Solution-Driven Adaptive Total Variation Regularization

Frank Lenzen and Johannes Berger

HCI & IPA, University of Heidelberg, Germany
`{frank.lenzen,johannes.berger}@iwr.uni-heidelberg.de`

Abstract. We consider solution-driven adaptive variants of Total Variation, in which the adaptivity is introduced as a fixed point problem. We provide existence theory for such fixed points in the continuous domain. For the applications of image denoising, deblurring and inpainting, we provide experiments which demonstrate that our approach in most cases outperforms state-of-the-art regularization approaches.

Keywords: regularization, inverse problems, adaptive total variation, solution-driven adaptivity, fixed point problems, image restoration

1 Introduction

In recent years, several adaptive variants of Total Variation (TV) for the task of image restoration have been proposed, see e.g. [3, 7, 8, 11–13, 16, 20]. When adaptivity is introduced, there are several alternatives how to provide the information which is required to steer the adaptivity. The standard approach is to use the (pre-smoothed) input data to estimate prominent image structures (referred to as *data-driven adaptivity* in the following). Another way is to let the adaptivity depend directly on the solution (*solution-driven adaptivity*). In this paper we follow an approach which we have proposed in [14]: The adaptivity is defined based on an arbitrary input image v. Fixing v yields a convex optimization problem with uniquely determined minimizer u. One then seeks for a fixed point of the mapping from v to u, which makes the adaptivity solution-driven.

In [15, 16] we have shown that in the **discrete** setting under sufficient conditions a fixed point of this approach exists. Uniqueness can also be obtained with more restrictive assumptions.

In the work presented here, we consider the **continuous** setting of this ansatz, for which we provide existence results for a class of adaptive TV regularizers. Such theory has not been provided in literature so far. As exemplary applications we consider image denoising, non-blind deblurring and inpainting. Our approach, however, is applicable as regularization method for any inverse problem under the conditions retrieved in this paper.

2 Motivation

We start with the classical ROF functional [17] for denoising images, in which the data to be denoised are represented as a function $f \in L^2(\Omega)$ for some

To appear in proc. of SSVM 2015. The final publication will be available at `http://link.springer.com/`.

updated on March 6, 2015.

open bounded domain $\Omega \subset \mathbb{R}^d$ with Lipschitz boundary. The denoised image is represented by a function u lying in $L^2(\Omega) \cap BV(\Omega)$, where $BV(\Omega)$ is the space of functions of bounded variation [18]. Function u is obtained as minimizer of the strictly convex functional

$$\mathcal{F}(u; f) := \mathcal{S}(u; f) + \alpha \, \mathrm{TV}(u), \tag{1}$$

where $\mathcal{S}(u; f) := \frac{1}{2}\|u - f\|^2_{L^2(\Omega)}$ is the data term, $TV(u) := \sup\{\int_\Omega \operatorname{div} \varphi \, u \, d\mathcal{L} \mid \varphi \in C^1_c(\Omega, \mathbb{R}^d), \|\varphi(x)\|_2 \leq 1\}$ is the total-variation semi-norm and $\alpha > 0$ is the regularization strength. We generalize this approach by replacing $\alpha \, \mathrm{TV}(u)$ by an regularization term $\mathcal{R}_v(u)$, i.e. $\mathcal{F}_v(u; f) := \frac{1}{2}\|u - f\|^2_{L^2} + \mathcal{R}_v(u)$, where v is a function steering the regularization. We assume an $\mathcal{R}_v(u)$ of the form

$$\mathcal{R}_v(u) := \sup\{\int_\Omega u(x) \operatorname{div}(A(x, v)\varphi(x)) \, d\mathcal{L} \mid \varphi \in \mathcal{D}\}, \tag{2}$$

where $\mathcal{D} := \{C^1_c(\Omega, \mathbb{R}^d), \|\varphi(x)\|_2 \leq 1\}$ and $A(x, v)$ is continuously differentiable w.r.t. x and maps to the space of symmetric matrices.

Example 1. Let $v \in C^1(\Omega, \mathbb{R})$. We consider an adaptive TV regularization, where the regularization strength α is reduced at locations, where the gradient of v is high. To this end, we set $A(x, v) = \alpha(x, v) \, \mathrm{Id}$ with $\alpha(x, v) = \widetilde{\max}(\alpha_0(1 - \kappa\|\nabla v(x)\|_2), \varepsilon_0)$, where $\widetilde{\max}$ is a smoothed version of the max operator and $\varepsilon_0, \alpha_0 > 0$ and $\kappa \geq 0$. ⋄

Let us assume that the functional $\mathcal{F}_v(u; f)$ attains a unique minimizer $\overline{u}$. We introduce the short notation $\overline{u} := T(v; f) := \arg\min_u \mathcal{F}_v(u; f)$.

In the following, we assume that v is also an image. A special case would be to set $v = f$ (or a pre-smoothed version of f to be robust against noise). We refer to such an approach as *data-driven* adaptivity. Here, we follow a different approach by searching for a fixed point u^* of the mapping $v \to T(v; f)$. For such a fixed point $u^* = \overline{u} = v$, the regularization term becomes $\mathcal{R}_{u^*}(u)$, i.e. the adaptivity is determined by the solution u^* itself. For our example, this means that the adaptivity is reduced at edges of u^*. We refer to this approach as a *solution-driven* adaptivity. It has to be noted that this approach is in general not equivalent to solving the non-convex problem $\arg\min_u \mathcal{F}_u(u; f)$.

Two open issues remain, which are the *well-posedness* of the problem $\overline{u} = \arg\min \mathcal{F}_v(u; f)$ (and operator $T(v; f)$), and the *existence of a fixed point* $u^* = \overline{u} = v$ of T. We address both issues in the next section.

