Variational Image Denoising with Adaptive Constraint Sets

Frank Lenzen[1,2], Florian Becker[1], Jan Lellmann[1], Stefania Petra[1], and Christoph Schnörr[1]

[1] HCI & IPA, Heidelberg University,
Speyerer Str. 6, 69115 Heidelberg, Germany
frank.lenzen@iwr.uni-heidelberg.de
{becker,lellmann,petra,schnoerr}@math.uni-heidelberg.de
http://hci.iwr.uni-heidelberg.de,
http://ipa.iwr.uni-heidelberg.de
[2] Intel Visual Computing Institute, Saarland University,
Campus E2 1, 66123 Saarbrücken, Germany
http://www.intel-vci.uni-saarland.de

Abstract. We propose a generalization of the total variation (TV) minimization method proposed by Rudin, Osher and Fatemi. This generalization allows for adaptive regularization, which depends on the minimizer itself. Existence theory is provided in the framework of quasi-variational inequalities. We demonstrate the usability of our approach by considering applications for image and movie denoising.

Keywords: solution dependent adaptivity, quasi-variational inequalities, spatio-temporal TV, anisotropic TV, image denoising

1 Introduction

One of the most widely used methods for image denoising is total variation (TV) minimization. The TV method proposed by Rudin, Osher and Fatemi in [13] (ROF) consists in minimizing the functional

$$\frac{1}{2}\|u-f\|^2 + \alpha\,\mathrm{TV}(u), \tag{1}$$

w.r.t. u over $\mathrm{BV}(\Omega), \Omega \subset \mathbb{R}^d$ for given noisy data f. Here $\mathrm{TV}(u)$ is the total variation semi-norm and α is a regularization parameter. We consider the formulation of the regularization term $\alpha\,\mathrm{TV}(u)$ based on constraint sets:

$$\alpha\,\mathrm{TV}(u) = \sigma_{\mathcal{C}}(u), \quad \mathcal{C} = \operatorname{div}\mathcal{D}, \quad \mathcal{D} = \{p \in C_c^\infty(\Omega;\mathbb{R}^d)\colon \|p(x)\| \le \alpha\}, \tag{2}$$

where $\sigma_{\mathcal{C}}$ is the support function of the set $\mathcal{C}$ and div is applied elementwise.

In this paper, we generalize the ROF functional (1) by introducing the dependency $\mathcal{C} = \mathcal{C}(u)$. This allows for variants of the TV method, where the set $\mathcal{C}(u)$ locally adapts to the image content *depending on the solution u itself.*

In the literature, *adaptive* TV methods have been proposed e.g. in [6, 9], with locally varying regularization parameter, and [1, 14, 15, 10], where anisotropic regularization is steered by local structures. Except for [10], these methods gather the required local information either in a preprocessing or as an additional unknown of the variational problem, not depending on the minimizer itself. The variational framework presented below differs from [10]; possible connections will be explored in future work. Another kind of denoising methods are non-local methods, cf. e.g. [11, 8, 3]. Although these methods can be applied in an iterated fashion, a dependency of the regularization on the minimizer is not modeled explicitly.

Our paper is organized as follows. In Sect. 2, the proposed generalization of the TV minimization functional (1) is described. The mathematical framework is presented in terms of variational inequalities. Considering a sequence of convex variational inequalities, we provide an existence result for fixed points (see Sect. 3), using only general assumptions on the convex set $\mathcal{C}(u)$. In particular, non-local information can be used in the definition of $\mathcal{C}(u)$. Moreover, we provide a first basic algorithm. The usability of this novel concept is supported by applications for image and movie denoising. In particular, we generalize the approach of anisotropic TV with double orientations, proposed by Steidl & Teuber [15] (Sect. 4), and present an anisotropic spatio-temporal TV method for denoising image sequences (Sect. 5). In Sect. 6 we provide experimental results. Concluding remarks are given in Sect. 7.

2 Problem

We begin with the primal TV denoising approach (1). Inserting (2) in (1) yields

$$\min_{u \in \mathrm{BV}(\Omega)} \left\{ \frac{1}{2}\|u - f\|^2 + \sigma_{\mathcal{C}}(u) \right\}. \tag{3}$$

We follow [4] to derive the corresponding dual problem. With the fact that $v \in \partial\sigma_{\mathcal{C}}(u) \Leftrightarrow u \in \partial(\sigma_{\mathcal{C}})^*(v) = \partial(\delta_{\mathcal{C}})(v)$ for the subdifferentials of the support function $\sigma_{\mathcal{C}}$ and the indicator function $\delta_{\mathcal{C}}$, where * denotes the Legendre-Fenchel transform, we find $\partial\sigma_{\mathcal{C}}(\overline{u}) = \{v \in L^2(\Omega) \colon \langle \overline{u}, v - u\rangle \geq 0,\ \forall u \in \mathcal{C}\}$. Thus the optimality condition for $\overline{u}$ minimizing (3) reads

$$f - \overline{u} \in \partial\sigma_{\mathcal{C}}(\overline{u}) \quad \Leftrightarrow \quad \langle \overline{u}, f - \overline{u} - u\rangle \geq 0, \quad \forall u \in \mathcal{C}.$$

