

Adaptive Second-Order Total Variation: An Approach Aware of Slope Discontinuities

Frank Lenzen¹, Florian Becker¹, and Jan Lellmann²

¹ Heidelberg Collaboratory for Image Processing (HCI), Heidelberg, Germany

² DAMTP, University of Cambridge, UK

Abstract. Total variation (TV) regularization, originally introduced by Rudin, Osher and Fatemi in the context of image denoising, has become widely used in the field of inverse problems. Two major directions of modifications of the original approach were proposed later on. The first concerns *adaptive* variants of TV regularization, the second focuses on *higher-order* TV models. In the present paper, we combine the ideas of both directions by proposing *adaptive second-order* TV models, including one *anisotropic* model. Experiments demonstrate that introducing adaptivity results in an improvement of the reconstruction error.

Keywords: second-order total variation, adaptive, anisotropic, directional, TV, TGV, slope discontinuities

1 Introduction

In 1992 Rudin, Osher and Fatemi [14] proposed to apply the total-variation (TV) semi-norm for regularization in a variational framework for image denoising. Their approach not only had a significant impact in the area of image restoration, but in the whole field of inverse problems. Since then, various modifications and improvements have been contributed by the community. Several publications have been devoted to the idea of adaptive TV regularization methods, where the regularization varies locally depending on the noise level or the image content [3, 4, 6, 7]. Non-local TV models (e.g. [10]), which have proven as effective variants, can also be regarded as adaptive methods, since they use image information to locally determine the regularization weights. Another subclass of TV approaches are the anisotropic or directional methods, where the regularization not only depends on the location but also on the local orientation of the signal to be reconstructed [1, 8, 12, 18]. TV regularization has the major benefit that it allows piecewise constant signals to be recovered. Recent works have shown that in certain cases it might be beneficial to assume even higher regularity of the signal, and thus introduced higher-order regularization schemes [2, 9, 11, 13, 15, 17].

Contribution. We combine adaptive and second-order TV approaches into one regularization framework. Such a combination has not been proposed up to now. Our approach uses information on local image structures, in particular on edges and slope discontinuities obtained from structure tensors applied to the image and its epigraph. We demonstrate, that our approach can be applied to the

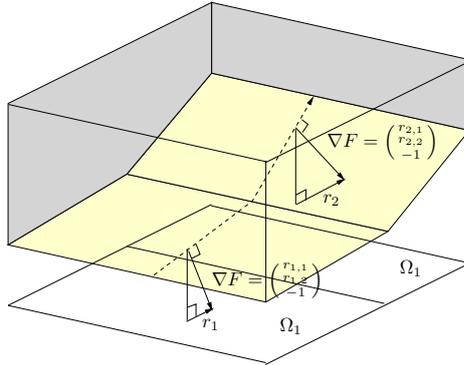


Fig. 1. Graph $\Gamma = (x, y, u(x, y))^\top$ (yellow) of a continuous and piecewise affine function u with a discontinuity in the gradient (interface between Ω_1 and Ω_2). The epigraph of u is the volume above Γ , represented as the super-level set of $F(x, y, z) = u(x, y) - z$. On the graph the gradient ∇F of F coincides with the surface normal of Γ .

standard second-order TV regularization as well as to regularization with total generalized variation (TGV) [2] and infimal convolution (IC) [17]. Moreover, we propose a new *anisotropic* second-order TV model and show its advantages over the isotropic models.

Paper organization. In Sect. 2 we describe how the information on image structures required to steer adaptive regularization is retrieved. In Sect. 3 we consider adaptive second-order TV models. Experiments are provided in Sect. 4.

2 Detecting Discontinuities in Piecewise Affine Functions

In this section we provide an approach to extract information about the direction and location of edges and the location of slope discontinuities from a given input image. The first task is already addressed in literature. We rely on the standard structure tensor and just briefly recall the required definitions. However, we will see that this approach is not suitable for detecting slope discontinuities (sharp bends, kinks). For this second task, we propose a new approach.

2.1 Edge detection

In the following, we represent an image as a function $u : \Omega \rightarrow \mathbb{R}$, $\Omega \subset \mathbb{R}^2$. For detecting edges in u we follow the standard approach and use the classical structure tensor (cf. [5]) to identify regions with high gradient magnitude. To this end, let

$$S_u(x, y) := (\nabla u_\sigma(x, y) \nabla u_\sigma(x, y)^\top)_\rho. \quad (1)$$

be the standard structure tensor calculated on u_σ , which is obtained from u by convolution with a Gaussian kernel with variance σ^2 . Furthermore, $(\cdot)_\rho$ denotes

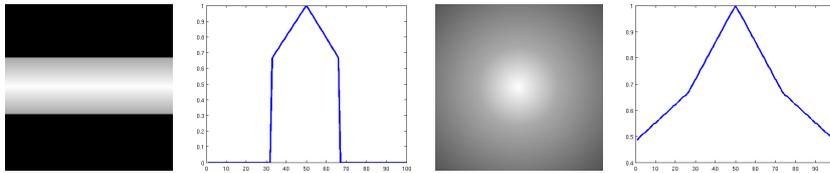


Fig. 2. Test images *roof* and *cone hat* for detecting slope discontinuities

a component-wise convolution of each entry with a Gaussian kernel with variance ρ^2 . We denote by $\lambda_1^S(x, y), \lambda_2^S(x, y)$ the eigenvalues of $S_u(x, y)$ ordered with decreasing value, i.e. $\lambda_1^S(x, y) \geq \lambda_2^S(x, y)$. Moreover, we consider the eigenvector v^S to the eigenvalue λ_1^S . It is known that along edges in the image, λ_1^S takes large values, whereas λ_2^S is almost zero. Thus, $d^S(x, y) := \lambda_1^S(x, y) - \lambda_2^S(x, y)$ indicates the presence of edges. We define $E^S : \Omega \rightarrow [0, 1]$ as $E^S(x, y) := \min\{c d^S(x, y), 1\}$ with some constant $c > 0$. In Sect. 3 we make use of the edge indicating function E^S together with the vector field v^S .

