J Math Imaging Vis (2014) 48:295–307
DOI 10.1007/s10851-013-0420-0

Fully Smoothed ℓ_1-TV Models: Bounds for the Minimizers and Parameter Choice

F. Baus · M. Nikolova · G. Steidl

Published online: 6 February 2013

Abstract We consider a class of convex functionals that can be seen as $\mathcal{C}^1$ smooth approximations of the ℓ_1-TV model. The minimizers of such functionals were shown to exhibit a qualitatively different behavior compared to the nonsmooth ℓ_1-TV model (Nikolova et al. in Exact histogram specification for digital images using a variational approach, 2012). Here we focus on the way the parameters involved in these functionals determine the features of the minimizers $\hat{u}$. We give explicit relationships between the minimizers and these parameters.

Given an input digital image f, we prove that the error $\|\hat{u} - f\|_\infty$ obeys $b - \varepsilon \leq \|\hat{u} - f\|_\infty \leq b$ where b is a constant independent of the input image. Further we can set the parameters so that $\varepsilon > 0$ is arbitrarily close to zero. More precisely, we exhibit explicit formulae relating the model parameters, the input image f and the values b and ε. Conversely, we can fix the parameter values so that the error $\|\hat{u} - f\|_\infty$ meets some prescribed b, ε. All theoretical results are confirmed using numerical tests on natural digital images of different sizes with disparate content and quality.

Keywords Parameter estimation for smoothed ℓ_1-TV model · Histogram specification · Quantisation noise · ℓ_∞ error · Convex optimization

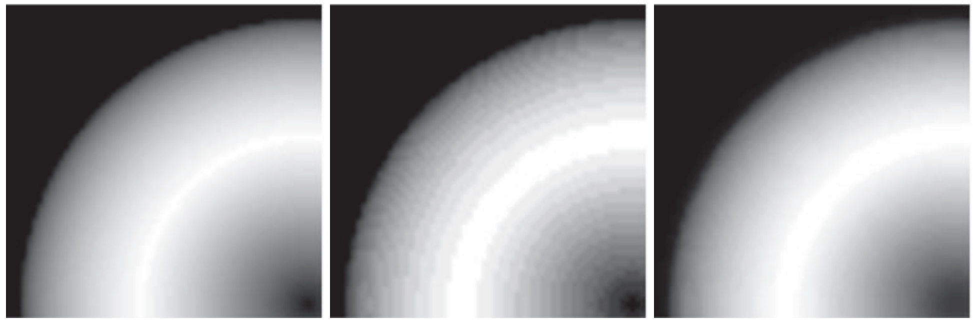
1 Introduction

In [11] a variational method using $\mathcal{C}^2$ smoothed ℓ_1-TV functionals were proposed. The goal was to process digital (quantized) images so that the obtained minimizer is quite close to the input digital image but its pixels are real-valued and can be ordered in a strict way. Indeed, the obtained minimizers were shown to enable faithful *exact histogram specification* outperforming the state-of-the-art methods [7, 12]. The intuition behind these functionals was that their minimizer can up to some degree remove some quantization noise and in this way yield an ordering of the pixels close to the unknown original real-valued image. Such an effect can be observed in Fig. 1 where a synthetic real-valued image is quantized and then restored using the proposed variational method. The *nonsmooth* L_1-TV model was originally studied in [5]. The main feature of its minimizers is that they contain parts that are equal to the data image and parts that are constant (living in a vanishing component of the TV term). Even though the model modification proposed in [11] might seem trivial, the minimizers of these $\mathcal{C}^2$ smoothed ℓ_1-TV functionals exhibit a *qualitatively* different behavior. Unlike the L_1-TV (ℓ_1-TV) minimizers, it was shown in [11] that the minimizers of the $\mathcal{C}^2$ smoothed ℓ_1-TV functionals generically do not have pixels equal to those of the data image and there are no equally valued pixels. Some of the authors of [11] observed that once the parameters of the model were fixed, for all kind of real-world digital images f, the residual error obeyed $\|\hat{u} - f\|_\infty = b$ where the constant b typically met $b < 0.5$. For this reason, they qualified

F. Baus
Fraunhofer ITWM, Fraunhofer Platz 1, 67663 Kaiserslautern, Germany

M. Nikolova
CMLA, ENS Cachan, CNRS, 61 av. du President Wilson, 94235 Cachan Cedex, France
e-mail: nikolova@cmla.ens-cachan.fr

G. Steidl (✉)
Department of Mathematics, University of Kaiserslautern, Paul-Ehrlich-Str. 31, 67663 Kaiserslautern, Germany
e-mail: steidl@mathematik.uni-kl.de



Real-valued image f—quantized on $\{0, \cdots, 15\}$ Restored $\hat{u}$

Fig. 1 The restored image is obtained by minimizing $J(\cdot, f)$ of the form (1) where $\psi(t) = \sqrt{t^2 + \alpha_1}$ and $\varphi(t) = \sqrt{t^2 + \alpha_2}$ for $\mathcal{N}8$

this variational approach as *detail preserving*. Therefore we were interested in monitoring the error $\|\hat{u} - f\|_\infty$.

In this paper we consider a wider class of $\mathcal{C}^1$ smoothed ℓ_1-TV functionals involving also ℓ_2 data fidelity terms. We give explicit relationships between the minimizers and the parameters tuning the model. The observation that $\|\hat{u} - f\|_\infty = b$, up to a small difference, is independent of the input image, is confirmed theoretically. Clear indications on the role of the parameter setting and the lower and upper bounds of $\|\hat{u} - f\|_\infty$ enable us to give restrictions on the parameter selection. All theoretical results are confirmed using numerical tests on a set of digital images of different sizes with disparate content and quality.

In spite of the progress in nonsmooth convex optimization [4], smooth approximations of nonsmooth objectives still remain a common approach in optimization [2]. Our results can help to design smooth approximations of ℓ_1/ℓ_2-TV functionals in a proper way.

The outline of this paper is as follows: In the next Sect. 2 we describe the variational model. Then, in Sect. 3 we estimate the ℓ_∞-error between the input image f and the minimizer of the functional. Section 4 provides explicit parameter estimates for the model. In Sect. 5 we give probability estimates for the behavior of neighboring pixels. Numerical tests demonstrate the quality of our estimates in Sect. 6. Finally, Sect. 7 finishes with conclusions and perspectives.

2 The Fully Smoothed ℓ_1-TV Model

We consider $M \times N$ digital images f with gray values in $\{0, \dots, L-1\}$. Let $n := MN$. To simplify the notation we reorder the image columnwise into a vector of size n and address the pixels by the index set $\mathbb{I}_n := \{1, \dots, n\}$. Further, we denote by $\mathbb{I}_n^{\text{int}} \subset \mathbb{I}_n$ the subset of all inner pixels, i.e., all pixels which are not boundary pixels.

We are interested in the minimizer $\hat{u}$ of a functional of the form

$$J(u, f) := \Psi(u, f) + \beta\Phi(u), \quad \beta > 0 \tag{1}$$

with

$$\Psi(u, f) := \sum_{i \in \mathbb{I}_n} \psi\big(u[i] - f[i]\big),$$

$$\Phi(u) := \sum_{i \in \mathbb{I}_n} \sum_{j \in \mathcal{N}_i} \varphi\big(\gamma_{i,j}\big(u[i] - u[j]\big)\big),$$

where $\mathcal{N}_i$ is a neighborhood of pixel i, the $\gamma_{i,j} > 0$ are weighting terms for the distance between neighbors, and the functions ψ and φ depend on a positive parameter, α_1 and α_2, respectively. To emphasize this dependence we use the notation $\psi(\cdot, \alpha_1)$ and $\varphi(\cdot, \alpha_2)$ when necessary. So $\psi : \mathbb{R} \times (0, +\infty) \to \mathbb{R}$ and $\varphi : \mathbb{R} \times (0, +\infty) \to \mathbb{R}$. The functions ψ and φ have to fulfill the properties stated below:

H0 The functions $t \mapsto \psi(t, \alpha_1)$ and $t \mapsto \varphi(t, \alpha_2)$ are continuously differentiable and even.
We denote

$$\psi'(t, \alpha_1) := \frac{d}{dt}\psi(t, \alpha_1) \quad \text{and}$$

$$\varphi'(t, \alpha_2) := \frac{d}{dt}\varphi(t, \alpha_2).$$

When it is clear from the context, we write $\psi'(t)$ for $\psi'(t, \alpha_1)$ and $\varphi'(t)$ for $\varphi'(t, \alpha_2)$. By H0, $\psi'(t)$ and $\varphi'(t)$ are continuous and odd functions.
These derivative functions have to satisfy certain conditions given next.

H1$^\psi$ $t \mapsto \psi'(t, \alpha_1) : \mathbb{R} \to (-Y, Y)$, where $Y > 0$, is a strictly increasing function for any fixed $\alpha_1 \in (0, +\infty)$ and maps onto $(-Y, Y)$.

