DOI 10.1007/s10958-016-3153-y
Journal of Mathematical Sciences, Vol. 219, No. 6, December, 2016

DENOISING AND INPAINTING OF IMAGES USING TV-TYPE ENERGIES: THEORETICAL AND COMPUTATIONAL ASPECTS

M. Bildhauer*
Universität des Saarlandes, Fachbereich 6.1 Mathematik
Postfach 15 11 50, D–66041 Saarbrücken, Germany
bibi@math.uni-sb.de

M. Fuchs
Universität des Saarlandes, Fachbereich 6.1 Mathematik
Postfach 15 11 50, D–66041 Saarbrücken, Germany
fuchs@math.uni-sb.de

J. Weickert
Mathematical Image Analysis Group
Faculty of Mathematics and Computer Sciences
Saarland University, Campus E1.7
D–66041 Saarbrücken, Germany
weickert@mia.uni-saarland.de

UDC 517.958

We discuss variational approaches towards the denoising of images and towards the image inpainting problem combined with simultaneous denoising. Our techniques are based on variants of the TV*-model, but in contrast to this case a complete analytical theory is available in our setting. At the same time, numerical experiments illustrate the advantages of our models in comparison with some established techniques. Bibliography*: 50 *titles. Illustrations*: 1 *figure.*

In this paper, we investigate the variational problem

$$J[u] := \int\limits_{\Omega - D} (f-u)^2\,dx + \alpha \int\limits_{\Omega} \Psi(|\nabla u|)\,dx \to \min\,, \tag{1}$$

which serves as a mathematical model either for the pure denoising of a greyscale image, if the case $D = \varnothing$ is considered, or for an image inpainting problem (with inpainting region D) combined with simultaneous denoising provided that we assume that $D \neq \varnothing$.

In both cases, we will mainly concentrate on densities Ψ of linear growth with respect to $|\nabla u|$, which means that we discuss variants of the TV-regularization. However, as it will be outlined below, we can include arbitrary growth rates of Ψ. Variational methods for the denoising of

* To whom the correspondence should be addressed.

Translated from *Problemy Matematicheskogo Analiza* **87**, October 2016, pp. 69-78.

images are nowadays well established and we refer the reader to the papers [1]–[12] and the references quoted therein.

At the same time, there is a variety of image inpainting techniques of local or non–local nature and using either variational arguments or PDE–type methods. For an overview the reader should consult the references [13]–[29].

Let us clarify our assumptions and notation concerning the problem (1). Suppose that Ω is a bounded Lipschitz domain in $\mathbb{R}^2$, for example, a rectangle, and let D denote an $\mathscr{L}^2$-measurable subset of Ω such that

$$0 \leqslant \mathscr{L}^2(D) < \mathscr{L}^2(\Omega)\,. \tag{2}$$

If we think of pure denoising, we just set $D = \varnothing$. We consider a greyscale image described by a function $u\colon\ \Omega \to [0,1]$, where $u(x)$ measures the intensity of the grey level at the point $x \in \Omega$. Suppose now that a certain part of the image represented by the region D is damaged and that the observed data described in terms of a given function $f\colon\ \Omega - D \to [0,1]$ are noisy. Then the restoration of the missing part $D \to [0,1]$ of the image together with simultaneous denoising of the observed data f might be achieved by looking at the problem (1), where $\alpha > 0$ denotes a positive parameter and the quality of data fitting is measured through the quadratic fidelity term

$$\int\limits_{\Omega - D} (f-u)^2\,dx.$$

Of essential importance is the structure of the regularizing quantity

$$\int\limits_{\Omega} \Psi(|\nabla u|)\,dx,$$

i.e., the behavior of the density $\Psi\colon\ [0,\infty) \to [0,\infty)$. We here look at the following family of densities: for a real parameter μ we set

$$\Psi(t) := \Phi_\mu(t) := \int\limits_0^t \int\limits_0^s (1+r)^{-\mu}\,dr\,ds, \quad t \geqslant 0, \tag{3}$$

and observe that for $\mu \neq 1$ and $\mu \neq 2$

$$\Phi_\mu(t) = \frac{t}{\mu-1} + \frac{1}{\mu-1}\frac{1}{\mu-2}(t+1)^{-\mu+2} - \frac{1}{\mu-1}\frac{1}{\mu-2}, \tag{4}$$

whereas

$$\Phi_1(t) = t\ln(1+t) + \ln(1+t) - t\,, \tag{5}$$

$$\Phi_2(t) = t - \ln(1+t)\,. \tag{6}$$

This shows the following:

(i) If $\mu < 1$, then Ψ is a strictly convex and strictly increasing function of growth order $p := 2-\mu > 1$. Thus, the variational problem (1) is well posed on the Sobolev space $W^1_p(\Omega)$ (cf., for example, [30] for a definition of this class).

(ii) In the case $\mu = 1$, the correct class for the problem (1) is the Orlicz–Sobolev space (cf. again [30]) $W^1_h(\Omega)$ generated by the N-function $h(t) = t\ln(1+t)$, $t \geqslant 0$.

