Expected absolute value estimators for a spatially adapted regularization parameter choice rule in L 1 -TV-based image restoration — Source link ↗

Michael Hintermüller, Michael Hintermüller, M. Monserrat Rincon-Camacho

Institutions: University of Graz, Humboldt University of Berlin

Related papers:

- Nonlinear total variation based noise removal algorithms

- Automated Regularization Parameter Selection in Multi-Scale Total Variation Models for Image Restoration

- Image recovery via total variation minimization and related problems

- TV Based Image Restoration with Local Constraints

- A TV Based Restoration Model with Local Constraints

EXPECTED ABSOLUTE VALUE ESTIMATORS FOR A SPATIALLY ADAPTED REGULARIZATION PARAMETER CHOICE RULE IN L^1-TV-BASED IMAGE RESTORATION

MICHAEL HINTERMÜLLER AND M. MONSERRAT RINCON-CAMACHO

ABSTRACT. A total variation (TV) model with a L^1-fidelity term and a spatially adapted regularization parameter is presented in order to reconstruct images contaminated by impulse noise. This model intends to preserve small details while homogeneous features still remain smooth. The regularization parameter is locally adapted according to a local expected absolute value estimator depending on the statistical characteristics of the noise. The numerical solution of the L^1-TV minimization problem with a spatially adapted parameter is obtained by a superlinearly convergent algorithm based on Fenchel-duality and inexact semismooth Newton techniques, which is stable with respect to noise in the data. Numerical results justifying the advantage of such a regularization parameter choice rule are presented.

1. INTRODUCTION

Images are often blurred and corrupted by different kinds of noise, such as Gaussian noise, random-valued impulse noise or salt-and-pepper noise; see Figure 1 (a) and (d) on page 16. The deblurring and denoising of such images are necessary before further image processing operations, such as edge detection, segmentation, or object recognition, are performed.

The image $\hat{u}$ is considered to be a real function defined on a bounded and piecewise smooth open subset Ω of $\mathbb{R}^2$ with range $[0, 1]$. The corrupted image z is then given by

$$(1.1) \qquad z = K\hat{u} + \rho_{\hat{u}}.$$

Here, $K \in \mathcal{L}(L^2(\Omega))$ is a blurring operator, which is assumed to be known, and by $\mathcal{L}(L^2(\Omega))$ we denote the space of linear and continuous operators from $L^2(\Omega)$ to $L^2(\Omega)$. The quantity $\rho_{\hat{u}}$ represents the noise, which may or may not depend on the original image. In their seminal work [18], Rudin, Osher and Fatemi proposed total variation regularization for image restoration. The corresponding minimization task is as follows:

$$(1.2) \qquad \begin{aligned} &\text{minimize } J(u) := \int_\Omega |Du| \text{ over } u \in BV(\Omega) \\ &\text{subject to (s.t.) } \int_\Omega Ku\, dx = \int_\Omega z\, dx, \quad \int_\Omega |Ku - z|^2 dx = \sigma^2 |\Omega|, \end{aligned}$$

where $BV(\Omega)$ denotes the space of functions of bounded variation, i.e. $u \in BV(\Omega)$ iff $u \in L^1(\Omega)$ and the BV-seminorm

$$\int_\Omega |Du| = \sup \left\{ \int_\Omega u \, \text{div} \vec{v} \, dx : \vec{v} \in (C_0^\infty(\Omega))^2, |\vec{v}|_{l^2} \leq 1 \right\}$$

is finite. In what follows we refer to (1.2) as the ROF-model. Usually, the ROF-model is solved via the following unconstrained optimization problem:

$$(1.3) \qquad \text{minimize} \ \int_\Omega |Du| + \frac{\lambda}{2} \int_\Omega |Ku - z|^2 dx \ \text{over} \ u \in BV(\Omega)$$

for a given $\lambda > 0$. For more details about the ROF-model; see, e.g., [4, 5, 7, 10, 16, 17, 21, 22, 24, 25]. In [5], conditions are specified such that (1.3) is equivalent to (1.2) for an adequate $\lambda \geq 0$. In order to cope with different image scales, in [12] the spatial adaptation of λ is studied. For this purpose, the following model with pointwise almost everywhere (a.e.) constraints, i.e.,