H2$^\psi$ There is a constant $T > 0$ such that for any fixed $t \in (0, T)$, the function $\alpha_1 \mapsto \psi'(t, \alpha_1)$ is strictly decreasing on $(0, +\infty)$.
Here the cases $T = +\infty$ and $Y = +\infty$ are included.

H1$^\varphi$ $t \mapsto \varphi'(t, \alpha_2)$ is an increasing function for any fixed $\alpha_2 \in (0, +\infty)$ satisfying

$$\lim_{t \to \infty} \varphi'(t, \alpha_2) = 1.$$

H2$^\varphi$ For any fixed $t > 0$, the function $\alpha_2 \mapsto \varphi'(t, \alpha_2)$ is continuous and decreasing on $(0, +\infty)$ and

$$\lim_{\alpha_2 \searrow 0} \varphi'(t, \alpha_2) = 1.$$

These properties imply further useful relations which are collected in the following remark.

Remark 1 (i) By H1$^\psi$ we know that ψ is strictly convex and monotone increasing on $(0, +\infty)$ and by H1$^\varphi$ that φ is convex. Therefore there exists a unique minimizer of (1). This minimizer can be computed, e.g. by using a Weiszfeld-like semi-implicit algorithm, or the nonlinear (preconditioned) conjugate gradient method, see [6, 11, 13], among other viable algorithms.

(ii) By H1$^\psi$ there exists the inverse function $(\psi')^{-1}(\cdot, \alpha_1) : (-Y, Y) \to \mathbb{R}$, and this function is also odd, continuous and strictly increasing.

Some relevant choices of functions θ obeying all properties H0, H1$^\psi$, H2$^\psi$, H1$^\varphi$ and H2$^\varphi$ are given in Table 1. For the latter functions, $t \mapsto \theta'(t, \alpha)$ maps onto $(-1, 1)$, i.e., $Y = 1$ and $T = +\infty$ for any $\alpha > 0$. A typical graph of such a function, its derivative and inverse derivative is depicted in Fig. 2.

Another choice for ψ fulfilling H0, H1$^\psi$ and H2$^\psi$ is the scaled ℓ_p-norm for $p = \alpha_1 + 1$:

$$\psi(t) := \frac{1}{\alpha_1 + 1} |t|^{\alpha_1 + 1} \quad \text{with}$$

Table 1 Options for functions θ obeying all the assumptions stated above. These functions are nearly affine beyond a neighborhood of zero. The size of the latter neighborhood is controlled by the parameter $\alpha > 0$

	θ	θ'	$(\theta')^{-1}$
$\Theta 1$	$\sqrt{t^2 + \alpha}$	$\frac{t}{\sqrt{t^2+\alpha}}$	$y\sqrt{\frac{\alpha}{1-y^2}}$
$\Theta 2$	$\lvert t\rvert - \alpha \log(1 + \frac{\lvert t\rvert}{\alpha})$	$\frac{t}{\alpha+\lvert t\rvert}$	$\frac{\alpha y}{1-\lvert y\rvert}$
$\Theta 3$	$\alpha \log(\cosh(\frac{t}{\alpha}))$	$\tanh(\frac{t}{\alpha})$	$\alpha \operatorname{atanh}(y)$

sign(y)

$$\psi'(t) = |t|^{\alpha_1} \operatorname{sign}(t), \quad (\psi')^{-1}(y) = |y|^{\frac{1}{\alpha_1}}, \quad \alpha_1 > 0. \tag{2}$$

Here ψ' maps onto $\mathbb{R}$ so that $Y = +\infty$. Moreover $\alpha_1 \mapsto \psi'(t, \alpha_1)$ is strictly monotone decreasing for $|t| < 1$ hence $T = 1$ in this case. An upper bound for $\|\hat{u} - f\|_\infty$ when $\alpha_1 = 1$ in (2) was derived in [10]. Some general results on the functionals J for $\alpha_1 = 1$ in (2) can be found in [1] in a continuous setting.

For φ we can also use the *scaled Huber function*

$$\varphi(t) := \begin{cases} \frac{t^2}{2\alpha_2} & \text{if } |t| \le \alpha_2, \\ |t| - \frac{\alpha_2}{2} & \text{if } |t| > \alpha_2 \end{cases} \quad \text{with}$$

$$\varphi'(t) = \begin{cases} \frac{t}{\alpha_2} & \text{if } |t| \le \alpha_2, \\ \operatorname{sign}(t) & \text{if } |t| > \alpha_2. \end{cases} \tag{3}$$

Note that the functions ψ and φ in Table 1 and (3) are nearly affine beyond a small neighborhood of the origin.

In this paper, we focus on the neighborhoods $\mathcal{N}4$ and $\mathcal{N}8$ depicted in Fig. 3 top. When taking the gradient of the functional in (1) we have to take into account that the pixel combination $u[i] - u[j]$ appears for $j \in \mathcal{N}_i^2$, where $\mathcal{N}_i^2$ denotes the "double" neighborhood associated with $\mathcal{N}_i$ in Fig. 3 bottom. The usual choices are (see e.g. [8])

$$\begin{aligned} \gamma_{i,j} &:= 1 && \text{for vertical and horizontal neighbors,} \\ \gamma_{i,j} &:= \frac{1}{\sqrt{2}} && \text{for diagonal neighbors.} \end{aligned} \tag{4}$$

In all cases we have $\gamma_{i,j} = \gamma_{j,i}$.

Functionals of the form (1) with functions $\psi, \varphi \in \mathcal{C}^s$, $s \ge 2$ having alike properties (e.g. all functions in Table 1) were successfully used in [11] to process digital images f so that the obtained minimizer $\hat{u}$ is quite close to the input digital image but its pixels can be ordered in a strict way. An analysis of the minimizers $\hat{u}$ of these functionals has shown that with a probability close to one, $\hat{u}$ has pixel values that are different from each other and different from the input pixels.

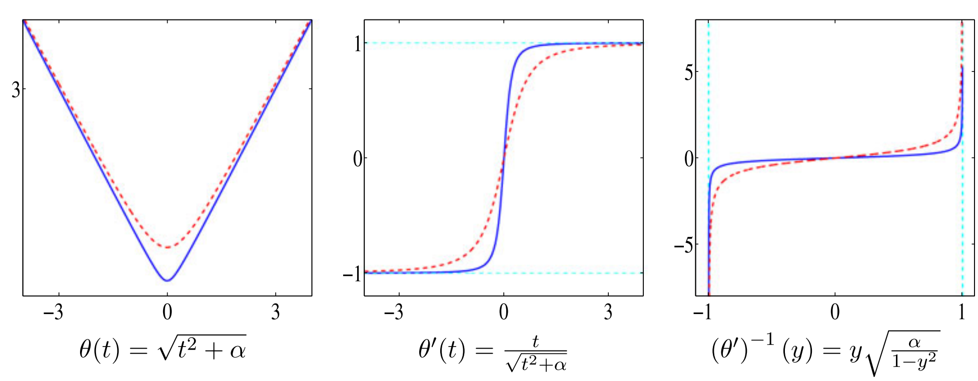


Fig. 2 The function $\Theta 1$ in Table 1, where the plots for $\alpha = 0.05$ are in *solid line* and for $\alpha = 0.5$ in *dashed line*

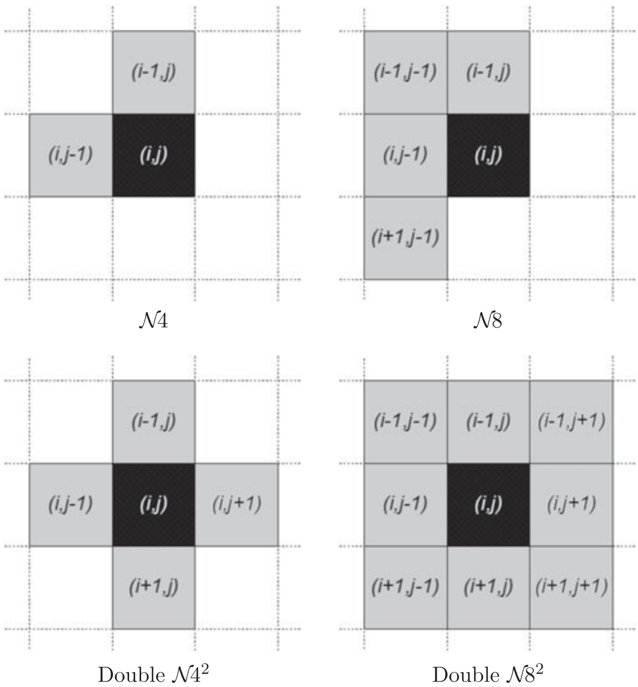


Fig. 3 Neighborhoods $\mathcal{N}4$ (*left*) and $\mathcal{N}8$ (*right*) of a pixel (i, j) are used to formulate $\Phi(u)$. The double neighborhoods $\mathcal{N}4^2$ and $\mathcal{N}8^2$ appear in the gradient of $\Phi(u)$, see (7)

3 Bounds for the ℓ_∞-Error

In this section, we give upper and lower estimates for the ℓ_∞-error between the input image f and the image $\hat{u}$ obtained by minimizing the functional $J(\cdot, f)$.