3 Theory

In the following, for the simplicity of notation, we omit the dependency of data term $\mathcal{S}$, functional $\mathcal{F}$ and operator T on the input data f. Moreover, we omit the argument x of functions, if it is clear from the context. We generalize the variational problem by considering a larger class of data terms $\mathcal{S}(u)$. The presentation of our approach concentrates on the two-dimensional case $\Omega \subset \mathbb{R}^2$. The generalization to a d-dimensional domain is straightforward.

First, we generalize the approach from last section to general inverse problems by considering arbitrary data terms $\mathcal{S}(u)$ with certain conditions stated below and additional constraints $u \in \mathcal{C}$, where $\mathcal{C} \subset BV(\Omega)$ is convex and weakly closed in $L^2(\Omega)$. Second, we state operator T more precisely:

$$T : L^2(\Omega) \to BV(\Omega) \cap L^2(\Omega) \quad T(v) := \arg\min_{u \in \mathcal{C}} \mathcal{F}_v(u). \tag{3}$$

3.1 Well-posedness of operator T

Remark 1. Recall that $\Omega \subset \mathbb{R}^d$ is open bounded with Lipschitz boundary. In the case $d = 2$, on which we focus here, the embedding from $BV(\Omega)$ to $L^2(\Omega)$ is continuous [18, Thm 9.78]. Thus $\mathcal{F}_v(u)$ is well-defined on $BV(\Omega)$. In the case $d > 2$ we have to restrict the optimization of $\mathcal{F}_v(u)$ to the space $L^2(\Omega) \cap BV(\Omega)$.

Assumption 1 *(i) $\mathcal{S} : L^2(\Omega) \to \mathbb{R}$ is strictly convex and lower semi-continuous w.r.t. the weak convergence in $L^2(\Omega)$.*
(ii) There exist constants $c_1, c_2 > 0$ such that $\|u\|_{L^2} \leq c_1\mathcal{S}(u) + c_2$ for any $u \in \mathcal{C}$.
(iii) Let $\lambda_i(x, v)$, $i = 1, 2$, be the eigenvalues of $A(x, v)$. We assume that there exist $C_{min}, C_{max} > 0$ such that $C_{min} \leq \lambda_i(x, v) \leq C_{max}$, $i = 1, 2$.

Proposition 1. *For $v \in L^2(\Omega)$, $\mathcal{F}_v(\cdot)$ has a unique minimizer $\overline{u} \in \mathcal{C} \subset BV(\Omega)$. Moreover, there exists an $R_{max} \geq 0$ independent from v, such that $\|\overline{u}\|_{BV} \leq R_{max}$. Thus $T(v) : L^2(\Omega) \to U := \{u \in \mathcal{C} \mid \|u\|_{BV} \leq R_{max}, \}$, is well defined.*

Remark 2 (Sketch of the proof of Prop. 1). The proof follows the standard proof for convex minimization problems. Starting with a minimizing sequence, one first shows that the sequence is bounded in $BV(\Omega)$: Using Ass. 1(iii), one can prove that $C_{min}TV(u) \leq \mathcal{R}(u, v)$. Together with Ass. 1(ii) boundedness of the sequence follows. We therefore can find a subsequence weakly* converging to some $\overline{u} \in BV(\Omega)$, also converging weakly in $L^2(\Omega)$. Analogously to the proof of Thm. 1 in [9, Sect. 5.2.1], we can show that $\mathcal{R}_v(u)$ is weakly* lower semi-continuous. This requires the upper bound C_{max} from Ass. 1(iii). Together with Ass. 1(i) we have that $\mathcal{F}_v(u)$ is weakly* lower semi-continuous, and thus $\overline{u}$ is a minimizer. From the strict convexity of $\mathcal{F}_v(u)$ it follows that $\overline{u}$ is unique.

3.2 Existence of a fixed point

Assumption 2 *$A(x, v)$ is continuously differentiable w.r.t. x and Lipschitz-continuous with constant w.r.t. v. Moreover, for any $u \in C^\infty(\Omega)$ and $v \in L^1(\Omega)$ there exists a function $a(x, u, v)$ continuous in x such that*

$$\sup_{\varphi \in \mathcal{D}} \int_\Omega \varphi(x)^\top A(x, v) \nabla u(x) \, d\mathcal{L} = \int_\Omega a(x, u, v) \|\nabla u(x)\|_2 \, d\mathcal{L}, \tag{4}$$

where $a(x, u, v)$ is Lipschitz-continuous w.r.t. v: there exists a constant $C_l > 0$ independent from u and v such that

$$\|a(x, u, v_1) - a(x, u, v_2)\|_{L^\infty} \leq C_l \|v_1 - v_2\|_{L^1}. \tag{5}$$

Theorem 1. *Let Ass. 1 and 2 be satisfied. Then, there exists a fixed point* u^* *of* $T(v)$*, i.e.* u^* *minimizes* $\mathcal{F}_{u^*}(u)$*.*

The proof of Thm.1 requires the following definition and proposition.

Definition 1. *We consider the weak topology in* $L^2(\Omega)$ *(cf. [2]). We call* $\mathcal{F} : L^2(\Omega) \to L^2(\Omega)$ *weakly upper semi-continuous if for any weakly closed subset* B *of* $im(\mathcal{F})$ *the set* $\mathcal{F}^{-1}(B)$ *is weakly closed.*

Proposition 2. *The mapping* $v \mapsto T(v)$ *from* $\mathcal{C}$ *to* $U \subset \mathcal{C} \subset L^2(\Omega)$ *is weakly upper semi-continuous w.r.t. the weak topology in* $L^2(\Omega)$*.*

The proof of Prop. 2 will be given below. First we state two required lemmas, the proofs of which are provided in the appendix.