$$(1.4) \quad \text{minimize} \ \int_\Omega |Du| \ \text{over} \ u \in BV(\Omega), \ \text{subject to} \ \ S(u)(x) \leq \sigma^2 \ \text{for almost every} \ \ x \in \Omega$$

was proposed in order to find an adequate function $\lambda \in L^\infty(\Omega)$ for the model

$$(1.5) \qquad \text{minimize} \ \int_\Omega |Du| + \frac{1}{2} \int_\Omega \lambda |Ku - z|^2 dx \ \text{over} \ u \in BV(\Omega).$$

In (1.4), $S(u)$ represents a local variance estimator which relies on some local filter such as the mean filter, the Gaussian filter or the Wiener filter. In the present work, a similar approach is proposed to improve the total variation model with an L^1 data-fidelity term, i.e.

$$(1.6) \qquad \text{minimize} \ \int_\Omega |Du| + \int_\Omega \lambda |Ku - z| \ dx \ \text{over} \ u \in BV(\Omega).$$

The standard formulation of (1.6) with a scalar λ (rather than an $L^\infty(\Omega)$-function) was studied, e.g., in [6], [19], [20] and [11] among other references.

Concerning the blurring by the known operator K we mention that in our tests we use a Gaussian convolution with a 9×9 window and a standard deviation of 1. Its discrete version is ill-conditioned, but invertible. We emphasize, however, that our subsequent theory and algorithms work as long as KK^* is invertible. Here, K^* denotes the adjoint of K. Numerically, we do not need to store the blurring operator, rather we store the convolution kernel. Consequently, for the restoration we rely on an iterative solver (bi-conjugate gradient stabilized) which only needs the application of the (discrete) blurring operator to a vector (discrete image).

The outline of the rest of the paper is as follows. In Section 2 descriptive statistics of Gaussian white noise, random-valued impulse noise and salt-and-pepper noise are listed. In Section 3, a relation between the constrained and unconstrained models with L^1 data-fidelity term and the existence of their solutions are studied. A primal-dual method for solving the problem (1.6) with $\lambda \in L^\infty(\Omega)$ is the subject of Section 4. We call this model spatially adapted total variation, denoted by $SA\text{-}TV$ for short. Its numerical solution relies on the algorithm presented in [11]. Section 5 focuses on how to automatically adapt the parameter by considering a local expected absolute value estimator, denoted by $LEAVE$. In Section 6 the proposed algorithm is briefly summarized. Numerical results for the restoration of perturbed images are presented in Section 7. The paper ends by conclusions and further outlook.

2. Statistical characteristics of the noise

In this section, the characteristics of the noise corrupting the image are described. In order to simplify the exposition, no blurring is considered. As it was stated before, the image is a function $\hat{u}$ defined over a domain $\Omega \subset \mathbb{R}^2$ which is corrupted by noise. At a point $x \in \Omega$, the resulting contaminated image $z(x) = \hat{u}(x) + \rho_{\hat{u}}(x)$ is a stochastic observation, whose random element $\rho_{\hat{u}}(x)$ depends on the underlying type of noise. For any two points of the domain $x, y \in \Omega$, $\rho_{\hat{u}}(x)$ and $\rho_{\hat{u}}(y)$ are independent.

Despite the abuse of notation, for the ease of exposition let ρ denote the random variable representing the noise corrupting the image, and let f be its probability density function. There are different measures that describe the behavior of the random variable (r.v.) ρ. An important

measure of location is the mean, here denoted by ϑ, which is given by

(2.1)
$$\vartheta = E(\rho) = \begin{cases} \int_{-\infty}^{\infty} \xi_\rho f(\xi_\rho) \, d\xi_\rho, & \text{if } \rho \text{ is a continuous random variable (c.r.v)}, \\ \sum_{\xi_\rho \in \mathcal{V}} \xi_\rho f(\xi_\rho), & \text{if } \rho \text{ is a discrete random variable (d.r.v) in a universe } \mathcal{V}. \end{cases}$$