2.2 Slope discontinuities

The standard structure tensor as considered so far is sufficient to identify discontinuities (edges) in u . We now focus on regions where u is continuous but has discontinuities in its first derivatives. In addition, we assume that u is piecewise affine. This assumption is in view of our ansatz in Sect. 3 to determine u as the solution of a second-order TV approach. For the sake of simplicity, let us consider a prototypical function model with only one discontinuity, which locally represents a part of a larger image: we assume that Ω can be divided into two segments $\Omega_i, i = 1, 2$ such that u is affine in each segment, i.e. u can be represented as

$$u(x, y) = \begin{cases} r_1^\top \begin{pmatrix} x \\ y \end{pmatrix} + b_1 & \text{if } (x, y) \in \Omega_1, \\ r_2^\top \begin{pmatrix} x \\ y \end{pmatrix} + b_2 & \text{if } (x, y) \in \Omega_2, \end{cases} \quad (2)$$

for Ω_i open, such that $\Omega_1 \cap \Omega_2 = \emptyset$ and $\overline{\Omega_1} \cup \overline{\Omega_2} = \Omega$, and $r_i \in \mathbb{R}^2, b_i \in \mathbb{R}$ for $i = 1, 2$. Fig. 1 illustrates such a prototypical function u .

The aim of this section is to derive a method to detect the case where $r_1 \neq r_2$. To this end, we consider the epigraph of u defined as the super-level set $\{(x, y, z) \mid F(x, y, z) \geq 0\}$ of $F(x, y, z) := u(x, y) - z$. In order to detect (surface) edges of the graph (i.e. locations, where the slope changes), we now apply the three-dimensional structure tensor to F , i.e. $((\nabla F)(\nabla F)^\top)_\rho$, where $\nabla F(x, y, z) = (\partial_x u^2 + \partial_y^2 u + 1)^{-\frac{1}{2}} (\partial_x u, \partial_y u, -1)^\top$. Note that ∇F is constant in z . Since we are only interested in edges of the graph $\Gamma := \{(x, y, z) \mid F(x, y, z) = 0\}$ (i.e. slope discontinuities), we restrict this structure tensor to Γ :

$$T_u(x, y) := \left((\nabla \tilde{F}(x, y)) (\nabla \tilde{F}(x, y))^\top \right)_\rho, \quad (3)$$

where $\nabla\tilde{F}_u(x, y) := \nabla F(x, y, u(x, y))$. We observe that $\nabla\tilde{F}_u(x, y)$ is the normal to the graph Γ at $(x, y, u(x, y))$.

Remark 1. The following two scenarios are of particular interest:

Within an affine region: For an affine function u , $T_u(x, y)$ has exactly one non-zero eigenvalue. This is due to the fact that in this case $\nabla\tilde{F}_u(x, y)$ is constant and convolution of $\nabla\tilde{F}_u\nabla\tilde{F}_u^\top$ does not change the rank.

Interface between two affine regions of different slope: For such u , $T_u(x, y)$ sums up two different directions $(r_{1,1}, r_{1,2}, -1)$ and $(r_{2,1}, r_{2,2}, -1)$: rewriting the convolution of the matrix entries as a weighted integral,

$$\begin{aligned} T_u(x, y) &= (\nabla\tilde{F}_u\nabla\tilde{F}_u^\top)_\rho = \int_{\Omega} w(x)\nabla\tilde{F}_u\nabla\tilde{F}_u^\top dx \\ &= w_1 \begin{pmatrix} r_{1,1} & r_{1,2} \\ r_{1,2} & -1 \end{pmatrix} (r_{1,1}, r_{1,2}, -1) + w_2 \begin{pmatrix} r_{2,1} & r_{2,2} \\ r_{2,2} & -1 \end{pmatrix} (r_{2,1}, r_{2,2}, -1) \end{aligned} \quad (4)$$

with $w_i := \int_{\Omega_i} w(x) dx$, we observe that in (4) two rank-1 matrices are added up. Each matrix has one non-zero eigenvalue $w_i \cdot \|(r_{i,1}, r_{i,2}, -1)\|_2^2$ with corresponding eigenvector $v_i = (r_{i,1}, r_{i,2}, -1)$. Since the eigenvectors are linear dependent only if $r_1 = r_2$, $T_u(x, y)$ has rank 2 near the discontinuity, where $r_1 \neq r_2$.

