lambda \geq 0$ is a field of weights which control the respective influence of the fidelity and the regularization terms. It can be either manually tuned beforehand [if $\lambda(u,v) \equiv \lambda$, λ can be seen as a "hyper-parameter"], or defined as a function of $\|\nabla z(u,v)\|$.

The choice of Ψ must be made accordingly to a desired smoothness of the restored image. The *quadratic penalty* $\Psi_{L_2}(s) = s^2$ will produce "smooth" images, while piecewise-constant images are obtained when choosing the *sparsity penalty* $\Psi_{L_0}(s) = 1 - \delta(s)$, with $\delta(s) = 1$ if $s = 0$ and $\delta(s) = 0$ otherwise. The latter approach preserves the edges, but the numerical solving is much more difficult, since the regularization term is non-smooth and non-convex. Hence, several choices of regularizers "inbetween" the quadratic and the sparsity ones have been suggested.

For instance, the total variation (TV) regularizer is obtained by setting $\Psi(s) = |s|$. Efficient numerical methods exist for solving this non-smooth, yet convex, problem. Examples include primal–dual methods [13], augmented Lagrangian approaches [23], and forward–backward splittings [40]. The latter can also be adapted to the case where the regularizer Ψ is non-convex, but smooth [41]. Such non-convex regularization terms were shown to be particularly effective for edge-preserving image restoration [22, 36,38].

Another strategy is to stick to quadratic regularization ($\Psi = \Psi_{L_2}$), but apply it in a non-uniform manner by tuning the field of weights λ in (4). For instance, setting $\lambda(u,v)$ in (4) inversely proportional to $\|\nabla z(u,v)\|$ yields the "anisotropic diffusion" model by Perona and Malik [44]. The discontinuity set K can also be automatically estimated and discarded by setting $\lambda(u,v) \equiv 0$ over K and $\lambda(u,v) \equiv \lambda$ over $\Omega \setminus K$, in the spirit of Mumford and Shah's segmentation method [37].

2.2 Notations

Although we chose for simplicity to write the variational problems in a continuous form, we are overall interested

in solving discrete problems. Two different discretization strategies exist. The first one consists in using variational calculus to derive the (continuous) necessary optimality condition, then discretize it by finite differences, and eventually solve the discretized optimality condition. The alternative method is to discretize the functional itself by finite differences, before solving the optimality condition associated with the discrete problem.

As shown in [20], the latter approach eases the handling of the boundary of Ω; hence, we use it as discretization strategy. The variational models hereafter will be presented using the continuous notations, because we find them more readable. The discrete notations will be used only when presenting the numerical solving. Yet, to avoid confusion, we will use calligraphic letters for the continuous energies (e.g., $\mathcal{E}$), and capital letters for their discrete counterparts (e.g., E). With these conventions, it should be clear whether an optimization problem is discrete or continuous. Hence, we will use the same notation $\nabla z = [\partial_u z, \partial_v z]^\top$ both for the gradient of z and its finite differences approximation.

2.3 Proposed Variational Framework

In this work, we show how to adapt the aforementioned variational models, originally designed for image restoration, to normal integration. Although both these inverse problems are formally very similar, they are somehow different, for the following reasons:

- The concept of edges in an image to restore is replaced by those of depth discontinuities and kinks.
- Unlike image processing functionals, our data consist in an estimate $\mathbf{g}$ of the *gradient* of the unknown z, in lieu of a corrupted version z^0 of z. Therefore, the fidelity term $\mathcal{F}(z)$ will apply to the difference between ∇z and $\mathbf{g}$, and it is the choice of this term which will or not allow depth discontinuities.
- Regularization terms are optional here: all the methods we discuss basically work even with $\mathcal{R}(z) \equiv 0$, but we may use this regularization term to allow introducing, if available, a prior on the surface (e.g., user-defined control points [30,33] or a rough depth estimate obtained using a low-resolution depth sensor [32]). Such feature "is appreciable, although not required" [48].

We will discuss methods seeking the depth z as the minimizer of an energy $\mathcal{E}(z)$ in the form (2), but with different choices for $\mathcal{F}(z)$ and $\mathcal{R}(z)$:

- $\mathcal{F}(z)$ now represents a fidelity term penalizing the difference between the gradient of the recovered depth map z and the datum $\mathbf{g}$:

$$\mathcal{F}(z) = \iint_{(u,v)\in\Omega} \Phi\left(\|\nabla z(u,v) - \mathbf{g}(u,v)\|\right) \mathrm{d}u\,\mathrm{d}v \quad (5)$$

- The regularization term $\mathcal{R}(z)$ now represents prior knowledge of the depth[1]:

$$\mathcal{R}(z) = \iint_{(u,v)\in\Omega} \lambda(u,v)\left[z(u,v) - z^0(u,v)\right]^2 \quad (6)$$

where z^0 is the prior, and $\lambda(u,v) \geq 0$ is a user-defined, spatially varying, regularization weight. In this work, we consider for simplicity only the case where λ does not depend on z.

2.4 Choosing λ and z^0

The main purpose of the regularization term $\mathcal{R}$ defined in (6) is to avoid numerical instabilities which may arise when considering solely the fidelity term (5): this fidelity term depends only on ∇z, and not on z itself; hence, the minimizer of (5) can be estimated only up to an additive ambiguity.

Besides, one may also want to impose one or several control points on the surface [30,33]. This can be achieved very simply within the proposed variational framework, by setting $\lambda(u,v) \equiv 0$ everywhere, except on the control points locations (u,v) where a high value for $\lambda(u,v)$ must be set and the value $z^0(u,v)$ is fixed.

Another typical situation is when, given both a coarse depth estimate and an accurate normal estimate, one would like to "merge" them in order to create a high-quality depth map. Such a problem arises, for instance, when refining the depth map of an RGB-D sensor (e.g., a Kinect) by means of shape-from-shading [42], photometric stereo [25] or shape-from-polarization [32]. In such cases, we may set z^0 to the coarse depth map, and tune λ so as to merge the coarse and fine estimates in the best way. Non-uniform weights may be used, in order to lower the influence of outliers in the coarse depth map [25].

Eventually, in the absence of such priors, we will use the regularization term only to fix the integration constant: this is easily achieved by setting an arbitrary prior [e.g., $z^0(u,v) \equiv 0$], along with a small value for λ [typically, $\lambda(u,v) \equiv \lambda = 10^{-6}$].

3 Smooth Surfaces

We first tackle the problem of recovering a "smooth" depth map z from a noisy estimate $\mathbf{g}$ of ∇z. To this end, we consider the quadratic variational problem:

---

[1] We consider only quadratic regularization terms: studying more robust ones (e.g., L^1 norm) is left as perspective.

$$\min_z \iint_{(u,v)\in\Omega} \|\nabla z(u,v) - \mathbf{g}(u,v)\|^2$$

$$+ \lambda(u,v) \left[z(u,v) - z^0(u,v) \right]^2 \mathrm{d}u \, \mathrm{d}v \qquad (7)$$

When $\lambda \equiv 0$, Problem (7) comes down to Horn and Brook's model [29]. In that particular case, an infinity of solutions $z \in W^{1,2}(\Omega)$ exist, and they differ by an additive constant.[2] On the other hand, the regularization term allows us to guarantee uniqueness of the solution as soon as λ is strictly positive almost everywhere.[3,4]

If the depth map z is further assumed to be twice differentiable, the necessary optimality condition associated with the continuous optimization problem (7) (Euler–Lagrange equation) is written:

$$-\Delta z + \lambda z = -\nabla \cdot \mathbf{g} + \lambda z^0 \quad \text{over } \Omega \qquad (8)$$
$$(\nabla z - \mathbf{g}) \cdot \boldsymbol{\eta} = 0 \quad \text{over } \partial\Omega \qquad (9)$$

with $\boldsymbol{\eta}$ a normal vector to the boundary $\partial\Omega$ of Ω, Δ the Laplacian operator, and $\nabla \cdot$ the divergence operator. This condition is a linear PDE in z which can be discretized using finite differences. Yet, providing a consistent discretization on the boundary of Ω is not straightforward [26], especially when dealing with non-rectangular domains Ω where many cases have to be considered [6]. Hence, we follow a different track, based on the discretization of the functional itself.

3.1 Discretizing the Functional

Instead of a continuous gradient field $\mathbf{g}: \Omega \to \mathbb{R}^2$ over an open set Ω, we are actually given a finite set of values $\{\mathbf{g}_{u,v} = [p_{u,v}, q_{u,v}]^\top, (u,v) \in \Omega\}$, where the (u,v) represent the *pixels* of a discrete subset Ω of a regular square 2D-grid.[5] Solving the discrete integration problem requires estimating a finite set of values, i.e., the $|\Omega|$ unknown depth values $z_{u,v}$, $(u,v) \in \Omega$ ($|\cdot|$ denotes the cardinality), which are stacked columnwise in a vector $\mathbf{z} \in \mathbb{R}^{|\Omega|}$.

For now, let us use a Gaussian approximation for the noise contained in $\mathbf{g}$.[6] That is, let us assume in the rest of this section that each datum $\mathbf{g}_{u,v}$, $(u,v) \in \Omega$, is equal to the gradient $\nabla z(u,v)$ of the unknown depth map z, taken at point (u,v), up to a zero-mean additive, homoskedastic (same variance at each location (u,v)), Gaussian noise:

$$\mathbf{g}_{u,v} = \nabla z(u,v) + \boldsymbol{\epsilon}(u,v) \qquad (10)$$

where $\boldsymbol{\epsilon}(u,v) \sim \mathcal{N}\left([0,0]^\top, \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}\right)$ and σ is unknown.[7] Now, we need to give a discrete interpretation of the gradient operator in (10), through finite differences.

In order to obtain a second-order accurate discretization, we combine forward and backward first-order finite differences, i.e., we consider that each measure of the gradient $\mathbf{g}_{u,v} = [p_{u,v}, q_{u,v}]^\top$ provides us with up to four independent and identically distributed (i.i.d.) statistical observations, depending on the neighborhood of (u,v). Indeed, its first component $p_{u,v}$ can be understood either in terms of both forward or backward finite differences (when both the bottom and the top[8] neighbors are inside Ω), by one of both these discretizations (only one neighbor inside Ω), or by none of these finite differences (no neighbor inside Ω). Formally, we model the p-observations in the following way:

$$p_{u,v} = \overbrace{z_{u+1,v} - z_{u,v}}^{\partial_u^+ z_{u,v}} + \epsilon_u^+(u,v),$$
$$\forall (u,v) \in \underbrace{\{(u,v) \in \Omega \mid (u+1,v) \in \Omega\}}_{\Omega_u^+} \qquad (11)$$

$$p_{u,v} = \overbrace{z_{u,v} - z_{u-1,v}}^{\partial_u^- z_{u,v}} + \epsilon_u^-(u,v),$$
$$\forall (u,v) \in \underbrace{\{(u,v) \in \Omega \mid (u-1,v) \in \Omega\}}_{\Omega_u^-} \qquad (12)$$

where $\epsilon_u^{+/-} \sim \mathcal{N}(0, \sigma^2)$. Hence, rather than considering that we are given $|\Omega|$ observations p, our discretization handles these data as $|\Omega_u^+| + |\Omega_u^-|$ observations, some of them being interpreted in terms of forward differences, some in terms of backward differences, some in terms of both forward and backward differences, the points without any neighbor in the u-direction being excluded.

---

[2] Proof: by developing the terms inside the integral in (7), and integrating by parts, Theorem 6.2.5 in [3] applies with $f := -\nabla \cdot \mathbf{g}$ and $g := \mathbf{g} \cdot \boldsymbol{\eta}$.

[3] Proof: by developing the terms inside the integral in (7) and integrating by parts, Theorem 6.2.2-(ii) in [3] applies with $f := -\nabla \cdot \mathbf{g} + \lambda z^0$ and $g := \mathbf{g} \cdot \boldsymbol{\eta}$.

[4] This condition makes the matrix of the associated discrete problem strictly diagonally dominant, see Sect. 3.2.

[5] To ease the comparison between the variational and the discrete problems, we will use the same notation Ω for both the open set of $\mathbb{R}^2$ and the discrete subset of the grid.

[6] In 3D-reconstruction applications such as photometric stereo [55], the assumption on the noise should rather be formulated on the images. This will be discussed in more details in Sect. 4.4.

[7] The assumptions of equal variance σ^2 for both components and of a diagonal covariance matrix are introduced only for consistency with the least-squares problem (7). They are discussed with more care in Sect. 4.4.

[8] The u-axis points "downward," the v-axis points "to the right" and the z-axis points from the surface to the camera, see Fig. 1.

Symmetrically, the second component q of $\mathbf{g}$ corresponds either to two, one or zero observations:

$$q_{u,v} = \overbrace{z_{u,v+1} - z_{u,v}}^{\partial_v^+ z_{u,v}} + \epsilon_v^+(u,v),$$
$$\forall (u,v) \in \underbrace{\{(u,v) \in \Omega \mid (u,v+1) \in \Omega\}}_{\Omega_v^+} \quad (13)$$

$$q_{u,v} = \overbrace{z_{u,v} - z_{u,v-1}}^{\partial_v^- z_{u,v}} + \epsilon_v^-(u,v),$$
$$\forall (u,v) \in \underbrace{\{(u,v) \in \Omega \mid (u,v-1) \in \Omega\}}_{\Omega_v^-} \quad (14)$$

where $\epsilon_v^{+/-} \sim \mathcal{N}(0, \sigma^2)$. Given the Gaussianity of the noises $\epsilon_{u/v}^{+/-}$, their independence, and the fact that they all share the same standard deviation σ and mean 0, the joint likelihood of the observed gradients $\{\mathbf{g}_{u,v}\}_{(u,v)}$ is:

$$L\left(\{\mathbf{g}_{u,v}, (u,v) \in \Omega\} \mid \{z_{u,v}, (u,v) \in \Omega\}\right)$$
$$= \prod_{(u,v) \in \Omega_u^+} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{[\partial_u^+ z_{u,v} - p_{u,v}]^2}{2\sigma^2}\right\}$$
$$\times \prod_{(u,v) \in \Omega_u^-} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{[\partial_u^- z_{u,v} - p_{u,v}]^2}{2\sigma^2}\right\}$$
$$\times \prod_{(u,v) \in \Omega_v^+} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{[\partial_v^+ z_{u,v} - q_{u,v}]^2}{2\sigma^2}\right\}$$
$$\times \prod_{(u,v) \in \Omega_v^-} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{[\partial_v^- z_{u,v} - q_{u,v}]^2}{2\sigma^2}\right\}, \quad (15)$$

and hence, the maximum-likelihood estimate for the depth values is obtained by minimizing:

$$F_{L_2}(\mathbf{z}) = \frac{1}{2}\left(\sum\sum_{(u,v) \in \Omega_u^+} [\partial_u^+ z_{u,v} - p_{u,v}]^2\right.$$
$$\left. + \sum\sum_{(u,v) \in \Omega_u^-} [\partial_u^- z_{u,v} - p_{u,v}]^2\right)$$
$$+ \frac{1}{2}\left(\sum\sum_{(u,v) \in \Omega_v^+} [\partial_v^+ z_{u,v} - q_{u,v}]^2\right.$$
$$\left. + \sum\sum_{(u,v) \in \Omega_v^-} [\partial_v^- z_{u,v} - q_{u,v}]^2\right) \quad (16)$$

where the $\frac{1}{2}$ coefficients are meant to ease the continuous interpretation: the integral of the fidelity term in (7) is approximated by $F_{L_2}(\mathbf{z})$, expressed in (16) as the mean of the forward and the backward discretizations.