Using the additive decomposition $f = \overline{u} + \overline{v}$, we find $\overline{v} = \mathit{\Pi}_{\overline{\mathcal{C}}}(f)$, where $\mathit{\Pi}_{\overline{\mathcal{C}}}$ denotes the projection onto the closure $\overline{\mathcal{C}}$ of $\mathcal{C}$. Finally, we end up with the dual problem

$$\inf_{p \in \mathcal{D}} F(p), \qquad F(p) := \frac{1}{2}\|f - \operatorname{div} p\|^2. \tag{4}$$

In the following, we study generalized adaptive denoising approaches that take into account dependencies of the primal and dual constraint sets $\mathcal{C}(u)$ and $\mathcal{D}(p)$, respectively, on the solutions themselves. To this end, let

$$\mathcal{C}(u) := \operatorname{div}\{p \in C_c^\infty(\Omega, \mathbb{R}^d) : p(x) \in \tilde{D}(x, u)\}, \tag{5}$$

where $\tilde{D}(x,u) : \Omega \times \mathrm{BV}(\Omega) \rightrightarrows \mathbb{R}^d$ is a set-valued mapping. In view of the dual problem (4), we define $D(x,p) := \tilde{D}(x, f - \operatorname{div} p)$ and

$$\mathcal{D}(p) := \{\tilde{p} \in C_c^\infty(\Omega, \mathbb{R}^d) : \tilde{p}(x) \in D(x,p)\}. \tag{6}$$

3 Approach

3.1 A Quasi-Variational Inequality

We approximate the space $C_c^\infty(\Omega, \mathbb{R}^d)$ by a finite-dimensional space $\mathbb{R}^{nd}$. Then the set-valued mapping $\mathcal{D}$ in (4) takes the form

$$\mathcal{D} : \mathbb{R}^{nd} \rightrightarrows \mathbb{R}^{nd}, \qquad p \mapsto \{\tilde{p} \in \mathbb{R}^{nd} : \tilde{p}_i \in D_i(p), i = 1, \dots, n\}. \tag{7}$$

Here, $D_i(p) : \mathbb{R}^d \rightrightarrows \mathbb{R}^d, i = 1, \dots, n$ is the discrete analogue of $D(x,p)$. Note that in the finite dimensional setting $\mathcal{D}(p)$ is compact. In order to show existence of a solution, if the constraint sets (5) and (6) vary, in analogy to [2, Prop. 4.7.1], we formulate our approach as a generalization of the variational inequality corresponding to the dual problem (4): find $\overline{p} \in \mathcal{D}(\overline{p})$ such that

$$\langle \nabla F(\overline{p}), p - \overline{p} \rangle \geq 0, \qquad \forall p \in \mathcal{D}(\overline{p}). \tag{8}$$

Notice the dependency of the dual constraint set on $\overline{p}$, that significantly generalizes the dual TV minimization problem.

3.2 Existence of Solutions

The existence of a solution to (8) can be shown under the following assumption:

Assumption 1. *$D_i(p) : \mathbb{R}^d \rightrightarrows \mathbb{R}^d, i = 1, \dots, n$ have the following properties:*

1. *For fixed p the set $D_i(p)$ is a closed convex subset of $\mathbb{R}^d$.*
2. *There exists $c > 0$, such that for all i, p: $\{0\} \subset D_i(p) \subset \overline{B_c(0)}$, where $\overline{B_c(0)}$ is the closed unit ball. In particular, $D_i(p)$ is non-empty.*
3. *The projection $\Pi_{D_i(p)}(q)$ of q onto $D_i(p)$ for a fixed q is continuous w.r.t. p.*

Proposition 1. *Let $F := \frac{1}{2}\|f - \operatorname{div} p\|^2$ and $\mathcal{D}$ be defined as in (7), such that $D_i(p), i = 1, \dots, n$ satisfy Assumption 1. Then the problem*

$$\textit{find } \overline{p} \in \mathbb{R}^{nd} \textit{ such that} \qquad \langle \nabla F(\overline{p}), p - \overline{p} \rangle \geq 0, \quad \forall p \in \mathcal{D}(\overline{p}) \tag{9}$$

has a solution.

The proof of Proposition 1 utilizes the following theorem and lemma.

Theorem 2. *(cf. Theorem 5.2 in [5]) Let $G\colon \mathbb{R}^m \to \mathbb{R}^m$ be a point-valued and $\mathcal{D}\colon \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ be a set-valued mapping. Suppose that there exists a nonempty compact convex set P such that*

1. $\mathcal{D}(P) = \cup_{p\in P}\mathcal{D}(p) \subseteq P$;
2. $\mathcal{D}$ *takes nonempty closed convex sets as values;*
3. $\mathcal{D}$ *is continuous, that is* $\mathcal{D}(p^k) \to \mathcal{D}(\overline{p})$ *whenever* $p^k \to \overline{p}$*, or in view of (2), denoting the projection onto* $\mathcal{D}$ *by* $\Pi_{\mathcal{D}}(p)$*, equivalently* $\Pi_{\mathcal{D}(p^k)}(p) \to \Pi_{\mathcal{D}(\overline{p})}(p)$ *for all* p

Then there exists $\overline{p} \in \mathbb{R}^{nd}$ *such that* $\langle G(\overline{p}), p - \overline{p}\rangle \geq 0$ *for all* $p \in \mathcal{D}(\overline{p})$.