In the following we denote by $\lambda_i^T(x, y)$, $i = 1, 2, 3$ the eigenvalues of $T_u(x, y)$ in decreasing order. As an indicator for the existence of slope discontinuities we propose to use $\lambda_2^T(x)$. This is motivated by the fact that, similar to the standard structure tensor in $2D$, $T_u(x, y)$ reveals two eigenvalues significantly larger than 0 at edges of the graph, while in regions of constant slope the second eigenvalue becomes 0. Therefore the magnitude of the second eigenvalue can be used to distinguish between both cases. We propose $E^T : \Omega \rightarrow [0, 1]$, $E^T(x) := \min(c\lambda_2^T(x), 1)$ with some constant $c > 0$ as an indicator for regions of slope discontinuities. In order to be less sensitive to edges, which are already covered by the standard structure tensor, we use an upwind scheme to compute the gradient in (3). In practice, it is advisable to use the pre-smoothed u_σ (cf. Sect. 2.1) instead of u to be robust against noise.

To demonstrate the benefits of using E^T to detect slope discontinuities, we compare our approach to one approach based on the standard structure tensor and one based on curvature, see Fig. 3. We observe that our approach detects slope discontinuities more reliably than the competitive methods.

3 Adaptive Second-Order Total Variation

In the following we discuss three state-of-the-art approaches for second-order total variation (TV) regularization. First, we focus on the straightforward approach of combining two TV semi-norms of first and second order [12, 15]. We generalize this approach to allow for anisotropic regularization with locally adaptive strength. In addition, we consider two alternative approaches – infimal convolution (IC) [17] and total generalized variation (TGV) [2] – and propose a spatially adaptive choice of the regularization parameters.

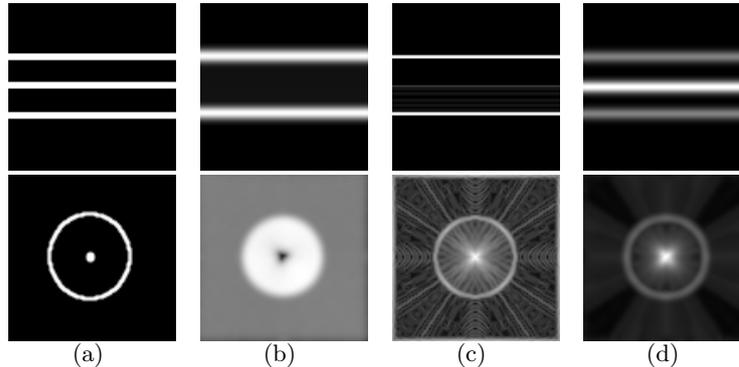


Fig. 3. Detecting slope discontinuities using the standard structure tensor (b), a curvature based approach (c), and the proposed method (d) in the test images depicted in Fig. 2 (black=0, white=1). In both cases the standard structure tensor fails to detect the slope discontinuities as shown in the ideal result (a) (middle line in the first image, ring and center point in the second image). Only the proposed approach detects the slope discontinuity in the first test image (top row). On the second test image (bottom row), the proposed approach provides a less noisy and more precise result than the curvature based approach.

3.1 Proposed Approach

Let $BV^2(\Omega)$ (Ω open, bounded, with Lipschitz boundary) be the space of functions with bounded first and second-order TV, i.e. $u \in BV^2(\Omega)$ iff $u \in L^1(\Omega)$ and

$$TV^l(u) := \sup \left\{ \int_{\Omega} u \operatorname{div}^l \varphi \, dx \mid \varphi \in C_c^\infty(\Omega, \mathbb{R}^{2l}), \forall x \in \Omega : \|\varphi(x)\|_2 \leq 1 \right\}, \quad (5)$$

is finite for $l = 1, 2$. Here, div^1 is the divergence operator and $\operatorname{div}^2 \varphi := \partial_{xx} \varphi_1 + \partial_{yx} \varphi_2 + \partial_{xy} \varphi_3 + \partial_{yy} \varphi_4$, where $\varphi = (\varphi_1, \varphi_2, \varphi_3, \varphi_4)^\top$. Note that for $u \in BV^2(\Omega)$ we have $\partial_x u, \partial_y u \in L^1(\Omega)$. For details on $BV^2(\Omega)$ we refer to [16, Chapter 9.8]. A standard denoising approach with first and second-order TV regularization consists in minimizing the functional

$$\mathcal{F}_{TV^2}(u) := \frac{1}{2} \|u - f\|_{L^2}^2 + \alpha TV(u) + \beta TV^2(u) \quad (6)$$

for given data $f \in L^2(\Omega)$ and regularization parameters $\alpha, \beta > 0$. We generalize this approach in two ways. Firstly, we allow α, β to vary depending on the location, i.e., $\alpha, \beta : \Omega \rightarrow \mathbb{R}_+$. Secondly, we allow anisotropic, i.e. directionally dependent regularization. To this end, we consider the optimization problem

$$\mathcal{F}(u) := \frac{1}{2} \|u - f\|_{L^2}^2 + \mathcal{R}_1(u) + \mathcal{R}_2(u) \quad (7)$$

with two regularization terms $\mathcal{R}_1(u)$ and $\mathcal{R}_2(u)$ defined as follows. For first-order TV, we use anisotropic TV regularization (cf. [7]) given as

$$\mathcal{R}_1(u) := \int_{\Omega} (\nabla u^\top(x) A(x) \nabla u(x))^{\frac{1}{2}} \, dx, \quad (8)$$

for some matrix-valued mapping $A : \Omega \rightarrow \mathbb{R}_{sym}^{2 \times 2}$, where $A(x)$ is symmetric and positive semi-definite at every x . Every such matrix $A(x)$ can be written as $A(x) = (v(x), v^\perp(x)) \begin{pmatrix} \alpha_1(x) & 0 \\ 0 & \alpha_2(x) \end{pmatrix} (v(x), v^\perp(x))^\top$ with some vector field $v(x)$, $\|v(x)\|_2 = 1$. We observe that (8) leads to an anisotropic regularization with strength $\alpha_1(x)$ in direction of $v(x)$ and $\alpha_2(x)$ in direction of $v^\perp(x)$.