Lemma 1. *Let* U *be a bounded subset of* $BV(\Omega)$ *and let* $v^k \to v$ *in* $L^1(\Omega)$*. Then, we have* $\mathcal{R}_{v^k}(u) \to \mathcal{R}_v(u)$ *and* $\mathcal{F}_{v^k}(u) \to \mathcal{F}_v(u)$ *uniformly for every* $u \in U$*.*

Lemma 2. *For* $v^k \to v^0$ *in* $L^1(\Omega)$ *and* $u^k := T(v^k) := \arg\min_u \mathcal{F}_{v^k}(u)$*:*

$$\liminf_{k\to\infty} \mathcal{F}_{v^0}(u^k) \le \limsup_{k\to\infty} \mathcal{F}_{v^k}(u^k). \tag{6}$$

Proof (Proof of Proposition 2). Let B be a weakly closed subset of U (recall that $\mathrm{im}(T) \subset U$). Let $(v^k)_k$ be a sequence in $T^{-1}(B) \subset V$ weakly converging in $L^2(\Omega)$ to some $v^0 \in V$, i.e. there exists a sequence $(u^k)_k \in B$ such that $u^k = T(v^k)$, and $v^k \overset{L^2}{\rightharpoonup} v^0$. Let $u^0 := T(v^0)$. Since U is weakly* pre-compact in $BV(\Omega)$ and weakly in $L^2(\Omega)$, there exists a subsequence also denoted by $(u^k)_k$, such that $u^k \overset{*}{\rightharpoonup} u$ for some $u \in BV_\Omega$ and $u^k \overset{L^2}{\rightharpoonup} u$. Since B is weakly closed in $L^2(\Omega)$, we find $u \in B$. Next, we show that $u = u^0$, i.e. u is the unique minimizer of $\mathcal{F}_{v^0}(\cdot)$. Since the embedding from $L^2(\Omega)$ to $L^1(\Omega)$ is compact, we can find a subsequence $(v^{k'})_{k'}$, of $(v^k)_k$, which converges to v strongly in $L^1(\Omega)$. Using the weakly* lower semi-continuity of $\mathcal{F}_{v^0}$ (cf. Rem. 2) together with Lem. 2 we find

$$\begin{aligned} 0 \le \mathcal{F}_{v^0}(u) - \mathcal{F}_{v^0}(u^0) &\le \liminf_{k'\to\infty} \mathcal{F}_{v^0}(u^{k'}) - \mathcal{F}_{v^0}(u^0) \\ &\le \limsup_{k'\to\infty} \mathcal{F}_{v^{k'}}(u^{k'}) - \mathcal{F}_{v^0}(u^0). \end{aligned} \tag{7}$$

Recall that $u^{k'}$ is the minimizer of $\mathcal{F}_{v^{k'}}(\cdot)$ for all $k' \ge 0$, which induces $\mathcal{F}_{v^{k'}}(u^{k'}) \le \mathcal{F}_{v^{k'}}(u^0)$. Using this fact, we obtain from (7) that

$$0 \le \mathcal{F}_{v^0}(u) - \mathcal{F}_{v^0}(u^0) \le \limsup_{k'\to\infty} \mathcal{F}_{v^{k'}}(u^0) - \mathcal{F}_{v^0}(u^0). \tag{8}$$

Lemma 1 guarantees that the r.h.s. of (8) tends to zero, thus $\mathcal{F}_{v^0}(u) = \mathcal{F}_{v^0}(u^0)$. On the other hand, u^0 by definition is the unique minimizer of $\mathcal{F}_{v^0}(\cdot)$, from which $u = u^0$ follows. Since $u \in B$ and $u = u^0 = T(v^0)$, we have shown that $v^0 \in T^{-1}(B)$ and thus $T^{-1}(B)$ is weakly closed. □

Proof (Proof of Thm. 1). The claim follows from Theorem 2.3. in [1], since $\mathcal{C}$ is convex and closed in $L^2(\Omega)$ and $v \to T(v)$ is weakly compact in $L^2(\Omega)$ (it maps to the pre-compact set U, thus $T(B)$ is pre-compact for any $B \subset \mathcal{C}$) and weakly upper semi-continuous in $L^2(\Omega)$ (Prop. 2). □

Fig. 1. Test images used for evaluation. Top row – denoising: with additive Gaussian noise (zero mean, standard deviation 0.1). Middle row – deblurring: blurred images with Gaussian noise (zero mean, standard deviation 0.01). Bottom row – inpainting: white regions mark the regions to be inpainted.

4 Image Restortation

Data terms We consider the applications of denoising, non-blind deblurring and inpatining. In the case of denoising, we utilize the standard L^2 data term, which is strictly convex. In the case of deblurring, we assume a kernel K with bounded support, such that $u \in L^2(\Omega)$ leads to data $f(x) := (K * u)(x)$ in $L^2(\Omega_0)$, $\Omega_0 \subset \Omega$. We set $\mathcal{S}(u) := \frac{1}{2}\|K * u - f\|_{L^2(\Omega_0)}$. By an additional regularization of the data term, i.e. by adding $\varepsilon\|u\|^2_{L^2(\Omega)}$ with small $\varepsilon > 0$, we assert strict convexity of $\mathcal{F}(u)$. For inpainting, we consider the domain $\Omega_0 \subset \Omega$ on which the data f are known and set $\mathcal{C} := \{u \in BV(\Omega) \mid u = f \text{ a.e. on } \Omega_0\}$. Moreover, to obtain a strictly convex problem we choose $\mathcal{S}(u) = \varepsilon\|u\|^2_{L^2(\Omega\setminus\overline{\Omega_0})}$. In all three applications Ass 1(i) and (ii) are satisfied.

Adaptivity We consider two examples of adaptive TV regularization. In both, we steer the adaptivity by choosing appropriate functions $A(x, v)$.