Lemma 1. *Let* $\mathcal{D}$ *be defined as in* (7)*. Assume that for every* $i = 1, \dots, n$ *and* $q \in \mathbb{R}^d$ *the projection* $\Pi_{D_i(p)}(q)$ *is continuous w.r.t.* p*. Then* $\Pi_{\mathcal{D}(p)}(q)$ *is continuous for fixed* $q \in \mathbb{R}^{nd}$*.*

Proof. $\Pi_{\mathcal{D}}(q)$ can be written as $\Pi_{\mathcal{D}(p)}(q) = (\Pi_{D_1(p)}(q_1), \dots, \Pi_{D_n(p)}(q_n))^\top$. Thus each component of $\Pi_{\mathcal{D}(p)}$ is continuous, from which the continuity of $\Pi_{\mathcal{D}}(p)$ follows immediately. □

Proof of Proposition 1: We apply Thm. 2. Conditions (1) and (2) follow from Assumption 1, that, in turn, has to be verified later, see Prop. 2, 3 and 4. Lemma 1 shows that also condition (3) holds. □

3.3 Algorithm

We propose an algorithm for solving (9). Let us first consider the case where $\mathcal{D}$ does not dependend on the dual variable p. The problem then can be solved by a projected gradient method:

$$p^{k+1} = \Pi_{\mathcal{D}}\big(p^k - \tau\nabla F(p^k)\big), \qquad 0 < \tau < 2/L,$$

where L denotes the Lipschitz-constant of ∇F. In order to adapt to the dependency of $\mathcal{D}$ on p, we propose to use

$$p^{k+1} = p^k - \frac{1}{\lambda}\Big(p^k - \Pi_{\mathcal{D}^k}\big(p^k - \tau\nabla F(p^k)\big)\Big), \quad \mathcal{D}^k := \mathcal{D}(p^k),$$

with sufficiently large $\lambda \in (0, 1)$. In practice, two nested iterations, one outer iteration for updating $\mathcal{D}$ and one inner iteration for updating p, are used. Providing convergence results will be part of our future work. However, our experiments show that this iteration converges for λ sufficiently large.

4 Adaptive Anisotropic TV Minimization for Image Denoising

In order to improve the image quality of TV methods for denoising, Steidl & Teuber [15], proposed an anisotropic TV method based on two independent orientations. In this section, we demonstrate how this approach can be modified in order to fit into the ansatz presented above. As a consequence, Prop. 1 provides a theoretical underpinning. Before discussing the approach in [15] (Sect. 4.2), we describe the required modifications by means of a simpler model (Sect. 4.1).

4.1 Anisotropic TV with a Single Direction

Consider $u \in L^2(\Omega), \Omega \subset \mathbb{R}^2$. The aim is to define a convex set $\mathcal{D}(p)$ (cf. (6)) satisfying Assumption 1 in order to derive an anisotropic TV measure.

In our approach, we are interested in a *local* description of the set $\mathcal{D}$. To this end, we define $D(x,p)$ for any $x \in \Omega$, based on local edge information obtained from $u = f - \operatorname{div} p$. To be precise, for an edge being present at a location x the set $D(x,p)$ will be defined as a square with one side parallel to the edge.

In order to detect edges, we utilize the structure tensor $J(x,u)$ defined as follows: Let

$$J_0(x,u) := \nabla u_\sigma(x) \nabla u_\sigma(x)^\top, \tag{10}$$

where $u_\sigma := K_\sigma * u$ is a smoothed version of u, obtained by convolution with a Gaussian kernel K_σ with standard deviation $\sigma > 0$. The structure tensor $J(x,u)$ is given as

$$J(x,u) := K_\rho * J_0(x,u), \tag{11}$$

with $\rho > 0$. (Here the convolution is applied componentwise). Moreover, let $v_i(x,u)$ and $\lambda_i(x,u), i = 1,2$ be the eigenvectors and eigenvalues of J, respectively. We assume w.l.o.g. that the eigenvalues of $J(x,u)$ are ordered, $\lambda_1(x,u) \geq \lambda_2(x,u) \geq 0$, with corresponding eigenvectors $v_1(x,u)$ and $v_2(x,u)$. For simplicity of notation, we omit the dependency of D, J, v_i and λ_i on x in the following.

Consider for a moment some arbitrary $r \in \mathbb{R}^2, \|r\| = 1$. We define the square $\mathcal{S}(r)$ with sides parallel to r and $r^\perp$ as

$$\mathcal{S}(r) := \{p \in \mathbb{R}^2 : |r^\top p| \leq \alpha, |(r^\perp)^\top p| \leq \alpha\}. \tag{12}$$

We would like to set $D(p) = \mathcal{S}(r(f - \operatorname{div} p))$ with $r(u) = v_1(u)$. But then the projection $\Pi_{D(p)}$ would not depend continuously on p, since the eigenvector $v_1(u)$ in general does not depend continuously on the entries of $J(u)$.