If $\hat{u}$ is a minimizer of $u \mapsto J(u, f)$ we denote by $h \in \mathbb{R}^n$ the vector with components

$$h[i] := \sum_{j\in\mathcal{N}_i^2} \gamma_{i,j}\varphi'\big(\gamma_{i,j}\big(\hat{u}[i]-\hat{u}[j]\big)\big), \quad i \in \mathbb{I}_n. \tag{5}$$

First we provide a lemma which gives a useful expression for $\|\hat{u} - f\|_\infty$.

Lemma 1 *Let* H0, H1$^\psi$ *and* H1$^\varphi$ *be satisfied. Let $\hat{u}$ be the minimizer of $u \mapsto J(u, f)$ and h be given by* (5). *Then*

$$\|\hat{u} - f\|_\infty = \big(\psi'\big)^{-1}\big(\beta\|h\|_\infty, \alpha_1\big). \tag{6}$$

Proof In this proof we can omit the parameter α_1. Using the definition of J and taking into account that φ' is odd, we have

$$\begin{aligned} \frac{\partial\Psi}{\partial u[i]} &= \psi'\big(u[i]-f[i]\big) \quad \text{and} \\ \frac{\partial\Phi}{\partial u[i]} &= \sum_{j\in\mathcal{N}_i^2} \gamma_{i,j}\varphi'\big(\gamma_{i,j}\big(u[i]-u[j]\big)\big). \end{aligned} \tag{7}$$

The minimizer $\hat{u}$ of $J(\cdot, f)$ has to satisfy $\nabla_u J(\hat{u}, f) = 0$ which can be rewritten as $\nabla_u \Psi(\hat{u}, f) = -\beta\nabla\Phi(\hat{u})$ or as

$$\begin{aligned} &\psi'\big(\hat{u}[i]-f[i]\big) \\ &\quad = -\beta \sum_{j\in\mathcal{N}_i^2} \gamma_{i,j}\varphi'\big(\gamma_{i,j}\big(\hat{u}[i]-\hat{u}[j]\big)\big), \quad i\in\mathbb{I}_n. \end{aligned}$$

Using (5), the latter is equivalent to

$$\psi'\big(\hat{u}[i]-f[i]\big) = -\beta h[i], \quad i \in \mathbb{I}_n.$$

Since ψ' is by H0 and H1$^\psi$ odd and strictly increasing,

$$\psi'\big(\big|\hat{u}[i]-f[i]\big|\big) = \big|\psi'\big(\hat{u}[i]-f[i]\big)\big| = \beta\big|h[i]\big|. \tag{8}$$

Using Remark 1(ii), we see that (8) is equivalent to

$$\big|\hat{u}[i]-f[i]\big| = \big(\psi'\big)^{-1}\big(\beta\big|h[i]\big|\big) \tag{9}$$

where $(\psi')^{-1}$ is strictly increasing, hence

$$\|\hat{u}-f\|_\infty = \max_{i\in\mathbb{I}_n}\big(\psi'\big)^{-1}\big(\beta\big|h[i]\big|\big) = \big(\psi'\big)^{-1}\big(\beta\|h\|_\infty\big). \qquad \square$$

For inner points $i \in \mathbb{I}_n^{\text{int}}$ we define

$$\eta := \sum_{j\in\mathcal{N}_i^2} \gamma_{i,j}. \tag{10}$$

Of course η does not depend on i but just on the choice of the neighborhood. If the weights are defined as in (4), we have

$$\eta = 4 \qquad \text{for } \mathcal{N}4,$$

$$\eta = 4 + \frac{4}{\sqrt{2}} = 6.8284 \quad \text{for } \mathcal{N}8.$$

For $i \in \mathbb{I}_n \setminus \mathbb{I}_n^{\text{int}}$ we have $\sum_{j\in\mathcal{N}_i^2}\gamma_{i,j} \le \eta$ whose value depends on the boundary conditions.

In order to extend the obtained result, we shall use a property of $(\psi')^{-1}$ which is stated below.

Lemma 2 *Let ψ satisfy* H0, H1$^\psi$ *and* H2$^\psi$. *Set*

$$\widetilde{Y} := \min\big\{Y, \psi'(T)\big\},$$

where $\psi'(T) := \lim_{t\to+\infty}\psi'(t)$ if $T = +\infty$. Then for any $y \in (0, \widetilde{Y})$, the function $\alpha_1 \mapsto (\psi')^{-1}(y, \alpha_1)$ is strictly increasing on $(0, +\infty)$.

Proof Let $0 < a_1 < a_2$ and $y \in (0, \widetilde{Y})$ be arbitrarily fixed. Since $t \mapsto \psi'(t, \alpha_1)$ is one-to-one and odd, there exist $t_1, t_2 \in (0, T)$ such that

$$\psi'(t_1, a_1) = y = \psi'(t_2, a_2). \tag{11}$$

Thus we have $(\psi')^{-1}(y,\alpha_1)=t_1$ and $(\psi')^{-1}(y,\alpha_2)=t_2$. From H1$^\psi$, $t\mapsto\psi'(t,\alpha_1)$ is strictly increasing for any fixed $\alpha_1>0$ and from H2$^\psi$, $\alpha_1\mapsto\psi'(t,\alpha_1)$ is strictly decreasing for any fixed $t\in(0,T)$. Therefore

$$t_2\le t_1 \quad\Rightarrow\quad y=\psi'(t_1,\alpha_1)>\psi'(t_1,\alpha_2)\ge\psi'(t_2,\alpha_2).$$

This contradicts (11). Consequently, $t_1<t_2$ which implies the assertion. □

For all functions in Table 1 and for ψ in (2) we have $\widetilde{Y}=1$.

The following theorem provides an upper bound for $\|\hat{u}-f\|_\infty$ which is independent of f as well as of the particular shape of $\varphi(t,\alpha_2)$ provided that the latter meets the relevant assumptions.

Theorem 1 *Assume that* H0, H1$^\psi$ *and* H1$^\varphi$ *are satisfied. Let $\beta\eta<Y$, where η is given in* (10). *Then the minimizer $\hat{u}$ of $u\mapsto J(u,f)$ satisfies*

$$\|\hat{u}-f\|_\infty\le\big(\psi'\big)^{-1}(\beta\eta,\alpha_1)=:b(\beta,\alpha_1). \tag{12}$$

If, in addition, ψ fulfills H2$^\psi$ *and $\beta\eta<\widetilde{Y}$, where $\widetilde{Y}=\min\{Y,\psi'(T)\}$, then $\alpha_1\mapsto b(\beta,\alpha_1)$ is strictly increasing on* $(0,+\infty)$.

Proof From H1$^\varphi$, φ' is increasing with $|\varphi'(t)|\le1$ for any $t\in\mathbb{R}$. Inserting this into the definition of h in (5) yields

$$\|h\|_\infty\le\eta. \tag{13}$$

Since $(\psi')^{-1}$ is by Remark 1(ii) strictly increasing on $(0,Y)$, we deduce from (6) and (13) for $\beta\eta<Y$ that

$$\|\hat{u}-f\|_\infty=\big(\psi'\big)^{-1}\big(\beta\|h\|_\infty,\alpha_1\big)\le\big(\psi'\big)^{-1}(\beta\eta,\alpha_1).$$

If ψ meets H2$^\psi$ and $\beta\eta<\widetilde{Y}$ we obtain by Lemma 2 that the function $\alpha_1\mapsto(\psi')^{-1}(\beta\eta,\alpha_1)$ is strictly increasing on $(0,+\infty)$. □

We clarify the statement of Theorem 1 below.

- By Remark 1, the function $\beta\mapsto b(\beta,\alpha_1)$ is strictly increasing since η is a fixed number.
- The equality in (12) can only be met if φ' attains the limit in H1$^\psi$, i.e., if $\varphi'(t)=1$ for some $t\in\mathbb{R}$. This is for example the case for the scaled Huber function in (3).
- The bound in (12) depends only on $\psi(\cdot,\alpha_1)$ and on β but it is independent of the selection of φ provided that H1$^\varphi$ holds.
- For all functions ψ listed in Table 1 we have $Y=1$ which limits the action of β to less than $1/\eta$. So H2$^\psi$ furnishes a flexible tool to control the upper bound $b(\beta,\alpha_1)$ by using α_1 under the condition that $\beta\eta<\widetilde{Y}$, where we remind that $\widetilde{Y}=1$ for all ψ in Table 1 and in (2).