For adaptive second-order TV regularization we propose a new approach, which we motivate by the smooth case $u \in C^2(\Omega)$: for arbitrary $\varphi \in C_c^\infty(\Omega, \mathbb{R}^4)$ we have

$$\int_{\Omega} (\operatorname{div}^2 \varphi) u \, dx = \int_{\Omega} \langle \varphi, \nabla^2 u \rangle \, dx, \quad (9)$$

where $\nabla^2 u := (\partial_{xx} u, \partial_{xy} u, \partial_{yx} u, \partial_{yy} u)^\top$. For a given normalized vector field $v(x) = (v_1(x), v_2(x))^\top \in \mathbb{R}^2$, $\|v(x)\|_2 = 1$, we represent φ as $\varphi = t_1 w_1 + t_2 w_2 + s_1 w_3 + s_2 w_4$, where $t, s \in \mathbb{R}^2$ and $w_1 := (v_1, v_2, 0, 0)^\top$, $w_2 := (0, 0, v_1, v_2)^\top$, $w_3 := (v_1^\perp, v_2^\perp, 0, 0)$ and $w_4 := (0, 0, v_1^\perp, v_2^\perp)$. Note that $\{w_i\}_i$ form an orthonormal basis of \mathbb{R}^4 . Then, standard calculus shows

$$\langle \varphi, \nabla^2 u \rangle = t^\top (Hu)v + s^\top (Hu)v^\perp \quad \text{for } Hu := \begin{pmatrix} \partial_{xx} u & \partial_{xy} u \\ \partial_{yx} u & \partial_{yy} u \end{pmatrix}. \quad (10)$$

Now we calculate $\beta_1 \|(Hu)v\|_2 + \beta_2 \|(Hu)v^\perp\|_2$ for some weighting constants $\beta_1, \beta_2 > 0$. To this end, we take in (10) the supremum over $t \in B_{\beta_1}(0)$ and $s \in B_{\beta_2}(0)$, where $B_r(0)$ denotes the ball centered at 0 with radius r , and derive

$$\sup_{t \in B_{\beta_1}(0), s \in B_{\beta_2}(0)} \langle \varphi, \nabla^2 u \rangle = \beta_1 \|(Hu)v\|_2 + \beta_2 \|(Hu)v^\perp\|_2. \quad (11)$$

Thus, we obtain in (11) the absolute values of the second order derivative of u in direction of v weighted by β_1 and in perpendicular direction weighted by β_2 . The above considerations motivate the following definition for arbitrary $u \in L^1(\Omega)$:

$$\mathcal{R}_2(u) := \sup \left\{ \int_{\Omega} (\operatorname{div}^2 \varphi) u \, dx \mid \varphi \in \mathcal{C} \right\}, \quad \text{with} \quad (12)$$

$$\mathcal{C} := \left\{ C_c^\infty(\Omega; \mathbb{R}^4), \forall x \in \Omega : \langle \varphi(x), w_1(x) \rangle^2 + \langle \varphi(x), w_2(x) \rangle^2 \leq (\beta_1(x))^2, \right. \\ \left. \langle \varphi(x), w_3(x) \rangle^2 + \langle \varphi(x), w_4(x) \rangle^2 \leq (\beta_2(x))^2 \right\}, \quad (13)$$

Existence Theory

We now show the existence of a unique minimizer of (7), where $\mathcal{R}_1(u)$ and $\mathcal{R}_2(u)$ are given by (8) and (12), respectively.

Proposition 1. *Assume that for every $x \in \Omega$ the eigenvalues $\lambda_i(x)$ of $A(x)$ are uniformly bounded by $0 < c_1 \leq \lambda_i(x) \leq c_2 < \infty$. Moreover, assume that $\|v(x)\|_2 = 1$ and that $\beta_i(x)$, $i = 1, 2$ are bounded by $0 < c_3 \leq \beta_i(x) \leq c_4 < \infty$. Then functional (7) attains a unique minimizer in $L^2(\Omega) \cap BV^2(\Omega)$.*

The proof of Prop. 1 utilizes the following two lemmas:

Lemma 1. *Under the assumptions of Prop. 1 we have*

$$c_1 \operatorname{TV}(u) \leq \mathcal{R}_1(u) \leq c_2 \operatorname{TV}(u), \quad c_3 \operatorname{TV}^2(u) \leq \mathcal{R}_2(u) \leq \sqrt{2} c_4 \operatorname{TV}^2(u). \quad (14)$$

Proof. The first claim follows, since $c_1\|v\|_2 \leq \sqrt{v^\top Av} \leq c_2\|v\|_2$ for any $v \in \mathbb{R}^2$. To show the inequalities for \mathcal{R}_2 , we note that $cTV^2(u) = \sup\{\int_\Omega (\operatorname{div}^2 \varphi)u \, dx \mid \varphi \in \mathcal{C}(c)\}$ with $\mathcal{C}(c) := \{\varphi \in C_c^\infty(\Omega, \mathbb{R}^4) \mid \|\varphi(x)\|_2 \leq c\}$. Since $\{w_i\}_i$ is an orthonormal basis of \mathbb{R}^4 and the set \mathcal{C} in (12) includes $\mathcal{C}(c_3)$, the first inequality follows. Moreover, \mathcal{C} is included in $\mathcal{C}(\sqrt{2}c_4)$, providing the second inequality. \square