On the other hand, for $\mathcal{S}(r)$ as defined in (12), the mapping $r \to \Pi_{\mathcal{S}(r)}$ is continuous, as the following lemma shows. Moreover, $u = f - \operatorname{div} p$ depends continuously on p. Thus, asserting the continuity of $r(u)$ is sufficient to guarantee the continuity of $\Pi_{D(p)}$.

Lemma 2. *Let $\mathcal{S}(r)$ be defined as in* (12)*. Then $\Pi_{\mathcal{S}(r)}(q)$ depends continuously on r for fixed but arbitrary q.*

Proof. For $q \in \mathcal{S}(r)$, we have $\Pi_{\mathcal{S}(r)}(q) = q$. For $q \notin \mathcal{S}(r)$ the projection onto $\mathcal{S}(r)$ can be calculated as follows: Let $j^* := \operatorname{argmin}_{j=1,\dots,4} \|q - \Pi_j(q)\|$, where Π_j is the projection on the j-th side of the square. Then $\Pi_{\mathcal{S}(r)}(q) = \Pi_{j^*}(q)$.

Each of the projections Π_j is a composition of the orthogonal projection onto a line and a projection from the line onto a line segment. Only the projection onto the line depends on the parameter r. Since the orthogonal projection Π onto a line $\{a + tb \mid t \in \mathbb{R}\}, \|b\| = 1$, which is given by $\Pi(q) = a + \langle q - a, b\rangle b$, depends continuously on a, b, the continuity of $\Pi_{\mathcal{S}(r)}(q)$ w.r.t. r follows. Obviously, the transition between the cases $q \in \mathcal{S}(r)$ and $q \notin \mathcal{S}(r)$ is continuous. □

In the following, we describe the construction of a vector $r(u)$ depending continuously on u, such that $r(u) = v_1(u)$, if $\lambda_1(u) \gg \lambda_2(u)$.

Note that the eigenvectors of $J(u) \in \mathbb{R}^{2\times 2}$ depend continuously on u, as long as the eigenvalues $\lambda_1(u)$ and $\lambda_2(u)$ differ (cf. Theorem 3 in [12]). We define

$$\mathrm{coh}(u) := \lambda_1(u) - \lambda_2(u) \geq 0.$$

Note that $\mathrm{coh}(u)$ depends continuously on u, since the eigenvalues depend continuously on $J(u)$ (cf. e.g. Theorem of Wielandt-Hoffman in [16]), and $J(u)$, which is a composition of convolution and differentiation, is continuous w.r.t. u.

Now let $g : \mathbb{R}_0^+ \to [0,1]$ be a continuous and increasing function, such that $g(0) = 0$ and $\lim_{x\to\infty} g(x) = 1$. Moreover, let $I(p,q,t) : S^1 \times S^1 \times [0,1] \to S^1$ be a continuous interpolation from p to q on the unit sphere S^1 with the properties, that $I(p,q,1) = p$, $I(p,q,0) = q$ and $\|I(p,q,t) - q\| \leq C\|t\|$ for some $C > 0$. (For example, a steady rotation of vector p onto q suffices.) We set

$$D(p) := \mathcal{S}(r(f - \mathrm{div}\, p)), \tag{13}$$

where $r(u) := I(v_1(u), (1,0)^\top, g(\mathrm{coh}(u)))$. The set $D(p)$ satisfies Assumption 1:

Proposition 2. *Let $D(p)$ be defined as in* (13)*. Then*

1. *$D(p)$ is closed, convex and satisfies $\{0\} \subset D(p) \subset \overline{B_{\sqrt{2}\alpha}(0)}$.*
2. *For fixed $q \in \mathbb{R}^2$, $u \to \Pi_{D(p)}(q)$ is continuous.*

The proof of Prop. (2) utilizes the following lemma:

Lemma 3. *Let q be fixed. Then $r(u) = I(v_1(u), q, g(\mathrm{coh}(u)))$ depends continuously on u.*

Proof. We distinguish between the cases $\mathrm{coh}(u) > 0$, and $\mathrm{coh}(u) = 0$. In the first case, $v_1(u)$ is an eigenvector to an isolated eigenvalue and thus depends continuously on $J(u)$, see [12]. Moreover $\mathrm{coh}(u)$ depends continuously on $J(u)$ (cf. [16]). Since $J(u)$ is a composition of convolutions and differentiation, it depends continuously on u; thus $\mathrm{coh}(u)$ and $v_1(u)$ are continuous. The continuity of $r(u)$ at $u, \mathrm{coh}(u) > 0$ then follows from the continuity of I and g.