The lower bound on $\|\hat{u}-f\|_\infty$ exhibited in the next Theorem 2 depends on $\varphi(t,\alpha_2)$ and on the input image f as well. In our formula, the reliance on f is expressed via the magnitude ν_f defined below:

$$\begin{aligned}\mathcal{I}&:=\Big\{i\in\mathbb{I}_n^{\text{int}}\;\Big|\;\operatorname{sign}\big(f[i]-f[j]\big)=\sigma,\ \forall j\in\mathcal{N}_i\ \text{where}\\ &\qquad\sigma\in\{-1,+1\}\Big\},\\ \nu_f&:=\max_{i\in\mathcal{I}}\min_{j\in\mathcal{N}_i}\big(\gamma_{i,j}\big|f[i]-f[j]\big|\big),\end{aligned}\tag{14}$$

where we set $\nu_f:=0$ if $\mathcal{I}=\emptyset$. The values of ν_f for some real-world images can be seen in Fig. 7.

Theorem 2 *Let* H0, H1$^\psi$, H2$^\psi$ *and* H1$^\varphi$, H2$^\varphi$ *be verified. Let $\beta\eta<Y$, where η is given in* (10). *Assume that $\nu_f>2b(\beta,\alpha_1)$. Then the minimizer $\hat{u}$ of $u\mapsto J(u,f)$ fulfills*

$$\|\hat{u}-f\|_\infty\ge\big(\psi'\big)^{-1}(c\beta\eta,\alpha_1)=:\ell(\beta,\alpha_1,\alpha_2,\nu_f), \tag{15}$$

where

$$c=c(\beta,\alpha_1,\alpha_2,\nu_f):=\varphi'\big(\nu_f-2b(\beta,\alpha_1),\alpha_2\big)\le1.$$

The function $\alpha_2\mapsto\ell(\beta,\alpha_1,\alpha_2,\nu_f)$ is decreasing on $(0,+\infty)$ and

$$\ell(\beta,\alpha_1,\alpha_2,\nu_f)\nearrow b(\beta,\alpha_1)\quad\text{as }\alpha_2\searrow0. \tag{16}$$

Moreover, for $\varepsilon>0$ arbitrarily close to zero, α_2 can be set so that

$$\|\hat{u}-f\|_\infty\ge\big(\psi'\big)^{-1}\big((1-\varepsilon)\beta\eta,\alpha_1\big). \tag{17}$$

Proof From the definition on ν_f, there exists $i\in\mathbb{I}_n^{\text{int}}$ such that

$$\gamma_{i,j}\big|f[i]-f[j]\big|\ge\nu_f,\quad\forall j\in\mathcal{N}_i.$$

We consider the case

$$\gamma_{i,j}\big(f[i]-f[j]\big)\ge\nu_f>2b(\beta,\alpha_1),\quad\forall j\in\mathcal{N}_i. \tag{18}$$

The opposite case, namely $\gamma_{i,j}(f[j]-f[i])\ge\nu_f>2b(\beta,\alpha_1)$, $\forall j\in\mathcal{N}_i$ can be handled in the same way. By Theorem 1, the minimizer $\hat{u}$ of $J(\cdot,f)$ meets

$$\begin{aligned}&-b(\beta,\alpha_1)\le\hat{u}[i]-f[i],\\ &-b(\beta,\alpha_1)\le f[j]-\hat{u}[j],\quad\forall j\in\mathcal{N}_i.\end{aligned}$$

Thus

$$\begin{aligned}&-2b(\beta,\alpha_1)\le\hat{u}[i]-\hat{u}[j]-\big(f[i]-f[j]\big),\quad\forall j\in\mathcal{N}_i,\\ &-2b(\beta,\alpha_1)+\big(f[i]-f[j]\big)\le\hat{u}[i]-\hat{u}[j],\quad\forall j\in\mathcal{N}_i.\end{aligned}\tag{19}$$

Combining (18) and (19) along with the fact that $\gamma_{i,j}\le1$ yields

$$0 < -2b(\beta,\alpha_1)+\nu_f \le -2b(\beta,\alpha_1)+\gamma_{i,j}\big(f[i]-f[j]\big) \le \gamma_{i,j}\big(\hat u[i]-\hat u[j]\big) \quad \forall j \in \mathcal{N}_i.$$

Since $t \mapsto \varphi'(t,\alpha_2)$ is increasing by H1$^\varphi$, the value $h[i]$ in (5) satisfies

$$h[i] \ge \sum_{j\in\mathcal{N}_i^2} \gamma_{i,j}\varphi'\big(\nu_f - 2b(\beta,\alpha_1),\alpha_2\big) = \eta c(\beta,\alpha_1,\alpha_2,\nu_f).$$

Using yet again that $y \mapsto (\psi')^{-1}(y,\alpha_1)$ is strictly increasing (Remark 1(ii)) we obtain by (9) that

$$\big|\hat u[i]-f[i]\big| \ge \big(\psi'\big)^{-1}(c\beta\eta,\alpha_1).$$

Since $\|\hat u - f\|_\infty \ge |\hat u[i]-f[i]|$, it follows that

$$\|\hat u - f\|_\infty \ge \big(\psi'\big)^{-1}(c\beta\eta,\alpha_1) = \ell(\beta,\alpha_1,\alpha_2,\nu_f).$$

Using H2$^\varphi$, the function $\alpha_2 \mapsto c(\beta,\alpha_1,\alpha_2,\nu_f)$ is continuous and decreasing on $(0,+\infty)$ and $\lim_{\alpha_2\searrow 0} c(\beta,\alpha_1,\alpha_2,\nu_f) = 1$. Combining the latter with Remark 1(ii) entails that $\alpha_2 \mapsto \ell(\beta,\alpha_1,\alpha_2,\nu_f)$ is decreasing on $(0,+\infty)$. Then the definition of $b(\beta,\alpha_1)$ in (12) leads to (16).

Finally, H2$^\varphi$ shows that for ε arbitrarily close to zero there is $\alpha_2 > 0$ such that $c(\beta,\alpha_1,\alpha_2,\nu_f) = (1-\varepsilon)$ and consequently $\|\hat u - f\|_\infty \ge (\psi')^{-1}((1-\varepsilon)\beta\eta)$. □

Some comments on Theorem 2 may be useful.

- The expression in (17) tells us that by decreasing α_2, the lower bound $\ell(\cdot)$ can be adjusted arbitrarily close to the upper bound $b(\cdot)$. The amount of decrease of α_2 needed to reach $(1-\varepsilon)$ depends on the input image f and can be calculated.
- If $t \mapsto \varphi'(t,\alpha_2)$ is *nonstrictly* increasing on $[0,+\infty)$, as the Huber function in (3), it is easy to see that there is α_2 such that $c(\beta,\alpha_1,\alpha_2,\nu_f) = 1$ and hence $\ell(\beta,\alpha_1,\alpha_2,\nu_f) = b(\beta,\alpha_1)$.

4 Explicit Parameter Estimates

In this section we want to use the error bounds from the previous section to give explicit parameter estimates of β, α_1 and α_2 for the functions ψ, φ mentioned in Sect. 2. More precisely, for a given β satisfying a constraint and for δ fixed, we exhibit the value $\alpha_1 = \widehat{\alpha}_1$ ensuring that $b(\beta,\widehat{\alpha}_1) = \delta$ and then calculate $\ell(\beta,\widehat{\alpha}_1,\alpha_2,\nu_f)$.

For the functions ψ in Table 1 and in (2) we have $\widetilde{Y} = 1$. When the weights $\gamma_{i,j}$ are chosen as in (4) and H2$^\psi$ holds, the assumption $\beta\eta < \widetilde{Y} = 1$ in Theorem 1 reads

$$\begin{aligned} &\beta < \frac{1}{4} = 0.25 && \text{for } \mathcal{N}4,\\ &\beta < \frac{1}{6.8284} = 0.1464 && \text{for } \mathcal{N}8. \end{aligned} \tag{20}$$

In the following we choose $\beta > 0$ such that $\beta < \frac{1}{\eta}$. For $\delta > 0$ fixed, let $\widehat{\alpha}_1$ solve the equation

$$b(\beta,\alpha_1) = \big(\psi'\big)^{-1}(\beta\eta,\alpha_1) = \delta. \tag{21}$$

Then we have by Theorem 1 that $\|\hat u - f\|_\infty \le \delta$ for all $\alpha_1 \in (0,\widehat{\alpha}_1]$ and there does not exist $\alpha_1 > \widehat{\alpha}_1$ such that $\|\hat u - f\|_\infty \le \delta$ holds true. In this sense we call $\widehat{\alpha}_1$ *optimal for* δ. This claim is ensured thanks to H2$^\psi$ which guarantees that $\alpha_1 \mapsto b(\beta,\alpha_1)$ is strictly increasing (see Lemma 2). The value c in Theorem 2 depends on φ and on f via ν_f. Given the input image f the constant ν_f is easy to compute. When

$$z := \nu_f - 2b(\beta,\alpha_1) > 0,$$

Theorem 2 indicates that the constant c reads

$$c = \varphi'(z,\alpha_2). \tag{22}$$

In our experiments on real-world digital images, we always had $z \gg 0$ for $\delta = 0.5$. By Theorem 2 a sharper lower bound requires a smaller value for α_2. According to Theorem 1 and Theorem 2, the upper and lower bounds for $\|f - \hat u\|_\infty$ and the optimal value for α_1 as defined in (21) for the functions ψ in Table 1 and in (2) are given in Table 2.