A dispersion measure of interest is the variance

(2.2) $\sigma^2 = \text{Var}(\rho) = E[(\rho - \vartheta)^2] = E(\rho^2) - \vartheta^2 = \begin{cases} \int_{-\infty}^{\infty} \xi_\rho^2 f(\xi_\rho) \, d\xi_\rho - \vartheta^2, & \text{if } \rho \text{ is a c.r.v.}, \\ \sum_{\xi_\rho \in \mathcal{V}} \xi_\rho^2 f(\xi_\rho) - \vartheta^2, & \text{if } \rho \text{ is a d.r.v.} \end{cases}$

In what follows, we mainly utilize the expected absolute value given by

(2.3) $\nu = \text{EAV}(\rho) = E(|\rho|) = \begin{cases} \int_{-\infty}^{\infty} |\xi_\rho| f(\xi_\rho) \, d\xi_\rho, & \text{if } \rho \text{ is a c.r.v.}, \\ \sum_{\xi_\rho \in \mathcal{V}} |\xi_\rho| f(\xi_\rho), & \text{if } \rho \text{ is a d.r.v.} \end{cases}$

2.1. **Gaussian noise.** In this case, ρ is normally distributed, with expectation 0 and variance σ^2. The probability density function is given by

(2.4)
$$f(\xi_\rho) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\xi_\rho^2}{2\sigma^2}\right)$$

and the mean, variance and expected absolute value are

(2.5)
$$E(\rho) = 0, \quad \text{Var}(\rho) = E(\rho^2) = \sigma^2 \text{ and } \text{EAV}(\rho) = E(|\rho|) = \sqrt{\frac{2}{\pi}}\sigma.$$

2.2. **Salt-and-pepper noise.** The salt-and-pepper characteristics are more difficult to describe. Here the random variable ρ depends on the value of the function. Thus, the associated conditional probability density function depends on the value u (once more abusing notation):

(2.6)
$$f(\xi_\rho \mid u) = \begin{cases} 1 - r & \text{if } \xi_\rho = 0, \\ \frac{r}{2} & \text{if } \xi_\rho = 1 - u, \\ \frac{r}{2} & \text{if } \xi_\rho = -u, \end{cases}$$

here $|$ stands for "given". As it can be expected, the mean and the variance depend on u:

(2.7)
$$E(\rho \mid u) = \frac{r}{2}(1 - 2u), \quad \text{Var}(\rho) = \frac{r}{2}(1 - 2u + 2u^2) - \frac{r^2}{4} + r^2 u(1 - u).$$

Since the range of u belongs to the interval $[0, 1]$, we have

$$1 - 2u + 2u^2 \le 1 \text{ and } r^2 u(1 - u) \le r^2,$$

which yield the following estimates:

(2.8)
$$E(\rho \mid u) \in \left[-\frac{r}{2}, \frac{r}{2}\right] \quad \text{and} \quad \text{Var}(\rho \mid u) \le \frac{r}{2} + \frac{3r^2}{4}.$$

However, the expected absolute value does not depend on the value u. It is given by

(2.9)
$$\text{EAV}(\rho \mid u) = E(|\rho| \mid u) = \frac{r}{2}.$$

2.3. **Random-valued impulse noise.** This kind of noise also depends on the value of the image, as in the case of salt-and-pepper noise. Its conditional probability density function is described in dependence on u by

(2.10)
$$f(\xi_\rho \mid u) = \begin{cases} 1 - r & \text{if } \xi_\rho = 0, \\ r & \text{if } \xi_\rho = y - u, \end{cases}$$

where y is a uniformly distributed random variable in the range $[0, 1]$. The mean and the variance depending on u are given by

(2.11)
$$E(\rho \mid u) = r\left(\frac{1}{2} - u\right) \quad \text{and} \quad \text{Var}(\rho) = r\left(\frac{1}{3} - u + u^2\right) - r^2\left(\frac{1}{4} - u + u^2\right).$$