In the second case, $\mathrm{coh}(u) = 0$, we find from the continuity of $\mathrm{coh}(u)$ that $\mathrm{coh}(u^k) \to 0$ for every sequence u^k converging to u. Then the continuity of $r(u)$ follows from

$$\begin{aligned}\|r(u^k) - r(u)\| &= \|I(u^k, q, g(\mathrm{coh}(u^k))) - I(u, q, g(\mathrm{coh}(u)))\| \\ &= \|I(u^k, q, g(\mathrm{coh}(u^k))) - q\| \leq Cg(\mathrm{coh}(u^k)) \to 0,\end{aligned}$$

using the properties of the interpolation I. □

Proof of Prop. 2: (i) The set $D(p)$ is a closed square with center 0 and sides of length $2\alpha > 0$. (ii) Lemma 3 provides the continuity of $r(u)$. Moreover, $u = f - \mathrm{div}\, p$ depends continuously on p. Together with Lemma 2 the continuity of $D(p)$ follows. □

4.2 Anisotropic TV with Double Directions

Steidl & Teuber [15] proposed an anisotropic TV method based on the estimation of two orientations $r_1, r_2 : \Omega \to \mathbb{R}^2$. They consider the variational problem:

$$\min_u \frac{1}{2}\|u - f\|^2 + \alpha \left(|r_1^\top \nabla u| + |r_2^\top \nabla u| \right), \tag{14}$$

where two models for obtaining r from the data f are proposed. As an alternative to (14), they propose to use infimal convolution. In the dual formulation of (14) the set $D = D(f)$ is a parallelogram with sides $r_i, i = 1, 2$:

$$\mathcal{P}(r_1, r_2) := \{p \in \mathbb{R}^2 : |r_1^\top p| \le \alpha,\ |r_2^\top p| \le \alpha\}. \tag{15}$$

In our considerations, we concentrate on the 'occlusion model' described in [15]. Moreover, we consider r_i depending on the unknown $u := f - \operatorname{div} p$ and, by introducing slight changes of the original approach, guarantee the applicability of the theoretical results of Sect. 3. The orientations r_i are obtained as follows.

Let $\nu(u) := \left((\partial_x u_\sigma)^2, \partial_x u_\sigma \partial_y u_\sigma, (\partial_y u_\sigma)^2\right)^\top$, where u_σ is defined as in Sect. 4.1. For the occlusion model, the following structure tensor is utilized:

$$J_0(u) := \nu(u)\nu^\top(u), \qquad J(u) := K_\rho * J_0(u),$$

where the convolution is applied componentwise.

Now let $\lambda_1(u) \ge \lambda_2(u) \ge \lambda_3(u) \ge 0$ denote the eigenvalues of $J(u)$, and $v_1(u)$, $v_2(u)$ and $v_3(u)$ the corresponding eigenvectors.

Analogously to the previous section, in view of the continuity of $v_i(u)$ we have to deal with non-isolated eigenvalues. To this end, we define

$$\mathrm{coh}_1(u) := \lambda_1(u) - \lambda_2(u), \quad \mathrm{coh}_2(u) := \lambda_2(u) - \lambda_3(u).$$

In order to define $r_1(u), r_2(u)$, we consider the following cases:
Case 1 & 2 – corners ($\mathrm{coh}_2(u) > 0$): Steidl & Teuber distinguish between the cases $v_{3,1} \neq 0$ and $v_{3,1} = 0$ ($v_{3,1}$ being the first entry of v_3). In the case $v_{3,1} \neq 0$, they propose to use the unit vectors $r_1^1(u) \parallel (v_{3,1}(u), y_1(u))^\top$ and $r_2^1(u) \parallel (v_{3,1}(u), y_2(u))^\top$, where $y_1(u), y_2(u)$ are the solutions of the quadratic equation $y^2 + v_{3,2}(u)\, y + v_{3,1}(u)\, v_{3,3}(u) = 0$. Otherwise, the unit vectors $r_1^2(u) \parallel (v_{3,2}(u), v_{3,3}(u))^\top$ and $r_2^2(u) \parallel (-v_{3,3}(u), v_{3,2}(u))^\top$ can be used.
Case 3 – edges ($\mathrm{coh}_2(u) \approx 0$, $\mathrm{coh}_1(u) > 0$): Since we can only guarantee that eigenvalue $\lambda_1(u)$ is isolated, we determine r_1, r_2 depending on the eigenvector $v_1(u)$. Along straight edges, the eigenvector v_1 is parallel to the normal of the edge. Therefore v_1 and $v_1^\perp$ are suitable for defining the orientation for anisotropic TV at edges. We set $r_1^3(u) \parallel (v_{1,1}(u), v_{1,2}(u))^\top$ and $r_2^3(u) \parallel (-v_{1,2}(u), v_{1,1}(u))^\top$.
Case 4 – homogeneous regions ($\mathrm{coh}_1(u) \approx \mathrm{coh}_2(u) \approx 0$): We use the default orientations $r_1^4(u) := (1, 0)^\top$ and $r_2^4(u) := (0, 1)^\top$.

In general, $r_1(u), r_2(u)$ have to be continuous interpolations between the above cases. For $i = 1, 2$ let

$$r_i(u) = I\Big(I\left(r_i^1(u), r_i^2(u), g(|v_{3,1}(u)|)\right), I\left(r_i^3(u), r_i^4(u), g(\mathrm{coh}_1(u))\right), g(\mathrm{coh}_2(u))\Big) \tag{16}$$

using g and I as defined in the previous section.