To obtain a more concise representation of this fidelity term, let us stack the data in two vectors $\mathbf{p} \in \mathbb{R}^{|\Omega|}$ and $\mathbf{q} \in \mathbb{R}^{|\Omega|}$. In addition, let us introduce four $|\Omega| \times |\Omega|$ differentiation matrices $\mathbf{D}_u^+$, $\mathbf{D}_u^-$, $\mathbf{D}_v^+$ and $\mathbf{D}_v^-$, associated with the finite differences operators $\partial_{u/v}^{+/-}$. For instance, the ith line of $\mathbf{D}_u^+$ reads:

$$\left(\mathbf{D}_u^+\right)_{i,\cdot}$$
$$= \begin{cases} \left[0,\ldots,0, \underbrace{-1}_{\text{Position } i}, \underbrace{1}_{\text{Position } i+1}, 0,\ldots,0\right] & \text{if } m(i) \in \Omega_u^+ \\ \mathbf{0}^\top & \text{otherwise} \end{cases}$$
$$(17)$$

where m is the mapping associating linear indices i with the pixel coordinates (u,v):

$$m: \{1,\ldots,|\Omega|\} \to \Omega$$
$$i \mapsto m(i) = (u,v) \quad (18)$$

Once these matrices are defined, (16) is equal to:

$$F_{L_2}(\mathbf{z}) = \frac{1}{2}\left(\left\|\mathbf{D}_u^+ \mathbf{z} - \mathbf{p}\right\|^2 + \left\|\mathbf{D}_u^- \mathbf{z} - \mathbf{p}\right\|^2\right)$$
$$+ \frac{1}{2}\left(\left\|\mathbf{D}_v^+ \mathbf{z} - \mathbf{q}\right\|^2 + \left\|\mathbf{D}_v^- \mathbf{z} - \mathbf{q}\right\|^2\right)$$
$$- \frac{1}{2}\left(\sum\sum_{(u,v) \in \Omega \setminus \Omega_u^+} p_{u,v}^2 + \sum\sum_{(u,v) \in \Omega \setminus \Omega_u^-} p_{u,v}^2\right)$$
$$- \frac{1}{2}\left(\sum\sum_{(u,v) \in \Omega \setminus \Omega_v^+} q_{u,v}^2 + \sum\sum_{(u,v) \in \Omega \setminus \Omega_v^-} q_{u,v}^2\right)$$
$$(19)$$

The terms in both the last rows of (19) being independent from the z-values, they do not influence the actual minimization and will thus be omitted from now on.

The regularization term (6) is discretized as:

$$R(\mathbf{z}) = \sum\sum_{(u,v) \in \Omega} \lambda_{u,v} \left[z_{u,v} - z_{u,v}^0\right]^2 = \left\|\boldsymbol{\Lambda}\left(\mathbf{z} - \mathbf{z}^0\right)\right\|^2 \quad (20)$$

with $\boldsymbol{\Lambda}$ a $|\Omega| \times |\Omega|$ diagonal matrix containing the values $\sqrt{\lambda_{u,v}}$, $(u,v) \in \Omega$.

Putting it altogether, our quadratic integration method reads as the minimization of the discrete functional:

$$E_{L_2}(\mathbf{z}) = \frac{1}{2}\left(\left\|\mathbf{D}_u^+ \mathbf{z} - \mathbf{p}\right\|^2 + \left\|\mathbf{D}_u^- \mathbf{z} - \mathbf{p}\right\|^2\right)$$
$$+ \frac{1}{2}\left(\left\|\mathbf{D}_v^+ \mathbf{z} - \mathbf{q}\right\|^2 + \left\|\mathbf{D}_v^- \mathbf{z} - \mathbf{q}\right\|^2\right) + \left\|\boldsymbol{\Lambda}\left(\mathbf{z} - \mathbf{z}^0\right)\right\|^2$$
$$(21)$$

3.2 Numerical Solution

The optimality condition associated with the discrete functional (21) is a linear equation in $\mathbf{z}$:

$$\mathbf{A}\mathbf{z} = \mathbf{b} \tag{22}$$

where $\mathbf{A}$ is a $|\Omega| \times |\Omega|$ symmetric matrix[9]:

$$\mathbf{A} = \overbrace{\frac{1}{2}\left[\mathbf{D}_u^{+\top}\mathbf{D}_u^+ + \mathbf{D}_u^{-\top}\mathbf{D}_u^- + \mathbf{D}_v^{+\top}\mathbf{D}_v^+ + \mathbf{D}_v^{-\top}\mathbf{D}_v^-\right]}^{\mathbf{L}} \\ + \Lambda^2 \tag{23}$$

and $\mathbf{b}$ is a $|\Omega| \times 1$ vector:

$$\mathbf{b} = \frac{1}{2}\overbrace{\left[\mathbf{D}_u^{+\top} + \mathbf{D}_u^{-\top}\right]}^{\mathbf{D}_u}\mathbf{p} + \frac{1}{2}\overbrace{\left[\mathbf{D}_v^{+\top} + \mathbf{D}_v^{-\top}\right]}^{\mathbf{D}_v}\mathbf{q} \\ + \Lambda^2\mathbf{z}^0 \tag{24}$$

The matrix $\mathbf{A}$ is sparse: it contains at most five nonzero entries per row. In addition, it is diagonal dominant: if $(\Lambda)_{i,i} = 0$, the value $(A)_{i,i}$ of a diagonal entry is equal to the opposite of the sum of the other entries $(A)_{i,j}$, $i \neq j$, from the same row i. It becomes strictly superior as soon as $(\Lambda)_{i,i}$ is strictly positive. Let us also remark that, when Ω describes a rectangular domain and the regularization weights are uniform ($\lambda(u,v) \equiv \lambda$), $\mathbf{A}$ is a Toeplitz matrix. Yet, this structure is lost in the general case where it can only be said that $\mathbf{A}$ is a sparse, symmetric, diagonal dominant (SDD) matrix with at most $5|\Omega|$ nonzero elements. It is positive semi-definite when $\Lambda = 0$, and positive definite as soon as one of the $\lambda_{u,v}$ is nonzero.

System (22) can be solved by means of the conjugate gradient algorithm. Initialization will not influence the actual solution, but it may influence the number of iterations required to reach convergence. In our experiments, we used z^0 as initial guess, yet more elaborate initialization strategies may yield faster convergence [6]. To ensure $\mathcal{P}_{\text{Fast}}$, we used the multigrid preconditioning technique [35], which has a negligible cost of computation and still bounds the computational complexity required to reach a ϵ relative accuracy[10] by:

$$O\left(5n\,\log(n)\,\log(1/\epsilon)\right) \tag{25}$$

where $n = |\Omega|$.[11] This complexity is inbetween the complexities of the approaches based on Sylvester equations [26] ($O(n^{1.5})$) and on DCT [53] ($O(n\log(n))$). Besides, these competing methods explicitly require that Ω is rectangular, while ours does not.

By construction, the integration method consisting in minimizing (21) satisfies the $\mathcal{P}_{\text{Robust}}$ property (it is the maximum-likelihood estimate in the presence of zero-mean Gaussian noise). The discretization we introduced does not assume any particular shape for Ω, neither treats the boundary in a specific manner; hence, $\mathcal{P}_{\text{FreeB}}$ and $\mathcal{P}_{\text{NoRect}}$ are also satisfied. We also showed that $\mathcal{P}_{\text{Fast}}$ could be satisfied, using a solving method based on the preconditioned conjugate gradient algorithm. Eventually, let us recall that tuning λ and/or manually fixing the values of the prior z_0 is necessary only to introduce a prior, but not in general. Hence, $\mathcal{P}_{\text{NoPar}}$ is also enforced. In conclusion, all the desired properties are satisfied, except $\mathcal{P}_{\text{Disc}}$. Let us now provide additional remarks on the connections between the proposed discrete approach and a fully variational one.

3.3 Continuous Interpretation

System (22) is nothing else than a discrete analogue of the continuous optimality conditions (8) and (9):

$$\underbrace{\mathbf{L}\mathbf{z}}_{\approx -\Delta z} + \underbrace{\Lambda^2\mathbf{z}}_{\approx \lambda z} = \underbrace{\mathbf{D}_u\mathbf{p} + \mathbf{D}_v\mathbf{q}}_{\approx -\nabla\cdot\mathbf{g}} + \underbrace{\Lambda^2\mathbf{z}^0}_{\approx \lambda z^0} \tag{26}$$

where the matrix–vector products are easily interpreted in terms of the differential operators in the continuous formula (8). One major advantage when reasoning from the beginning in the discrete setting is that one does not need to find out how to discretize the *natural*[12] boundary condition (9), which was already emphasized in [20,26]. Yet, the identifications in (26) show that both the discrete and continuous approaches are equivalent, provided that an appropriate discretization of the continuous optimality condition is used. It is thus possible to derive $O(5n\,\log(n)\,\log(1/\epsilon))$ algorithms based on the discretization of the Euler–Lagrange equation, contrarily to what is stated in [26]. The real drawback of such approaches does not lie in complexity, but in the difficult discretization of the boundary condition. This is further explored in the next subsection.

---

[9] $\mathbf{A}$ and $\mathbf{b}$ are purposely divided by two in order to ease the continuous interpretation of Sect. 3.3.

[10] In our experiments, the threshold of the stopping criterion is set to $\epsilon = 10^{-4}$.

[11] In (25), the factor $5n$ is nothing else than the number of nonzero elements in $\mathbf{A}$. Therefore, exploiting sparsity is not as "fruitless" as argued in [26] when it comes to solving large linear systems faster than using Gaussian elimination (complexity $O(n^3)$).

[12] As stated in [26], homogeneous Neumann boundary conditions of the type $\nabla z \cdot \boldsymbol{\eta} = 0$, used, e.g., in [1], should be avoided.
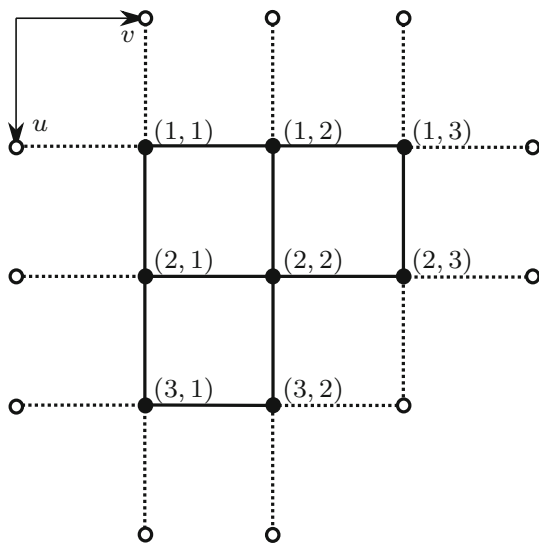
Fig. 1 Example of non-rectangular domain Ω (solid dots) inside a 3×3 grid. When invoking the continuous optimality condition, the discrete approximations of the Laplacian and of the divergence near the boundary involve several points inside $\partial\Omega$ (circles) for which no data are available. The first-order approximation of the *natural* boundary condition (9) is thus required. Relying only on discrete optimization simplifies a lot the boundary handling

3.4 Example

To clarify the proposed discretization of the integration problem, let us consider a non-rectangular domain Ω inside a 3×3 grid, like the one depicted in Fig. 1.

The vectorized unknown depth $\mathbf{z}$ and the vectorized components $\mathbf{p}$ and $\mathbf{q}$ of the gradient write in this case:

$$\mathbf{z} = \begin{bmatrix} z_{1,1} \\ z_{2,1} \\ z_{3,1} \\ z_{1,2} \\ z_{2,2} \\ z_{3,2} \\ z_{1,3} \\ z_{2,3} \end{bmatrix} \quad \mathbf{p} = \begin{bmatrix} p_{1,1} \\ p_{2,1} \\ p_{3,1} \\ p_{1,2} \\ p_{2,2} \\ p_{3,2} \\ p_{1,3} \\ p_{2,3} \end{bmatrix} \quad \mathbf{q} = \begin{bmatrix} q_{1,1} \\ q_{2,1} \\ q_{3,1} \\ q_{1,2} \\ q_{2,2} \\ q_{3,2} \\ q_{1,3} \\ q_{2,3} \end{bmatrix} \tag{27}$$

The sets $\Omega_{u/v}^{+/-}$ all contain five pixels:

$$\Omega_u^+ = \{(1,1),(2,1),(1,2),(2,2),(1,3)\} \tag{28}$$
$$\Omega_u^- = \{(2,1),(3,1),(2,2),(3,2),(2,3)\} \tag{29}$$
$$\Omega_v^+ = \{(1,1),(2,1),(3,1),(1,2),(2,2)\} \tag{30}$$
$$\Omega_v^- = \{(1,2),(2,2),(3,2),(1,3),(2,3)\} \tag{31}$$

so that the differentiation matrices $\mathbf{D}_{u/v}^{+/-}$ have five nonzero rows according to their definition (17). For instance, the matrix associated with the forward finite differences operator ∂_u^+ reads:

$$\mathbf{D}_u^+ = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{32}$$

The negative Laplacian matrix $\mathbf{L}$ defined in (23) is worth:

$$\mathbf{L} = \begin{bmatrix} 2 & -1 & 0 & -1 & 0 & 0 & 0 & 0 \\ -1 & 3 & -1 & 0 & -1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 & 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 3 & -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 & 4 & -1 & 0 & -1 \\ 0 & 0 & -1 & 0 & -1 & 2 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & 2 \end{bmatrix} \tag{33}$$

One can observe that this matrix describes the connectivity of the graph representing the discrete domain Ω: the diagonal elements $(\mathbf{L})_{i,i}$ are the numbers of neighbors connected to the ith point, and the off-diagonals elements $(\mathbf{L})_{i,j}$ are worth -1 if the ith and jth points are connected, 0 otherwise.

Eventually, the matrices $\mathbf{D}_u$ and $\mathbf{D}_v$ defined in (24) are equal to:

$$\mathbf{D}_u = \frac{1}{2}\begin{bmatrix} -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \tag{34}$$

$$\mathbf{D}_v = \frac{1}{2}\begin{bmatrix} -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \tag{35}$$

Let us now show how these matrices relate to the discretization of the continuous optimality condition (8). Using second-order central finite differences approximations of the Laplacian ($\Delta z_{u,v} \approx z_{u,v-1} + z_{u-1,v} + z_{u+1,v} + z_{u,v+1} - 4z_{u,v}$) and of the divergence operator ($\nabla \cdot \mathbf{g}_{u,v} \approx \frac{1}{2}(p_{u+1,v} - p_{u-1,v}) + \frac{1}{2}(q_{u,v+1} - q_{u,v-1})$), we obtain:

$$\left[4z_{u,v} - z_{u,v-1} - z_{u-1,v} - z_{u+1,v} - z_{u,v+1}\right] + \lambda_{u,v}z_{u,v}$$
$$= \frac{1}{2}\left[p_{u-1,v} - p_{u+1,v}\right] + \frac{1}{2}\left[q_{u,v-1} - q_{u,v+1}\right] + \lambda_{u,v}z_{u,v}^0 \tag{36}$$

The pixel $(u, v) = (2, 2)$ is the only one whose four neighbors are inside Ω. In that case, (36) becomes:

$$\underbrace{\left[4z_{2,2} - z_{2,1} - z_{1,2} - z_{3,2} - z_{2,3}\right]}_{=(\mathbf{L})_{5,\cdot}\mathbf{z}} + \underbrace{\lambda_{2,2}z_{2,2}}_{=(\boldsymbol{\Lambda}^2)_{5,\cdot}\mathbf{z}}$$
$$= \underbrace{\frac{1}{2}\left[p_{1,2} - p_{3,2}\right]}_{=(\mathbf{D}_u)_{5,\cdot}\mathbf{p}} + \underbrace{\frac{1}{2}\left[q_{2,1} - q_{2,3}\right]}_{=(\mathbf{D}_v)_{5,\cdot}\mathbf{q}} + \underbrace{\lambda_{2,2}z_{2,2}^0}_{=(\boldsymbol{\Lambda}^2)_{5,\cdot}\mathbf{z}^0} \tag{37}$$

where we recognize the fifth equation of the discrete optimality condition (26). This shows that, for pixels having all four neighbors inside Ω, both the continuous and the discrete variational formulations yield the same discretizations.