(iii) For $\mu > 1$ the density Ψ is still strictly increasing and strictly convex, but now of linear growth approximating the TV-density $|\nabla u|$ in the sense that (cf. (4))

$$\lim_{\mu\to\infty}(\mu-1)\Phi_\mu(t) = t, \quad t \geqslant 0. \tag{7}$$

Now we have to work in the space $\mathrm{BV}(\Omega)$ of functions having finite total variation (cf., for example, [31]), which means that the quantity

$$\int_\Omega \Psi(|\nabla u|)\,dx$$

has to be replaced by the expression

$$\int_\Omega \Phi_\mu(|\nabla u|) := \int_\Omega \Phi_\mu(|\nabla^a u|)\,dx + \frac{1}{\mu-1}|\nabla^s u|(\Omega) \tag{8}$$

for functions $u \in \mathrm{BV}(\Omega)$. Here, $\nabla u = \nabla^a u \llcorner \mathscr{L}^2 + \nabla^s u$ is the decomposition of the vector measure ∇u in the regular and singular parts with respect to the Lebesgue measure. The reader should note that the definition (8) is in accordance with the notation of a convex function of a measure introduced, for example, in [32].

In the sequel, we will describe the theoretical results valid for the problem (1) for various choices of μ. We emphasize the following.

The density $\Psi = \Phi_\mu$ occurring in formulas (3)–(6) can be replaced by

$$\int_0^t\int_0^s (\varepsilon + r)^{-\mu}\,dr\,ds \quad \text{or} \quad \int_0^t\int_0^s (\varepsilon + r^2)^{-\mu/2}\,dr\,ds$$

with an arbitrary parameter $\varepsilon > 0$. More generally, we can even consider μ-elliptic energies in the spirit, for example, of [26].

All the results stated below remain valid in the case of pure denoising for which the hypothesis (2) is replaced by the requirement $D = \varnothing$.

Theorem 1. *Consider a measurable function $f : \Omega - D \to [0,1]$ and fix a parameter $\alpha > 0$. Then the following assertions hold.*

(i) (*power growth*) *Let $\mu < 1$. Define $p := 2 - \mu$. Then the variational problem* (1) *with Ψ defined in* (3) *admits a unique solution u in the space $W^1_p(\Omega)$. The solution u satisfies $0 \leqslant u \leqslant 1$ and belongs to the class $C^{1,\beta}(\Omega)$ for any $\beta \in (0,1)$.*

(ii) (*logarithmic growth*) *If $\mu = 1$, then we have the results of* (i) *with $W^1_p(\Omega)$ replaced by $W^1_h(\Omega)$, $h(t) := t\ln(1+t)$.*

Proof. The statements of (i) follow by standard arguments as outlined, for example, in the textbook [33]. For (ii) we refer to [12, Theorems 1.1–1.3] and [26, Theorems 1.1 and 1.2]. ☐

We pass to the linear growth case for which we obtained the following result, according to [12, Theorems 1.4–1.8], [26, Theorems 1.3. and 1.4], [27, Theorems 1.2 and Corollary 1.1], and [29].

Theorem 2. *Consider f, α as in Theorem 1 and assume that $\mu > 1$. Then the following assertions hold.*

(i) *The problem (1) has at least one solution $u \in BV(\Omega)$, and each solution u satisfies $0 \leqslant u(x) \leqslant 1$ almost everywhere on Ω.*

(ii) *If u and $\widetilde{u}$ are J-minimizing in $BV(\Omega)$, then $u = \widetilde{u}$ almost everywhere on $\Omega - D$, $\nabla^a u = \nabla^a \widetilde{u}$ on Ω, and $|\nabla^s u|(\Omega) = |\nabla^s \widetilde{u}|(\Omega)$.*

(iii) $\inf\limits_{W_1^1(\Omega)} J = \inf\limits_{BV(\Omega)} J$.

(iv) *Let $\mathscr{M}$ denote the set of all L^1-cluster points of J-minimizing sequences from $W_1^1(\Omega)$. Then $\mathscr{M}$ coincides with the set of all* BV*-solutions of the problem* (1).

(v) *For any $u \in \mathscr{M}$ there is an open set $D_u \subset D$ with $\mathscr{L}^2(D - D_u) = 0$ and $u \in C^{1,\beta}(D_u)$ for any $\beta \in (0,1)$.*

(vi) *Let $\mu \in (1,2)$. Then* (1) *admits exactly one minimizer u being in addition of class $W_1^1(\Omega) \cap C^{1,\beta}(\Omega)$ for all $\beta \in (0,1)$.*

(vii) *If there exists $v \in \mathscr{M}$ belonging to the space $W_1^1(\Omega)$, then it follows that $\mathscr{M} = \{v\}$.*

(viii) *For $u, v \in \mathscr{M}$ we have the estimate*

$$\|u - v\|_{L^2(\Omega)} = \|u - v\|_{L^2(D)} \leqslant \frac{1}{2\sqrt{\pi}} |\nabla^s(u - v)|(\overline{D})\,.$$