Proposition 3. *Let $D(p) = \mathcal{P}(r_1(f - \operatorname{div} p), r_2(f - \operatorname{div} p))$ with $\mathcal{P}(r_1, r_2)$ being the parallelogram defined in (15) and $r_i(u)$, $i = 1, 2$ defined as in (16).*

1. *$D(p)$ is closed, convex and satisfies $\{0\} \subset D(p) \subset \overline{B_{2\alpha}(0)}$.*
2. *$\Pi_{D(p)}(q)$ for fixed q depends continuously on p.*

In particular, $D(p)$ satisfies Assumption 1.

Proof. The first claim follows from the fact that $D(p)$ is a closed parallelogram with sides of length α. For the continuity of $\Pi_{D(p)}(q)$, we observe that the vectors $r_i^j(u), i = 1, 2,\ j = 1, \ldots, 4$ are defined in a way that they depend continuously on $u = f - \operatorname{div} p$. The continuity of $r_i(u), i = 1, 2$ is guaranteed by smooth interpolation (c.f. proof of Proposition 3). The proof of the continuity of $\mathcal{P}(r_1, r_2)$ is analogous to the proof of Lemma 2. □

5 Anisotropic Spatio-Temporal TV Minimization

In the following we describe a spatio-temporal TV minimization approach. We interpret time as third coordinate, thus $u, f : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$.

To obtain directional information, we utilize the three-dimensional structure tensor $J_\rho(u)$ defined analogously to (10) and (11). Let $\lambda_1(u) \geq \lambda_2(u) \geq \lambda_3(u) > 0$ denote the eigenvalues and $v_1(u), v_2(u), v_3(u)$ the eigenvectors of $J_\rho(u)$.

Let us assume that a two-dimensional surface is present in $u_\sigma(x)$. Then $\lambda_1(u) \gg \lambda_2(u)$ and $v_1(u)$ approximates the normal to this surface. The idea is to penalize variations mainly in directions tangential to the surface. To this end we set

$$D(p) := \mathcal{E}(v_1(f - \operatorname{div} p), \alpha, \beta),$$

where $\mathcal{E}(r, \alpha, \beta) := \{q \in \mathbb{R}^3 : |r^\top q|^2/\beta^2 + \|q - rr^\top q\|^2/\alpha^2 \leq 1\},\ 0 < \beta \ll \alpha$.

In homogeneous regions, where a unique orientation r can not be estimated, we choose $D(p) := B_\alpha(0)$. A continuous transition between both cases is obtained by defining

$$D(p) := \mathcal{E}(r(f - \operatorname{div} p), \tilde{\alpha}(f - \operatorname{div} p), \beta), \tag{17}$$

where

$$\begin{aligned} \operatorname{coh}_1(u) &:= \lambda_1(u) - \lambda_2(u) \gg 0, \\ r(u) &:= I\left(v_1(u), (0, 0, 1)^\top, g(\operatorname{coh}_1(u))\right), \\ \tilde{\alpha}(u) &:= g(\operatorname{coh}_1(u))\alpha + (1 - g(\operatorname{coh}_1(u)))\beta. \end{aligned}$$

In order to remove speckles and similar kinds of distortions, an adaptation of (17) is required. This is due to the fact that at speckles, $v_1(u)$ is in direction of $(0, 0, 1)^\top$. Using (17) with the above $\tilde{\alpha}$ then would lead to a penalization of ∇u mainly in spatial directions, which is not suitable for removing distortions of medium/large scale in spatial directions. Instead we propose to use (17) with

$$\tilde{\alpha}(u) = g(\operatorname{coh}_1(u))g(\phi(u))\alpha + (1 - g(\operatorname{coh}_1(u))g(\phi(u)))\beta, \tag{18}$$

where $\phi(u)$ is the angle between $v_1(u)$ and $(0, 0, 1)^\top$. The above modification leads to stronger smoothing of surfaces parallel to the x_1, x_2-axes.

Fig. 1. 2D anisotropic TV filtering of artificial test image. Left: noisy test images, middle: filtering with the standard ROF model, right: anisotropic filtering with double directions. All images are scaled with respect to the intensity range of the original test image. Undesirable smoothing effects are considerably reduced on the right.

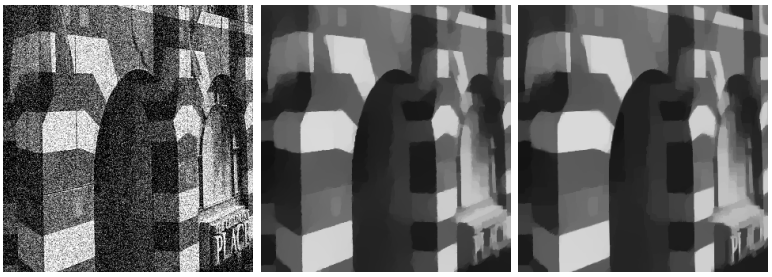

Fig. 2. 2D anisotropic filtering of real-world test image. Left: noisy test images, middle: result of standard ROF minimization, right: result of anisotropic TV minimization with double directions. Undesirable smoothing effects are considerably reduced on the right, see Fig. 3 for detailed views.