If $\delta = 0.5$ then $\hat u$ has the important property that it preserves the order of the pixel values in a digital image $f \in \{0,\dots,L-1\}^n$. The corresponding values $\widehat{\alpha}_1$ and β are presented in Table 3.

Remark 2 Equation (21) offers several other exploits than only fixing the optimal $\widehat{\alpha}_1$. For any $\beta < \frac{Y}{\eta}$ one can also

- calculate δ when α_1 and β are given—this can be useful e.g. when ℓ_1-TV or ℓ_2-TV are approximated by a fully smooth functional;
- determine the optimal β for fixed α_1 and δ—we remind that from Remark 1, $\beta \mapsto b(\beta,\alpha_1)$ is strictly increasing, hence this value of β is unique.

5 Probability Estimates for Pixel Neighborhoods

Consider that the assumptions H0, H1$^\psi$, H1$^\varphi$ and H2$^\varphi$ are met and that the parameters $\beta < Y/\eta$, α_1 and α_2 are fixed. From Theorem 2 we know that the upper bound $b(\beta,\alpha_1)$ in Theorem 1 provides a nearly perfect approximation of the true error $\|\hat u - f\|_\infty$ when $c = \varphi'(\nu_f - 2b,\alpha_2)$ is close to one, which by H1$^\varphi$ means that ν_f is large enough. In order to get an intuition—even though very rough—on the behaviour of ν_f, we assume in this section that the values of f are realizations of a discrete random variable X taking values in $\{0,\dots,L-1\}$ whose probability density function (pdf) p_X is specialized to real-world digital images. Figure 4

Table 2 Bounds and parameter $\widehat{\alpha}_1$ for various functions ψ in Table 1 and in (2). The parameter c depends on φ' by (22). The allowed values for β by Theorem 1 are given in (20)

$\psi(t)$	$b(\beta,\alpha_1)$	$\ell(\beta,\alpha_1,\alpha_2,\nu_f)$	$\widehat{\alpha}_1$
$\sqrt{t^2+\alpha_1}$	$\sqrt{\frac{\alpha_1(\beta\eta)^2}{1-(\beta\eta)^2}}$	$\sqrt{\frac{\alpha_1(c\beta\eta)^2}{1-(c\beta\eta)^2}}$	$\delta^2(\frac{1}{\beta^2\eta^2}-1)$
$\lvert t\rvert-\alpha_1\log(1+\frac{\lvert t\rvert}{\alpha_1})$	$\frac{\alpha_1\beta\eta}{1-\beta\eta}$	$\frac{\alpha_1 c\beta\eta}{1-c\beta\eta}$	$\delta(\frac{1}{\beta\eta}-1)$
$\alpha_1\log(\cosh(\frac{t}{\alpha_1}))$	$\alpha_1\,\mathrm{atanh}(\beta\eta)$	$\alpha_1\,\mathrm{atanh}(c\beta\eta)$	$\frac{\delta}{\mathrm{atanh}(\beta\eta)}$
$\frac{1}{\alpha_1+1}\lvert t\rvert^{\alpha_1+1}$	$(\beta\eta)^{\frac{1}{\alpha_1}}$	$(c\beta\eta)^{\frac{1}{\alpha_1}}$	$\frac{\ln(\beta\eta)}{\ln\delta}$

Table 3 Allowed values $\beta<1/\eta$ and the optimal $\widehat{\alpha}_1$ for $\delta=b(\beta,\hat{\alpha}_1)=0.5$

$\psi(t)$	$\sqrt{t^2+\alpha_1}$		$\lvert t\rvert-\alpha_1\log(1+\frac{\lvert t\rvert}{\alpha_1})$	
neighborhood	β	$\widehat{\alpha}_1$	β	$\widehat{\alpha}_{\cancel{2}}\ 1$
$\mathcal{N}4$	0.2	0.1406	0.2	0.1250
$\mathcal{N}4$	0.1	1.3125	0.1	0.7500
$\mathcal{N}8$	0.1	0.2862	0.1	0.2322
$\mathcal{N}8$	0.05	1.8947	0.05	0.9645

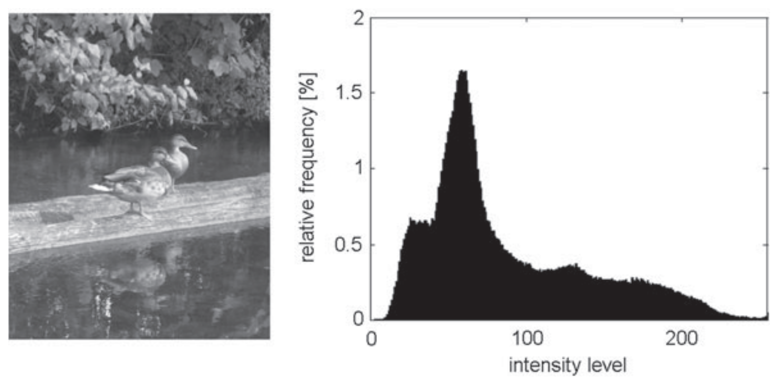


Fig. 4 *Left*: Duck image. *Right*: Histogram of "duck image" furnishing an empirical estimate of the corresponding pdf

shows an image together with its histogram which furnishes an empirical estimate of the corresponding pdf.

First, we ask for the probability that an inner image pixel $i\in\mathbb{I}_n^{\mathrm{int}}$ fulfills

$$\big|f[i]-f[j]\big|\geq a \quad\text{and}\quad \mathrm{sign}\big(f[i]-f[j]\big)=\sigma,\quad \forall j\in\mathcal{N}_i \tag{23}$$

where $\sigma\in\{-1,+1\}$ and $a>0$ is fixed.

Lemma 3 *Let $X, X_i, i=1,\dots,k$ be independent and identically distributed (iid) discrete random variables taking values in $\{0,\dots,L-1\}$. Then it holds for $a>0$ that*

$$\begin{aligned} q(X,k,a) &:= P(X-X_1\geq a,\dots,X-X_k\geq a)\\ &= \sum_{i=0}^{L-1}\big(P(X\leq i-a)\big)^k P(X=i). \end{aligned} \tag{24}$$

Proof Since the random variables are iid we obtain

$$\begin{aligned} &P(X-X_1\geq a,\dots,X-X_k\geq a)\\ &= \sum_{i=0}^{L-1} P(i-X_1\geq a,\dots,i-X_k\geq a, X=i)\\ &= \sum_{i=0}^{L-1}\big(P(X\leq i-a)\big)^k P(X=i). \end{aligned}$$

□

A case relevant to our context is when X is a given inner pixel and X_i for $i\in\{1,\dots,k\}$ are the pixels in the "double" neighborhood of X, see Fig. 3. Then the setting of Lemma 3 considers neighborhoods where the central pixel X is bigger than all its neighbors by at least the amount of a. It is clear that the opposite case (when $X-X_i\leq -a$ for all $i\in 1,\dots,k$) is of the same interest and appears with the same probability $P(X-X_1\leq -a,\dots,X-X_k\leq -a)=q(X,k,a)$. Of course the "iid" assumption is not realistic for natural images.

For $k=1$, the probabilities $P(X-X_1\geq a)$ and $P(X-X_1\leq -a)$ can be easily exemplified. Let X and X_1 follow independently the same pdf p_X. In order to obtain the joint pdf of X and X_1, one has to compute $P(X=i_1)P(X=i_2)$ for all gray levels i_1,i_2 obeying $|i_1-i_2|\geq a$ and then

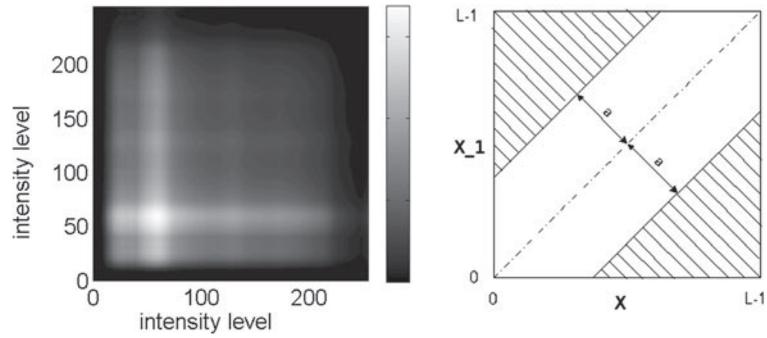


Fig. 5 *Left*: Joint pdf of two iid random variables X, X_1 where X and X_1 follow the pdf of the "ducks image" in Fig. 4 right. Here *light areas* correspond to high probability. *Right*: Areas where $|i_1 - i_2| \geq a$, $i_1, i_2 \in \{0, \dots, L-1\}$. The value $2q(X, 1, a)$ is the sum of the probabilities in the *shaded areas*

take their sum. Figure 5 (left) shows for example the joint pdf of X and X_1 when X and X_1 are iid random variables following the pdf p_X of the "ducks image" in Fig. 4 left. At position $(i_1, i_2) \in \{0, \dots, 255\}^2$ the probability $P(X = i_1)P(X_1 = i_2)$ is visualized as a gray value where lighter areas correspond to higher probability.