Now, let us consider a pixel near the boundary, for instance pixel $(1, 1)$. Using the same second-order differences, (36) reads:

$$\left[4z_{1,1} - z_{1,0} - z_{0,1} - z_{2,1} - z_{1,2}\right] + \lambda_{1,1}z_{1,1}$$
$$= \frac{1}{2}\left[p_{0,1} - p_{2,1}\right] + \frac{1}{2}\left[q_{1,0} - q_{1,2}\right] + \lambda_{1,1}z_{1,1}^0 \tag{38}$$

which involves the values $z_{1,0}$ and $z_{0,1}$ of the depth map, which we are not willing to estimate, and the values $p_{0,1}$ and $q_{1,0}$ of the gradient field, which are not provided as data. To eliminate these four values, we need to resort to boundary conditions on z, p and q. The discretizations, using first-order forward finite differences, of the natural boundary condition (9), at locations $(1, 0)$ and $(0, 1)$, read:

$$z_{1,1} - z_{1,0} = q_{1,0} \tag{39}$$
$$z_{1,1} - z_{0,1} = p_{0,1}; \tag{40}$$

hence, the unknown depth values $z_{1,0}$ and $z_{0,1}$ can be eliminated from Eq. (38):

$$\left[2z_{1,1} - z_{2,1} - z_{1,2}\right] + \lambda_{1,1}z_{1,1}$$
$$= \frac{1}{2}\left[-p_{0,1} - p_{2,1}\right] + \frac{1}{2}\left[-q_{1,0} - q_{1,2}\right] + \lambda_{1,1}z_{1,1}^0 \tag{41}$$

Eventually, the unknown values $p_{0,1}$ and $q_{1,0}$ need to be approximated. Since we have no information at all about the values of $\mathbf{g}$ outside Ω, we use homogeneous Neumann boundary conditions[13]:

---

[13] This assumption is weaker than the homogeneous Neumann boundary condition $\nabla z \cdot \boldsymbol{\eta} = 0$ used by Agrawal et al. [1].

$$\nabla p \cdot \boldsymbol{\eta} = 0 \quad \text{over } \partial\Omega \tag{42}$$
$$\nabla q \cdot \boldsymbol{\eta} = 0 \quad \text{over } \partial\Omega \tag{43}$$

Discretizing these boundary conditions using first-order forward finite differences, we obtain:

$$p_{0,1} = p_{1,1} \tag{44}$$
$$q_{1,0} = q_{1,1} \tag{45}$$

Using these identifications, the discretized optimality condition (41) is given by:

$$\underbrace{\left[2z_{1,1} - z_{2,1} - z_{1,2}\right]}_{=(\mathbf{L})_{1,\cdot}\mathbf{z}} + \underbrace{\lambda_{1,1}z_{1,1}}_{=(\boldsymbol{\Lambda}^2)_{1,\cdot}\mathbf{z}}$$
$$= \underbrace{\frac{1}{2}\left[-p_{1,1} - p_{2,1}\right]}_{=(\mathbf{D}_u)_{1,\cdot}\mathbf{p}} + \underbrace{\frac{1}{2}\left[-q_{1,1} - q_{1,2}\right]}_{=(\mathbf{D}_v)_{1,\cdot}\mathbf{q}} + \underbrace{\lambda_{1,1}z_{1,1}^0}_{=(\boldsymbol{\Lambda}^2)_{1,\cdot}\mathbf{z}^0} \tag{46}$$

which is exactly the first equation of the discrete optimality condition (26).

Using a similar rationale, we obtain equivalence of both formulations for the eight points inside Ω. Yet, let us emphasize that discretizing the continuous optimality condition requires treating, on this example with a rather "simple" shape for Ω, not less than seven different cases (only pixels $(3, 2)$ and $(2, 3)$ are similar). More general shapes bring out to play even more particular cases (points having only one neighbor inside Ω). Furthermore, boundary conditions must be invoked in order to approximate the depth values and the data outside Ω. On the other hand, the discrete functional provides exactly the same optimality condition, but without these drawbacks. The boundary conditions can be viewed as *implicitly* enforced; hence, $\mathcal{P}_{\text{FreeB}}$ is satisfied.

3.5 Empirical Evaluation

We first consider the smooth surface from Fig. 2, whose normals are analytically known [26], and compare three discrete least-squares methods which all satisfy $\mathcal{P}_{\text{Fast}}$, $\mathcal{P}_{\text{Robust}}$ and $\mathcal{P}_{\text{FreeB}}$: the DCT solution [53], the Sylvester equations method [26], and the proposed one. As shown in Figs. 2 and 3, our solution is slightly more accurate. Indeed, the bias near the boundary induced by the DCT method is corrected. On the other hand, we believe the reason why our method is more accurate than that from [26] is because we use a combination of forward and backward finite differences, while [26] relies on central differences. Indeed, when using central differences to discretize the gradient, the second-order operator (Laplacian) appearing in the Sylvester equations from [26] involves none of the direct neighbors, which may be nonrobust for noisy data (see, for instance, Appendix 3 in [4]).
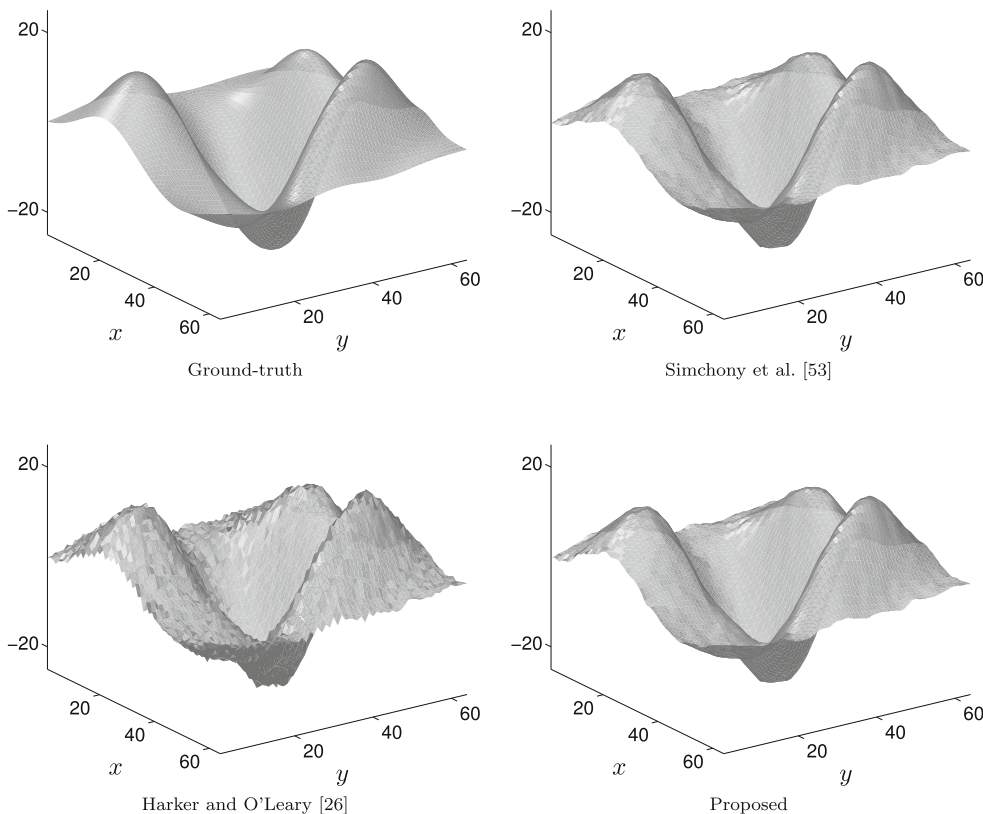
Fig. 2 Qualitative evaluation of the $\mathcal{P}_{\text{Robust}}$ property. An additive, zero-mean, Gaussian noise with standard deviation $0.1\|\mathbf{g}\|_\infty$ was added to the (analytically known) gradient of the ground-truth surface, before integrating this gradient by three least-squares methods. Ours qualitatively provides better results than the Sylvester equations method from Harker and O'Leary [26]. It seems to provide similar robustness as the DCT solution from Simchony et al. [53], but the quantitative evaluation from Fig. 3 shows that our method is actually more accurate

For instance, let us consider a 1D domain Ω with 7 pixels. Then, the following differentiation matrix is advocated in [26]:

$$\mathbf{D}_u = \frac{1}{2}\begin{bmatrix} -3 & 4 & -1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & -4 & 3 \end{bmatrix} \quad (47)$$

The optimality condition (Sylvester equation) in [26] involves the following second-order operator $\mathbf{D}_u{}^\top\mathbf{D}_u$:

$$\mathbf{D}_u{}^\top\mathbf{D}_u = \frac{1}{4}\begin{bmatrix} 10 & -12 & 2 & 0 & 0 & 0 & 0 \\ -12 & 17 & -4 & -1 & 0 & 0 & 0 \\ 2 & -4 & 3 & 0 & -1 & 0 & 0 \\ \mathbf{0} & \mathbf{-1} & \mathbf{0} & \mathbf{2} & \mathbf{0} & \mathbf{-1} & \mathbf{0} \\ 0 & 0 & -1 & 0 & 3 & -4 & 2 \\ 0 & 0 & 0 & -1 & -4 & 17 & -12 \\ 0 & 0 & 0 & 0 & 2 & -12 & 10 \end{bmatrix}$$
$$(48)$$

The bolded values of this matrix indicate that computation of the second-order derivatives for the fourth pixel does not involve the third and fifth pixels. On the other hand, with the proposed operator defined in Eq. (24), the second-order operator always involves the "correct" neighborhood:

$$\mathbf{D}_u{}^\top\mathbf{D}_u = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{-1} & \mathbf{2} & \mathbf{-1} & \mathbf{0} & \mathbf{0} \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}$$
$$(49)$$

In addition, as predicted by the complexity analysis in Sect. 3.2, our solution relying on preconditioned conjugate gradient iterations has an asymptotic complexity $(O(5n \log(n) \log(1/\epsilon)))$ which is inbetween that of the
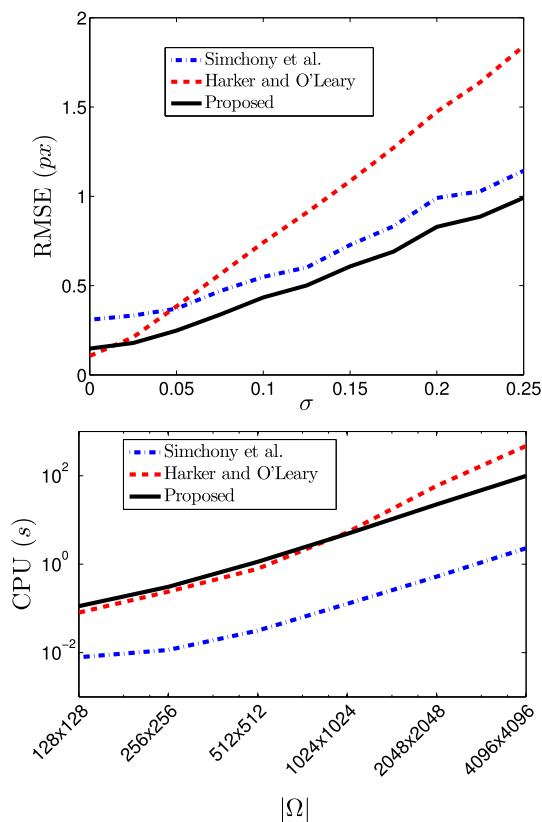
Fig. 3 Quantitative evaluation of the $\mathcal{P}_{\text{Robust}}$ (top) and $\mathcal{P}_{\text{Fast}}$ (bottom) properties. Top: RMSE between the depth ground truth and the ones reconstructed from noisy gradients (adding a zero-mean Gaussian noise with standard deviation $\sigma \|\mathbf{g}\|_{\infty}$, for several values of σ). Bottom: computation time as a function of the size $|\Omega|$ of the reconstruction domain Ω. The method we put forward has a complexity which is inbetween those of the methods of Simchony et al. [53] (based on DCT) and of Harker and O'Leary [26] (based on Sylvester equations), while being slightly more accurate than both of them

Sylvester equations approach [26] ($O(n^{1.5})$) and of DCT [53] ($O(n \log(n))$). The CPU times of our method and of the DCT solution, measured using MATLAB codes running on a recent i7 processor, actually seem proportional: according to this complexity analysis, we guess the proportionality factor is around $5 \log(1/\epsilon)$. Indeed, with $\epsilon = 10^{-4}$, which is the value we used in our experiments, $5 \log(1/\epsilon) \approx 46$, which is consistent with the second graph in Fig. 3.

Besides its improved accuracy, the major advantage of our method over [26,53] is its ability to handle non-rectangular domains ($\mathcal{P}_{\text{NoRect}}$). This makes possible the 3D-reconstruction of piecewise-smooth surfaces, provided that a user segments the domain into pieces where z is smooth beforehand (see Fig. 4). Yet, if the segmentation is not performed *a priori*, artifacts are visible near the discontinuities, which get smoothed, and Gibbs phenomena appear near the continuous, yet non-differentiable kinks. We will discuss in the next section several strategies for removing such artifacts.
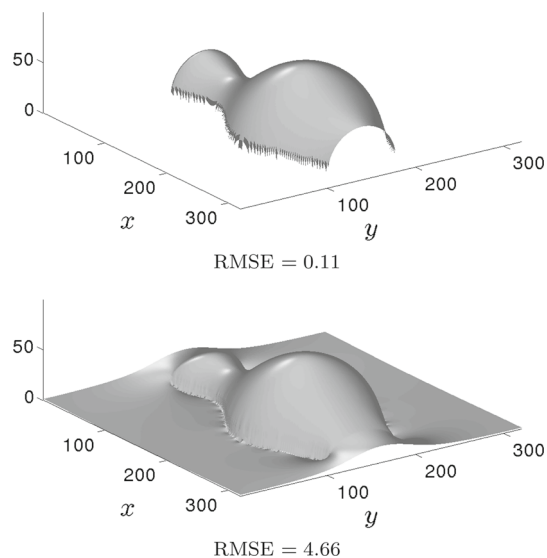
Fig. 4 3D-reconstruction of surface $\mathcal{S}_{\text{vase}}$ (see Figure 3 in [48]) from its (analytically known) normals, using the proposed discrete least-squares method. Top: when Ω is restricted to the image of the vase. Bottom: when Ω is the whole rectangular grid. Quadratic integration smooths the depth discontinuities and produces Gibbs phenomena near the kinks

4 Piecewise-Smooth Surfaces

We now tackle the problem of recovering a surface which is smooth only *almost* everywhere, i.e., everywhere except on a "small" set where discontinuities and kinks are allowed. Since all the methods discussed hereafter rely on the same discretization as in Sect. 3, they inherit its $\mathcal{P}_{\text{FreeB}}$ and $\mathcal{P}_{\text{NoRect}}$ properties, which will not be discussed in this section. Instead, we focus on the $\mathcal{P}_{\text{Fast}}$, $\mathcal{P}_{\text{Robust}}$, $\mathcal{P}_{\text{NoPar}}$, and of course $\mathcal{P}_{\text{Disc}}$ properties.

4.1 Recovering Discontinuities and Kinks

In order to clarify which variational formulations may provide robustness to discontinuities, let us first consider the 1D-example of Fig. 5, with Dirichlet boundary conditions. As illustrated in this example, least-squares integration of a noisy normal field will provide a smooth surface. Replacing the least-squares estimator $\Phi_{L_2}(s) = s^2$ by the sparsity one $\Phi_{L_0}(s) = 1 - \delta(s)$ will minimize the cardinality of the difference between $\mathbf{g}$ and ∇z, which provides a surface whose gradient is almost everywhere equal to $\mathbf{g}$. As a consequence, robustness to noise is lost, yet discontinuities may be preserved.

These estimators can be interpreted as follows: least squares assume that *all* residuals defined by $\|\nabla z(u, v) - \mathbf{g}(u, v)\|$ are "low," while sparsity assumes that *most of* them are "zero." The former is commonly used for "noise," and the latter for "outliers." In the case of normal integration, outliers may occur when: (1) $\nabla z(u, v)$ exists but its estimate
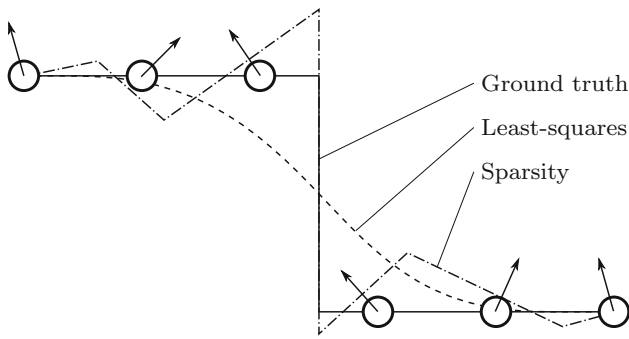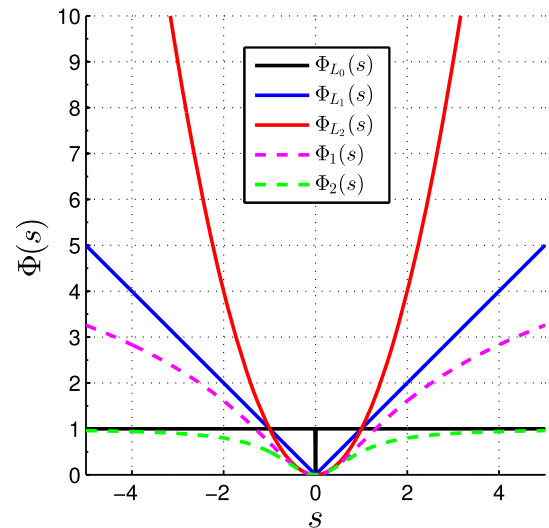
Fig. 5 1D-illustration of integration of a noisy normal field (arrows) over a regular grid (circles), in the presence of discontinuities. The least-squares approach is robust to noise, but smooths the discontinuities. The sparsity approach preserves the discontinuities, but is not robust to noise. An ideal integration method would inherit robustness from least squares, and the ability to preserve discontinuities from sparsity

Fig. 6 Graph of some robust estimators. The ability of Φ_{L_2} to handle noise (small residuals) comes from its over-linear behavior around zero, while that of Φ_{L_0} to preserve discontinuities (large residuals) is induced by its sublinear behavior in $+\infty$. An estimator holding both these properties is necessarily non-convex (e.g., Φ_1 and Φ_2, whose graphs are shown with $\beta = \gamma = 1$), although Φ_{L_1} may be an acceptable convex compromise

$\mathbf{g}(u, v)$ is not reliable; (2) $\nabla z(u, v)$ is not defined because (u, v) lies within the vicinity of a discontinuity or a kink. Considering that situation (1) should rather be handled by robust estimation of the gradient [31], we deal only with the second one, and use the terminology "discontinuity" instead of "outlier," although this also covers the concept of "kink."