Proposition 4. *The set $D(p)$ defined in* (17) *with the above definitions of $\tilde{\alpha}(u)$ satisfies Assumption 1.*

Proof. The set $D(p)$ is a closed ellipsoid and therefore is convex. Its half-axes are bounded by $\max\{\alpha, \beta\}$, thus $0 \subset D(p) \subset B_{\max\{\alpha,\beta\}}(0)$. The projection onto the ellipsoid $\mathcal{E}(r, \tilde{\alpha}, \beta)$ can be expressed as a continuous function of r, $\tilde{\alpha}$, β and one distinct root of a rational function, see [7]. In a surrounding of this root, the function depends continuously on the half-axes. Thus the root depends continuously on r, $\tilde{\alpha}$ and β. r and $\tilde{\alpha}$ depend continuously on $u = f - \operatorname{div} p$, as $\mathrm{coh}_1(u)$ and $\phi(u)$ do. Moreover, u depends continuously on p. □

6 Experiments

6.1 Anisotropic TV Minimization with Double Directions

We present experimental results for the anisotropic TV model with $D(p)$ as defined in (15) and r_1, r_2 as defined in (16). We compare this method with standard ROF minimization, using the same regularization parameter α. We consider two different test images, both with artificial noise.

For the first test image (cf. Fig. 1, left) we use $\alpha = 0.6$ and 10 outer iteration steps. The results of the standard and anisotropic TV model are shown

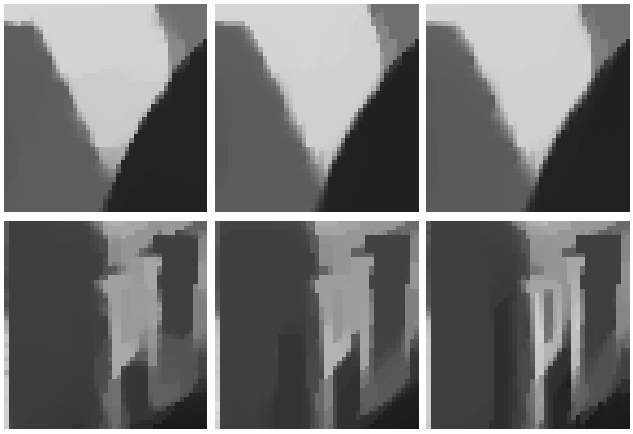

Fig. 3. Zoom into two regions of the filtered images shown in Fig. 2. Left: standard ROF, middle: anisotropic TV minimization with $D = D(f)$, right: adaptive anisotropic TV minimization with $D = D(p)$. It can be observed that adaptivity of the TV regularization improves with increasing number of iterations.

in Fig. 1, middle and right, respectively. A comparison shows, that anisotropic TV minimization better reconstructs corners of parallelogram and produces less smoothing at corners (as already demonstrated in [15]).

The second test image is a real world image with artificial noise, cf. Fig. 2, left. The result of standard ROF and anisotropic TV minimization for $\alpha = 0.4$ and 10 outer iteration steps is depicted in Fig. 2, middle and right, respectively.

In order to highlight differences, we zoom into two regions of the image: Fig. 3 shows the results for the standard ROF model (left), the result of applying anisotropic TV minimization with double directions, where the constraint set depends only on the data f, i.e. $D = D(f)$ (middle), and the result of anisotropic TV with the constraint set depending on the solution, $D = D(p)$ (right). It can be observed that anisotropic filtering leads to an improved and more regular reconstruction of edges and less stair-casing. If the constraint sets depend on the solution itself, an adaption to local structures can be observed during the iterations, see Fig. 3, bottom right. Here, the reconstruction of the characters improves when using fully adaptive constraint sets.

6.2 Adaptive Motion-Based TV Minimization for Image Sequences

In our example for spatio-temporal TV minimization, we consider an image sequence taken with a time-of-flight (ToF) camera, see Fig. 4 (4 frames out of the whole sequence). ToF cameras provide a depth map of the captured scene. The noise and speckles, which can be observed in the original data, are introduced by the camera system.

For filtering, we propose to use spatio-temporal anisotropic TV with $D(p)$ as defined in (17) and $\tilde{\alpha}$ defined as in (18). As parameters, we chose $\alpha = 0.3$,

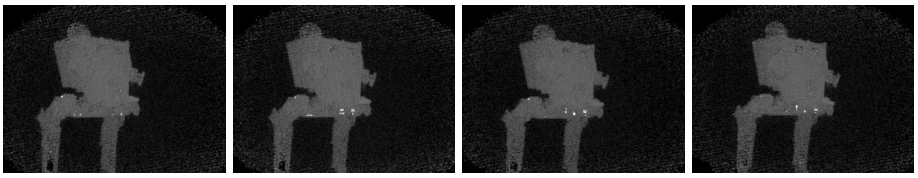

Fig. 4. Four exemplarily selected frames of a sequence of depth maps taken with a time-of-flight camera.