An alternative approach towards the linear growth case $\mu > 1$ consists in an analysis of the dual variational problem. As in [27], we define the Lagrangian

$$l(v, \tau) := \alpha \int\limits_{\Omega} [\tau \cdot \nabla v - \Phi_\mu^*(|\tau|)]\, dx + \int\limits_{\Omega - D} (v - f)^2\, dx,$$

where $(v, \tau) \in W_1^1(\Omega) \times L^\infty(\Omega, \mathbb{R}^2)$. Here, Φ_μ^* denotes the conjugate function of Φ_μ. We consider the variational problem

$$R[\tau] := \inf_{v \in W_1^1(\Omega)} l(v, \tau) \to \max \ \text{in } L^\infty(\Omega, \mathbb{R}^2) \tag{9}$$

in duality to the problem (1). In [27, Theorems 1.4 and 1.5], we proved the following results (cf. [34] for further details).

Theorem 3. *Let $\mu > 1$, and let f, α be as in Theorem 1. Then the following assertions hold.*

(i) *The problem* (9) *admits a unique solution σ.*

(ii) *$\sigma \in W_{2,\mathrm{loc}}^1(\mathrm{int}\,(D), \mathbb{R}^2)$ as well as $\sigma = DF(\nabla^a u)$ almost everywhere on Ω. Here, we have abbreviated $F(\xi) = \Phi_\mu(|\xi|), \xi \in \mathbb{R}^2$, and u denotes any* BV*-minimizer of the problem* (1).

(iii) *The* inf-sup *relation holds, i.e.,* $\inf\limits_{W_1^1(\Omega)} J = \sup\limits_{L^\infty(\Omega, \mathbb{R}^2)} R$.

(v) *In the case $\mu \in (1,3)$, we have $\sigma \in C^{0,\beta}(\mathrm{int}(D), \mathbb{R}^2)$ for any $\beta \in (0,1)$.*

Up to now, the exponent μ occurring in formula (3) denotes a fixed real number. However, from the point of view of applications, it might be helpful to work with different values of μ

on prescribed subregions of Ω leading to a solution that is smooth on the zone $[\mu < 2]$ and with "irregular behavior" on parts of Ω with large values of $\mu(x)$. To be precise, we discuss this idea in the context of pure denoising working with densities involving variable exponents $\mu(x)$ generating functionals of linear growth. Assume that we are given a function $\mu = \mu(x)$ of class $C^1(\overline{\Omega})$ such that

$$\mu(x) \in (1, \infty), \quad x \in \overline{\Omega}. \tag{10}$$

We define $\Phi_{\mu(x)}(t)$ according to (3) and observe the validity of (4)–(7) for each $x \in \overline{\Omega}$. Let $F(x, \xi) = \Phi_{\mu(x)}(|\xi|)$. It is easy to check that the density F satisfies (i)–(iv) in [35, p. 312]. The reader should observe that from (10) together with the requirement $\mu \in C^1(\overline{\Omega})$ it actually follows that $\mu(x) \in [\mu_1, \mu_2]$ with suitable numbers $1 < \mu_1 \leqslant \mu_2$. For $u \in \mathrm{BV}(\Omega)$ we set

$$K[u] := \int\limits_\Omega F(x, \nabla^a u)\, dx + \int\limits_\Omega F_\infty\Big(x, \frac{\nabla^s u}{|\nabla^s u|}\Big) d|\nabla^s u|,$$

$$F_\infty(x, \xi) := \lim_{t\to\infty} \frac{F(x, t\xi)}{t}.$$

Then we have the following lower semicontinuity result (cf. Theorem 5.54 and the remarks on p. 313 in [35]): $K[u] \leqslant \liminf\limits_{n\to\infty} K[u_n]$ for each sequence $u_n \in \mathrm{BV}(\Omega)$ converging in $L^1(\Omega)$ to the BV-function u. In our particular case, we have $F_\infty(x, \xi) = \frac{1}{\mu(x)-1}|\xi|$, which is a direct consequence of formulas (4) and (6). Altogether we therefore look at the following variational problem as a model for pure denoising with energies of linear growth involving variable exponents:

$$I[u] := \int\limits_\Omega (f-u)^2\, dx + \alpha \int\limits_\Omega \Phi_{\mu(x)}(|\nabla^a u|)\, dx + \alpha \int\limits_\Omega \frac{1}{\mu(x)-1} d|\nabla^s u| \to \min \text{ in BV } (\Omega). \tag{11}$$

Theorem 4. *Consider a measurable function $f\colon \Omega \to [0,1]$, fix $\alpha > 0$, and assume that $\mu \in C^1(\overline{\Omega})$ satisfies* (10)*. Then the following assertions hold.*

(i) *The problem* (11) *admits a unique solution $u \in BV(\Omega)$ such that $0 \leqslant u(x) \leqslant 1$ for almost all $x \in \Omega$.*

(ii) *Let $\Omega_2 := x \in \Omega : \mu(x) < 2$ (possibly empty depending on the choice of μ). Then $u \in C^{1,\beta}(\Omega_2)$ for any $\beta \in (0,1)$.*

Proof. (i) The existence of an I-minimizer can be deduced from the preliminary remarks, the uniqueness is a consequence of the strict convexity of $v \mapsto \int\limits_\Omega (v-f)^2\, dx$, and the "maximum-principle" follows as in [12].