We are looking for an estimator which combines the robustness of least squares to noise, and that of sparsity to discontinuities. These abilities are actually due to their asymptotic behaviors. Robustness of least squares to noise comes from the quadratic behavior around 0, which ensures that "low" residuals are considered as "good" estimates, while this quadratic behavior becomes problematic in $\pm\infty$: discontinuities yield "high" residuals, which are over-penalized. The sparsity estimator has the opposite behavior: treating the high residuals (discontinuities) exactly as the low ones ensures that discontinuities are not over-penalized, yet low residuals (noise) are. A good estimator would thus be quadratic around zero, but sublinear around $\pm\infty$. Obviously, only non-convex estimators hold both these properties. We will discuss several choices "inbetween" the quadratic estimator Φ_{L_2} and the sparsity one Φ_{L_0} (see Fig. 6): the convex compromise $\Phi_{L_1}(s) = |s|$ is studied in Sect. 4.2, and the non-convex estimators $\Phi_1(s) = \log(s^2 + \beta^2)$ and $\Phi_2(s) = \frac{s^2}{s^2 + \gamma^2}$, where β and γ are hyper-parameters, in Sect. 4.3.

Another strategy consists in keeping least squares as basis, but using it in a non-uniform manner. The simplest way would be to remove the discontinuity points from the integration domain Ω and then to apply our quadratic method from the previous section, since it is able to manage non-rectangular domains. Yet, this would require detecting the discontinuities beforehand, which might be tedious. It is actually more convenient to introduce weights in the least-squares functionals, which are inversely proportional to the probability of lying on a discontinuity [47,50]. We discuss this weighted least-squares approach in Sect. 4.4, where a statistical inter-

pretation of the Perona and Malik's anisotropic diffusion model [44] is also exhibited. Eventually, an extreme case of weighted least squares consists in using binary weights, where the weights indicate the presence of discontinuities. This is closely related to Mumford and Shah's segmentation method [37], which simultaneously estimates the discontinuity set and the surface. We show in Sect. 4.5 that this approach is the one which is actually the most adapted to the problem of integrating a noisy normal field in the presence of discontinuities.

4.2 Total Variation-Like Integration

The problem of handling outliers in a noisy normal field has been tackled by Du, Robles-Kelly, and Lu, who compare in [18] the performances of several M-estimators. They conclude that regularizers based on the L_1 norm are the most effective ones. We provide in this subsection several numerical considerations regarding the discretization of the L_1 fidelity term:

$$\mathcal{F}_{L_1}(z) = \iint_{(u,v)\in\Omega} \|\nabla z(u, v) - \mathbf{g}(u, v)\|_1 \mathrm{d}u\,\mathrm{d}v$$
$$= \iint_{(u,v)\in\Omega} \Big\{ |\partial_u z(u, v) - p(u, v)|$$
$$+ |\partial_v z(u, v) - q(u, v)| \Big\} \mathrm{d}u\,\mathrm{d}v \qquad (50)$$

When $p(u, v) \equiv 0$ and $q(u, v) \equiv 0$, (50) is the so-called anisotropic total variation (anisotropic TV) regularizer, which tends to favor piecewise-constant solutions while allowing discontinuity jumps. Considering the discontinuities and kinks as the equivalent of edges in image restoration, it seems natural to believe that the fidelity term (50) may be useful for discontinuity-preserving integration.

This fidelity term is not only convex, but also decouples the two directions u and v, which allows fast ADMM-based (Bregman iterations) numerical schemes involving shrinkages [24,47]. On the other hand, it is not so natural to use such a decoupling: if the value of p is not reliable at some point (u, v), usually that of q is not reliable either. Hence, it may be worthwhile to use instead a regularizer adapted from the "isotropic TV." This leads us to adapt the well-known model from Rudin et al. [49] to the integration problem:

$$
\mathcal{E}_{\mathrm{TV}}(z) = \iint_{(u,v)\in\Omega} \|\nabla z(u, v) - \mathbf{g}(u, v)\|
$$
$$
+ \lambda(u, v) \left[z(u, v) - z^0(u, v) \right]^2 \mathrm{d}u \, \mathrm{d}v \qquad (51)
$$

Discretization Since the term $\|\nabla z(u, v) - \mathbf{g}(u, v)\|$ can be interpreted in different manners, depending on the neighborhood of (u, v), we need to discretize it appropriately. Let us consider all four possible first-order discretizations of the gradient ∇z, associated with the four following sets of pixels:

$$
\Omega^{UV} = \Omega_u^U \cap \Omega_v^V, \ (U, V) \in \{+, -\}^2 \qquad (52)
$$

The discrete functional to minimize is thus given by:

$$
E_{\mathrm{TV}}(\mathbf{z})
$$
$$
= \frac{1}{4} \Bigg(\sum_{(u,v)\in\Omega^{++}} \sqrt{\left[\partial_u^+ z_{u,v} - p_{u,v}\right]^2 + \left[\partial_v^+ z_{u,v} - q_{u,v}\right]^2}
$$
$$
+ \sum_{(u,v)\in\Omega^{+-}} \sqrt{\left[\partial_u^+ z_{u,v} - p_{u,v}\right]^2 + \left[\partial_v^- z_{u,v} - q_{u,v}\right]^2}
$$
$$
+ \sum_{(u,v)\in\Omega^{-+}} \sqrt{\left[\partial_u^- z_{u,v} - p_{u,v}\right]^2 + \left[\partial_v^+ z_{u,v} - q_{u,v}\right]^2}
$$
$$
+ \sum_{(u,v)\in\Omega^{--}} \sqrt{\left[\partial_u^- z_{u,v} - p_{u,v}\right]^2 + \left[\partial_v^- z_{u,v} - q_{u,v}\right]^2} \Bigg)
$$
$$
+ \sum_{(u,v)\in\Omega} \lambda_{u,v} \left[z_{u,v} - z_{u,v}^0 \right]^2 \qquad (53)
$$

Minimizing (53) comes down to solving the following constrained optimization problem:

$$
\min_{\mathbf{z}, \{\mathbf{r}^{UV}\}} \frac{1}{4} \sum_{(U,V)\in\{+,-\}^2} \sum_{(u,v)\in\Omega^{UV}} \|\mathbf{r}_{u,v}^{UV}\|
$$
$$
+ \sum_{(u,v)\in\Omega} \lambda_{u,v} \left[z_{u,v} - z_{u,v}^0 \right]^2
$$
$$
\text{s.t.} \quad \mathbf{r}_{u,v}^{UV} = \nabla^{UV} z_{u,v} - \mathbf{g}_{u,v} \qquad (54)
$$

where we denote $\nabla^{UV} = [\partial_u^U, \partial_v^V]^\top$, $(U, V) \in \{+, -\}^2$, the discrete approximation of the gradient corresponding to domain Ω^{UV}.

Numerical Solution We solve the constrained optimization problem (54) by the augmented Lagrangian method, through an ADMM algorithm [21] (see [9] for a recent overview of such algorithms). This algorithm reads:

$$
\mathbf{z}^{(k+1)} = \operatorname*{argmin}_{\mathbf{z}\in\mathbb{R}^{|\Omega|}} \frac{\alpha}{8} \sum_{(U,V)\in\{+,-\}^2} \sum_{(u,v)\in\Omega^{UV}} \Big\| \nabla^{UV} z_{u,v}
$$
$$
- \left(\mathbf{g}_{u,v} + \mathbf{r}_{u,v}^{UV(k)} - \mathbf{b}_{u,v}^{UV(k)} \right) \Big\|^2
$$
$$
+ \sum_{(u,v)\in\Omega} \lambda_{u,v} \left[z_{u,v} - z_{u,v}^0 \right]^2 \qquad (55)
$$

$$
\mathbf{r}_{u,v}^{UV(k+1)} = \operatorname*{argmin}_{\mathbf{r}\in\mathbb{R}^2} \frac{\alpha}{8} \left\| \mathbf{r} - \left(\nabla^{UV} z_{u,v}^{(k+1)} - \mathbf{g}_{u,v} + \mathbf{b}_{u,v}^{UV(k)} \right) \right\|^2
$$
$$
+ \|\mathbf{r}\| \qquad (56)
$$
$$
\mathbf{b}_{u,v}^{UV(k+1)} = \mathbf{b}_{u,v}^{UV(k)} + \nabla^{UV} z_{u,v}^{(k+1)} - \mathbf{g}_{u,v} - \mathbf{r}_{u,v}^{UV(k+1)} \qquad (57)
$$

where the $\mathbf{b}^{UV}$ are the scaled dual variables, and $\alpha > 0$ corresponds to a descent stepsize, which is supposed to be fixed beforehand. Note that the choice of this parameter influences only the convergence rate, not the actual minimizer. In our experiments, we used $\alpha = 1$.

The z-update (55) is a linear least-squares problem similar to the one which was tackled in Sect. 3. Its solution $\mathbf{z}^{(k+1)}$ is the solution of the following SDD linear system:

$$
\mathbf{A}_{\mathrm{TV}} \mathbf{z}^{(k+1)} = \mathbf{b}_{\mathrm{TV}}^{(k)} \qquad (58)
$$

with:

$$
\mathbf{A}_{\mathrm{TV}} = \frac{\alpha}{8} \sum_{(U,V)\in\{+,-\}^2} \left[\mathbf{D}_u^{U\top} \mathbf{D}_u^U + \mathbf{D}_v^{V\top} \mathbf{D}_v^V \right] + \mathbf{\Lambda}^2 \qquad (59)
$$

$$
\mathbf{b}_{\mathrm{TV}}^{(k)} = \frac{\alpha}{8} \sum_{(U,V)\in\{+,-\}^2} \left[\mathbf{D}_u^{U\top} \mathbf{p}^{UV(k)} + \mathbf{D}_v^{V\top} \mathbf{q}^{UV(k)} \right] + \mathbf{\Lambda}^2 \mathbf{z}^0
$$
$$
\qquad (60)
$$

where the $\mathbf{D}_{u/v}^{U/V}$ matrices are defined as in (17), the $\mathbf{\Lambda}$ matrix as in (20), and where we denote $\mathbf{p}^{UV(k)}$ and $\mathbf{q}^{UV(k)}$ the components of $\mathbf{g} + \mathbf{r}^{UV(k)} - \mathbf{b}^{UV(k)}$.

The solution of System (58) can be approximated by conjugate gradient iterations, choosing at each iteration the previous estimate $\mathbf{z}^{(k)}$ as initial guess (setting $\mathbf{z}^{(0)}$, for instance, as the least-squares solution from Sect. 3). In addition, the matrix $\mathbf{A}_{TV}$ is always the same: this allows computing the preconditioner only once.

Eventually, the **r**-updates (56), $(u, v) \in \Omega$, are basis pursuit problems [17], which admit the following closed-form solution (generalized shrinkage):
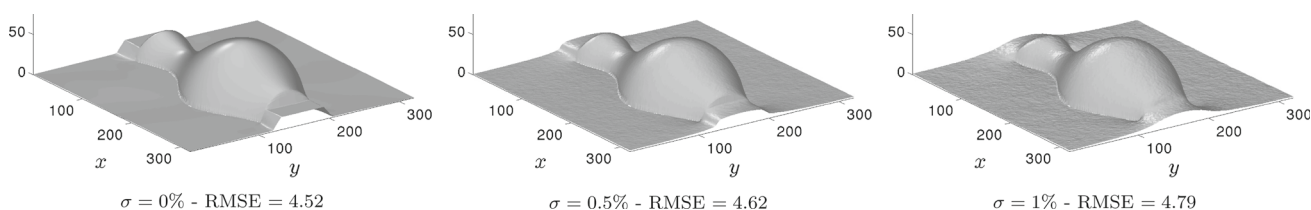
$$\mathbf{r}_{u,v}^{UV(k+1)} = \max \left\{ \|\mathbf{s}_{u,v}^{UV(k+1)}\| - \frac{4}{\alpha}, 0 \right\} \frac{\mathbf{s}_{u,v}^{UV(k+1)}}{\|\mathbf{s}_{u,v}^{UV(k+1)}\|} \qquad (61)$$

with:

$$\mathbf{s}_{u,v}^{UV(k+1)} = \nabla^{UV} z_{u,v}^{(k+1)} - \mathbf{g}_{u,v} + \mathbf{b}_{u,v}^{UV(k)} \qquad (62)$$

Discussion This TV-like approach has two main advantages: apart from the stepsize α which controls the speed of convergence, it does not depend on the choice of a parameter, and it is convex. The initialization has influence only on the speed of convergence and not on the actual minimizer: convergence toward the global minimum is guaranteed [51]. It can be shown that the convergence rate of this scheme is ergodic, and this rate can be improved rather simply [23]. We cannot consider that $\mathcal{P}_{Fast}$ is satisfied since, in comparison with the quadratic method from Sect. 3, yet the TV approach is "reasonably" fast. Possibly faster algorithms could be employed, as for instance the FISTA algorithm from Beck and Teboulle [7], or primal–dual algorithms [13], but we leave such improvements as future work.

On the other hand, according to the results from Fig. 7, discontinuities are recovered in the absence of noise, although staircasing artifacts appear (such artifacts are partly due to the non-differentiability of TV in zero [38]). Yet, the recovery of discontinuities is deceiving when the noise level increases. On noisy datasets, the only advantage of this approach over least squares is thus that it removes the Gibbs phenomena around the kinks, i.e., where the surface is continuous, but non-differentiable (e.g., the sides of the vase).

Because of the staircasing artifacts and of the lack of robustness to noise, we cannot find this first approach satisfactory. Yet, since turning the quadratic functional into a non-quadratic one seems to have positive influence on discontinuities recovery, we believe that exploring non-quadratic models is a promising route. Staircasing artifacts could probably be reduced by replacing total variation by total generalized variation [10], but we rather consider now non-convex models.

4.3 Non-convex Regularization

Let us now consider non-convex estimators Φ in the fidelity term (5), which are often referred to as "Φ-functions" [4]. As discussed in Sect. 4.1, the choice of a specific Φ-function should be made according to several principles:

- Φ should have a quadratic behavior around zero, in order to ensure that the integration is guided by the "good" data. The typical choice ensuring this property is $\Phi_{L_2}(s) = s^2$, which is discussed in Sect. 3;
- Φ should have a sublinear behavior at infinity, so that outliers do not have a predominant influence, and also to preserve discontinuities and kinks. The typical choice is the sparsity estimator $\Phi_{L_0}(s) = 0$ if $s = 0$ and $\Phi_{L_0}(s) = 1$ otherwise;
- Φ should ideally be a convex function.

Obviously, it is not possible to simultaneously satisfy these three properties. The TV-like fidelity term introduced in Sect. 4.2 is a sort of "compromise": it is the only convex function being (over-) linear in 0 and (sub-) linear in $\pm\infty$. Although it does not depend on the choice of any hyperparameter, we saw that it has the drawback of yielding the so-called staircase effect and that discontinuities were not recovered so well in the presence of noise. If we accept to lose the convexity of Φ, we can actually design estimators which better fit both other properties. Although there may then be several minimizers, such non-convex estimators were recently shown to be very effective for image restoration [36].