Lemma 2 (Weakly*-semi-continuity). *Let $u^k \in BV^2(\Omega)$ be weakly*-converging to u^* , i.e. $\|u^k - u^*\|_{L^1} \rightarrow 0$, $\|\partial_{x_i} u^k - \partial_{x_i} u^*\|_{L^1} \rightarrow 0$, $i = 1, 2$, and $\sup_k TV^2(u^k) < \infty$. Then, again under the assumptions of Prop. 1, we have*

$$\mathcal{R}_1(u^*) \leq \liminf_{k \rightarrow +\infty} \mathcal{R}_1(u^k) \quad \text{and} \quad \mathcal{R}_2(u^*) \leq \liminf_{k \rightarrow +\infty} \mathcal{R}_2(u^k). \quad (15)$$

Proof. Semi-continuity of $\mathcal{R}_1(u^k)$: note that $\mathcal{R}_1(u^k) \leq c_2\|\nabla u^k\|_{L^1} < \infty$. Since $\|\nabla u^k - \nabla u^*\|_{L^1} \rightarrow 0$, there exists a subsequence ∇u^{k_l} converging pointwise almost everywhere to ∇u^* . From the continuity of the mapping $v \mapsto (v^\top Av)^{\frac{1}{2}}$ it follows that $((\nabla u^{k_l})^\top A \nabla u^{k_l})^{\frac{1}{2}}(x) \rightarrow ((\nabla u^*)^\top A \nabla u^*)^{\frac{1}{2}}(x)$ almost everywhere. Since any converging subsequence of ∇u^k converges to ∇u^* (Lebesgue thm.), we find $\liminf_{k \rightarrow +\infty} ((\nabla u^k)^\top A \nabla u^k)^{\frac{1}{2}}(x) = ((\nabla u^*)^\top A \nabla u^*)^{\frac{1}{2}}(x)$ almost everywhere. The claim then follows from Fatou's Lemma.

Semi-continuity of $\mathcal{R}_2(u^k)$: For $\varphi \in \mathcal{C}$ we have

$$\int_\Omega (\operatorname{div}^2 \varphi)u^* \, dx = - \int_\Omega (\partial_x \varphi_1 + \partial_y \varphi_2) \partial_x u^* + (\partial_x \varphi_3 + \partial_y \varphi_4) \partial_y u^* \, dx \quad (16)$$

$$= - \lim_{k \rightarrow +\infty} \int_\Omega (\partial_x \varphi_1 + \partial_y \varphi_2) \partial_x u^k + (\partial_x \varphi_3 + \partial_y \varphi_4) \partial_y u^k \, dx \quad (17)$$

$$= \lim_{k \rightarrow +\infty} \int_\Omega (\operatorname{div}^2 \varphi)u^k \, dx \leq \liminf_{k \rightarrow +\infty} \mathcal{R}_2(u^k). \quad (18)$$

Thus

$$\mathcal{R}_2(u^*) = \sup \left\{ \int_\Omega (\operatorname{div}^2 \varphi)u^* \, dx \mid \varphi \in \mathcal{C} \right\} \leq \liminf_{k \rightarrow +\infty} \mathcal{R}_2(u^k). \quad (19)$$

\square

Proof (of Prop. 1).

Since $\mathcal{F}(u)$ is bounded from below, we have $F_{inf} := \inf_{u \in BV^2(\Omega)} \mathcal{F}(u) > -\infty$. We consider a minimizing sequence $\{u^k\}_k$, $\mathcal{F}(u^k) \rightarrow F_{inf}$. Due to Lemma 1, $\mathcal{F}(u)$ is finite on $BV^2(\Omega)$, thus $\sup_k \mathcal{F}(u^k) \leq C < +\infty$ for some $C > 0$. We show that $\{u^k\}_k$ is bounded in $L^2(\Omega) \cap BV^2(\Omega)$ due to coercivity of \mathcal{F} : from $C \geq \mathcal{F}(u^k) \geq \frac{1}{2}\|u^k - f\|_{L^2}^2$ it follows that $\{u^k\}_k$ is bounded in $\|\cdot\|_{L^2}$ and, since Ω is bounded, also in $\|\cdot\|_{L^1}$; $C \geq \mathcal{F}(u^k) \geq \mathcal{R}_l(u^k)$ and Lemma 1 provide that the minimizing sequence is bounded in $TV^l(\cdot)$, $l = 1, 2$. From boundedness follows by Theorem 9.83 in [16] that a weakly*-converging subsequence in $L^2(\Omega) \cap BV^2(\Omega)$ with some limit u^* exists. We denote the subsequence also by $\{u^k\}_k$. We have $\frac{1}{2}\|u^k - f\|_{L^2}^2 \rightarrow \frac{1}{2}\|u^* - f\|_{L^2}^2$, and due to Lemma 2, $\mathcal{R}_l(u^*) \leq \liminf_{k \rightarrow +\infty} \mathcal{R}_l(u^k)$, $l = 1, 2$. Thus $\mathcal{F}(u^*) \leq \liminf_{k \rightarrow +\infty} \mathcal{F}(u^k) = F_{inf}$, i.e. u^* is a minimizer of $\mathcal{F}(u)$. Uniqueness follows from the strict convexity of $\mathcal{F}(u)$. \square

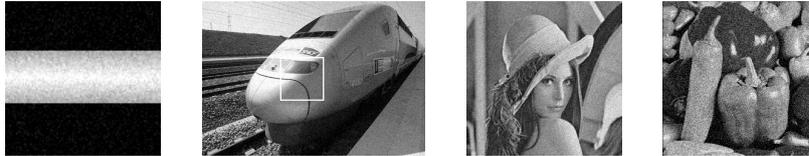


Fig. 4. Test images used in the comparison in Table 1.