(ii) On Ω_2, the continuity of $\dfrac{\partial F}{\partial \xi}(\cdot, \nabla u)$ follows with similar arguments as applied in [26] leading to the continuity of ∇u on Ω_2 modulo a small singular set. In order to rule out these singularities, the reader should consult [29] applying minor adjustments: as usual, the approach towards the regularity of the I-minimizing function $u \in \mathrm{BV}(\Omega)$ is based on the analysis of a suitable regularizing sequence $\{u_\delta\}$, i.e., for $\delta > 0$ one studies the problem

$$I_\delta[w] := \frac{\delta}{2} \int\limits_\Omega |\nabla w|^2\, dx + I[w] \to \min \text{ in } W^1_2(\Omega)$$

with a unique solution u_δ showing a nice behavior. In order to justify $u_\delta \to u$ in $L^1(\Omega)$ as $\delta \to 0$, we first observe that there exists $\overline{u}$ $\in$BV(Ω) such that $u_\delta \to \overline{u}$ in $L^1(\Omega)$ at least for a subsequence. We claim the validity of $u = \overline{u}$. In fact, the lower semicontinuity of the functional I with respect to the BV-convergence gives $I[\overline{u}] \leqslant \liminf\limits_{\delta\to 0} I[u_\delta]$. Hence, by the I-minimality of u, we find $I[u] \leqslant I[\overline{u}] \leqslant \liminf\limits_{\delta\to 0} I_\delta[u_\delta]$. By [36, Proposition 2.3], we can choose a sequence $\{u_m\}$ from $C^\infty(\Omega) \cap W^1_1(\Omega)$ such that

$$u_m \to u \text{ in } L^1(\Omega), \qquad \int\limits_\Omega \sqrt{1 + |\nabla u_m|^2}\, dx \to \int\limits_\Omega \sqrt{1 + |\nabla u|^2}\,.$$

Actually, the sequence $\{u_m\}$ can be taken from the space $C^\infty(\overline{\Omega})$ as it outlined in [34, Lemma 2.2].

The I_δ- minimality of u_δ shows that $I_\delta[u_\delta] \leqslant I_\delta[u_m]$. Therefore,

$$\liminf_{\delta\to 0} I_\delta[u_\delta] \leqslant I[u_m]\,,$$

which means $I[u] \leqslant I[\overline{u}] \leqslant I[u_m]$. We note (cf., for example, [37, 38] or [39, Theorem 4.1]) that the functional K is continuous with respect to the convergence $u_m \to u$. Thus, $I[u] = I[\overline{u}]$ since

$$\int\limits_\Omega (u_m - f)^2\, dx \to \int\limits_\Omega (u - f)^2\, dx$$

follows from $u_m \to u$ in $L^1(\Omega)$ and the observation that due to $0 \leqslant u \leqslant 1$ we may assume the validity of the same inequality for u_m which is a direct consequence of the definition of u_m. This implies $u = \overline{u}$ by the uniqueness of a minimizer. □

We note that the assumption $\mu \in C^1(\overline{\Omega})$ can be weakened, but we omit details.

The idea of denoising an observed image f: $\Omega \to [0, 1]$ in such a way that the solution $u : \Omega \to [0, 1]$ shows a different degree of smoothness on prescribed subregions of Ω can also be made precise by considering the energy density

$$G(x, \xi) := \eta(x)|\xi| + (1 - \eta(x))\Phi_{\mu(x)}(|\xi|), \quad x \in \overline{\Omega},\ \xi \in \mathbb{R}^2\,, \tag{12}$$

where η denotes some function in $C^1(\overline{\Omega})$ such that $0 \leqslant \eta(x) \leqslant 1$. In place of (11) we have to discuss the problem (using the natural definition of $\int\limits_\Omega G(x, \nabla u)$ on BV (Ω))

$$\begin{aligned}&\int\limits_\Omega (u - f)^2\, dx + \alpha \int\limits_\Omega \eta(x) d|\nabla u| + \alpha \int\limits_\Omega (1 - \eta(x))\Phi_{\mu(x)}(|\nabla^a u|)\, dx \\ &+ \alpha \int\limits_\Omega \frac{1 - \eta(x)}{\mu(x) - 1} d|\nabla^s u| \to \min \text{ in BV } (\Omega)\,,\end{aligned} \tag{13}$$

which means that on the set $[\eta = 1]$ we actually apply a TV-regularization. Along the lines of Theorem 4 the following results can be established:

Theorem 5. *Consider f, α, μ as in Theorem 4 and define G according to (12) with η as before. Then the problem*

$$\int_\Omega (f-u)^2\,dx + \alpha \int_\Omega G(x,\nabla u) \to \min \ \textit{ in } \ \mathrm{BV}(\Omega)$$

as stated in (13) admits a unique solution u such that $0 \leqslant u \leqslant 1$ almost everywhere. Moreover, $u \in C^{1,\beta}(\Omega^*)$ for any $\beta \in (0,1)$, where Ω^* denotes the interior of the set $\{x \in \Omega : \eta(x) = 0 \text{ and } \mu(\mathrm{x}) < 2\}$.