First, we revisit Example 1, where we locally adapt the regularization strength. Assuming that v lies in $BV(\Omega)$, since it is expected to be the fixed point of operator T, we face the problem that the gradient of v cannot be interpreted as a function on Ω. To circumvent this problem, we introduce a pre-smoothing of v as follows: Let $K_\sigma(x)$ be a Gaussian kernel with variance σ^2. We define $v_\sigma := K_\sigma * v$, where $*$ is the convolution operator. We choose suitable boundary conditions to accommodate to the fact that v is only defined on Ω. We then set

$$A(x, v) := \alpha(x, v_\sigma)\,\mathrm{Id}, \text{ with } \alpha(x, v_\sigma) := \widetilde{\max}(\alpha_0(1 - \kappa\|\nabla v_\sigma\|_2), \varepsilon_0), \tag{9}$$

where $\widetilde{\max}$ is a smoothed version of the max-operator and $\alpha_0, \varepsilon_0 > 0$, $\kappa \geq 0$.

	Method	Train	Cameraman	Lena	Peppers	Boat
Denoising	std. TV [17]	0.789	0.808	0.793	0.794	0.718
	TGV [4]	0.791	0.814	0.798	0.810	0.723
	BM3D [6]	**0.840**	**0.856**	**0.845**	**0.834**	**0.762**
	adapt. TV (dd)	0.789	0.813	0.793	0.794	0.718
	adapt. TV (sd)	0.790 (4)	0.822 (4)	0.793 (3)	0.794 (4)	0.718 (4)
	anisotr. TV(dd)	0.793	0.808	0.796	0.794	0.718
	anisotr. TV(sd)	0.803 (4)	0.820 (4)	0.819 (4)	0.798 (4)	0.723 (4)
Deblurring	std. TV	0.774	0.766	0.760	0.795	0.644
	TGV	0.775	0.747	0.749	0.799	0.691
	Schmidt [19]	0.571	0.612	0.697	0.682	0.705
	adapt. TV(dd)	0.774	0.757	0.762	0.794	0.631
	adapt. TV(sd)	0.813 (3)	0.829 (4)	0.808 (2)	0.811 (2)	0.715 (4)
	anisotr. TV(dd)	0.774	0.783	0.762	0.797	0.660
	anisotr. TV(sd)	**0.839** (4)	**0.841** (4)	**0.848** (4)	**0.864** (4)	**0.776** (4)
Inpainting	std. TV	0.930	0.945	0.940	0.951	0.933
	TGV	0.958	0.958	0.958	0.963	0.954
	Garcia [10]	0.957	0.966	0.966	**0.969**	**0.961**
	adapt. TV(dd)	0.938	0.945	0.940	0.951	0.933
	adapt. TV(sd)	0.972 (4)	0.969 (4)	0.957 (4)	0.964 (4)	0.950 (4)
	anisotr. TV(dd)	0.947	0.946	0.942	0.952	0.936
	anisotr. TV(sd)	**0.976** (4)	**0.973** (4)	**0.971** (4)	0.968 (4)	0.958 (4)

Table 1. Similarity [21] to ground truth for the different applications, methods and test images. (dd=data-driven, sd= solution-driven, numbers in brackets are the iterations of the outer loop.) Except for the application of denoising, our approaches in most cases provide the best results compared to the other methods.

We refer to this example as solution-driven *adaptive* TV regularization. Note that $\alpha(x) \in [\varepsilon_0, \alpha_0]$, such that Assumption 1(iii) is satisfied. The following lemma shows, that Assumption 2 also holds.

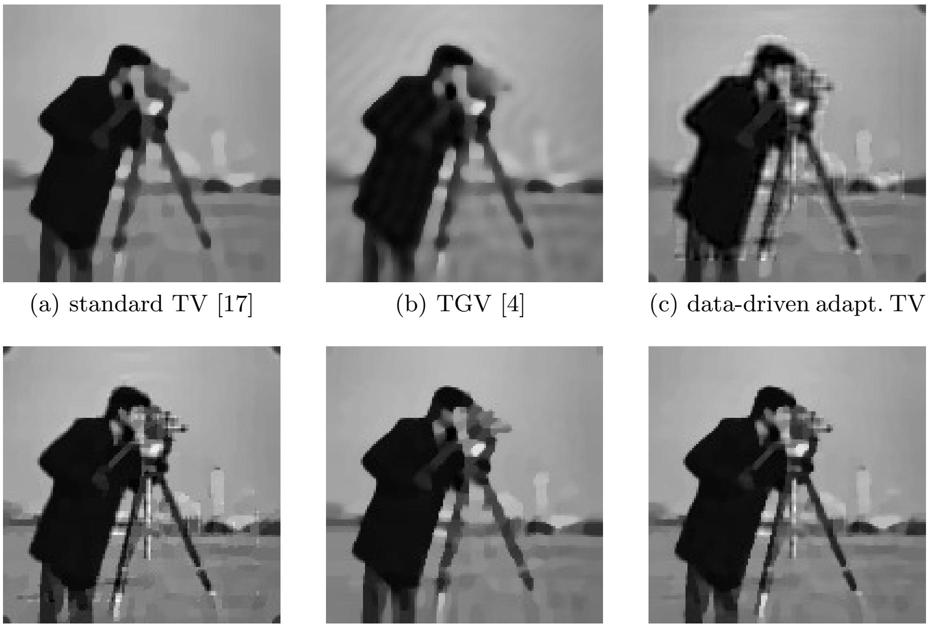
Lemma 3. *Let $A(x,v)$ be defined as in (9). Then, Assumption 2 is satisfied with $a(x,u,v) = \alpha(x, v_\sigma)$.*

Proof. For $u \in C^\infty(\Omega)$ we have

$$\sup_{\varphi\in\mathcal{D}} \int_\Omega \varphi^\top A(v)\nabla u \, d\mathcal{L} = \sup_{\varphi\in\mathcal{D}} \int_\Omega \varphi^\top \alpha(x,v)\nabla u(x) \, d\mathcal{L} \le \int_\Omega \alpha(x,v)\|\nabla u(x)\|_2 \, d\mathcal{L}. \tag{10}$$