Choice of Regularization Parameters

It remains to choose appropriate regularization parameters $\alpha_i(x), \beta_i(x)$, $i = 1, 2$, and directions $v(x)$. For the vector field $v(x)$ we choose $v^S(x)$ as defined in Sect. 2.1. Recall that $v^S(x)$ provides a smoothed version of the image gradient, which at edges coincides with the edge normals. To avoid a loss of contrast at edges and over-smoothing at slope discontinuities, we reduce α_i and β_i at edges and slope discontinuities using the indicator function $E(x) := \max(E^S(x), E^T(x))$ based on structure tensors S_f , and T_f applied to data f . We propose

$$\begin{aligned} \alpha_1(x) &:= E(x)\underline{\alpha} + (1 - E(x))\bar{\alpha}, & \alpha_2(x) &:= \bar{\alpha}, \\ \beta_1(x) &:= E(x)\underline{\beta} + (1 - E(x))\bar{\beta}, & \beta_2(x) &:= \bar{\beta}, \end{aligned} \quad (20)$$

with four free parameters $\underline{\alpha}, \bar{\alpha}, \underline{\beta}, \bar{\beta} > 0$ to be chosen appropriately. We propose a weak smoothing at edges and slope discontinuities with small $\underline{\alpha}, \underline{\beta}$. These parameters can be chosen fairly independent from the image content or noise level. Similar to other second-order TV approaches, it remains to choose two appropriate values for $\bar{\alpha}, \bar{\beta}$ depending mainly on the noise level of the image.

3.2 Remarks on Alternative Approaches

Regarding second-order approaches based on total generalized variation (TGV) [2] and infimal convolution (IC) [17], which both require two regularization parameters α, β , we observe that both approaches can be extended to be spatially adaptive by locally varying these parameters. We propose to choose

$$\alpha(x) := E(x)\underline{\alpha} + (1 - E(x))\bar{\alpha}, \quad \beta(x) := E(x)\underline{\beta} + (1 - E(x))\bar{\beta}, \quad (21)$$

with suitable $\underline{\alpha}, \bar{\alpha}, \underline{\beta}, \bar{\beta}$ and $E : \Omega \rightarrow [0, 1]$ as defined in Sect. 3.1.

4 Experiments

In this section we perform a quantitative comparison of the total generalized variation (TGV) approach, infimal convolution (IC), their adaptive counterparts as proposed in Sect. 3.2, and the proposed anisotropic second-order TV model (Sect. 3.1)³. For TGV and IC we use the original codes, which were kindly provided by the authors of [2, 17]. In addition, we consider an anisotropic second-order TV, where the adaptivity is determined only by the standard structure

³ Computational speed: 17 sec for a MATLAB implementation on an 256x256 image using an Intel i7 processor.

Example	Roof	Train (part)	Lena	Peppers
2nd order TV with std. struct. tensor	4.6946e-4	2.4249e-4	1.0703e-4	1.8858e-4
TGV	0.8857e-4	2.3644e-4	0.9362e-4	1.3883e-4
Adaptive TGV	0.8703e-4	2.3364e-4	0.8985e-4	1.3258e-4
IC	1.0405e-4	2.3968e-4	0.9519e-4	1.3822e-4
Adaptive IC	0.9861e-4	2.3693e-4	0.9205e-4	1.3589e-4
Proposed method	0.5703e-4	2.2560e-4	0.8749e-4	1.2997e-4

Table 1. Mean squared errors (MSE) to the noise-free image for the different methods. For each method, the approximate optimal parameters were retrieved by grid search. Independent of the model, introducing adaptivity always improves the error. The results of the proposed *anisotropic* method show the lowest reconstruction error.

tensor, i.e. $E(x) = E^S(x)$. As test images we use the image *roof*, cf. Fig. 2, left, a part of the *train* image from [2] and the *Lena* and *peppers* image, adding 5% zero mean Gaussian noise, cf. Fig. 4. For each image, the approximate optimal parameters for each method $(\underline{\alpha}, \underline{\beta}, \bar{\alpha}, \bar{\beta})$ were determined via a grid search minimizing the mean squared error (MSE) to the noise-free image. While other error norms are also applicable, we have chosen the MSE as it is the most commonly used. Table 1 shows the errors for each method. We observe that by introducing *adaptivity*, we are able to decrease the error compared to the non-adaptive methods. The proposed method achieves the smallest error across all instances, showing the advantage of introducing *anisotropic* regularization. Moreover, it becomes clear that using solely the standard structure tensor to steer the anisotropy does not suffice, justifying our approach of also taking slope discontinuities into account. Figs. 5 and 6 depict the results of the methods on the *roof* and *train* images. For the latter, we observe that the results still contain some amount of the original noise. It seems that minimizing the MSE by grid search favors such residual noise rather than to strongly smooth the results. Since human users generally prefer a stronger smoothing, we provide in Fig. 7 results with manually adapted parameters for TGV and the proposed methods (due to space constraints, we omit IC here). The increased smoothing removes some image structures, as can be seen in the difference images. The proposed method preserves edges better than the competitive approaches.