$\sigma = 0\%$ - RMSE = 4.52 $\qquad$ $\sigma = 0.5\%$ - RMSE = 4.62 $\qquad$ $\sigma = 1\%$ - RMSE = 4.79

Fig. 7 Depth estimated after 1000 iterations of the TV-like approach, in the presence of additive, zero-mean, Gaussian noise with standard deviation equal to $\sigma \|\mathbf{g}\|_\infty$. The indicated RMSE is computed on the whole domain. In the absence of noise, both discontinuities and kinks are restored, although staircasing artifacts appear. In the presence of noise, the discontinuities are smoothed. Yet, the 3D-reconstruction near the kinks is still more satisfactory than the least-squares one: Gibbs phenomena are not visible, unlike in the second row of Fig. 4

We will consider two classical Φ-functions, whose graphs are plotted in Fig. 6:

$$\begin{cases} \Phi_1(s) = \log(s^2 + \beta^2) \\ \Phi_2(s) = \dfrac{s^2}{s^2 + \gamma^2} \end{cases} \Rightarrow \begin{cases} \Phi_1'(s) = \dfrac{2s}{s^2 + \beta^2} \\ \Phi_2'(s) = \dfrac{2\gamma^2 s}{(s^2 + \gamma^2)^2} \end{cases} \quad (63)$$

Let us remark that these estimators were initially introduced in [19] in this context and that other non-convex estimators can be considered, based for instance on L^p norms, with $0 < p < 1$ [5].

Let us now show how to numerically minimize the resulting functionals:

$$\mathcal{E}_\Phi(z) = \iint_{(u,v) \in \Omega} \Phi\left(\|\nabla z(u,v) - \mathbf{g}(u,v)\|\right)$$
$$+ \lambda(u,v)\left[z(u,v) - z^0(u,v)\right]^2 \mathrm{d}u \, \mathrm{d}v \quad (64)$$

Discretization We consider the same discretization strategy as in Sect. 4.2, aiming at minimizing the discrete functional:

$$E_\Phi(\mathbf{z}) = \frac{1}{4} \sum\sum_{(U,V) \in \{+,-\}^2} \sum\sum_{(u,v) \in \Omega^{UV}} \Phi\left(\|\nabla^{UV} z_{u,v} - \mathbf{g}_{u,v}\|\right)$$
$$+ \sum\sum_{(u,v) \in \Omega} \lambda_{u,v}\left[z_{u,v} - z_{u,v}^0\right]^2 \quad (65)$$

which resembles the TV functional defined in (53), and where ∇^{UV} represents the finite differences approximation of the gradient used over the domain Ω^{UV}, with $\{U,V\} \in \{+,-\}^2$.

Introducing the notations:

$$f(\mathbf{z}) = \frac{1}{4} \sum\sum_{(U,V) \in \{+,-\}^2} \sum\sum_{(u,v) \in \Omega^{UV}} \Phi\left(\|\nabla^{UV} z_{u,v} - \mathbf{g}_{u,v}\|\right) \quad (66)$$

$$g(\mathbf{z}) = \|\boldsymbol{\Lambda}(\mathbf{z} - \mathbf{z}^0)\|^2 \quad (67)$$

the discrete functional (65) is rewritten:

$$E_\Phi(\mathbf{z}) = f(\mathbf{z}) + g(\mathbf{z}) \quad (68)$$

where f is smooth, but *non-convex*, and g is *convex* (and smooth, although non-smooth functions g could be handled).

Numerical Solution The problem of minimizing a discrete energy like (68), yielded by the sum of a convex term g and a non-convex, yet smooth term f, can be handled by forward–backward splitting. We use the "iPiano" iterative algorithm by Ochs et al. [41], which reads:

$$\mathbf{z}^{(k+1)} = (\mathbf{I} + \alpha_1 \partial g)^{-1}\left(\mathbf{z}^{(k)} - \alpha_1 \nabla f(\mathbf{z}^{(k)})\right.$$
$$\left. + \alpha_2\left(\mathbf{z}^{(k)} - \mathbf{z}^{(k-1)}\right)\right) \quad (69)$$

where α_1 and α_2 are suitable descent stepsizes (in our implementation, α_2 is fixed to 0.8, and α_1 is chosen by the "lazy backtracking" procedure described in [41]), $(\mathbf{I} + \alpha_1 \partial g)^{-1}$ is the *proximal operator* of g, and $\nabla f(\mathbf{z}^{(k)})$ is the gradient of f evaluated at current estimate $\mathbf{z}^{(k)}$. We detail hereafter how to evaluate the proximal operator of g and the gradient of f.

The proximal operator of g writes, using (67):

$$(\mathbf{I} + \alpha_1 \partial g)^{-1}(\widehat{\mathbf{x}}) = \underset{\mathbf{x} \in \mathbb{R}^{|\Omega|}}{\operatorname{argmin}} \frac{\|\mathbf{x} - \widehat{\mathbf{x}}\|}{2} + \alpha_1 g(\mathbf{x}) \quad (70)$$

$$= \left(\mathbf{I} + 2\alpha_1 \boldsymbol{\Lambda}^2\right)^{-1}\left(\widehat{\mathbf{x}} + 2\alpha_1 \boldsymbol{\Lambda} \mathbf{z}^0\right) \quad (71)$$

where the inversion is easy to compute, since the matrix involved is diagonal.

In order to obtain a closed-form expression of the gradient of f defined in (66), let us rewrite this function in the following manner:

$$f(\mathbf{z}) = \frac{1}{4} \sum\sum_{(U,V) \in \{+,-\}^2} \sum\sum_{(u,v) \in \Omega^{UV}} \Phi\left(\|\mathbf{D}_{u,v}^{UV}\mathbf{z} - \mathbf{g}_{u,v}\|\right) \quad (72)$$

where $\mathbf{D}_{u,v}^{UV}$ is a $2 \times |\Omega|$ finite differences matrix used for approximating the gradient at location (u,v), using the finite differences operator ∇^{UV}, $\{U,V\} \in \{+,-\}^2$:

$$\mathbf{D}_{u,v}^{UV} = \begin{bmatrix} \left(\mathbf{D}_u^U\right)_{m^{-1}(u,v),\cdot} \\ \left(\mathbf{D}_v^V\right)_{m^{-1}(u,v),\cdot} \end{bmatrix} \quad (73)$$

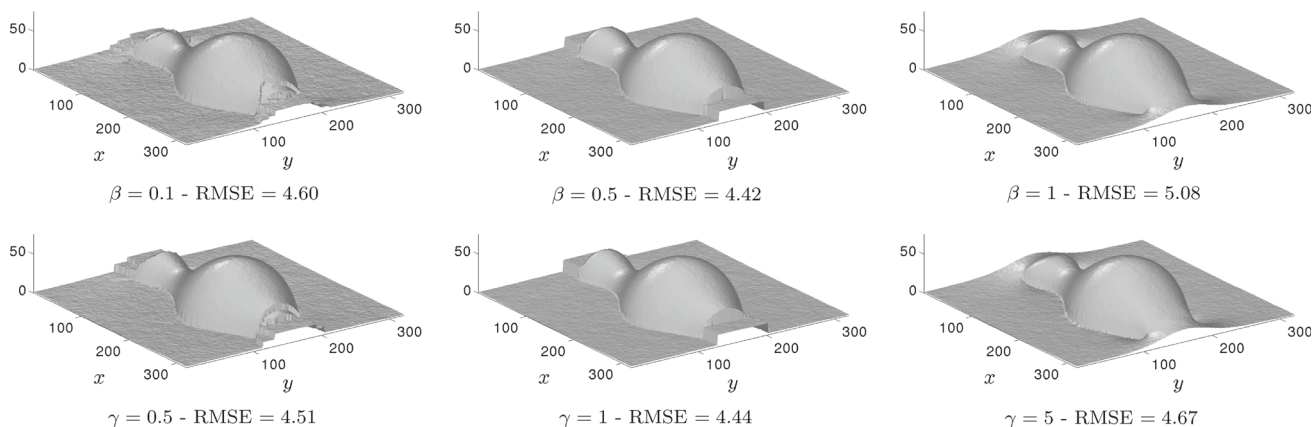where we recall that the mapping m associates linear indices with pixel coordinates [see Eq. (18)].

The gradient of f is thus given by:

$$\nabla f(\mathbf{z}) = \frac{1}{4} \sum\sum_{(U,V) \in \{+,-\}^2} \sum\sum_{(u,v) \in \Omega^{UV}} \left\{\mathbf{D}_{u,v}^{UV\top}\left(\mathbf{D}_{u,v}^{UV}\mathbf{z} - \mathbf{g}_{u,v}\right)\right.$$
$$\left. \times \frac{\Phi'\left(\|\mathbf{D}_{u,v}^{UV}\mathbf{z} - \mathbf{g}_{u,v}\|\right)}{\|\mathbf{D}_{u,v}^{UV}\mathbf{z} - \mathbf{g}_{u,v}\|}\right\} \quad (74)$$

Given the choices (63) for the Φ-functions, this can be further simplified:

$$\nabla f_1(\mathbf{z}) = \frac{1}{2} \sum\sum_{(U,V) \in \{+,-\}^2} \sum\sum_{(u,v) \in \Omega^{UV}} \frac{\mathbf{D}_{u,v}^{UV\top}\left(\mathbf{D}_{u,v}^{UV}\mathbf{z} - \mathbf{g}_{u,v}\right)}{\|\mathbf{D}_{u,v}^{UV}\mathbf{z} - \mathbf{g}_{u,v}\|^2 + \beta^2} \quad (75)$$

$$\nabla f_2(\mathbf{z}) = \frac{1}{2} \sum\sum_{(U,V) \in \{+,-\}^2} \sum\sum_{(u,v) \in \Omega^{UV}} \frac{\gamma^2 \mathbf{D}_{u,v}^{UV\top}\left(\mathbf{D}_{u,v}^{UV}\mathbf{z} - \mathbf{g}_{u,v}\right)}{\left(\|\mathbf{D}_{u,v}^{UV}\mathbf{z} - \mathbf{g}_{u,v}\|^2 + \gamma^2\right)^2} \quad (76)$$

$\beta = 0.1$ - RMSE = 4.60 $\beta = 0.5$ - RMSE = 4.42 $\beta = 1$ - RMSE = 5.08

$\gamma = 0.5$ - RMSE = 4.51 $\gamma = 1$ - RMSE = 4.44 $\gamma = 5$ - RMSE = 4.67

Fig. 8 Non-convex 3D-reconstructions of surface $\mathcal{S}_{\text{vase}}$, using Φ_1 (top) or Φ_2 (bottom). An additive, zero-mean, Gaussian noise with standard deviation $\sigma \|\mathbf{g}\|_\infty$, $\sigma = 1\%$, was added to the gradient field. The non-convex approaches depend on the tuning of a parameter (β or γ), but they are able to reconstruct the discontinuities in the presence of noise, unlike the TV approach. Staircasing artifacts indicate the presence of local minima (we used as initial guess $z^{(0)}$ the least-squares solution)

Discussion Contrarily to the TV-like approach (see Sect. 4.2), the non-convex estimators require setting one hyper-parameter (β or γ). As shown in Fig. 8, the choice of this parameter is crucial: when it is too high, discontinuities are smoothed, while setting a too low value leads to strong staircasing artifacts. Inbetween, the values $\beta = 0.5$ and $\gamma = 1$ seem to preserve discontinuities, even in the presence of noise (which was not the case using the TV-like approach).

Yet, staircasing artifacts are still present. Despite their non-convexity, the new estimators Φ_1 and Φ_2 are differentiable; hence, these artifacts do not come from a lack of differentiability, as this was the case for TV. They rather indicate the presence of local minima. This is illustrated in Fig. 9, where the 3D-reconstruction of a "Canadian tent"-like surface, with additive, zero-mean, Gaussian noise ($\sigma = 10\%$), is presented. When using the least-squares solution as initial guess $z^{(0)}$, the 3D-reconstruction is very close to the genuine surface. Yet, when using the trivial initialization $z^{(0)} \equiv 0$, we obtain a surface whose slopes are "almost everywhere" equal to the real ones, but unexpected discontinuity jumps appear. Since only the initialization differs in these experiments, this clearly shows that the artifacts indicate the presence of local minima.

Although local minima can sometimes be avoided by using the least-squares solution as initial guess (e.g., Fig. 9), this is not always the case (e.g., Fig. 8). Hence, the non-convex estimators perform overall better than the TV-like approach, but they are still not optimal. We now follow other routes, which use least squares as basis estimator, yet in a non-uniform manner, in order to allow discontinuities.

4.4 Integration by Anisotropic Diffusion

Both previous methods (total variation and non-convex estimators) replace the least-squares estimator by another one,

Ground-truth

$z^{(0)} = $ least-squares solution - RMSE = 0.78

$z^{(0)} \equiv 0$ - RMSE = 13.16
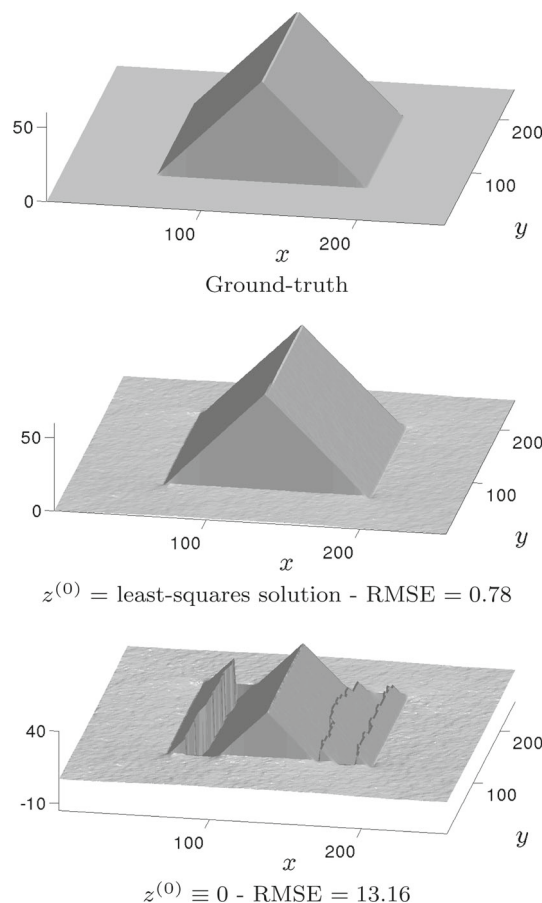
Fig. 9 3D-reconstruction of a "Canadian tent"-like surface from its noisy gradient ($\sigma = 1\%$), by the non-convex integrator Φ_1 ($\beta = 0.5$, 12,000 iterations), using two different initializations. The objective function being non-convex, the iterative scheme may converge toward a local minimum

assumed to be robust to discontinuities. Yet, it is possible to proceed differently: the 1D-graph in Fig. 5 shows that most of

data are corrupted only by noise, and that the discontinuity set is "small." Hence, applying least squares everywhere except on this set should provide an optimal 3D-reconstruction. To achieve this, a first possibility is to consider weighted least squares:

$$\min_z \iint_{(u,v)\in\Omega} \left\| \mathbf{W}(u,v) \left[\nabla z(u,v) - \mathbf{g}(u,v) \right] \right\|^2$$
$$+ \lambda(u,v) \left[z(u,v) - z^0(u,v) \right]^2 \mathrm{d}u\, \mathrm{d}v \qquad (77)$$

where $\mathbf{W}$ is a $\Omega \to \mathbb{R}^{2\times 2}$ tensor field, acting as a weight map designed to reduce the influence of discontinuity points. The weights can be computed beforehand according to the integrability of $\mathbf{g}$ [47], or by convolution of the components of $\mathbf{g}$ by a Gaussian kernel [1]. Yet, such approaches are of limited interest when $\mathbf{g}$ contains noise. In this case, the weights should rather be set as a function inversely proportional to $\|\nabla z(u,v)\|$, e.g.,

$$\mathbf{W}(u,v) = \frac{1}{\sqrt{\left(\frac{\|\nabla z(u,v)\|}{\mu}\right)^2 + 1}} \mathbf{I}_2 \qquad (78)$$

with μ a user-defined hyper-parameter. The latter tensor is the one proposed by Perona and Malik [44]: the continuous optimality condition associated with (77) is related to their "anisotropic diffusion model."[14] Such tensor fields $\mathbf{W}: \Omega \to \mathbb{R}^{2\times 2}$ are called "diffusion tensors": we refer the reader to [54] for a complete overview.

The use of diffusion tensors for the integration problem is not new [47], but we provide hereafter additional comments on the statistical interpretation of such tensors. Interestingly, the diffusion tensor (78) also appears when making different assumptions on the noise model than those we considered so far. Up to now, we assumed that the input gradient field $\mathbf{g}$ was equal to the gradient ∇z of the depth map z, up to an additive, zero-mean, Gaussian noise: $\mathbf{g} = \nabla z + \boldsymbol{\epsilon}$, $\boldsymbol{\epsilon} \sim \mathcal{N}\left([0,0]^\top, \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}\right)$. This hypothesis may not always be realistic. For instance, in 3D-reconstruction scenarii such as photometric stereo [55], one estimates the normal field $\mathbf{n}: \Omega \to \mathbb{R}^3$ pixelwise, rather than the gradient $\mathbf{g}: \Omega \to \mathbb{R}^2$, from a set of images. Hence, the Gaussian assumption should rather be made on these images. In this case, and provided that a maximum-likelihood for the normals is used, it may be assumed that the estimated normal field is the genuine one, up to an additive Gaussian noise. Yet, this does not imply that the noise in the gradient field $\mathbf{g}$ is Gaussian-distributed. Let us clarify this point.