Similarly as above, the following estimates are obtained

$$(2.12) \qquad E(\rho \mid u) \in \left[-\frac{r}{2}, \frac{r}{2}\right] \quad \text{and} \quad \text{Var}(\rho \mid u) \leq \frac{r}{3} + \frac{r^2}{4}.$$

Here, the expected absolute value depends also on the value u:

$$(2.13) \qquad \text{EAV}(\rho \mid u) = E(|\rho| \mid u) = r\left(u^2 - u + \frac{1}{2}\right).$$

Since the range of u is between 0 and 1, we have that $\text{EAV}(\rho \mid u) \in \left[\frac{r}{4}, \frac{r}{2}\right]$.

3. Constrained and unconstrained L^1-TV models

As it was mentioned in the introduction, a model commonly used to restore images contaminated by Gaussian white noise is the ROF-model

$$(3.1) \qquad \text{minimize} \int_\Omega |Du| + \frac{\lambda}{2} \int_\Omega |Ku - z|^2 dx \text{ over } u \in BV(\Omega),$$

with $K \in \mathcal{L}(L^2(\Omega))$ and $z \in L^2(\Omega)$. Recently, in [12] the locally constrained model
(3.2)
$$\text{minimize} \int_\Omega |Du| \text{ over } u \in BV(\Omega) \text{ s.t. } \int_\Omega w(x,y)|Ku - z|^2(y)dy \leq \sigma^2 \text{ for almost every } x \in \Omega$$

was considered, where w is the mean filter defined as

$$(3.3) \qquad w(x,y) = \begin{cases} \frac{1}{\omega_\delta^2} & \text{if } |y - x|_{l^\infty} \leq \frac{\omega}{2}, \\ \delta & \text{else,} \end{cases}$$

with $x \in \Omega$ fixed. The quantity $\omega > 0$ is assumed sufficiently small and represents the essential width of the filter window and $0 < \delta \ll 1$. Moreover, we have ω_δ such that $\int_\Omega \int_\Omega w(x,y)dydx = 1$. Thus, the Lagrange multiplier associated with the constraints in the model (3.2) is used to provide an adequate function λ when solving the unconstrained problem

$$(3.4) \qquad \text{minimize} \int_\Omega |Du| + \frac{1}{2} \int_\Omega \lambda |Ku - z|^2 dx \text{ over } u \in BV(\Omega).$$

The present text focusses on the L^1-total-variation restoration model

$$(3.5) \qquad \text{minimize} \int_\Omega |Du| + \int_\Omega \lambda |Ku - z| \, dx \text{ over } u \in BV(\Omega),$$

where $\lambda \in L^\infty(\Omega)$ with ess inf $\lambda \geq \underline{\lambda} > 0$ and its locally constrained variant
(3.6)
$$\text{minimize} \int_\Omega |Du| \text{ over } u \in BV(\Omega) \text{ s.t. } \int_\Omega w(x,y)|Ku - z|(y)dy \leq \nu \text{ for almost every } x \in \Omega.$$

Here, the value of $\nu \in \mathbb{R}$ is the expected absolute value that depends on the type of noise, such as Gaussian white noise, salt-and-pepper noise or random-valued impulse noise. Note that here we assume that in the case of random-valued impulse noise, ν is a fixed number in the interval $[\frac{r}{4}, \frac{r}{2}]$. In our numerics, however, we report on results when ν is selected empirically in dependence on some approximation of the true image.

For the existence proof for (3.5) and (3.6) a few prerequisits are needed. Assume that w is a normalized filter, i.e. $w \in L^\infty(\Omega \times \Omega)$, and $w \geq 0$ on $\Omega \times \Omega$ with

$$(3.7) \qquad \int_\Omega \int_\Omega w(x,y) \, dy \, dx = 1 \text{ and } \int_\Omega \int_\Omega w(x,y)|\phi(y)| \, dy \, dx \geq \epsilon \|\phi\|_{L^1(\Omega)} \ \forall \phi \in L^1(\Omega)$$

for some $\epsilon > 0$ independent of ϕ. An instance satisfying (3.7) is the mean filter defined in (3.3). The w-smoothed version of $|Ku - z|$ is denoted by $S(u)$ and is given by

$$(3.8) \qquad S(u)(x) := \int_\Omega w(x,y)|Ku - z|(y) \, dy \text{ for } x \in \Omega.$$