We show equality in (10). Let $\varphi(x) := \nabla u(x)/\|\nabla u(x)\|_2$ if $\nabla u(x) \neq 0$ and $\varphi(x) := 0$ otherwise. Then $\|\nabla u\|_2 = \varphi^\top \nabla u$. Let $\varphi_\varepsilon \in C_c^\infty(\Omega, \mathbb{R}^2) \to \varphi$ in $L^2(\Omega; \mathbb{R}^2)$ and

$$\int_\Omega \left|\varphi_\varepsilon^\top \alpha(x,v)\nabla u - \alpha(x,v)\|\nabla u\|_2\right| \, d\mathcal{L} \le \|\alpha(x,v)\|_{L^\infty} \|\varphi_\varepsilon - \varphi\|_{L^2} \|\nabla u\|_{L^2} \to 0. \tag{11}$$



(a) standard TV [17] (b) TGV [4] (c) data-driven adapt. TV

(d) sol.-driven adapt. TV (e) data-driven aniso. TV (f) sol.-driven aniso. TV

Fig. 2. Deblurring: best performing methods on the cameraman image. The solution-driven variants outperform the data-driven variants. Solution-driven *anisotropic* TV gives the best result.

The Lipschitz continuity of $v \to \alpha(x, v)$ follows form the continuous differentiability of α w.r.t. v_σ together with $\|\nabla(v_1)_\sigma - \nabla(v_2)_\sigma\|_{L^\infty} \leq C\|v_1 - v_2\|_{L^1}$ as a property of convolutions. □

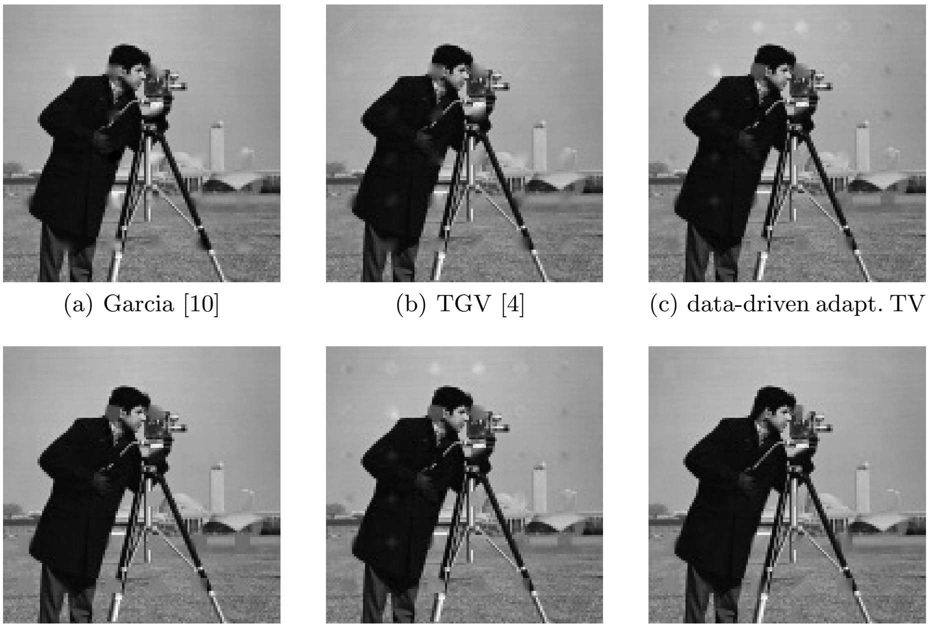
Our second example is a solution-driven *anisotropic* TV variant:

$$A(x, v) := \alpha(x)r(x)r(x)^\top + \beta(x)(r^\perp(x))(r^\perp(x))^\top \tag{12}$$

for some vector field $r(x) : \Omega \to \mathbb{R}^2$ and scalar functions $\alpha(x), \beta(x) : \Omega \to \mathbb{R}_+$. To determine $r(x)$ we consider the structure tensor $J_\rho(v_\sigma) := (\nabla v_\sigma \nabla v_\sigma^\top)_\rho$, where $(M)_\rho$ denotes the elementwise convolution of matrix M with kernel K_ρ. Let $\mu_1(x) \geq \mu_2(x) \geq 0$ be the eigenvalues of $J_\rho(v_\sigma)(x)$. We choose $r(x)$ as the normalized eigenvector to eigenvalue $\mu_1(x)$. $\alpha(x)$ and $\beta(x)$ are chosen as

$$\alpha(x) = g(\mu_1(x) - \mu_2(x))\alpha_0 + (1 - g(\mu_1(x) - \mu_2(x)))\beta_0, \qquad \beta(x) = \beta_0,$$

with $g(s) = \min(c \cdot s, 1)$, $c > 0$. For $c_{min} := \min\{\alpha_0, \beta_0\}$, $C_{max} := \max\{\alpha_0, \beta_0\}$ we have $0 < c_{min} \leq \alpha(x), \beta(x) \leq C_{max} < \infty$, so that Ass. 1(iii) follows. Due to space constraints we omit the proof that Assumption 2 is satisfied.



(a) Garcia [10] (b) TGV [4] (c) data-driven adapt. TV

(d) sol.-driven adapt. TV (e) data-driven aniso. TV (f) sol.-driven aniso. TV

Fig. 3. Inpainting: best performing methods on the cameraman image. The solution-driven variants outperform the data-driven variants. Solution-driven *anisotropic* TV gives the best result.

5 Experiments

Numerically, we solve the problem of finding a fixed point of T by two nested iterations. In the outer iteration, function v is updated; in the inner iteration we solve the discretized convex problem $\arg\min_u \mathcal{F}_v(u)$ with a primal-dual method [5].