In Fig. 5 (right) the shaded areas show the points where the pixel difference $|i_1 - i_2|$ is larger or equal to a. The sum of the probabilities corresponding to these areas is $2q(X, 1, a)$.

Theorem 3 *Assume that the $M \times N$ image f is the realization of a discrete iid random vector $(X_i)_{i=1}^n$ with iid components X_i as X, where $n = MN$. Let ν_f be defined as in* (14) *with respect to $\mathcal{N}4$. Then the probability that $\nu_f \geq a > 0$ is not smaller than*

$$1 - \big(1 - 2q(X, 4, a)\big)^m, \tag{25}$$

where q is defined in (24) *and $m = \lfloor M/3 \rfloor \times \lfloor N/3 \rfloor$.*

For $\mathcal{N}8$ we have to replace q by $\tilde{q}(X, 4, a) := \sum_{i=0}^{L-1} (P(X \leq i - a))^4 (P(X \leq i - \sqrt{2}a))^4 P(X = i)$.

Proof We consider only inner pixels i with non-overlapping neighborhoods as depicted in Fig. 6. Then, by Lemma 3, the probability that one of these pixels does not verify (23) is given by $1 - 2q(X, 4, a)$. Hence the probability that all these inner pixels do not fulfill (23) is $(1 - 2q(X, 4, a))^m$ and the probability that at least one of these pixel satisfies (23) is $1 - (1 - 2q(X, 4, a))^m$. □

Note that for $q(X, 4, a) > 0$ the probability in (25) is indeed very close to 1 even for moderate sizes of m. For instance, if the random variables are *uniformly* iid, we have

$$\begin{aligned} &q(X, 4, a) \\ &\quad = \frac{1}{L} \sum_{i=a}^{L-1} \left(\frac{i - a + 1}{L} \right)^4 \\ &\quad = \frac{(L-a)(L-a+1)(2(L-a)+1)(3(L-a)^2 + 3(L-a) - 1)}{30L^5}. \end{aligned}$$

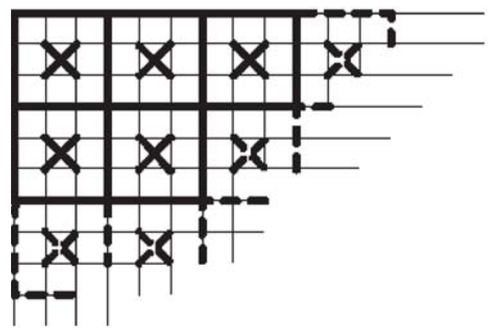


Fig. 6 Disjoint 3×3-adjacencies with center pixels "x"

For $a = 137$ and $L = 256$ this formula gives $q(X, 4, a) \approx 0.0044$ and for $M = N = 128$ further $1 - (1 - q(X, 4, a))^m \approx 1 - 10^{-7}$.

6 Numerical Tests

The bounds on $\|\hat{u} - f\|_\infty$ with respect to the model parameters were tested on a wide amount of images. Here we present the results on 15 digital images of different sizes, with gray values in $\{0, \dots, 255\}$, available at http://sipi.usc.edu/database/. In our selection the images have various quality and content (presence or quasi-absence of edges, textures, nearly flat regions). They are displayed in Fig. 7. The values of ν_f for $\mathcal{N}8$ under each image shows that the assumption $\nu_f - 2b(\beta, \alpha_1) > 0$ in Theorem 2 is generously satisfied in all these cases as far as we are interested to fix $b(\beta, \alpha_1) \leq 0.5$. We also performed tests with 10^4 random 256×256 images with pixel values uniformly distributed in $\{0, \dots, 255\}$. For $\mathcal{N}4$ we obtained $\mathrm{mean}(\nu_f) = 224.2267$ and for $\mathcal{N}8$, $\mathrm{mean}(\nu_f) = 137.7871$.

We tested two functionals $J(\cdot, f)$ as described in Sect. 2: the first corresponds to $\psi = \Theta 1$ and $\varphi = \Theta 1$ and the second to $\psi = \Theta 2$ and $\varphi = \Theta 1$ as given in Table 1. In all tests, $\mathcal{N}8$ was adopted with the weights $\gamma_{i,j}$ given in (4). Two choices for β satisfying (20) were considered along with different values for α_1 and α_2. The minimizers $\hat{u}$ were computed using Polak-Ribière conjugated gradients [3] with

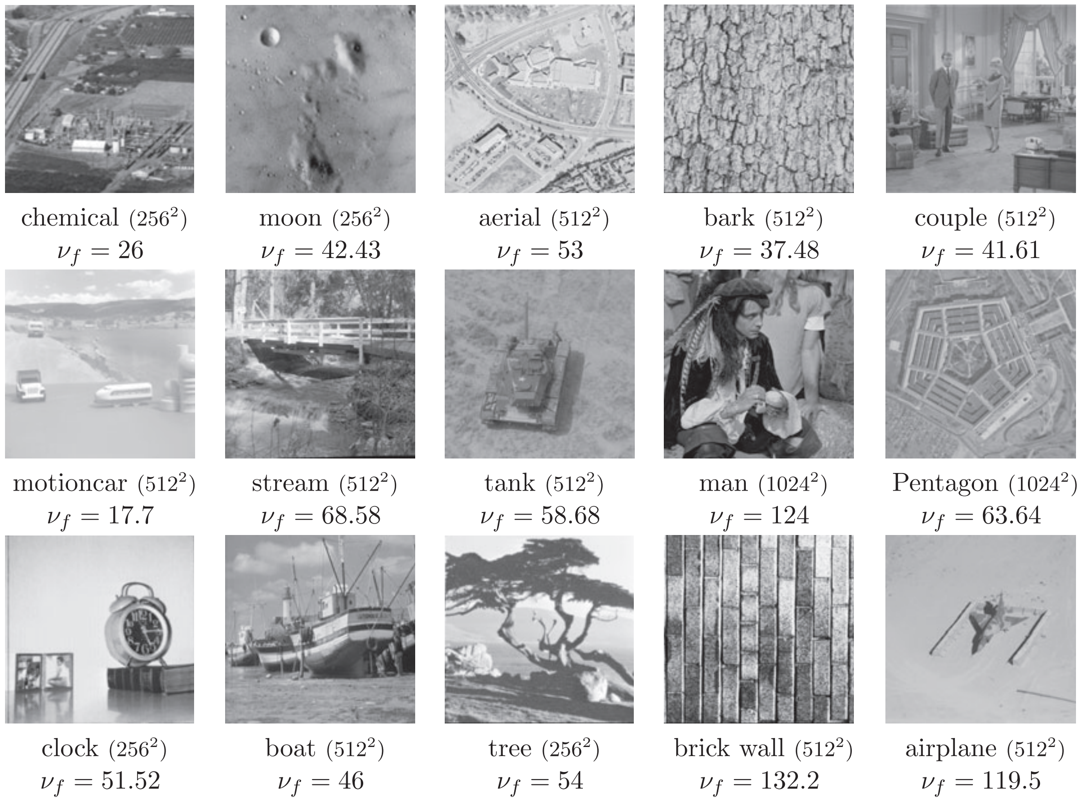


Fig. 7 The set of images used in the tests provided in this section. The values of ν_f are computed according to (14) in the case $\mathcal{N}8$ for the weights in (4)

high numerical precision. For each restored image we computed $\|\hat{u} - f\|_\infty$ and present the distance between the theoretical upper bound $b(\beta, \alpha_1)$ and the obtained $\|\hat{u} - f\|_\infty$:

$$b(\beta, \alpha_1) - \|\hat{u} - f\|_\infty.$$

The tables show also the difference between the upper and the lower theoretical bounds on $\|\hat{u} - f\|_\infty$:

$$b - \ell := b(\beta, \alpha_1) - \ell(\beta, \alpha_1, \alpha_2, \nu_f),$$

computed using the explicit formulae given in Sect. 4. Furthermore, we evaluate the amount of pixels that closely approach the ℓ_∞ norm:

$$q = \#\big\{i \in \mathbb{I}_n \mid \|\hat{u} - f\|_\infty - \big|\hat{u}[i] - f[i]\big| < \varepsilon\big\} \quad \text{and}$$

$$Q\ \% = 100\frac{q}{n},$$

where # stands for cardinality and $\varepsilon \gtrapprox 0$ in order to account for numerical errors. In the experiments, we set $\varepsilon := 10^{-3}$.