Since $Ku - z \in L^1(\Omega)$ and $w \in L^\infty(\Omega \times \Omega)$, we have $S(u) \in L^\infty(\Omega)$. Moreover, note that $S : L^2(\Omega) \to L^\infty(\Omega)$ is continuous. The closed and convex feasible set of (3.6) is given by

(3.9) $$\mathcal{U} = \{u \in BV(\Omega) \, : \, S(u) \leq \nu \text{ a.e. in } \Omega\}.$$

In order to prove the existence of a solution to the problems (3.5) and (3.6), a technique similar to the one of [1] is utilized in this section. We start by establishing a BV-coercivity result for the functional

(3.10) $$\mathcal{Q}(u) = J(u) + \int_\Omega S(u)(x) \, dx,$$

where $J(u) := \int_\Omega |Du|$.

Proposition 1. *Assume that K does not annihilate constant functions, i.e. $K\chi_\Omega \neq 0$ where $\chi_\Omega(x) = 1$ for $x \in \Omega$. Then $\|u\|_{BV(\Omega)} \to \infty$ implies $\mathcal{Q}(u) \to \infty$.*

Proof. Any $u \in BV(\Omega)$ can be decomposed according to

$$u = t + v, \quad \text{with} \quad t = \left(\frac{\int_\Omega u \, dx}{|\Omega|}\right) \chi_\Omega \quad \text{and} \quad \int_\Omega v \, dx = 0.$$

This yields

$$\|u\|_{BV} \leq \|v\|_{BV} + \|t\|_{BV} = \int_\Omega |v| \, dx + J(v) + \int_\Omega |t| \, dx + J(t)$$
$$= J(v) + \|v\|_{L^1(\Omega)} + \|t\|_{L^1(\Omega)}$$
$$\leq (1 + C_1)J(v) + \|t\|_{L^1(\Omega)}.$$

here the inequality $\|v\|_{L^1(\Omega)} \leq C_1 J(v)$ results from the Sobolev inequality $\|v\|_{L^2(\Omega)} \leq C_4 J(v)$ with $C_4 > 0$; see [15, page 24]. Thus, we infer

$$\mathcal{Q}(u) \geq J(u) + \epsilon \|Kv - z + Kt\|_{L^1(\Omega)}$$
$$\geq J(u) + \epsilon \|Kt\|_{L^1(\Omega)} - \epsilon \|Kv - z\|_{L^1(\Omega)}$$
$$\geq J(u) + \epsilon \|Kt\|_{L^1(\Omega)} - \epsilon(C_3 \|v\|_{L^2(\Omega)} + \|z\|_{L^1(\Omega)})$$
$$\geq J(v) + \epsilon \|Kt\|_{L^1(\Omega)} - \epsilon C_3 C_4 J(v) - \epsilon \|z\|_{L^1(\Omega)}$$
$$\geq J(v) + \epsilon C_2 \|t\|_{L^1(\Omega)} - \epsilon C_3 C_4 J(v) - \epsilon \|z\|_{L^1(\Omega)}$$
$$\geq (1 - \epsilon C_3 C_4)J(v) + \epsilon C_2 \|t\|_{L^1(\Omega)} - \epsilon \|z\|_{L^1(\Omega)},$$

where C_2, C_3 and C_4 are positive constants. Here we used the fact that K does not annihilate constant functions, i.e. C_2 is such that $\|Kt\|_{L^1(\Omega)} \geq C_2 \|t\|_{L^1(\Omega)}$. The above estimate yields