After this overview on the theoretical properties, let us discuss some numerical aspects and pass to a computational experiment. The goal is to illustrate the behavior of the model (1), (3) for different parameters μ and compare it to alternative approaches used in the literature. We focus on the pure denoising problem, i.e., $D = \varnothing$.

A minimizer of the energy (1) with penalizing function $\Psi(t) = \Phi_\mu(t)$ satisfies necessarily an Euler–Lagrange equation of type

$$u - f - \alpha \,\mathrm{div}\,(g_\mu(|\nabla u|)\,\nabla u) = 0 \tag{14}$$

with the homogeneous Neumann boundary condition

$$\partial_n u = 0 \quad \text{on } \partial\Omega, \tag{15}$$

where n denotes the normal vector to the image boundary $\partial\Omega$. For $\mu \neq 1$ and $\mu \neq 2$ the diffusivity function $g_\mu(t) := \dfrac{\Phi'_\mu(t)}{2t}$ in (14) is given by

$$g_\mu(t) = \begin{cases} 1/2, & t = 0, \\ \dfrac{1-(1+t)^{1-\mu}}{2\,(\mu-1)\,t} & \text{otherwise.} \end{cases} \tag{16}$$

For $\mu = 1$

$$g_1(t) = \begin{cases} 1/2, & t = 0, \\ \dfrac{\ln(1+t)}{2t} & \text{otherwise,} \end{cases} \tag{17}$$

and $\mu = 2$ yields

$$g_2(t) = \frac{1}{2\,(1+t)}\,. \tag{18}$$

We discretize this problem with a finite difference method [40, 41] on a regular pixel grid of size h in both directions. This leads to a nonlinear system of equations with the following structure:

$$(\boldsymbol{I} - \alpha \boldsymbol{A}(\boldsymbol{u}))\,\boldsymbol{u} = \boldsymbol{f}, \tag{19}$$

where N denotes the number of pixels, $\boldsymbol{I} \in \mathbb{R}^{N\times N}$ is the unit matrix, the unknown vector $\boldsymbol{u} \in \mathbb{R}^N$ contains the grey values of the processed image u in all pixels, and $\boldsymbol{f} \in \mathbb{R}^N$ is a discretization of the noisy image f. The term $\boldsymbol{A}(\boldsymbol{u})\,\boldsymbol{u}$ represents the discrete counterpart of $\mathrm{div}\,(g_\mu(|\nabla u|)\,\nabla u)$, where the matrix-valued function $\boldsymbol{A} : \mathbb{R}^N \to \mathbb{R}^{N\times N}$ depends in a nonlinear way on the unknown image $\boldsymbol{u}$. It also incorporates the homogeneous Neumann boundary conditions.

In order to solve the nonlinear system (19) numerically, we use an iterative algorithm that replaces it by a sequence of linear systems of equations:

$$\boldsymbol{u}^{(0)} = \boldsymbol{f}, \tag{20}$$

$$(\boldsymbol{I} - \alpha \boldsymbol{A}(\boldsymbol{u}^{(k-1)}))\boldsymbol{u}^{(k)} = \boldsymbol{f}, \quad k = 1, 2, ..., \tag{21}$$