---

[14] Although (78) actually yields an isotropic diffusion model, since it "utilizes a scalar-valued diffusivity and not a diffusion tensor" [54].

Assuming orthographic projection, the relationship between $\mathbf{n} = [n_1, n_2, n_3]^\top$ and ∇z is written, in every point (u,v) where the depth map z is differentiable:

$$\mathbf{n}(u,v) = \frac{1}{\sqrt{\|\nabla z(u,v)\|^2 + 1}} \left[-\nabla z(u,v)^\top, 1 \right]^\top \qquad (79)$$

which implies that $[-\frac{n_1}{n_3}, -\frac{n_2}{n_3}]^\top = [\partial_u z, \partial_v z]^\top = \nabla z$. If we denote $\bar{\mathbf{n}} = [\bar{n}_1, \bar{n}_2, \bar{n}_3]^\top$ the estimated normal field, it follows from (79) that $[-\frac{\bar{n}_1}{\bar{n}_3}, -\frac{\bar{n}_2}{\bar{n}_3}]^\top = [p, q]^\top = \mathbf{g}$.

Let us assume that $\bar{\mathbf{n}}$ and $\mathbf{n}$ differ according to an additive, zero-mean, Gaussian noise:

$$\bar{\mathbf{n}}(u,v) = \mathbf{n}(u,v) + \boldsymbol{\epsilon}(u,v) \qquad (80)$$

where :

$$\boldsymbol{\epsilon}(u,v) \sim \mathcal{N}\left([0,0,0]^\top, \begin{bmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix}\right) \qquad (81)$$

Since $\bar{n}_3$ is unlikely to take negative values (this would mean that the estimated surface is not oriented toward the camera), the following Geary–Hinkley transforms:

$$t_1 = \frac{n_3 \left(\frac{\bar{n}_1}{\bar{n}_3}\right) - n_1}{\sqrt{\sigma^2 \left(\left(\frac{\bar{n}_1}{\bar{n}_3}\right)^2 + 1\right)}} \qquad (82)$$

$$t_2 = \frac{n_3 \left(\frac{\bar{n}_2}{\bar{n}_3}\right) - n_2}{\sqrt{\sigma^2 \left(\left(\frac{\bar{n}_2}{\bar{n}_3}\right)^2 + 1\right)}} \qquad (83)$$

both follow standard Gaussian distribution $\mathcal{N}(0,1)$ [27]. After some algebra, this can be rewritten as:

$$\frac{1}{\sigma\sqrt{1+p^2}\sqrt{\|\nabla z\|^2 + 1}} [\partial_u z - p] \sim \mathcal{N}(0,1) \qquad (84)$$

$$\frac{1}{\sigma\sqrt{1+q^2}\sqrt{\|\nabla z\|^2 + 1}} [\partial_v z - q] \sim \mathcal{N}(0,1) \qquad (85)$$

This rationale suggests the use of the following fidelity term:

$$\mathcal{F}_{\mathrm{PM}}(z) = \iint_{(u,v)\in\Omega} \left\| \mathbf{W}(u,v) \left[\nabla z(u,v) - \mathbf{g}(u,v) \right] \right\|^2 \mathrm{d}u\, \mathrm{d}v \qquad (86)$$

where $\mathbf{W}(u,v)$ is the following 2×2 anisotropic diffusion tensor field:

$$\mathbf{W}(u, v) = \frac{1}{\sqrt{\|\nabla z(u, v)\|^2 + 1}} \begin{bmatrix} \frac{1}{\sqrt{1+p(u,v)^2}} & 0 \\ 0 & \frac{1}{\sqrt{1+q(u,v)^2}} \end{bmatrix} \tag{87}$$

Unfortunately, we experimentally found with the choice (87) for the diffusion tensor field, discontinuities were not always recovered. Instead, following the pioneering ideas from Perona and Malik [44], we introduce two parameters μ and ν to control the respective influences of the terms depending on the gradient of the unknown $\|\nabla z\|$ and on the input gradient (p, q). The new tensor field is then given by:

$$\mathbf{W}(u, v) = \frac{1}{\sqrt{\left(\frac{\|\nabla z(u,v)\|}{\mu}\right)^2 + 1}} \begin{bmatrix} \frac{1}{\sqrt{1+\left(\frac{p(u,v)}{\nu}\right)^2}} & 0 \\ 0 & \frac{1}{\sqrt{1+\left(\frac{q(u,v)}{\nu}\right)^2}} \end{bmatrix} \tag{88}$$

Replacing the matrix in (88) by $\mathbf{I}_2$ yields exactly the Perona–Malik diffusion tensor (78), which reduces the influence of the fidelity term on locations (u, v) where $\|\nabla z(u, v)\|$ increases, which are likely to indicate discontinuities. Yet, our diffusion tensor (88) also reduces the influence of points where p or q is high, which are also likely to correspond to discontinuities. In our experiments, we found that $\nu = 10$ could always be used, yet the choice of μ has more influence on the actual results.

Discretization Using the same discretization strategy as in Sects. 4.2 and 4.3 leads us to the following discrete functional:

$$E_{\mathrm{PM}}(\mathbf{z}) = \frac{1}{4} \sum_{(U,V)\in\{+,-\}^2} \sum \left\{ \left\| \mathbf{A}^{UV}(\mathbf{z}) \left(\mathbf{D}_u^U \mathbf{z} - \mathbf{p}\right) \right\|^2 \right.$$
$$\left. + \left\| \mathbf{B}^{UV}(\mathbf{z}) \left(\mathbf{D}_v^V \mathbf{z} - \mathbf{q}\right) \right\|^2 \right\} + \left\| \boldsymbol{\Lambda} \left(\mathbf{z} - \mathbf{z}^0\right) \right\|^2 \tag{89}$$

where the $\mathbf{A}^{UV}(\mathbf{z})$ and $\mathbf{B}^{UV}(\mathbf{z})$ are $|\Omega| \times |\Omega|$ diagonal matrices containing the following values:

$$a_{u,v}^{UV} = \frac{1}{\sqrt{1 + \left(\frac{p_{u,v}}{\nu}\right)^2} \sqrt{\frac{(\partial_u^U z_{u,v})^2 + (\partial_v^V z_{u,v})^2}{\mu^2} + 1}} \tag{90}$$

$$b_{u,v}^{UV} = \frac{1}{\sqrt{1 + \left(\frac{q_{u,v}}{\nu}\right)^2} \sqrt{\frac{(\partial_u^U z_{u,v})^2 + (\partial_v^V z_{u,v})^2}{\mu^2} + 1}} \tag{91}$$

with $(U, V) \in \{+, -\}^2$.

Numerical Solution Since the coefficients $a_{u,v}^{UV}$ and $b_{u,v}^{UV}$ depend on a nonlinear way on the unknown values $z_{u,v}$, it is difficult to derive a closed-form expression for the minimizer of (89). To deal with this issue, we use the following fixed point scheme, which iteratively updates the anisotropic diffusion tensors and the z-values:

$$\mathbf{z}^{(k+1)} = \underset{\mathbf{z}\in\mathbb{R}^{|\Omega|}}{\mathrm{argmin}} \frac{1}{4} \sum_{(U,V)\in\{+,-\}^2} \sum \left\{ \left\| \mathbf{A}^{UV}(\mathbf{z}^{(k)}) \left(\mathbf{D}_u^U \mathbf{z} - \mathbf{p}\right) \right\|^2 \right.$$
$$\left. + \left\| \mathbf{B}^{UV}(\mathbf{z}^{(k)}) \left(\mathbf{D}_v^V \mathbf{z} - \mathbf{q}\right) \right\|^2 \right\} + \left\| \boldsymbol{\Lambda} \left(\mathbf{z} - \mathbf{z}^0\right) \right\|^2 \tag{92}$$

Now that the diffusion tensor coefficients are fixed, each optimization problem (92) is reduced to a simple linear least-squares problem. In our implementation, we solve the corresponding optimality condition using Cholesky factorization, which we experimentally found to provide more stable results than conjugate gradient iterations.

Discussion We first experimentally verify that the proposed anisotropic diffusion approach is indeed a statistically meaningful approach in the context of photometric stereo. As stated in [39], "in previous work on photometric stereo, noise is [wrongly] added to the gradient of the height function rather than camera images." Hence, we consider the images from the "Cat" dataset presented in [52] and add a zero-mean, Gaussian noise with standard deviation $\sigma \|I\|_\infty$, $\sigma = 5\%$, to the images, where $\|I\|_\infty$ is the maximum gray-level value. The normals were computed by photometric stereo [55] over the part representing the cat. Then, since only the normals ground truth is provided in [52], and not the depth ground truth, we a posteriori computed the final normal maps by central finite differences. This allows us to calculate the angular error, in degrees, between the real surface and the reconstructed one. The mean angular error (MAE) can eventually be computed over the set of pixels for which central finite differences make sense (boundary and background points are excluded).

Figure 10 shows that the 3D-reconstruction obtained by anisotropic diffusion outperforms that obtained by least square: discontinuities are partially recovered, and robustness to noise is improved (see Fig. 11). However, although the diffusion tensor (87) does not require any parameter tuning, the restoration of discontinuities is not as sharp as with the non-convex integrators, and artifacts are visible along the discontinuities.

Although the parameter-free diffusion tensor (87) seems able to recover discontinuities, this is not always the case. For instance, we did not succeed in recovering the discontinuities of the surface S_{vase}. For this dataset, we had to use the tensor (88). The results from Fig. 12 show that with an
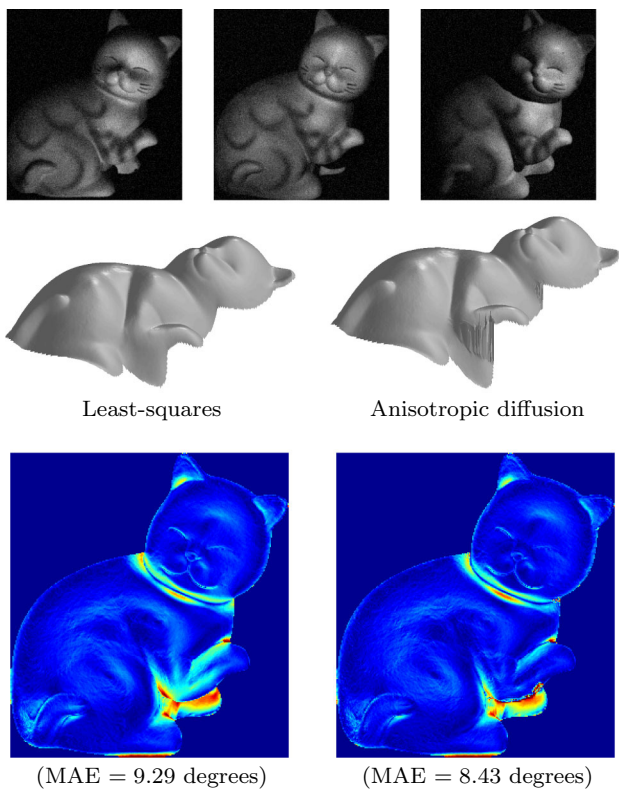
Fig. 10 Top row: three out of the 96 input images used for estimating the normals by photometric stereo [55]. Middle row, left: 3D-reconstruction by least-squares integration of the normals (see Sect. 3). Bottom row, left: angular error map (blue is 0°, red is 60°). The estimation is biased around the occluded areas. Middle and bottom rows, right: same, using anisotropic diffusion integration with the tensor field defined in (87). The errors remain confined in the occluded parts, and do not propagate over the discontinuities
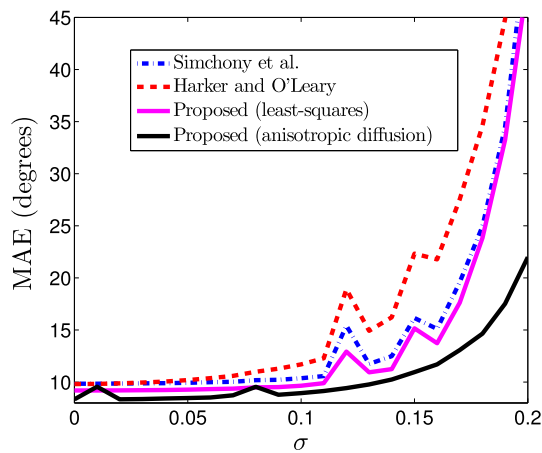
Fig. 11 Mean angular error (in degrees) as a function of the standard deviation $\sigma \|I\|_\infty$ of the noise which was added to the photometric stereo images. The anisotropic diffusion approach always outperforms least squares. For the methods [26,53], the gradient field was filled with zeros outside the reconstruction domain, which adds even more bias

appropriate tuning of μ, discontinuities are recovered and Gibbs phenomena are removed, without staircasing artifact.
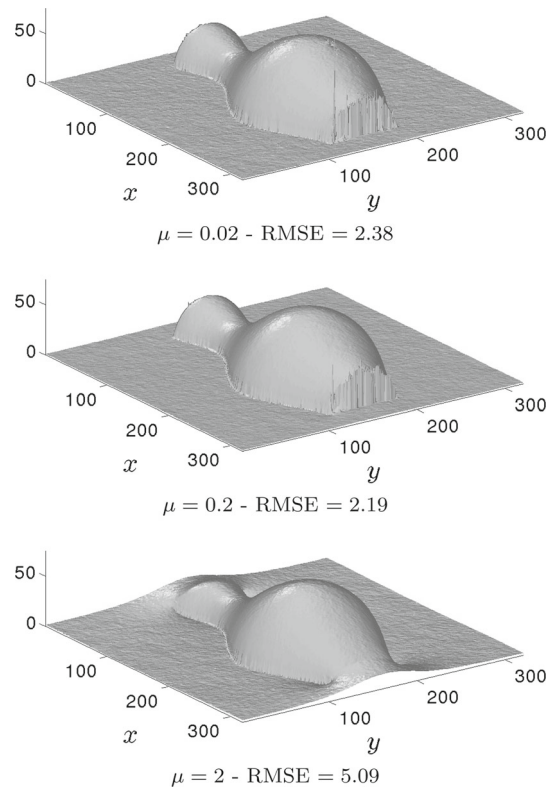
Fig. 12 Integration of the noisy gradient of $\mathcal{S}_{\text{vase}}$ ($\sigma = 1\%$) by anisotropic diffusion. As long as μ is small enough, discontinuities are recovered. Besides, no staircasing artifact is visible. Yet, the restored discontinuities are not perfectly sharp

Yet, as in the experiment of Fig. 10, the discontinuities are not very sharp. Such artifacts were also observed by Badri et al. [5], when experimenting with the anisotropic diffusion tensor from Agrawal et al. [1]. Sharper discontinuities could be recovered by using binary weights: this is the spirit of the Mumford–Shah segmentation method, which we explore in the next subsection.

4.5 Adaptation of the Mumford and Shah Functional

Let $z^0 \colon \Omega \to \mathbb{R}$ be a noisy image to restore. In order to estimate a denoised image z while preserving the discontinuities of the original image, Mumford and Shah suggested in [37] to minimize a quadratic functional only over a subset $\Omega \setminus K$ of Ω, while automatically estimating the discontinuity set K according to some prior. A reasonable prior is that the length of K is "small," which leads to the following optimization problem:

$$
\min_{z,K} \quad \mu \iint_{(u,v)\in\Omega\setminus K} \|\nabla z(u,v)\|^2 \, \mathrm{d}u \, \mathrm{d}v + \int_K \mathrm{d}\sigma
$$
$$
+ \lambda \iint_{(u,v)\in\Omega\setminus K} \left[z(u,v) - z^0(u,v)\right]^2 \, \mathrm{d}u \, \mathrm{d}v \tag{93}
$$

where λ and μ are positive constants, and $\int_K \mathrm{d}\sigma$ is the length of the set K. See [4] for a detailed introduction to this model and its qualitative properties.