In all tests, given $0 < \beta < 1/\eta$, we fixed $\alpha_1 = \widehat{\alpha}_1$ so that

$$b(\beta, \widehat{\alpha}_1) = \delta \quad \text{for } \delta = \frac{1}{2}.$$

The numerical outcomes confirm the theoretical results on $\|\hat{u} - f\|_\infty$ established in Sects. 3 and 4. From Tables 4, 5 and 6 the following observations can be drawn:

- Decreasing $\alpha_2 > 0$ towards 0 enables to make the difference between the upper and the lower bounds on $\|\hat{u} - f\|_\infty$ arbitrarily small which leads to $\|u - f\|_\infty \approx b(\beta, \alpha_1)$.
  In this case a large percentage of the pixels i meet $|\hat{u}[i] - f[i]| \approx b(\beta, \alpha_1)$.
- An important increase of $\alpha_2 > 0$ entails a decrease of the lower bound $\ell(\beta, \alpha_1, \alpha_2, \nu_f)$. Moreover, the number of pixels i verifying $|\hat{u}[i] - f[i]| \approx b(\beta, \alpha_1)$ is reduced to a few ones.
  Such a situation may be preferable when one wishes that there are not too many pixels close to the upper bound.

Tables 7 and 8 show yet again that the gap between the upper bound $b(\beta, \alpha_1)$ and the lower bound $\ell(\beta, \alpha_1, \alpha_2, \nu_f)$ vanishes when α_2 is close to zero and that it increases when α_2 increases. For α_2 fixed, we see that $b(\beta, \alpha_1) - \ell(\beta, \alpha_1, \alpha_2, \nu_f)$ tends to decrease along with β.

Figure 8 shows the histograms of the differences $\{f[i] - \hat{u}[i], i \in \mathbb{I}_n\}$ relevant to "moon", where the upper bound was set to $b(\beta, \alpha_1) = 0.5$, for an increasing set of values of α_2.

Table 4 Results for $\psi = \Theta 1$, $\varphi = \Theta 1$, $\beta = 0.1$ and a small and large value of α_2, respectively. Over the whole set of these images, for $\alpha_2 = 0.02$ we have mean$(0.5 - \|\hat{u} - f\|_\infty) = 2.968 \times 10^{-6}$ and mean$(0.5 - \ell(\beta, \alpha_1, \alpha_2, \nu_f)) = 6.0678 \times 10^{-6}$. For $\alpha_2 = 100$ these values read mean$(0.5 - \|\hat{u} - f\|_\infty) = 1.307 \times 10^{-2}$ and mean$(0.5 - \ell(\beta, \alpha_1, \alpha_2, \nu_f)) = 2.491 \times 10^{-2}$

image	$\alpha_2 = 0.02$			$\alpha_2 = 100$		
	$b - \|\hat{u} - f\|_\infty \mid \times 10^{-6}$	$b - \ell \mid \times 10^{-6}$	Q %	$b - \|\hat{u} - f\|_\infty \mid \times 10^{-3}$	$b - \ell \mid \times 10^{-2}$	q
chemical	4.764	14.90	4.04	22.85	6.143	2
moon	2.438	5.459	9.27	12.49	2.525	1
aerial	2.066	3.465	3.46	6.949	1.647	1
bark	2.977	7.041	6.57	13.44	3.188	1
couple	2.485	2.568	3.25	12.77	2.619	4
motioncar	19.98	33.68	0.18	77.56	11.35	1
stream	0.918	2.051	7.14	5.412	0.995	2
tank	1.960	2.815	6.95	9.297	1.351	1
man	0.025	0.619	4.94	1.581	0.307	8
Pentagon	1.181	2.388	9.12	6.368	1.153	1
clock	2.079	3.671	2.88	6.884	1.740	1
boat	1.707	4.626	6.04	8.425	2.164	2
tree	1.202	3.325	5.27	8.026	1.584	1
brick wall	0.334	0.544	11.8	1.842	0.270	43
airplane	0.412	0.667	1.73	2.089	0.330	1

$\mathcal{N}8$, $\psi(t) = \sqrt{t^2 + \alpha_1}$ for $\alpha_1 = 0.2862$, $\beta = 0.1$ hence $b = 0.5$, $\varphi(t) = \sqrt{t^2 + \alpha_2}$

Table 5 Results for $\psi = \Theta 1$, $\varphi = \Theta 1$, $\beta = 0.05$ and a small and large value of α_2, respectively. For $\alpha_2 = 0.02$ we have mean$(0.5 - \|\hat{u} - f\|_\infty) = 1.777 \times 10^{-6}$ and mean$(0.5 - \ell(\beta, \alpha_1, \alpha_2, \nu_f)) = 3.666 \times 10^{-6}$. For $\alpha_2 = 100$, we find mean$(0.5 - \|\hat{u} - f\|_\infty) = 8.265 \times 10^{-3}$ and mean$(0.5 - \ell(\beta, \alpha_1, \alpha_2, \nu_f)) = 1.610 \times 10^{-2}$

image	$\alpha_2 = 0.02$			$\alpha_2 = 100$		
	$b - \|\hat{u} - f\|_\infty \mid \times 10^{-6}$	$b - \ell \mid \times 10^{-6}$	Q %	$b - \|\hat{u} - f\|_\infty \mid \times 10^{-3}$	$b - \ell \mid \times 10^{-2}$	q
chemical	2.561	9.055	4.54	14.17	3.993	2
moon	1.580	3.300	10.2	7.649	1.572	1
aerial	0.872	2.093	3.92	4.229	1.015	2
bark	1.673	4.254	6.82	8.239	2.000	1
couple	1.642	3.432	3.25	7.830	1.632	4
motioncar	12.39	20.35	0.28	51.43	7.847	1
stream	0.727	1.240	7.19	3.291	0.608	3
tank	1.020	1.701	8.31	5.678	0.829	1
man	0.162	0.374	6.00	0.968	0.186	11
Pentagon	0.871	1.442	10.2	3.877	0.706	1
clock	1.013	2.220	2.88	4.193	1.073	1
boat	0.799	2.795	7.14	5.136	1.342	2
tree	0.993	2.009	6.06	4.895	0.975	2
brick wall	0.125	0.329	11.9	1.115	0.164	99
airplane	0.228	0.403	3.48	1.274	0.200	1

$\mathcal{N}8$, $\psi(t) = \sqrt{t^2 + \alpha_1}$ for $\alpha_1 = 1.895$, $\beta = 0.05$ hence $b = 0.5$, $\varphi(t) = \sqrt{t^2 + \alpha_2}$

These histograms were plotted for 100 bins equally spaced in $[-0.5, +0.5]$. For very small values of α_2, there are many pixels meeting $|f[i] - \hat{u}[i]| \approx \|f - \hat{u}\|_\infty$. When α_2 increases, such pixels become more and more rare and the differences $|f[i] - \hat{u}[i]|$ become centered near zero. However they never reach zero: see the value of μ defined in the cap-

Table 6 Results for $\psi = \Theta 2$, $\varphi = \Theta 1$, $\beta = 0.05$ and a small and large value of α_2, respectively. For $\alpha_2 = 0.05$ we have mean$(0.5 - \|\hat{u} - f\|_\infty) = 5.441 \times 10^{-6}$ and mean$(0.5 - \ell(\beta, \alpha_1, \alpha_2, \nu_f)) = 10.29 \times 10^{-6}$. For $\alpha_2 = 100$, we find mean$(0.5 - \|\hat{u} - f\|_\infty) = 1.09 \times 10^{-2}$ and mean$(0.5 - \ell(\beta, \alpha_1, \alpha_2, \nu_f)) = 2.11 \times 10^{-2}$

image	$\alpha_2 = 0.05$			$\alpha_2 = 100$		
	$b - \|\hat{u} - f\|_\infty \mid \times 10^{-6}$	$b - \ell \mid \times 10^{-6}$	Q %	$b - \|\hat{u} - f\|_\infty \mid \times 10^{-3}$	$b - \ell \mid \times 10^{-2}$	q
chemical	0.101	0.304	2.79	18.81	5.236	2
moon	5.347	11.06	7.03	10.22	2.090	1
aerial	2.670	7.019	2.63	5.663	1.354	2
bark	5.843	14.26	5.55	11.01	2.653	1
couple	5.369	11.51	3.25	10.46	2.170	4
motioncar	41.36	68.23	0.09	66.99	0.101	1
stream	1.687	4.155	6.66	4.404	0.813	3
tank	3.8069	5.703	4.45	7.592	1.107	1
man	0.673	1.255	3.14	1.298	0.249	10
Pentagon	2.723	4.837	6.55	5.188	0.943	1
clock	2.622	7.437	2.88	5.610	1.431	1
boat	3.879	9.373	3.97	6.874	1.786	2
tree	4.070	6.737	4.18	6.549	1.301	2
brick wall	0.721	1.102	11.3	1.710	0.219	61
airplane	0.682	1.352	0.74	4.983	0.268	1

$\mathcal{N}8$, $\psi(t) = |t| - \alpha_1 \log(1 + \frac{|t|}{\alpha_1})$ for $\alpha_1 = 0.9645$, $\beta = 0.05$, hence $b = 0.5$, $\varphi(t) = \sqrt{t^2 + \alpha_2}$

Table 7 The mean value of the difference $b(\beta, \alpha_1) - \ell(\beta, \alpha_1, \alpha_2, \nu_f)$ was computed over the selection of images shown in Fig. 7. Here we consider the $\mathcal{N}8$ neighborhood for the weights in (4)