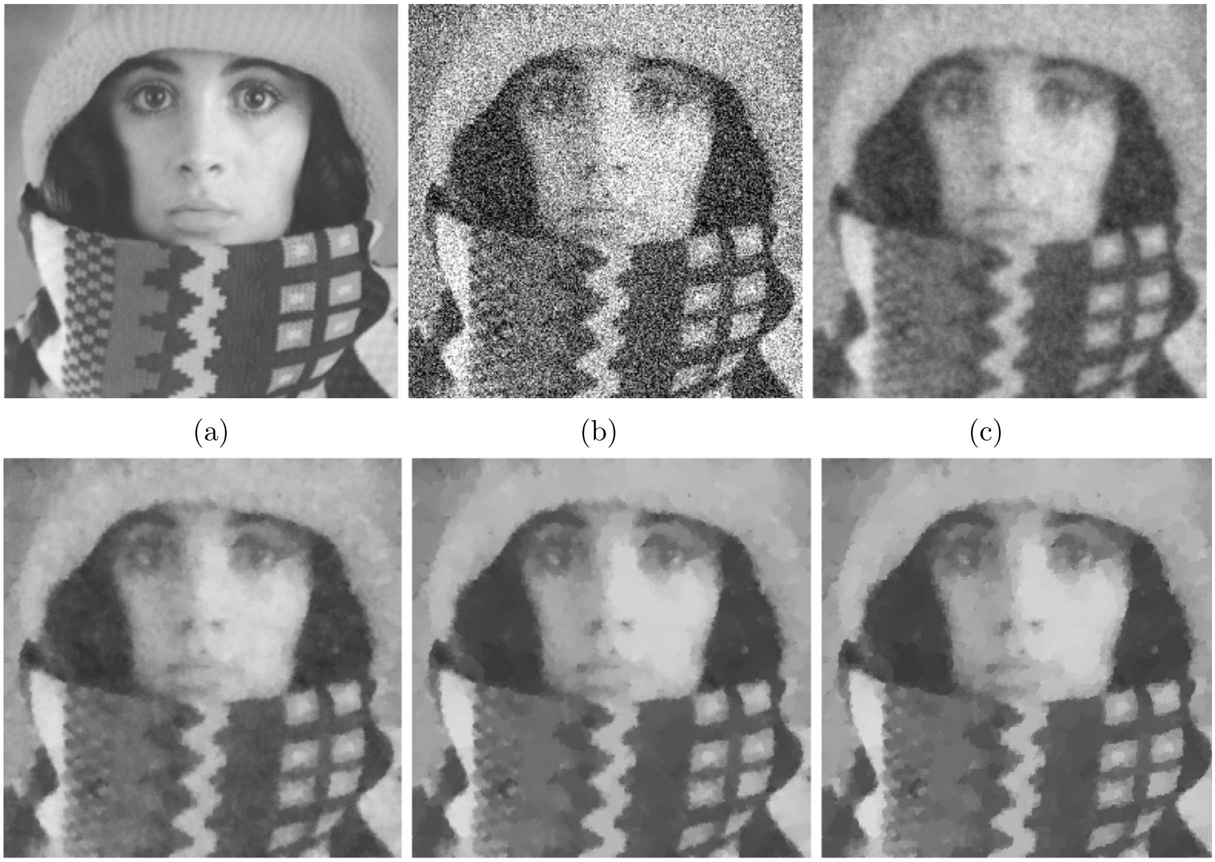
where $\boldsymbol{u}^{(k)}$ denotes the solution of the kth iteration step. Approaches of this type are known under the name *Kačanov method* [42] or *lagged diffusivity method*. As is shown in [43], they are equivalent to the half-quadratic regularization algorithms of Geman and Yang [44], and an analysis of their convergence properties can be found, for example, in [45]. With Gershgorin's theorem [46] it follows that for any $\alpha \geqslant 0$ and any vector $\boldsymbol{u}^{(k-1)}$ the system matrix of (21) is positive definite. Hence each linear system has a unique solution. Since we use a 4-neighborhood for the discretization of $\operatorname{div}(g_\mu(|\nabla u|)\,\nabla u)$, the system matrices have a sparse structure with at most five nonvanishing entries in each row. A classical iterative solver such as the Gauss–Seidel method constitutes a simple numerical algorithm that benefits from this sparsity and is guaranteed to converge for positive definite system matrices [41]. We stop our Gauss–Seidel iterations when the Euclidean norm of the residual of (21) has decreased by a specified factor ε. We choose $\varepsilon := 10^{-4}$. In a similar way, we can check the residual of the nonlinear system (19) in order to stop the outer Kačanov iterations.

Figure 1 shows an application of our approach to the denoising of a digital greyscale image. Since each grey value of this image is encoded by a single byte, we consider the range $[0, 255]$ instead of $[0, 1]$. Of course, all theoretical results that we presented before an unaffected by this rescaling. For our denoising experiment we add Gaussian noise with zero mean to the original image, and we evaluate different parameter settings in order to recover the noise-free image as good as possible. As a measure of the approximation quality of some restoration $\boldsymbol{u} = (u_i)_{i=1}^N \in \mathbb{R}^N$ with respect to the noise-free image $\boldsymbol{v} = (v_i)_{i=1}^N \in \mathbb{R}^N$ we use the *mean squared error* (MSE)

$$\mathrm{MSE}(\boldsymbol{u}, \boldsymbol{v}) := \frac{1}{N}\sum_{i=1}^{N}(u_i - v_i)^2. \tag{22}$$

For each chosen parameter μ we optimize the regularization parameter α such that the MSE is minimized. Figure 1 (c)–(e) shows the results of our approach for $\mu = 0$, 1.5, and 20. The case $\mu = 0$ represents the classical Whittaker–Tikhonov regularization [47]–[49]. It minimizes a quadratic energy, which comes down to a linear Euler–Lagrange equation with constant diffusivity $g_0 = 1/2$. Its MSE is given by 155.82. Such linear methods are well known to suffer from blurring of semantically important structures such as edges. Using a nonlinear approach with $\mu = 1.5$ allows a better preservation of edges since its diffusivity decreases with $|\nabla u|$. This is rewarded by a lower MSE of 122.96. These results can be improved by allowing larger values for μ: Choosing $\mu = 20$ yields an MSE of 114.70. Increasing μ any further, however, would deteriorate the results again. Interestingly, our reconstructions for $\mu = 1.5$ and $\mu = 20$ also outperform the popular TV-regularization [10] that uses the nondifferentiable density $\Psi(|\nabla u|) = |\nabla u|$ which corresponds to the singular diffusivity $g(|\nabla u|) = \frac{1}{2|\nabla u|}$: An implementation of TV-regularization by means of the FISTA algorithm [50] gives a worse MSE of 124.48. The result is depicted in

Figure 1 (f). Thus, our experiment illustrates that staying a bit away from the TV limit can be beneficial for obtaining better restorations.