5 Conclusion & Future Work

We proposed a way to modify state-of-the-art second-order TV models by introducing spatial adaptivity. Moreover, we introduced a new anisotropic second-order TV model. Experiments show that the modifications lead to an improved reconstruction performance. Since all considered methods exhibit over-smoothing in textured regions, future work will focus on how adaptive approaches can be improved by including texture information. New insight into this problem could also possibly close the conceptual gap to non-local regularization approaches.

Acknowledgements We thank Tanja Teuber and Kristian Bredies for kindly providing their codes. The work of J.L. was supported by Award No. KUK-I1-007-43, made by King Abdullah University of Science and Technology (KAUST), EPSRC first grant EP/J009539/1, and EPSRC/Isaac Newton Trust Small Grant.

References

1. B. Berkels, M. Burger, M. Droske, O. Nemitz, and M. Rumpf. Cartoon extraction based on anisotropic image classification. In *Vision, Modeling, and Visualization Proceedings*, pages 293–300, 2006.
2. K. Bredies, K. Kunisch, and T. Pock. Total Generalized Variation. *SIAM J. Imaging Sciences*, 3(3):492–526, 2010.
3. Q. Chen, P. Montesinos, Q. S. Sun, P. A. Heng, and D. S. Xia. Adaptive total variation denoising based on difference curvature. *Image Vision Comput.*, 28(3):298–306, 2010.
4. Y. Dong and M. Hintermüller. Multi-scale vectorial total variation with automated regularization parameter selection for color image restoration. In *proceedings SSVM '09*, pages 271–281, 2009.
5. W. Förstner and E. Gülch. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In *Proc. ISPRS conf. on fast processing of photogrammetric data*, pages 281–305, 1987.
6. K. Frick, P. Marnitz, and A. Munk. Statistical multiresolution estimation for variational imaging: With an application in Poisson-biophotonics. *J. Math. Imaging Vis.*, pages 1–18, 2012.
7. M. Grasmair. Locally adaptive total variation regularization. In *Proceedings SSVM '09*, volume 5567, pages 331–342, 2009.
8. M. Grasmair and F. Lenzen. Anisotropic Total Variation Filtering. *Appl. Math. Optim.*, 62(3):323–339, 2010.
9. Y. Hu and M. Jacob. Higher degree total variation (HDTV) regularization for image recovery. *IEEE Trans. Image Processing*, 21:2559–2571, 2012.
10. S. Kindermann, S. Osher, and P.W. Jones. Deblurring and denoising of images by nonlocal functionals. *Multiscale Model. Simul.*, 4(4):1091–1115 (electronic), 2005.
11. S. Lefkimmiatis, A. Bourquard, and M. Unser. Hessian-based norm regularization for image restoration with biomedical applications. *Image Processing, IEEE Transactions on*, 21(3):983–995, 2012.
12. F. Lenzen, F. Becker, J. Lellmann, S. Petra, and C. Schnörr. A class of quasi-variational inequalities for adaptive image denoising and decomposition. *Comput. Optim. Appl.*, 2012. online first.
13. M. Lysaker and X.-C. Tai. Iterative image restoration combining total variation minimization and a second-order functional. *Int. J. Comp. Vis.*, 66:5–18, 2006.
14. L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1–4):259–268, 1992.
15. O. Scherzer. Denoising with higher order derivatives of bounded variation and an application to parameter estimation. *Computing*, 60:1–27, 1998.
16. O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational Methods in Imaging*. Springer, 2009.
17. S. Setzer, G. Steidl, and T. Teuber. Infimal convolution regularizations with discrete l1-type functionals. *Comm. Math. Sci.*, 9:797–872, 2011.
18. G. Steidl and T. Teuber. Anisotropic smoothing using double orientations. In *Proceedings SSVM '09*, pages 477–489, 2009.