We consider three different applications: denoising, non-blind deblurring and inpainting. As test images, we consider the five images depicted in Fig. 1. (*Train* image by courtesy of Kristian Bredies.) All ground truth images are scaled to the range $[0, 1]$. We compare our method to regularization with standard TV, Total Generalized Variation (TGV) [4] and the data-driven variants (i.e. using $T(f)$, where f are the input data) of our approach. Moreover, we compare with BM3D [6] for denoising, the method by Schmidt [19] for deblurring and to Garcia's method [10] for inpainting. For comparison, we utilize the similarity measure from [21] applied to result and ground truth. To find the optimal parameters for each method w.r.t this measure we applied a hierarchical grid search. The similarity measures for each application and each method are listed in Table 3.2. For the task of denoising, we observe that we cannot cope with the BM3D method, but obtain better results than the other regularization approaches. Concerning deblurring and inpainting, we obtain in most cases the best results compared to the other methods. For deblurring and inpainting, the results of the best performing methods on the cameraman image are depicted in Figs. 2 and 3. (The other results are omitted due to space constraints.)

6 Conclusion

We have considered solution-driven adaptive variants of Total Variation. In our ansatz the adaptivity leads to a fixed point problem, for which we provided existence theory. Our experiments demonstrated for the applications of image denoising, deblurring and inpainting, that our approach in most cases outperforms state-of-the-art *regularization* approaches. In future work we will study the issue of uniqueness of fixed points in the continuous setting.

Appendix

The following lemma is required in the proof of Lemma 1.

Lemma 4. *For any* $u \in BV(\Omega)$ *there exists* $(u^k)_k \in C^\infty(\Omega)$ *such that* $u^k \xrightarrow{L^1} u$, $\mathrm{TV}(u^k) \to TV(u)$ *and* $\mathcal{R}_{v_i}(u^k) \to \mathcal{R}_{v_i}(u)$ *simultaneously for a finite set of* $\{v_i\}_i$.

Proof. For fixed v the proof follows the proof of [9, Thm. 2, Sect. 5.2], with some modifications. Recall $\mathcal{D} = \{C_c^1(\Omega, \mathbb{R}^2), \|\varphi(x)\|_2 \leq 1\}$. Let $\varepsilon > 0$ be fixed. For a $m \in \mathbb{N}, m > 0$ and $k \in \mathbb{N}$ we define open sets $\Omega_k := \{x \in \Omega \mid \mathrm{dist}(x, \partial\Omega) > \frac{1}{m+k}\}$ and choose m large enough to guarantee

$$\mathrm{TV}(u)(\Omega \setminus \Omega_1) < \varepsilon, \tag{13}$$

where $\mathrm{TV}(u)(B)$ is the variation measure of u evaluated on the set B. Set $\Omega_0 := \emptyset$ and define $V_k := \Omega_{k+1} \setminus \overline{\Omega}_{k-1}$. Please note that each $x \in \Omega$ is contained in at most three sets V_k. Moreover, let $\{\zeta_k\}_{k=1}^\infty$ be a sequence of smooth functions with $\zeta_k \in C_c^\infty(V_k)$, $0 \leq \zeta_k \leq 1$ and $\sum_{k=1}^\infty \zeta_k = 1$ on Ω. Let η be a mollifier as in [9, Sect. 4.2.1]). For each k, select $\varepsilon_k > 0$ small enough such that for $\eta_{\varepsilon_k} := \frac{1}{\varepsilon_k^2}\eta(\frac{x}{\varepsilon_k})$

$$\begin{aligned} \mathrm{supp}(\eta_{\varepsilon_k} * (u\zeta_k)) \subset V_k, \quad \int_\Omega |\eta_{\varepsilon_k} * (u\zeta_k) - u\zeta_k| \, d\mathcal{L} &< \frac{\varepsilon}{2^k}, \\ \int_\Omega \|\eta_{\varepsilon_k} * (u\nabla\zeta_k) - u\nabla\zeta_k\|_2 \, d\mathcal{L} &< \frac{\varepsilon}{2^k}. \end{aligned} \tag{14}$$

Define $u_\varepsilon := \sum_{k=1}^\infty \eta_{\varepsilon_k} * (u\zeta_k)$. In this sum there are only finitely many terms, which are non-zero on a neighborhood of each $x \in \Omega$. Thus, $u_\varepsilon \in C^\infty(\Omega)$. Since also $u = \sum_{k=1}^\infty u\zeta_k$, we find from (14) that

$$\|u_\varepsilon - u\|_{L^1(\Omega)} \leq \sum_{k=1}^\infty \int_\Omega |\eta_{\varepsilon_k} * (u\zeta_k) - u\zeta_k| \, d\mathcal{L} < \varepsilon. \tag{15}$$

Thus, $u_\varepsilon \to u$ in $L^1(\Omega)$. Moreover, for $\varphi \in \mathcal{D}$, we have

$$\int_\Omega u_\varepsilon \operatorname{div} \varphi \, d\mathcal{L} = \int_\Omega u \left(\sum_k \zeta_k \operatorname{div}(\eta_{\varepsilon_k} * \varphi) \right) \, d\mathcal{L} \leq 3 \operatorname{TV}(u), \tag{16}$$

where we used that for every $x \in \Omega$ there exist at most three $\zeta_k(x) > 0$. Taking the supremum over all such φ we find $\operatorname{TV}(u_\varepsilon) \leq 3 \operatorname{TV}(u) < \infty$. Thus $u_\varepsilon \overset{*}{\rightharpoonup} u$. Then, due to the weak* lower semi-continuity of $\mathcal{R}_v(u)$ (cf. Rem. 2) we have

$$\mathcal{R}_v(u) \leq \liminf_{\varepsilon \to 0} \mathcal{R}_v(u_\varepsilon). \tag{17}$$