(a) (b) (c)

(d) (e) (f)

FIGURE 1. Denoising experiment. (a) Test image, 256×256 pixel, range $[0, 255]$. (b) Degraded by additive Gaussian noise with zero mean and standard deviation $\sigma = 49.49$. (c) Whittaker–Tikhonov regularization ($\mu = 0$) of the noisy image (b) with regularization parameter $\alpha = 5.8$, yielding an MSE of 155.82. (d) Result for $\mu = 1.5$ and $\alpha = 54$. MSE = 122.96. (e) Result for $\mu = 20$ and $\alpha = 1410$. MSE = 114.70. (f) Result for TV-regularization with $\alpha = 86$. MSE = 124.48.

References

1. R. Acar and C. R., Vogel, "Analysis of bounded variation penalty methods for ill-posed problems," *Inverse Probl.* **10**, No. 6, 1217–1229 (1994).

2. G. Aubert and L. Vese, "A variational method in image recovery," *SIAM J. Numer. Anal.* **34**, No. 5, 1948–1979 (1997).

3. P. Blomgren, T. F. Chan, P. Mulet, L. Vese, and W. L. Wan, "Variational PDE models and methods for image processing," In: *Numerical Analysis 1999*, pp. 43–67, Chapman and Hall/CRC, Boca Raton, FL (2000).

4. V. Caselles, A. Chambolle, and M. Novaga, "Regularity for solutions of the total variation denoising problem," *Rev. Mat. Iberoam.* **27**, No. 1, 233–252 (2011).

5. A. Chambolle and P.-L. Lions, "Image recovery via total variation minimization and related problems," *Numer. Math.* **76**, No. 2, 167–188 (1997).

6. T. F. Chan and S. Esedoglu, "Aspects of total variation regularized L^1 function approximation," *SIAM J. Appl. Math.* **65**, No. 5, 1817–1837 (2005).

7. T. Chan, J. Shen, and L. Vese, "Variational PDE models in image processing," *Notices Am. Math. Soc.* **50**, No. 1, 14–26 (2003).

8. Y. Chen, S. Levine, and M. Rao, "Variable exponent, linear growth functionals in image restoration," *SIAM J. Appl. Math.* **66**, No. 4, 1383–1406 (2006).

9. B. Kawohl, "Variational versus PDE-based approaches in mathematical image processing," *CRM Proc. Lect. Notes* **44**, 113–126 (2008).

10. L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D* **60**, 259–268 (1992).

11. L. Vese, "A study in the BV space of a denoising-deblurring variational problem," *Appl. Math. Optimization* **44**, No. 2, 131–161 (2001).

12. M. Bildhauer and M. Fuchs, "A variational approach to the denoising of images based on different variants of the TV-regularization," *Appl. Math. Optim.* **66**, No. 3, 331–361 (2012).

13. P. Arias, V. Caselles, G. Facciolo, V. Lazcano, and R. Sadek, "Nonlocal variational models for inpainting and interpolation," *Math. Models Methods Appl. Sci.* **22**, Suppl.2, 1230003 (2012).

14. P. Arias, V. Caselles, and G. Sapiro, "A variational framework for non-local image inpainting," In: *Energy Minimization Methods in Computer Vision and Pattern Recognition*, pp. 345–358, Springer, Berlin (2009).

15. P. Arias, G. Facciolo, V. Caselles, and G. Sapiro, "A variational framework for exemplar-based image inpainting," *Int. J. Comput. Vis.* **93**, No. 3, 319–347 (2011).

16. G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing*, Springer, New York (2002).

17. M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," In: *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 417–424,

18. M. Burger, L. He, and C.-B. Schönlieb, "Cahn-Hilliard inpainting and a generalization for grayvalue images," *SIAM J. Imaging Sci.* **2**, No. 4, 1129–1167 (2009).

19. T. F. Chan, S. H. Kang, and J. Shen, "Euler's elastica and curvature based inpaintings," *SIAM J. Appl. Math.* **63**, No. 2, 564–592 (2002).

20. T. F. Chan and J. Shen, "Nontexture inpainting by curvature-driven diffusions," *J. Visual Commun. Image Represen.* **12**, No. 4, 436–449 (2001).

21. T. F. Chan and J. Shen, "Mathematical models for local nontexture inpaintings," *SIAM J. Appl. Math.* **62**, No. 3, 1019–1043 (2001).

22. S. Esedoglu and J. Shen, "Digital inpainting based on the Mumford-Shah-Euler image model," *European J. Appl. Math.* **13**, No. 4, 353–370 (2002).