$$\|t\|_{L^1(\Omega)} \leq \frac{1}{\epsilon C_2}[\mathcal{Q}(u) + \epsilon \|z\|_{L^1(\Omega)} + (\epsilon C_3 C_4 - 1)J(v)]$$
$$\leq \frac{1}{\epsilon C_2}[\mathcal{Q}(u) + \epsilon C_3 C_4 \mathcal{Q}(u) + \epsilon \|z\|_{L^1(\Omega)}]$$
$$\leq C(\epsilon)\mathcal{Q}(u) + \frac{1}{C_2}\|z\|_{L^1(\Omega)}$$

where $C(\epsilon) = \max\left(\frac{1}{\epsilon C_2}, \frac{C_3 C_4}{C_2}\right)$. From this we obtain that

$$\|u\|_{BV} \leq \|t\|_{L^1(\Omega)} + (1 + C_1)J(v)$$
$$\leq C(\epsilon)\mathcal{Q}(u) + \frac{1}{C_2}\|z\|_{L^1(\Omega)} + (1 + C_1)J(v)$$
$$\leq [C(\epsilon) + 1 + C_1]\mathcal{Q}(u) + \frac{1}{C_2}\|z\|_{L^1(\Omega)},$$

which yields the assertion. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

The existence of a solution of the problem (3.6) is argued as follows.

Theorem 2. *Assume that $K \in \mathcal{L}(L^2(\Omega))$ does not annihilate constant functions. Then problem (3.6) admits a solution.*

Proof. Let $\{u_n\} \subset \mathcal{U}$ be a minimizing sequence. Since $\{u_n\}$ must satisfy the constraints in (3.6), we have that $\mathcal{Q}(u_n) \le J(u_n) + |\Omega|\nu$ where $\mathcal{Q}$ is given in (3.10) and $|\Omega|$ is the volume of the domain Ω. Hence, Proposition 1 implies that $\{u_n\}$ is bounded in $BV(\Omega)$. According to Theorem 2.6 in [1] there exists a subsequence $\{u_{n_k}\}$ which converges weakly in $L^2(\Omega)$ to some $\tilde{u} \in L^2(\Omega)$. The functional J is weakly lower semicontinuous with respect to the $L^2(\Omega)$ topology; see Theorem 2.3 in [1]. Thus, $\tilde{u} \in BV(\Omega)$. The sequence $\{Du_{n_k}\}$ converges weakly as a measure to $D\tilde{u}$, by Lemma 2.5 in [1]. Hence, we have

$$(3.11) \qquad J(\tilde{u}) \le \liminf_{k \to \infty} J(u_{n_k}) = \inf_{u \in \mathcal{U}} J(u).$$

Since K is a continuous linear operator, $\{Ku_{n_k}\}$ converges weakly to $K\tilde{u}$ in $L^2(\Omega)$. Moreover, since $\mathcal{U}$ is convex and closed (and, thus, weakly closed), we have $S(\tilde{u}) \le \nu$ a.e. in Ω. Therefore, $\tilde{u}$ is a solution to the problem. $\qquad \square$

Due to the convexity (only) of the problem, there is no uniqueness result. Next, we study the relation between the problems (3.5) and (3.6). As it is done in [12], the following penalized problem is considered:

$$(3.12) \qquad \text{minimize} \quad \mathcal{Q}_\gamma(u) := J(u) + \frac{\gamma}{2} \int_\Omega \max(S(u) - \nu, 0)^2 dx \quad \text{over} \quad u \in BV(\Omega),$$

where $\gamma > 0$ denotes a penalty parameter. The proof of the following proposition is similar to the one of Proposition 5 in [12].

Proposition 3. *Let the assumptions of Theorem 2 be satisfied. Then problem (3.12) admits a solution $u_\gamma \in BV(\Omega)$ for every $\gamma > 0$. Moreover, for $\gamma \to +\infty$, $\{u_\gamma\}$ converges weakly along a subsequence in $L^2(\Omega)$ to a solution of (3.6).*