	$\alpha_2 = 0.01$		$\alpha_2 = 100$	
	$\beta = 0.1$	$\beta = 0.05$	$\beta = 0.1$	$\beta = 0.05$
$\psi = \Theta 1, \varphi = \Theta 1$	3.034×10^{-6}	1.833×10^{-6}	2.491×10^{-2}	1.610×10^{-2}
$\psi = \Theta 2, \varphi = \Theta 1$	5.106×10^{-6}	2.459×10^{-6}	3.985×10^{-2}	2.112×10^{-2}
$\psi(t) = \frac{1}{\alpha_1 + 1}\|t\|^{\alpha_1 + 1}, \varphi = \Theta 1$	2.994×10^{-6}	1.045×10^{-6}	2.542×10^{-2}	0.941×10^{-2}

mean $(b(\beta, \alpha_1) - \ell(\beta, \alpha_1, \alpha_2, \nu_f))$, $b(\beta, \alpha_1) = 0.5$, $\mathcal{N}8$

Table 8 The neighborhood here is $\mathcal{N}4$ with the weights given in (4). The mean is calculated over the set of images in Fig. 7

	$\alpha_2 = 0.01$		$\alpha_2 = 100$	
	$\beta = 0.2$	$\beta = 0.1$	$\beta = 0.2$	$\beta = 0.1$
$\psi = \Theta 1, \varphi = \Theta 1$	2.980×10^{-6}	1.278×10^{-6}	2.253×10^{-2}	1.104×10^{-2}
$\psi = \Theta 2, \varphi = \Theta 1$	5.364×10^{-6}	1.788×10^{-6}	3.780×10^{-2}	1.504×10^{-2}
$\psi(t) = \frac{1}{\alpha_1 + 1}\|t\|^{\alpha_1 + 1}, \varphi = \Theta 1$	3.333×10^{-6}	0.812×10^{-6}	2.718×10^{-2}	0.722×10^{-2}

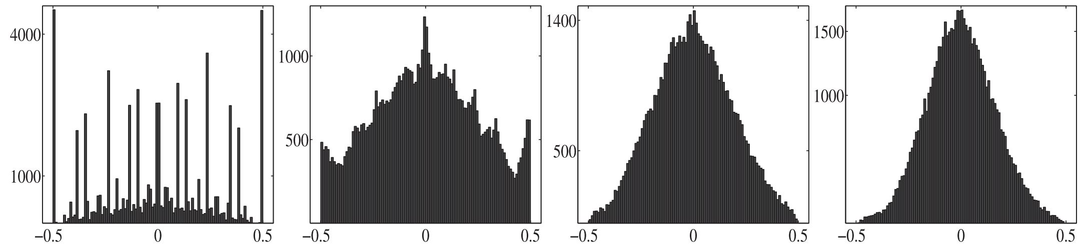
mean $(b(\beta, \alpha_1) - \ell(\beta, \alpha_1, \alpha_2, \nu_f))$, $b(\beta, \alpha_1) = 0.5$, $\mathcal{N}4$

tion of the figure. Here again, the numerical tests were done with a high precision.

7 Conclusions and Open Questions

ℓ_1-TV and ℓ_2-TV functionals have been often minimized using a smoothed version of the form we consider in this paper with ad hoc chosen smoothing parameters ("very small"). The results established in our work enable to clearly evaluate the resulting approximation.

The functions (ψ, φ) studied here have a lot of similarities. However, they produce different image restorations. The question of what couple of functions (ψ, φ) would give



$\alpha_2 = 0.02$, $q = 6700$ | $\mu = 6.6 \times 10^{-9}$
$\alpha_2 = 5$, $q = 48$ | $\mu = 5.556 \times 10^{-8}$
$\alpha_2 = 50$, $q = 3$ | $\mu = 1.523 \times 10^{-6}$
$\alpha_2 = 100$, $q = 1$ | $\mu = 2.249 \times 10^{-6}$

Fig. 8 Histograms of $\{f[i] - \hat{u}[i], i \in \mathbb{I}_n\}$ for "moon" restored using $\psi = \Theta 1$, $\varphi = \Theta 1$, $\mathcal{N}8$, $\beta = 0.05$ and for different values of α_2. The parameter $\alpha_1 = 1.8947$ was set so that $b(\beta, \alpha_1) = 0.5$. The image has $n = 65536$ pixels. The value μ is defined by $\mu := \min_{i \in \mathbb{I}_n} |f[i] - \hat{u}[i]|$

a better result in the framework of a given application, remains open.

Extension to the rotational-invariant (in a discrete sense) smoothed TV, i.e. $\Phi(u) = \sum_{i,j} \varphi(\|\nabla_{i,j} u\|)$, where $\nabla_{i,j} u \in \mathbb{R}^2$ stands for a discrete approximation of the gradient of u at pixel (i, j), deserves attention.

Extensions to cases when f are the coefficients of the expansion of the input image using an orthogonal transform as the discrete cosine transform or a frame transform as the curvelet transform, see, e.g., [9] are of interest.

Applications to quantization noise reduction should be envisaged.

Acknowledgements The work of M. Nikolova was supported in part by the "FMJH Program Gaspard Monge in optimization and operation research", and by the support to this program from EDF.

References

1. Aubert, G., Kornprobst, P.: Mathematical Problems in Image Processing, 2nd edn. Springer, Berlin (2006)
2. Beck, A., Teboulle, M.: Smoothing and first order methods: a unified framework. SIAM J. Optim. **22**(2), 557–580 (2012)
3. Bonnans, J.F., Gilbert, J.-C., Lemaréchal, C., Sagastizábal, C.A.: Numerical Optimization (Theoretical and Practical Aspects). Springer, Berlin (2003)
4. Combettes, P.L., Pesquet, J.-C.: Proximal splitting methods in signal processing. In: Bauschke, H.H., Burachik, R.S., Combettes, P.L., Elser, V., Luke, D.R., Wolkowicz, H. (eds.) Fixed-Point Algorithms for Inverse Problems in Science and Engineering, pp. 185–212. Springer, Berlin (2011)
5. Chan, T., Esedoglu, S.: Aspects of total variation regularized l^1 function approximation. SIAM J. Appl. Math. **65**, 1817–1837 (2005)
6. Chan, T., Mulet, P.: On the convergence of the lagged diffusivity fixed point method in total variation image restoration. SIAM J. Numer. Anal. **36**, 354–367 (1999)
7. Coltuc, D., Bolon, P.: Exact histogram specification. IEEE Trans. Image Process. **15**, 1143–1152 (2006)
8. Geman, D., Reynolds, G.: Constrained restoration and recovery of discontinuities. IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-14**, 367–383 (1992)
9. Mallat, S.: A Wavelet Tour of Signal Processing. Academic Press, San Diego (2009)
10. Nikolova, M.: Analytical bounds on the minimizers of least squares. Inverse Probl. Imaging **1**(4), 661–677 (2007)
11. Nikolova, M., Wen, Y.-W., Chan, R.: Exact histogram specification for digital images using a variational approach. J. Math. Imaging Vis. (2012). doi:10.1007/s10851-012-0401-8
12. Wan, Y., Shi, D.: Joint exact histogram specification and image enhancement through the wavelet transform. IEEE Trans. Image Process. **16**, 2245–2250 (2007)
13. Weiszfeld, E.: Sur le point pour lequel la somme des distances de n points donnés est minimum. Tôhoku Math. J. **43**, 355–386 (1937)

F. Baus finished her Bachelor's degree in Mathematics at the University of Kaiserslautern (Germany) in 2012. Currently, she is enrolled in the Master's programme at the same university. Since 2009 she works as a student assistant at the Fraunhofer Institute for Industrial Mathematics (ITWM) in Kaiserslautern.

M. Nikolova received the Ph.D. degree from the Université de Paris-Sud, France, in 1995. She got a Habilitation to direct research in 2006. Currently, she is Research Director with the National Center for Scientific Research (CNRS), France. She performs herduty as a full-time researcher at CMLA (Centre de Mathématiques et de Leurs Applications), Department of Mathematics, ENS Cachan, France. Her research interests are in mathematical Image and signal reconstruction, Inverse problems, Regularization and variational methods and the properties of their solutions, Scientific computing.

G. Steidl received her PhD and Habilitation in Mathematics from the University of Rostock, Germany, in 1988 and 1991, respectively. From 1993 to 1996, she held a position as assistant professor of Applied Mathematics at the University of Darmstadt, Germany. From 1996 to 2010, she was full professor at the Department of Mathematics and Computer Science, University of Mannheim, Germany. Since 2011, she is Professor at the Department of Mathematics at the University of Kaiserslautern, Germany. She was a visiting researcher at the University of Debrecen (Hungary), the University of Zürich (Switzerland), the Banach Center Warsaw (Poland), the ENS Cachan and the University Paris Est (France). Her research interests include applied and computational harmonic analysis and convex analysis with applications in image processing.