23. K. Papafitsoros, B. Sengul, and C.-B. Schönlieb, *Combined First and Second Order Total Variation Impainting Using Split Bregman*, IPOL Preprint (2012).

24. J. Shen, "Inpainting and the fundamental problem of image processing," *SIAM News* **36**, No. 5, 1–4 (2003).

25. M. Bildhauer and M. Fuchs, "Image inpainting with energies of linear growth. A collection of proposals," *J. Math. Sci., New York* **196**, No. 4, 490–497 (2014).

26. M. Bildhauer and M. Fuchs, "On some perturbations of the total variation image inpainting method. Part I: Regularity theory," *J. Math. Sci., New York* **202**, No. 2, 154–169 (2014).

27. M. Bildhauer and M. Fuchs, "On some perturbations of the total variation image inpainting method. Part II: Relaxation and dual variational formulation," *J. Math. Sci., New York* **202**, No. 2, 121–140 (2015).

28. M. Bildhauer and M. Fuchs, "On some perturbations of the total variation image inpainting method. Part III: Minimization among sets with finite perimeter," *J. Math. Sci., New York* **207**, No. 2, 142–146 (2015).

29. M. Bildhauer, M. Fuchs, and C. Tietz, "$C^{1,\alpha}$-interior regularity for minimizers of a class of variational problems with linear growth related to image inpainting," *St. Petersb. Math. J.* **27**, No. 3, 381–392 (2016).

30. R. A. Adams, *Sobolev Spaces*, Academic Press, New York (1975),

31. E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*, Birkhäuser, Basel etc. (1984).

32. F. Demengel and R. Temam, "Convex functions of a measure and applications," *Ind. Univ. Math. J.* **33**, 673–709 (1984).

33. O. A. Ladyzhenskaya and N. N. Ural'tseva, *Linear and Quasilinear Elliptic Equations* [in Russian], Nauka, Moscow (1964); English transl.: Academic Press, New York (1968).

34. M. Fuchs and C. Tietz, "Existence of generalized minimizers and of dual solutions for a class of variational problems with linear growth related to image recovery," *J. Math. Sci., New York* **210**, No. 4, 458–475 (2015).

35. L. Ambrosio, N. Fusco, and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Clarendon Press, Oxford (2000).

36. G. Anzellotti and M. Giaquinta, "Convex functionals and partial regularity," *Arch. Rat. Mech. Anal.* **102**, 243–272 (1988).

37. Yu. G. Reshetnyak, "Weak convergence of completely additive vector functions on a set" [in Russian], *Sib. Mat. Zh.* **9**, No. 6, 1386–1394 (1968); English transl.: *Sib. Math. J.* **9**, No. 6, 1039–1045 (1968).

38. D. Spector, "Simple proofs of some results of Reshetnyak," *Proc. Am. Math. Soc.* **139**, No. 5, 1681–1690 (2011).

39. L. Beck and T. Schmidt, "On the Dirichlet problem for variational integrals in BV," *J. Reine Angew. Math.* **674**, 113–194 (2013).

40. R. J. LeVeque, *Finite Difference Methods for Ordinary and Partial Differential Equations*, SIAM, Philadelphia, PA (2007).

41. K. W. Morton and L. M. Mayers, *Numerical Solution of Partial Differential Equations. An Introduction*, Cambridge University Press, Cambridge (1994).

42. S. Fučik, A. Kratochvil, and J. Nečas, “Kačanov–Galerkin method,” *Commentat. Math. Univ. Carol.* **14**, 651–659 (1973).

43. J. Weickert, J. Heers, C. Schnörr, K. J. Zuiderveld, O. Scherzer, and H. S. Stiehl, “Fast parallel algorithms for a broad class of nonlinear variational diffusion approaches,” *Real-Time Imaging* **7**, No. 1, 31–45 (2001).

44. D. Geman and C. Yang, “Nonlinear image recovery with half-quadratic regularization,” *IEEE Trans. Image Process. 4*, 932–945 (1995).

45. E. Zeidler, *Nonlinear Functional Analysis and its Applications II/A: Linear Monotone Operators*, Springer, New York etc. (1990).

46. S. Gerschgorin, Fehlerabschätzung für das Differenzenverfahren zur Lösung Partieller Differentialgleichungen. *Z. Angew. Math.* **10**, 373–382 (1930).

47. E. T. Whittaker, “A new method of graduation,” *Proc. Edinburgh Math. Soc.* **41**, 65–75 (1923).

48. A. N. Tikhonov, “Solution of incorrectly formulated problems and the regularization method” [in Russian], *Dokl. Akad. Nauk SSSR*, **151**, 501–504 (1963); English transl.: *Sov. Math., Dokl.* **5**, 1035–1038 (1963).

49. M. Bertero, T. A. Poggio, and V. Torre, “Ill-posed problems in early vision,” *Proc. IEEE* **76**, 869–889 (1988).

50. A. Beck and M. Teboulle, “Fast iterative shrinkage-thresholding algorithm for linear inverse problems,” *SIAM J. Imaging Sci.* **2**, No. 1, 183–202 (2009).

Submitted on June 13, 2016