Proof. Since $S : L^2(\Omega) \to L^\infty(\Omega)$ and $\max(\cdot, 0) : L^2(\Omega) \to L^2(\Omega)$ are continuous and convex, respectively, and since $J(u)$ is weakly lower semicontinuous, then $\mathcal{Q}_\gamma : BV(\Omega) \to \mathbb{R}$ is weakly lower semicontinuous. Let $\{u_n\} \subset BV(\Omega)$ be a minimizing sequence, and let $\tilde{u}$ be a solution of (3.6). For sufficiently large n, $\mathcal{Q}_\gamma(u_n) \le \mathcal{Q}_\gamma(\tilde{u}) + 1 = J(\tilde{u}) + 1$. Further, since $S(u) \ge 0$ a.e. in Ω for any $u \in BV(\Omega)$ there exists a constant C (independent of n and γ) such that $\|S(u_n)\|_{L^2(\Omega)} \le C$ for all $n \in \mathbb{N}$. By Proposition 1, $\{u_n\}$ is bounded in $BV(\Omega)$. Observing the lower-semicontinuity of the penalty term, the existence of a solution of (3.12) for fixed γ can be argued similarly as in the proof of Theorem 2. Moreover, it is clear that $\{u_\gamma\}$ is bounded in $BV(\Omega)$. By weak lower semicontinuity

$$J(\tilde{u}_\gamma) \le \liminf_{\gamma \to +\infty} \mathcal{Q}_\gamma(u_\gamma) \le J(\tilde{u}) = \inf_{u \in \mathcal{U}} J(u),$$

where $\tilde{u}_\gamma$ is a weak limit in $L^2(\Omega)$ of a subsequence of $\{u_\gamma\}$ in $L^2(\Omega)$ (still denoted by $\{u_\gamma\}$). It is necessary to show that $\tilde{u}_\gamma \in \mathcal{U}$. For all $\gamma > 0$ we have

$$\frac{\gamma}{2} \int_\Omega \max(S(u_\gamma) - \nu, 0)^2 \, dx \le J(\tilde{u}).$$

Thus, as $\gamma \to \infty$ we infer

$$\int_\Omega \max(S(u_\gamma) - \nu, 0)^2 dx \to 0$$

and by the continuity of K, weak lower semicontinuity and Fatou's lemma we have $S(\tilde{u}_\gamma) \le \nu$ a.e. in Ω. $\qquad \square$

In the same way, it can be obtained that

$$(3.13) \qquad \| \max(S(u_{\gamma_n}) - \nu, 0)\|_{L^2(\Omega)} = o(1/\sqrt{\gamma_n}),$$

where $o(a_n)/a_n \to 0$ for a sequence $\{a_n\}$ with $a_n \to 0$.

For subsequent results, let us define

$$(3.14) \qquad \lambda_\gamma^\circ \quad := \quad \gamma \max(S(u_\gamma) - \nu, 0),$$

$$(3.15) \qquad \lambda_\gamma \quad := \quad \int_\Omega w(x,y) \lambda_\gamma^\circ(y) \, dy.$$

Utilizing the proof technique of Theorem 6 of [12], we obtain the following result.

Theorem 4. *Let the assumptions of Theorem 2 hold true. Moreover, we assume that there exists $C > 0$ such that $\gamma_n \| \max(S(u_{\gamma_n}) - \nu, 0)\|_{L^1(\Omega)} \leq C$ for all $n \in \mathbb{N}$. Then there exist $\tilde{\lambda} \in L^\infty(\Omega)$, a bounded Borel measure $\tilde{\lambda}^\circ$ and a subsequence $\{\gamma_{n_k}\}$ such that the following properties hold true:*

(i) *$\int_\Omega \lambda_{\gamma_{n_k}} g \, dx \to \int_\Omega \tilde{\lambda} g \, dx$ for all $g \in L^1(\Omega)$ and $\tilde{\lambda} \geq 0$ a.e. in Ω.*

(ii) *$\int_\Omega \varphi \lambda_{\gamma_{n_k}}^\circ \, dx \to \int_\Omega \varphi d\tilde{\lambda}^\circ$ for all $\varphi \in C(\bar{\Omega})$, $\tilde{\lambda}^\circ \geq 0$ and $\int_\Omega \lambda_{\gamma_n}^\circ (S(u_{\gamma_n}) - \nu) \, dx \to 0$.*