Several approaches have been proposed to numerically minimize the Mumford–Shah functional: finite differences scheme [12], piecewise-constant approximation [14], primal–dual algorithms [46], etc. Another approach consists in using elliptic functionals. An auxiliary function $w : \Omega \to \mathbb{R}$ is introduced. This function stands for $1 - \chi_K$, where χ_K is the characteristic function of the set K. Ambrosio and Tortorelli have proposed in [2] to consider the following optimization problem:

$$
\min_{z,w} \mu \iint_{(u,v)\in\Omega} w(u,v)^2 \, \|\nabla z(u,v)\|^2 \, \mathrm{d}u \, \mathrm{d}v
$$
$$
+ \iint_{(u,v)\in\Omega} \left[\epsilon \, \|\nabla w(u,v)\|^2 + \frac{1}{4\epsilon} \, [w(u,v) - 1]^2 \right] \mathrm{d}u \, \mathrm{d}v
$$
$$
+ \lambda \iint_{(u,v)\in\Omega} \left[z(u,v) - z^0(u,v) \right]^2 \mathrm{d}u \, \mathrm{d}v \qquad (94)
$$

By using the theory of Γ-convergence, it is possible to show that (94) is a way to solve (93) when $\epsilon \to 0$.

We modify the above models, so that they fit our integration problem. Considering $\mathbf{g}$ as basis for least-squares integration everywhere except on the discontinuity set K, we obtain the following energy:

$$
\mathcal{E}_{\mathrm{MS}}(z, K)
$$
$$
= \mu \iint_{(u,v)\in\Omega\backslash K} \|\nabla z(u,v) - \mathbf{g}(u,v)\|^2 \, \mathrm{d}u \, \mathrm{d}v + \int_K \mathrm{d}\sigma
$$
$$
+ \iint_{(u,v)\in\Omega\backslash K} \lambda(u,v) \left[z(u,v) - z^0(u,v) \right]^2 \mathrm{d}u \, \mathrm{d}v \quad (95)
$$

for the Mumford–Shah functional, and the following Ambrosio–Tortorelli approximation:

$$
\mathcal{E}_{\mathrm{AT}}(z, w)
$$
$$
= \mu \iint_{(u,v)\in\Omega} w(u,v)^2 \, \|\nabla z(u,v) - \mathbf{g}(u,v)\|^2 \, \mathrm{d}u \, \mathrm{d}v
$$
$$
+ \iint_{(u,v)\in\Omega} \left[\epsilon \, \|\nabla w(u,v)\|^2 + \frac{1}{4\epsilon} \, [w(u,v) - 1]^2 \right] \mathrm{d}u \, \mathrm{d}v
$$
$$
+ \iint_{(u,v)\in\Omega} \lambda(u,v) \left[z(u,v) - z^0(u,v) \right]^2 \mathrm{d}u \, \mathrm{d}v \qquad (96)
$$

where $w : \Omega \to \mathbb{R}$ is a smooth approximation of $1 - \chi_K$.

Numerical Solution We use the same strategy as in Sect. 3 for discretizing $\nabla z(u, v)$ inside Functional (96), i.e., all the possible first-order discrete approximations of the differential operators are summed. Since discontinuities are usually "thin" structures, it is possible that a forward discretization contains the discontinuity while a backward discretization does not. Hence, the definition of the weights w should be made accordingly to that of ∇z. Thus, we define four fields $w_{u/v}^{+/-} : \Omega \to \mathbb{R}$, associated with the finite differences operators $\partial_{u/v}^{+/-}$. This leads to the following discrete analogue of Functional (96):

$$
E_{\mathrm{AT}}(\mathbf{z}, \mathbf{w}_u^+, \mathbf{w}_u^-, \mathbf{w}_v^+, \mathbf{w}_v^-)
$$
$$
= \frac{\mu}{2} \bigg(\left\| \mathbf{W}_u^+ \left(\mathbf{D}_u^+ \mathbf{z} - \mathbf{p} \right) \right\|^2 + \left\| \mathbf{W}_u^- \left(\mathbf{D}_u^- \mathbf{z} - \mathbf{p} \right) \right\|^2
$$
$$
+ \left\| \mathbf{W}_v^+ \left(\mathbf{D}_v^+ \mathbf{z} - \mathbf{q} \right) \right\|^2 + \left\| \mathbf{W}_v^- \left(\mathbf{D}_v^- \mathbf{z} - \mathbf{q} \right) \right\|^2 \bigg)
$$
$$
+ \frac{\epsilon}{2} \left(\left\| \mathbf{D}_u^+ \mathbf{w}_u^+ \right\|^2 + \left\| \mathbf{D}_u^- \mathbf{w}_u^- \right\|^2 + \left\| \mathbf{D}_v^+ \mathbf{w}_v^+ \right\|^2 \right.
$$
$$
+ + \left\| \mathbf{D}_v^- \mathbf{w}_v^- \right\|^2 \Big) + \frac{1}{8\epsilon} \left(\left\| \mathbf{w}_u^+ - \mathbf{1} \right\|^2 + \left\| \mathbf{w}_u^- - \mathbf{1} \right\|^2 \right.
$$
$$
+ \left\| \mathbf{w}_v^+ - v\mathbf{1} \right\|^2 + \left\| \mathbf{w}_v^- - \mathbf{1} \right\|^2 \Big) + \left\| A \left(\mathbf{z} - \mathbf{z}^0 \right) \right\|^2
$$
$$
\tag{97}
$$

where $\mathbf{w}_{u/v}^{+/-} \in \mathbb{R}^{|\Omega|}$ is a vector containing the values of the discretized field $w_{u/v}^{+/-}$, and $\mathbf{W}_{u/v}^{+/-} = \mathrm{Diag}(\mathbf{w}_{u/v}^{+/-})$ is the $|\Omega| \times |\Omega|$ diagonal matrix containing these values.

We tackle the nonlinear problem (97) by an alternating optimization scheme:

$$
\mathbf{z}^{(k+1)} = \operatorname*{argmin}_{\mathbf{z}\in\mathbb{R}^{|\Omega|}} E_{\mathrm{AT}} \left(\mathbf{z}, \mathbf{w}_u^{+(k)}, \mathbf{w}_u^{-(k)}, \mathbf{w}_v^{+(k)}, \mathbf{w}_v^{-(k)} \right)
$$
$$
\tag{98}
$$
$$
\mathbf{w}_u^{+(k+1)} = \operatorname*{argmin}_{\mathbf{w}\in\mathbb{R}^{|\Omega|}} E_{\mathrm{AT}} \left(\mathbf{z}^{(k+1)}, \mathbf{w}, \mathbf{w}_u^{-(k)}, \mathbf{w}_v^{+(k)}, \mathbf{w}_v^{-(k)} \right)
$$
$$
\tag{99}
$$

and similar straightforward updates for the other indicator functions. We can choose as initial guess, for instance, the smooth solution from Sect. 3 for $\mathbf{z}^{(0)}$, and $\mathbf{w}_u^{+(0)} = \mathbf{w}_u^{-(0)} = \mathbf{w}_v^{+(0)} = \mathbf{w}_v^{-(0)} \equiv \mathbf{1}$.

At each iteration (k), updating the surface and the indicator functions requires solving a series of linear least-squares problems. We achieve this by solving the resulting linear systems (normal equations) by means of the conjugate gradient algorithm. Contrarily to the approaches that we presented so far, the matrices involved in these systems are modified at each iteration. Hence, it is not possible to compute the preconditioner beforehand. In our experiments, we did not
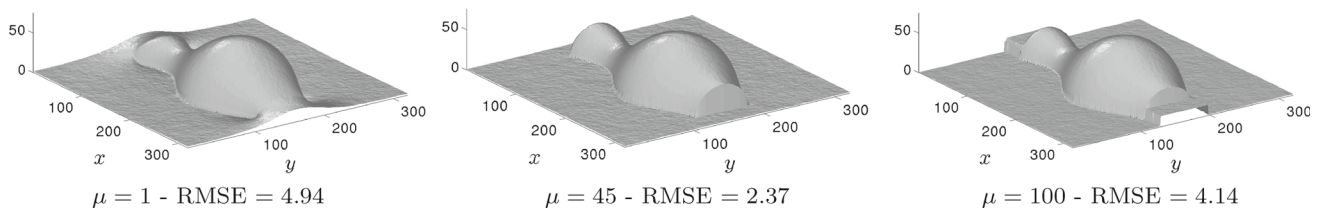
Fig. 13 3D-reconstructions from the noisy gradient of $\mathcal{S}_{\text{vase}}$ ($\sigma = 1\%$), using the Mumford–Shah integrator. If μ is tuned appropriately, sharp discontinuities can be restored, without staircasing artifacts

consider any preconditioning strategy at all. Thus, the proposed scheme could obviously be accelerated.

Discussion Let us now check experimentally, on the same noisy gradient of surface $\mathcal{S}_{\text{vase}}$ as in previous experiments, whether the Mumford–Shah integrator satisfies the expected properties. In the experiment of Fig. 13, we performed 50 iterations of the proposed alternating optimization scheme, with various choices for the hyper-parameter μ. The ϵ parameter was set to $\epsilon = 0.1$ (this parameter is not critical: it only has to be "small enough," in order for the Ambrosio–Tortorelli approximation to converge toward the Mumford–Shah functional). As it was already the case with other non-convex regularizers (see Sect. 4.3), a bad tuning of the parameter leads either to over-smoothing (low values of μ) or to staircasing artifacts (high values of μ), which indicate the presence of local minima. Yet, by appropriately setting this parameter, we obtain a 3D-reconstruction which is very close to the genuine surface, without staircasing artifact.

The Mumford–Shah functional being non-convex, local minima may exist. Yet, as shown in Fig. 14, the choice of the initialization may not be as crucial as with the non-convex estimators from Sect. 4.3. Indeed, the 3D-reconstruction of the "Canadian tent" surface is similar using as initial guess the least-squares solution or the trivial initialization $z^{(0)} \equiv 0$.

Hence, among all the variational integration methods we have studied, the adaptation of the Mumford–Shah model is the approach which provides the most satisfactory 3D-reconstructions in the presence of sharp features: it is possible to recover discontinuities and kinks, even in the presence of noise, and with limited artifacts. Nevertheless, local minima may theoretically arise, as well as staircasing if the parameter μ is not tuned appropriately.

5 Conclusion and Perspectives

We proposed several new variational methods for solving the normal integration problem. These methods were designed to satisfy the largest subset of properties that were identified in a companion survey paper [48] entitled *Normal Integration: A Survey*.

$z^{(0)} = $ least-squares solution - RMSE = 0.74

$z^{(0)} \equiv 0$ - RMSE = 1.84
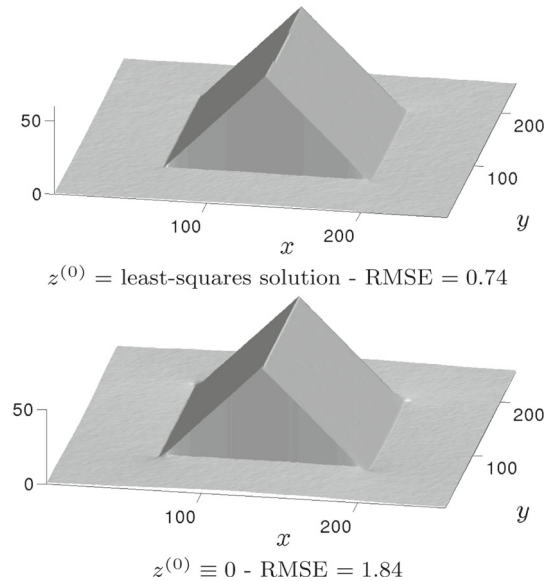
Fig. 14 3D-reconstructions of the "Canadian tent" surface from its noisy gradient ($\sigma = 1\%$), by the Mumford–Shah integrator ($\mu = 20$), using two different initializations. The initialization matters, but not as much as with the non-convex estimators from Sect. 4.3

We first detailed in Sect. 3 a least-squares solution which is fast, robust, and parameter-free, while assuming neither a particular shape for the integration domain nor a particular boundary condition. However, discontinuities in the surface can be handled only if the integration domain is first segmented into pieces without discontinuities. Therefore, we discussed in Sect. 4 several non-quadratic or non-convex variational formulations aiming at appropriately handling discontinuities. As we have seen, the latter property can be satisfied only if (slow) iterative schemes are used and / or one critical parameter is tuned. Therefore, there is still room for improvement: a fast, parameter-free integrator, able to handle discontinuities remains to be proposed.

Table 1 summarizes the main features of the five new integration methods proposed in this article. Contrarily to Table 1 in [48], which recaps the features of state-of-the-art methods, this time we use a more nuanced evaluation than binary features $+/-$. Among the new methods, we believe that the least-squares method discussed in Sect. 3 is the best if speed is the most important criterion, while the Mumford–

Table 1 Main features of the five methods of integration proposed in this paper

Method	$\mathcal{P}_{Fast}$	$\mathcal{P}_{Robust}$	$\mathcal{P}_{FreeB}$	$\mathcal{P}_{Disc}$	$\mathcal{P}_{NoRect}$	$\mathcal{P}_{NoPar}$	Local minima	Staircasing
Quadratic	+ + +	+	+	−	+	+	No	No
Total variation	+	+	+	+	+	+	No	Yes
Non-convex regularization	−	+	+	+ + +	+	−	Yes	Yes
Anisotropic diffusion	−	+	+	+ +	+	− −	No	No
Mumford–Shah	−	+	+	+ + +	+	− −	Yes	Yes

The quadratic method has all desirable properties, except $\mathcal{P}_{Disc}$. The others lose $\mathcal{P}_{Fast}$ but hold $\mathcal{P}_{Disc}$. Sharpest features are recovered by using non-convex regularization or the Mumford–Shah approach, yet staircasing artifacts and local minima may appear. In addition, all discontinuity-preserving methods except TV require tuning at least one hyper-parameter. Yet, TV is not able to recover discontinuities in the presence of noise. Overall, we recommend using: quadratic integration if speed is the most important issue; the Mumford–Shah approach if recovering discontinuities is the most important issue; and anisotropic diffusion if discontinuities are present, but limited
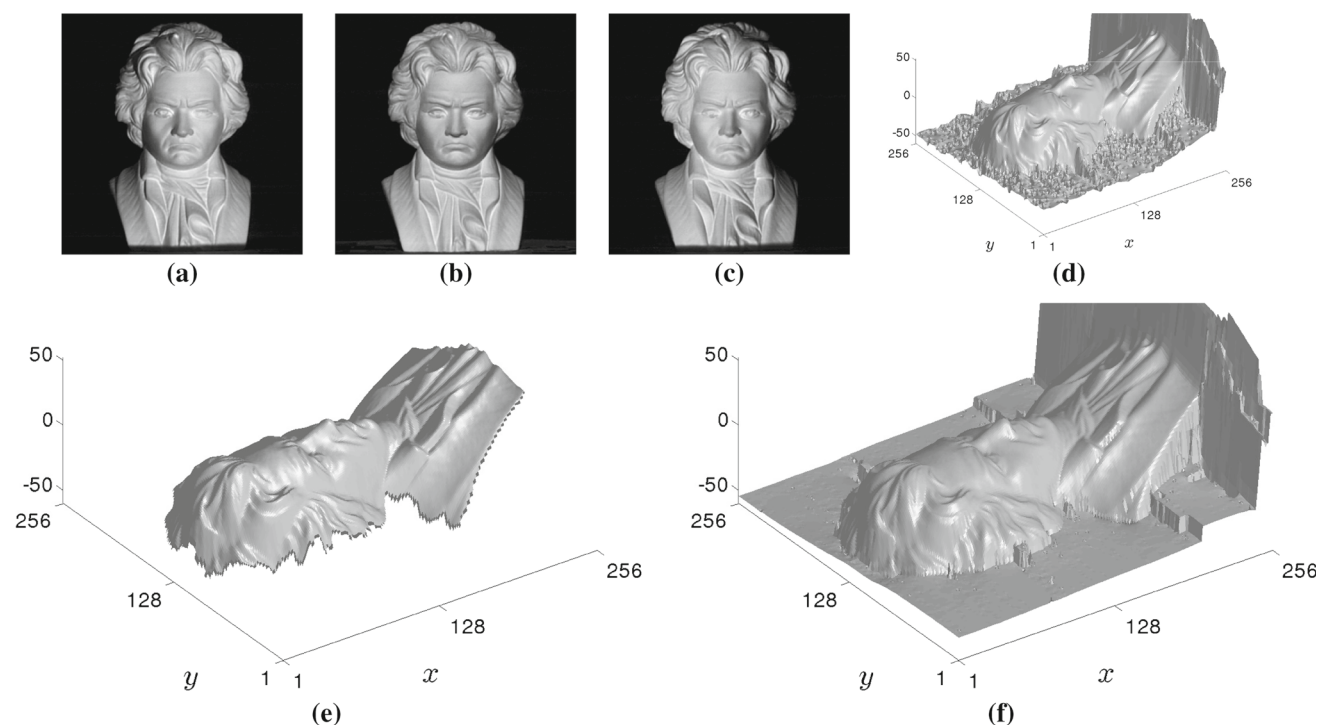
Fig. 15 3D-reconstruction using photometric stereo. **a–c** All (real) input images, **d** 3D-reconstruction by least squares on the whole grid, **e** 3D-reconstruction by least squares on the non-rectangular reconstruction domain corresponding to the images of the bust, **f** 3D-reconstruction using the Mumford–Shah approach, on the whole grid. When discontinuities are handled, it is possible to perform photometric stereo without prior segmentation of the object

Shah approach discussed in Sect. 4.5 is the most appropriate one for recovering discontinuities and kinks. Inbetween, the anisotropic diffusion approach from Sect. 4.4 represents a good compromise.