For any $\varphi \in \mathcal{D}$

$$\begin{aligned}
&\int_\Omega u_\varepsilon \operatorname{div}(A(v)\varphi) \, d\mathcal{L} = \sum_{k=1}^\infty \int_\Omega (\eta_{\varepsilon_k} * (u\zeta_k)) \operatorname{div}(A(v)\varphi) \, d\mathcal{L} \\
&= \sum_{k=1}^\infty \int_\Omega u\zeta_k \operatorname{div}(\eta_{\varepsilon_k} * (A(v)\varphi)) \, d\mathcal{L} \\
&= \underbrace{\sum_{k=1}^\infty \int_\Omega u \operatorname{div}(\zeta_k(\eta_{\varepsilon_k} * (A(v)\varphi))) \, d\mathcal{L}}_{:=I_1^\varepsilon} - \underbrace{\sum_{k=1}^\infty \int_\Omega \varphi^\top A(v)(\eta_{\varepsilon_k} * (u\nabla\zeta_k) - u\nabla\zeta_k) \, d\mathcal{L}}_{=:I_2^\varepsilon},
\end{aligned} \tag{18}$$

where we used the fact $\sum_{k=1}^\infty \nabla\zeta_k = 0$ on Ω. Since $\|A(x,v)\varphi(x)\|_2 \leq C_{max}$, we find that $\|\zeta_k(x)(\eta_{\varepsilon_k} * (A(v)\varphi))(x)\|_2 \leq C_{max}$. Thus, we can bound I_1^ε by

$$\begin{aligned}
|I_1^\varepsilon| = |&\int_\Omega u \operatorname{div}(\zeta_1(\eta_{\varepsilon_1} * (A(v)\varphi))) \, d\mathcal{L} + \sum_{k=2}^\infty \int_\Omega u \operatorname{div}(\zeta_k(\eta_{\varepsilon_k} * (A(v)\varphi))) \, d\mathcal{L}| \\
&\leq \mathcal{R}_v(u) + C_\eta L(v)\varepsilon + \sum_{k=2}^\infty C_{max} \operatorname{TV}(u)(V_k),
\end{aligned} \tag{19}$$

where $L(v)$ is the Lipschitz-constant of $x \to A(x,v)$ and where we use that

$$\|\eta_{\varepsilon_k} * (A(v)\varphi) - A(v)(\eta_{\varepsilon_k} * \varphi)\| \leq C_\eta L(v) \|\varphi\|_{L^\infty} \varepsilon \tag{20}$$

(which follows from standard calculus) with $C_\eta > 0$ only depending on η.

Since each point in Ω belongs to at most three sets V_k, together with $V_k \subset \Omega \setminus \Omega_1$ and (13), (19) can be bounded by

$$|I_1^\varepsilon| \leq \mathcal{R}_v(u) + C_\eta L(v)\varepsilon + 3C_{max} \operatorname{TV}(u)(V_k) < \mathcal{R}_v(u) + (3C_{max} + C_\eta L(v))\varepsilon. \tag{21}$$

Since $\|A(x,v)\varphi(x)\|_2 \leq C_{max}$ we obtain from (14) that

$$|I_2^\varepsilon| = |\sum_{k=1}^\infty \int_\Omega \varphi^\top A(v)(\eta_{\varepsilon_k} * (u\nabla\zeta_k) - u\nabla\zeta_k) \, d\mathcal{L}| \leq C_{max} \sum_{k=1}^\infty \frac{\varepsilon}{2^k} = C_{max}\varepsilon. \tag{22}$$

Combining (18), (21) and (22), we find

$$\int_{\Omega} u_\varepsilon \operatorname{div}(A(v)\varphi)\, d\mathcal{L} < \mathcal{R}_v(u) + (4C_{max} + C_\eta L(v))\varepsilon. \tag{23}$$

Taking in (23) the supremum over all $\varphi \in \mathcal{D}$, we obtain $\mathcal{R}_v(u_\varepsilon) < \mathcal{R}_v(u) + (4C_{max} + C_\eta L(v))\varepsilon$. Together with (17) we have

$$|\mathcal{R}_v(u_\varepsilon) - \mathcal{R}_v(u)| < (4C_{max} + C_\eta L(v))\varepsilon \tag{24}$$

for ε small enough. Since the choice of $\Omega_k, V_k, \varepsilon_k, \zeta_k$ and η_{ε_k} was independent from v, we get a simultaneous convergence for a finite set of v_{ii}. Moreover, as we can express $\mathrm{TV}(u)$ as $\mathcal{R}_v(u)$ with $A(v) = \mathrm{Id}$, (24) provides also $|\,\mathrm{TV}(u_\varepsilon) - \mathrm{TV}(u)| < (4C_{max} + C_\eta)\varepsilon$. □

Proof (of Lemma 1). We show that for arbitrary $u, v_1, v_2 \in BV(\Omega)$

$$|\mathcal{R}_{v_1}(u) - \mathcal{R}_{v_2}(u)| \le C_l\, R_{max}\, \|v_1 - v_2\|_{L^1}, \tag{25}$$

from which the claim for $\mathcal{R}_v(u)$ follows. Consider first a fixed $u \in C^\infty(\Omega)$. Then, by Assumption 2

$$\begin{aligned} |\mathcal{R}_{v_1}(u) - \mathcal{R}_{v_2}(u)| &= \int_{\Omega} |a(x,u,v_1) - a(x,u,v_2)|\, \|\nabla u(x)\|_2\, d\mathcal{L} \\ &\le \|a(x,u,v_1) - a(x,u,v_2)\|_{L^\infty} \mathrm{TV}(u) \le C_l \|v_1 - v_2\|_{L^1} \mathrm{TV}(u). \end{aligned} \tag{26}$$

Now, let $u \in U$ be arbitrary. Using Lemma 4 we can find for any $\varepsilon > 0$ a function $\tilde{u} \in C^\infty(\Omega)$ such that

$$|\mathcal{R}_{v_i}(u) - \mathcal{R}_{v_i}(\tilde{u})| \le \varepsilon \text{ for } i = 1,2, \quad |\,\mathrm{TV}(u) - \mathrm{TV}(\tilde{u})| \le \varepsilon. \tag{27}$$