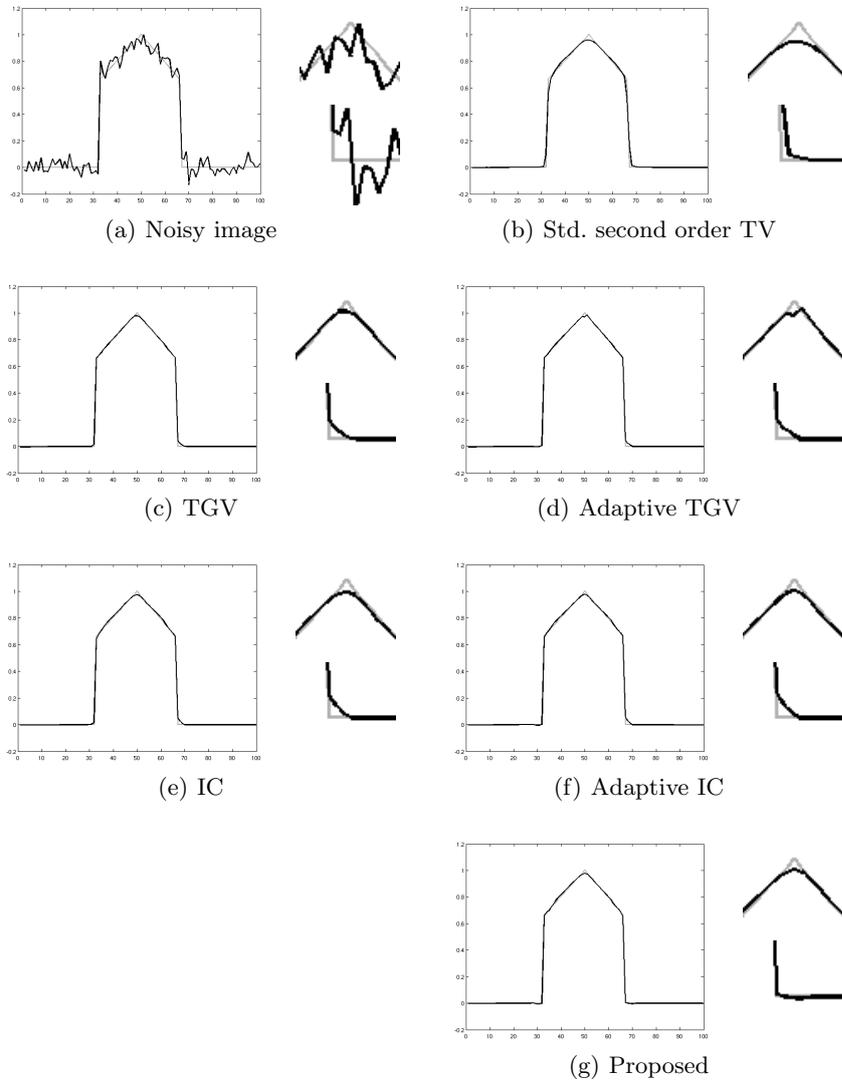


Fig. 5. Cross-section of the results (black lines) of TGV, IC, their adaptive variants and the proposed method on the *roof* image and detailed views of the peak and the left step. The noise-free data is shown in gray. We remark that standard second-order TV (b), cf. (6), significantly flattens the peak. All considered approaches avoid such a flattening to varying degrees. The TGV variants provide the sharpest reconstruction of the peak. The proposed approach provides a sharp reconstruction of both kinks.

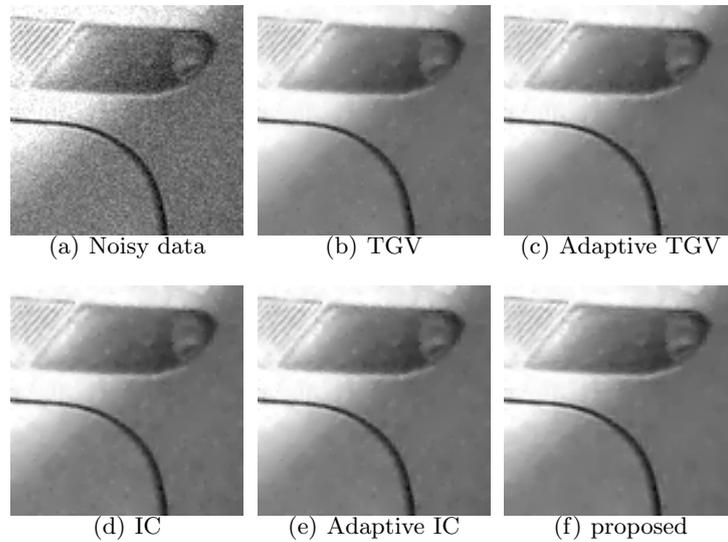


Fig. 6. Results of the tested methods on the *train* image. For each method the parameters were selected by a grid search minimizing the mean squared error (MSE). As a consequence, all methods preserve some noise. Visually, the results are very similar.

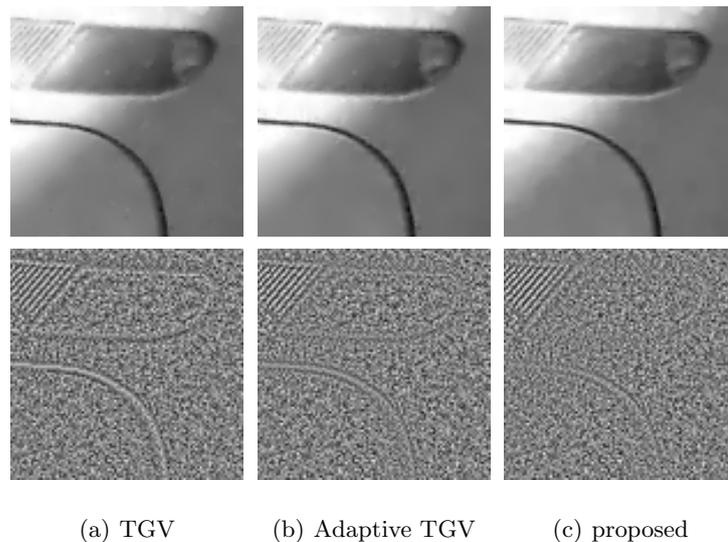


Fig. 7. Denoising results using manually chosen parameters (top row) and difference image to noisy data (bottom row). TGV shows a strong smoothing effect, with the drawback, that also edges become smoother. Adaptive methods preserve edge structures better, as can be seen from the weaker edges in the difference images. In textured regions, all methods partly remove the texture. The proposed method shows the smallest amount of structures in the difference image.