Future research directions may include accelerating the numerical schemes and proving their convergence when this is not trivial (e.g., for the non-convex integrators). We also believe that introducing additional smoothness terms inside the functionals may be useful for eliminating the artifacts in anisotropic diffusion integration. Quadratic (Tikhonov) smoothness terms were suggested in [26]: to enforce surface smoothness while preserving the discontinuities, we should rather consider non-quadratic ones. In this view, higher-order functionals (e.g., total generalized variation methods [10]) may reduce not only these artifacts, but also staircasing. Indeed, as shown in Fig. 15, such artifacts may be visible when performing photometric stereo [55] without prior segmentation. Yet, this example also shows that the artifacts are visible only over the background, and do not seem to affect the relevant part.

3D-reconstruction is not the only application where efficient tools for gradient field integration are required.
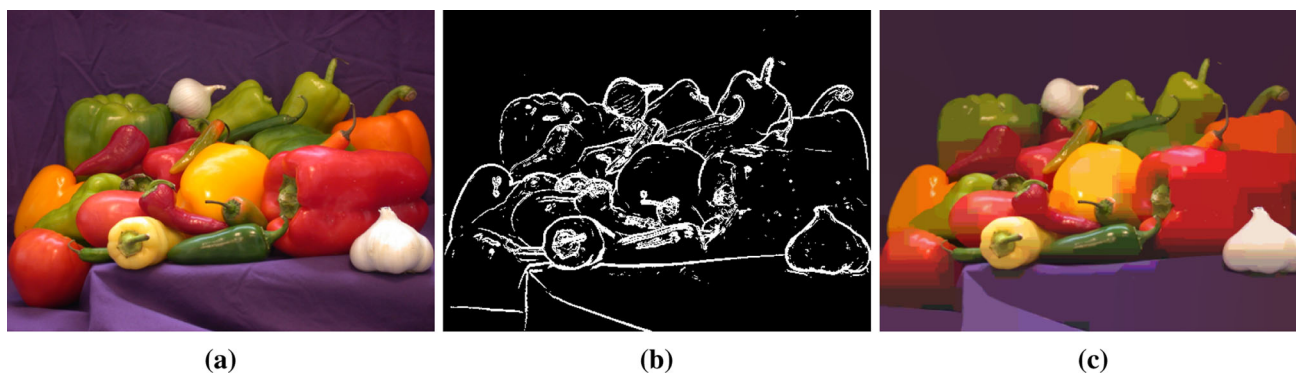
Fig. 16 Application to image compression/image editing. **a** Reference image, **b** control points (where the RGB values and their gradients are kept), **c** restored image obtained by considering the proposed Mumford–Shah integrator as a piecewise-constant interpolation method. A reasonable piecewise-constant restoration of the initial image can be obtained from as few as 10% of the initial information

Although the assumption on the noise distribution may differ from one application to another, PDE-based imaging problems such as Laplace image compression [45] or Poisson image editing [43] also require an efficient integrator. In this view, the ability of our methods to handle control points may be useful. We illustrate in Fig. 16 an interesting application. From an RGB image I, we selected the points where the norm of the gradient of the luminance (in the CIE-LAB color space) was the highest (conserving only 10% of the points). Then, we created a gradient field $\mathbf{g}$ equal to zero everywhere, except on the control points, where it was set to the gradient of the color levels. The prior z^0 was set to a null scalar field, except on the control points where we retained the original color data. Eventually, λ is set to an arbitrary small value ($\lambda = 10^{-9}$) everywhere, except on the control points ($\lambda = 10$). The integration of each color channel gradient is performed independently, using the Mumford–Shah method to extrapolate the data from the control points to the whole grid. Using this approach, we obtain a nice piecewise-constant approximation of the image, in the spirit of the "texture-flattening" application presented in [43]. Besides, by selecting the control points in a more optimal way [8,28], this approach could easily be extended to image compression, reaching state-of-the-art lossy compression rates. In fact, existing PDE-based methods can already compete with the compression rate of the well-known JPEG 2000 algorithm [45]. We believe that the proposed edge-preserving framework may yield even better results.

Eventually, some of the research directions already mentioned in the conclusion section of our survey paper [48] were ignored in this second paper, but they remain of important interest. One of the most appealing examples is multiview normal field integration [15]. Indeed, discontinuities represent a difficulty in our case because they are induced by occlusions, yet more information would be obtained near the occluding contours by using additional views.

References

1. Agrawal, A., Raskar, R., Chellappa, R.: What is the range of surface reconstructions from a gradient field? In: Proceedings of the 9th European Conference on Computer Vision (vol. 1), Lecture Notes in Computer Science, vol. 3951, pp. 578–591. Graz (2006)
2. Ambrosio, L., Tortorelli, V.M.: Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence. Commun. Pure Appl. Math. **43**, 999–1036 (1990)
3. Attouch, H., Buttazzo, G., Michaille, G.: Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization, 2nd edn. SIAM (2014)
4. Aubert, G., Kornprobst, P. (eds.): Mathematical problems in image processing. In: Applied Mathematical Sciences, vol. 147, 2nd edn. Springer, Berlin (2006)
5. Badri, H., Yahia, H., Aboutajdine, D.: Robust surface reconstruction via triple sparsity. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2291–2298. Columbus (2014)
6. Bähr, M., Breuß, M., Quéau, Y., Boroujerdi, A.S., Durou, J.D.: Fast and accurate surface normal integration on non-rectangular domains. Comput. Vis. Media **3**, 107–129 (2017)
7. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imaging Sci. **2**(1), 183–202 (2009)
8. Belhachmi, B., Bucur, D., Burgeth, B., Weickert, J.: How to choose interpolation data in images. SIAM J. Appl. Math. **70**(1), 333–352 (2009)
9. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends Mach. Learn. **3**(1), 1–122 (2011)
10. Bredies, K., Holler, M.: A TGV-based framework for variational image decompression, zooming, and reconstruction. Part I: analytics. SIAM J. Imaging Sci. **8**(4), 2814–2850 (2015)
11. Catté, F., Lions, P.L., Morel, J.M., Coll, T.: Image selective smoothing and edge detection by nonlinear diffusion. SIAM J. Numer. Anal. **29**(1), 182–193 (1992)
12. Chambolle, A.: Image segmentation by variational methods: Mumford and Shah functional and the discrete approximation. SIAM J. Appl. Math. **55**(3), 827–863 (1995)

13. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. J. Math. Imaging Vis. **40**(1), 120–145 (2010)

14. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Trans. Image Process. **10**(2), 266–277 (2001)

15. Chang, J.Y., Lee, K.M., Lee, S.U.: Multiview normal field integration using level set methods. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Workshop on Beyond Multiview Geometry: Robust Estimation and Organization of Shapes from Multiple Cues. Minneapolis (2007)

16. Charbonnier, P., Blanc-Féraud, L., Aubert, G., Barlaud, M.: Deterministic edge-preserving regularization in computed imaging. IEEE Trans. Image Process. **6**(2), 298–311 (1997)

17. Chen, S.S., Donoho, D.L., Saunders, M.A.: Atomic decomposition by basis pursuit. SIAM J. Sci. Comput. **20**(1), 33–61 (1998)

18. Du, Z., Robles-Kelly, A., Lu, F.: Robust surface reconstruction from gradient field using the L1 norm. In: Proceedings of the 9th Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications, pp. 203–209. Glenelg (2007)

19. Durou, J.D., Aujol, J.F., Courteille, F.: Integration of a normal field in the presence of discontinuities. In: Proceedings of the 7th International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition, Lecture Notes in Computer Science, vol. 5681, pp. 261–273. Bonn (2009)

20. Durou, J.D., Courteille, F.: Integration of a normal field without boundary condition. In: Proceedings of the 11th IEEE International Conference on Computer Vision, 1st Workshop on Photometric Analysis for Computer Vision. Rio de Janeiro (2007)

21. Gabay, D., Mercier, B.: A dual algorithm for the solution of nonlinear variational problems via finite element approximation. Comput. Math. Appl. **2**(1), 17–40 (1976)

22. Geman, D., Reynolds, G.: Constrained restoration and recovery of discontinuities. IEEE Trans. Pattern Anal. Mach. Intell. **14**(3), 367–383 (1992)

23. Goldstein, T., O'Donoghue, B., Setzer, S., Baraniuk, R.: Fast alternating direction optimization methods. SIAM J. Imaging Sci. **7**(3), 1588–1623 (2014)

24. Goldstein, T., Osher, S.: The split Bregman method for L1-regularized problems. SIAM J. Imaging Sci. **2**(2), 323–343 (2009)

25. Haque, S.M., Chatterjee, A., Govindu, V.M.: High quality photometric reconstruction using a depth camera. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2283–2290. Columbus (2014)

26. Harker, M., O'Leary, P.: Regularized reconstruction of a surface from its measured gradient field. J. Math. Imaging Vis. **51**(1), 46–70 (2015)

27. Hayya, J., Armstrong, D., Gressis, N.: A note on the ratio of two normally distributed variables. Manag. Sci. **21**(11), 1338–1341 (1975)

28. Hoeltgen, L., Setzer, S., Weickert, J.: An optimal control approach to find sparse data for Laplace interpolation. In: Proceedings of the 9th International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition, Lecture Notes in Computer Science, vol. 8081, pp. 151–164. Lund (2013)

29. Horn, B.K.P., Brooks, M.J.: The variational approach to shape from shading. Comput. Vis. Graph. Image Process. **33**(2), 174–208 (1986)

30. Horovitz, I., Kiryati, N.: Depth from gradient fields and control points: bias correction in photometric stereo. Image Vis. Comput. **22**(9), 681–694 (2004)

31. Ikehata, S., Wipf, D., Matsushita, Y., Aizawa, K.: Photometric stereo using sparse Bayesian regression for general diffuse surfaces. IEEE Trans. Pattern Anal. Mach. Intell. **36**(9), 1816–1831 (2014)

32. Kadambi, A., Taamazyan, V., Shi, B., Raskar, R.: Polarized 3D: high-quality depth sensing with polarization cues. In: Proceedings

33. Kimmel, R., Yavneh, I.: An algebraic multigrid approach for image analysis. SIAM J. Sci. Comput. **24**(4), 1218–1231 (2003)

34. Kornprobst, P., Aubert, G.: Image sequence analysis via partial differential equations. J. Math. Imaging Vis. **11**(1), 5–26 (1999)

35. Koutis, I., Miller, G.L., Peng, R.: A nearly-m log n time solver for SDD linear systems. In: Proceedings of the IEEE Annual Symposium on Foundations of Computer Science, pp. 590–598. Palm Springs (2011)

36. Lanza, A., Morigi, S., Sgallari, F.: Convex image denoising via nonconvex regularization with parameter selection. J. Math. Imaging Vis. **56**(2), 195–220 (2016)

37. Mumford, D., Shah, J.: Optimal approximations by piecewise smooth functions and associated variational problems. Commun. Pure Appl. Math. **42**(5), 577–685 (1989)

38. Nikolova, M.: Local strong homogeneity of a regularized estimator. SIAM J. Appl. Math. **61**(2), 633–658 (2000)

39. Noakes, L., Kozera, R.: Nonlinearities and noise reduction in 3-source photometric stereo. J. Math. Imaging Vis. **18**(2), 119–127 (2003)

40. Ochs, P., Brox, T., Pock, T.: iPiasco: inertial proximal algorithm for strongly convex optimization. J. Math. Imaging Vis. **53**(2), 171–181 (2015)

41. Ochs, P., Chen, Y., Brox, T., Pock, T.: iPiano: inertial proximal algorithm for nonconvex optimization. SIAM J. Imaging Sci. **7**(2), 1388–1419 (2014)

42. Or-el, R., Rosman, G., Wetzler, A., Kimmel, R., Bruckstein, A.M.: RGBD-Fusion: real-time high precision depth recovery. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5407–5416. Boston (2015)

43. Pérez, P., Gangnet, M., Blake, A.: Poisson image editing. ACM Trans. Graph. **22**(3), 313–318 (2003)

44. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell. **12**(7), 629–639 (1990)

45. Peter, P., Hoffmann, S., Nedwed, F., Hoeltgen, L., Weickert, J.: Evaluating the true potential of diffusion-based inpainting in a compression context. Signal Process. Image Commun. **46**, 40–53 (2016)

46. Pock, T., Cremers, D., Bischof, H., Chambolle, A.: An algorithm for minimizing the Mumford–Shah functional. In: Proceedings of the 12th IEEE International Conference on Computer Vision, pp. 1133–1140. Kyoto (2009)

47. Quéau, Y., Durou, J.D.: Edge-preserving integration of a normal field: weighted least squares, TV and L1 approaches. In: Proceedings of the 5th International Conference on Scale Space and Variational Methods in Computer Vision, Lecture Notes in Computer Science, vol. 9087, pp. 576–588. Lège Cap-Ferret (2015)

48. Quéau, Y., Durou, J.D., Aujol, J.F.: Normal integration: a survey. J. Math. Imaging Vis. (2017). https://doi.org/10.1007/s10851-017-0773-x

49. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Phys. D Nonlinear Phenom. **60**(1–4), 259–268 (1992)

50. Saracchini, R.F.V., Stolfi, J., Leitão, H.C.G., Atkinson, G.A., Smith, M.L.: A robust multi-scale integration method to obtain the depth from gradient maps. Comput. Vis. Image Underst. **116**(8), 882–895 (2012)

51. Shefi, R., Teboulle, M.: Rate of convergence analysis of decomposition methods based on the proximal method of multipliers for convex minimization. SIAM J. Optim. **24**(1), 269–297 (2014)

52. Shi, B., Wu, Z., Mo, Z., Duan, D., Yeung, S.K., Tan, P.: A benchmark dataset and evaluation for non-Lambertian and uncalibrated photometric stereo. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas (2016)

53. Simchony, T., Chellappa, R., Shao, M.: Direct analytical methods for solving Poisson equations in computer vision problems. IEEE Trans. Pattern Anal. Mach. Intell. **12**(5), 435–446 (1990)
54. Weickert, J.: Anisotropic Diffusion in Image Processing. Teubner, Stuttgart (1998)
55. Woodham, R.J.: Photometric method for determining surface orientation from multiple images. Opt. Eng. **19**(1), 139–144 (1980)

Yvain Quéau is a postdoctoral researcher in Technical University Munich. He received his Ph.D. in Computer Science from INP-ENSEEIHT, Université de Toulouse, in 2015. His research interests include 3D-reconstruction by photometric techniques (shape-from-shading and photometric stereo), as well as variational methods for solving computer vision and image processing problems.

Jean-François Aujol studied Mathematics at the Ecole Normale Supérieure (ENS) in Cachan. He received the Ph.D. degree in Mathematics from Nice-Sophia-Antipolis University in June 2004 and the Habilitation thesis in Applied Mathematics from ENS Cachan in May 2009. In 2004–2005, he was successively an Assistant Researcher in UCLA (Applied Mathematics Department) and then a Research Engineer at Telecom ParisTech. In 2005–2011, Jean-François Aujol was a scientist researcher with CNRS at ENS Cachan with CMLA (until 2009) and then in Marseille with LATP. Since March 2011, he has been a full professor of applied mathematics in Bordeaux with IMB. Since September 2013, he has been a member of Institut Universitaire de France (IUF). He is interested in mathematical problems related to image processing.

Jean-Denis Durou received the Ph.D. degree in Computer Science from the Université Paris Sud—Orsay in 1993 and the "Habilitation à Diriger les Recherches" from the Université Toulouse III—Paul Sabatier in 2007. He is an Assistant Professor at the Université Toulouse III since 1994 and a member of the VORTEX team at the IRIT laboratory. His main research interest is 3D-vision. He is more specifically interested in photometric 3D-reconstruction, i.e., shape-from-shading and photometric stereo.