Then,

$$\begin{aligned} |\mathcal{R}_{v_1}(u) - \mathcal{R}_{v_2}(u)| &\overset{(27)}{\le} |\mathcal{R}_{v_1}(\tilde{u}) - \mathcal{R}_{v_2}(\tilde{u})| + 2\varepsilon \overset{(26)}{\le} C_l \|v_1 - v_2\|_{L^1} \mathrm{TV}(\tilde{u}) + 2\varepsilon \\ &\overset{(27)}{\le} C_l \|v_1 - v_2\|_{L^1} (\mathrm{TV}(u) + \varepsilon) + 2\varepsilon \overset{u \in U}{\le} C_l \|v_1 - v_2\|_{L^1} (R_{max} + \varepsilon) + 2\varepsilon. \end{aligned}$$

Since we can find $\tilde{u}$ such that ε becomes arbitrary small, (25) follows for fixed u. Since the r.h.s. of (25) does not depend on u, we achieve an uniform convergence of $\mathcal{R}_{v^k}(\cdot) \to \mathcal{R}_v(\cdot)$ on U for $v^k \to v$ in $L^1(\Omega)$. Since $\mathcal{R}_v(u)$ and $\mathcal{F}_v(u)$ differ by $\mathcal{S}(u)$ not depending on v, the uniform convergence $\mathcal{F}_{v^k}(\cdot) \to \mathcal{F}_v(\cdot)$ follows. □

Proof (of Lemma 2). Let $v^k \to v^0$ in $L^1(\Omega)$. For the sequence $u^k := T(v^k)$, let $\liminf_{k\to\infty} \mathcal{F}_{v_0}(u^k) := c$. For any $\varepsilon > 0$ Lemma 1 guarantees the existence of a $K > 0$ such that $|\mathcal{F}_{v^k}(u) - \mathcal{F}_{v^0}(u)| \le \frac{\varepsilon}{2}$ for all $u \in U$ and all $k \ge K$. Moreover, we can find a $k' \ge K$ such that $|\mathcal{F}_{v^0}(u^{k'}) - c| \le \frac{\varepsilon}{2}$. From both together, we find

$$|\mathcal{F}_{v^{k'}}(u^{k'}) - c| \le |\mathcal{F}_{v^{k'}}(u^{k'}) - \mathcal{F}_{v^0}(u^{k'})| + |\mathcal{F}_{v^0}(u^{k'}) - c| \le \varepsilon. \tag{28}$$

In other words, there exists a sequence $k' \to \infty$, such that $\mathcal{F}_{v^{k'}}(u^{k'}) \to c$ and thus $\liminf_{k\to\infty} \mathcal{F}_{v^0}(u^k) = c \le \limsup_{k\to\infty} \mathcal{F}_{v^k}(u^k)$. □

References

1. R. Agarwal and D. O'Regan. Fixed-point theory for weakly sequentially upper-semicontinuous maps with applications to differential inclusions. *Nonlinear Oscillations*, 5(3), 2002.
2. H. W. Alt. *Linear functional analysis. An application oriented introduction.* Springer, 2006.
3. B. Berkels, M. Burger, M. Droske, O. Nemitz, and M. Rumpf. Cartoon extraction based on anisotropic image classification. In *VMV*, 2006.
4. K. Bredies, K. Kunisch, and T. Pock. Total Generalized Variation. *SIAM J. Imaging Sciences*, 3(3):492–526, 2010.
5. A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.
6. K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, et al. BM3D image denoising with shape-adaptive principal component analysis. In *SPARS*, 2009.
7. Y. Dong, M. Hintermüller, and M. M. Rincon-Camacho. Automated regularization parameter selection in multi-scale total variation models for image restoration. *Journal of Mathematical Imaging and Vision*, 40(1):82–104, 2011.
8. V. Estellers, S. Soato, and X. Bresson. Adaptive regularization with the structure tensor. 2014. preprint.
9. L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions*, volume 5. CRC press, 1992.
10. D. Garcia. Robust smoothing of gridded data in one and higher dimensions with missing values. *Computational statistics & data analysis*, 54(4):1167–1178, 2010.
11. M. Grasmair. Locally adaptive total variation regularization. In *SSVM*, 2009.
12. M. Grasmair and F. Lenzen. Anisotropic total variation filtering. *Applied Mathematics & Optimization*, 62:323–339, 2010.
13. S. Lefkimmiatis, A. Roussos, M. Unser, and P. Maragos. Convex generalizations of total variation based on the structure tensor with applications to inverse problems. In *SSVM*. Springer, 2013.
14. F. Lenzen, F. Becker, J. Lellmann, S. Petra, and C. Schnörr. Variational image denoising with adaptive constraint sets. In *SSVM*, 2011.
15. F. Lenzen, F. Becker, J. Lellmann, S. Petra, and C. Schnörr. A class of Quasi-Variational Inequalities for adaptive image denoising and decomposition. *Computational Optimization and Applications*, 54(2):371–398, 2013.
16. F. Lenzen, J. Lellmann, F. Becker, and C. Schnörr. Solving Quasi-Variational Inequalities for image restoration with adaptive constraint sets. *SIAM Journal on Imaging Sciences (SIIMS)*, 7:2139–2174, 2014.
17. L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
18. O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational methods in imaging*, volume 167 of *Applied Mathematical Sciences*. Springer, 2009.
19. U. Schmidt, K. Schelten, and S. Roth. Bayesian deblurring with integrated noise estimation. In *CVPR*, 2011.
20. G. Steidl and T. Teuber. Anisotropic smoothing using double orientations. In *SSVM*, 2009.
21. Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: from error visibility to structural similarity. *Image Processing, IEEE Transactions on*, 13(4):600–612, 2004.