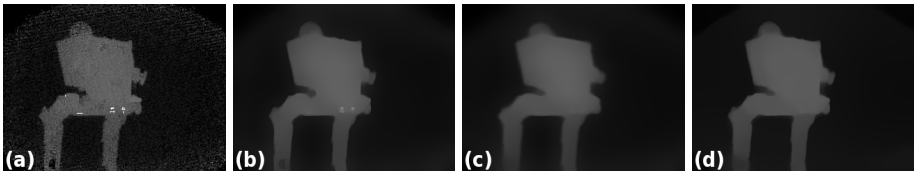

Fig. 5. (a) one of the original frames with real noise. (b) frame filtered with standard 2D ROF. (c) frame filtered with standard 3D ROF. (d) frame filtered with proposed adaptive TV minimization. Only the spatio-temporal methods are able to remove both noise and speckles. Anisotropic TV keeps the result sharper than isotropic 3D TV minimization.

$\beta = 0.001$ and 10 steps for the outer iteration. The result for one specific frame is depicted in Fig. 5, right. We compare this method with standard 2D ROF (Fig. 5, second left) and 3D ROF in the spatio-temporal domain (Fig. 5, second right), using the same parameter $\alpha = 0.3$. Additionally, we zoom into two image regions, see Fig. 6. We observe that standard 2D ROF filtering provides a good noise removal with preserving edges, but is not able to remove the speckles. 3D ROF filtering removes both noise and speckles, but introduces some blurring of edges, which is due to the stair-casing effect in 3D. The proposed adaptive anisotropic TV comprises both the advantages of the 2D and 3D isotropic model: it removes noise and speckles, while edges in each individual frame are kept sharp.

7 Conclusion

In this work we have presented a general approach for adaptive total variation. Existence results as well as a first basic algorithm have been provided. Several applications demonstrate the usability of our concept. As future work, we will support our framework with convergence results and investigate efficient numerical solvers.

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.

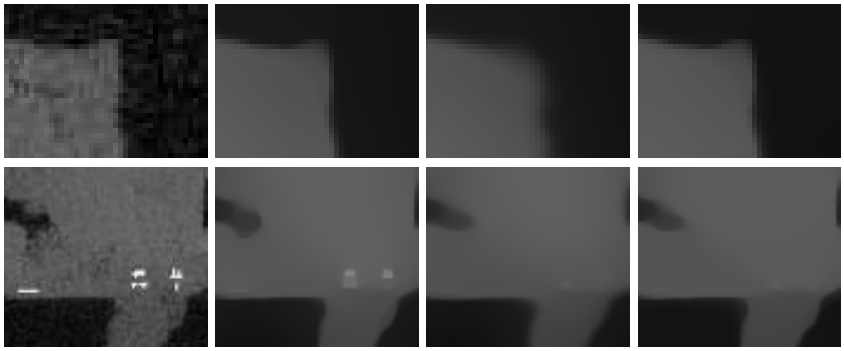

Fig. 6. Zoom into two regions of the depth map shown in Fig. 5. First column: original data, second column: result of standard 2D ROF filtering, third column: 3D ROF filtering, fourth column: proposed adaptive TV minimization. Only the spatio-temporal methods are able to remove both noise and speckles. Anisotropic TV keeps the result sharper than 3D ROF minimization.

2. D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar. *Convex Analysis and Optimization*. 2003.
3. A. Buades, B. Coll, and J.M. Morel. A review of image denoising algorithms, with a new one. *Multiscale Model. Simul.*, 4(2):490–530, 2005.
4. A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imaging Vision*, 20(1–2):89–97, 2004.
5. D. Chan and T.S. Pang. The generalized quasi-variational inequality problem. *Math. Operat. Res.*, 7(2):211–222, 1982.
6. 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.
7. D. Eberly. Distance from a point to an ellipse in 2D. Technical report, 2002.
8. Guy Gilboa and Stanley Osher. Nonlocal operators with applications to image processing. *Multiscale Model. Simul.*, 7(3):1005–1028, 2008.
9. M. Grasmair. Locally adaptive total variation regularization. In *proceedings SSVM '09*, volume 5567, pages 331–342, 2009.
10. M. Grasmair and F. Lenzen. Anisotropic Total Variation Filtering. *Appl. Math. Optim.*, 62(3):323–339, 2010.
11. 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.
12. F. Rellich. Störungstheorie der Spektralzerlegung, I. *Math. Ann.*, 1936.
13. L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1–4):259–268, 1992.
14. S. Setzer, G. Steidl, and T. Teuber. Restoration of images with rotated shapes. *Numerical Algorithms*, 48:49–66, 2008.
15. G. Steidl and T. Teuber. Anisotropic smoothing using double orientations. In *proceedings SSVM '09*, pages 477–489, 2009.
16. J.H. Wilkinson. *The algebraic eigenvalue problem*. Numerical Mathematics and Scientific Computation. Oxford University Press, 1988.