Proof. The assumption of the existence of a constant $C > 0$ such that $\gamma_n \| \max(S(u_{\gamma_n}) - \nu, 0)\|_{L^1(\Omega)} \leq C$ implies that there exists a constant $C' > 0$ independent of γ_n such that

$$\|\lambda_{\gamma_n}\|_{L^\infty(\Omega)} \leq \gamma_n \|w\|_{L^\infty(\Omega \times \Omega)} \| \max(S(u_{\gamma_n}) - \nu, 0)\|_{L^1(\Omega)} \leq C'.$$

Thus, the weak* sequential compactness of the closed unit ball in $L^\infty(\Omega)$ yields the first part of (i). The non-negativity of $\tilde{\lambda}$ is implied of the definition of λ_γ. This proves (i).

Concerning (ii) we observe that

$$\left| \int_\Omega \lambda_{\gamma_n}^\circ \max(S(u_{\gamma_n}) - \nu, 0) dx \right| = \gamma_n \| \max(S(u_{\gamma_n}) - \nu, 0)\|_{L^2(\Omega)}^2.$$

Thus, (3.13) implies the last relation of (ii). The first limit in (ii) follows immediately from the boundedness of $\gamma_n \| \max(S(u_{\gamma_n}) - \nu, 0)\|_{L^1(\Omega)}$ and, hence, of $|\lambda_{\gamma_n}^\circ|$. The non-negativity of $\tilde{\lambda}^\circ$ is an immediate consequence of the definition of λ_γ° $\qquad \square$

We note that if (3.13) holds true with $o(1/\sqrt{\gamma})$ replaced by $\mathcal{O}(1/\gamma)$, i.e. there exists a constant C such that

$$(3.16) \qquad \| \max(S(u_{\gamma_n}) - \nu, 0)\|_{L^2(\Omega)} \leq \frac{\tilde{C}}{\gamma_n},$$

then $\tilde{\lambda}^\circ \in L^2(\Omega)$. In this case, the system of Theorem 4(ii) becomes

$$\tilde{\lambda}^\circ \geq 0 \text{ a.e. in } \Omega, \quad S(\tilde{u}) \leq \nu \text{ a.e. in } \Omega, \quad \lim_{n \to \infty} \int_\Omega \lambda_{\gamma_n}^\circ (S(u_{\gamma_n}) - \nu) \, dx = 0.$$

If the last relation above holds as $\int_\Omega \tilde{\lambda}^\circ (S(\tilde{u}) - \nu) \, dx = 0$, then we may equivalently write

$$(3.17) \qquad \tilde{\lambda}^\circ \geq 0 \text{ a.e. in } \Omega, \quad \tilde{\lambda}^\circ = \tilde{\lambda}^\circ + \tau \max(S(\tilde{u}) - \nu, 0),$$

with $\tau > 0$ arbitrary, but fixed.

4. Primal-Dual Method for L^1 TV Image Restoration

In this section we propose a primal-dual method for solving (3.5). We start by recalling the Fenchel duality theorem in infinite-dimensional spaces; see [14].

Theorem 5. *Let V and Y be Banach spaces with topological duals denoted by V^* and Y^*, respectively. Let $\Lambda \in \mathcal{L}(V, Y)$, with Λ^* its adjoint, and let $\Psi : V \to \mathbb{R} \cup \{\infty\}$, $\Phi : Y \to \mathbb{R} \cup \{\infty\}$ be convex lower semicontinuous functionals not identically equal to ∞, and assume that there exists $v_0 \in V$ such that $\Psi(v_0) < \infty$, $\Phi(\Lambda v_0) < \infty$, and Φ is continuous at Λv_0. Then*

$$(4.1) \qquad \inf_{v \in V} \{\Psi(v) + \Phi(\Lambda v)\} = \sup_{q \in Y^*} \{-\Psi^*(\Lambda^* q) - \Phi^*(-q)\},$$

where $\Psi^* : V^* \to \mathbb{R} \cup \{\infty\}$ denotes the conjugate of Ψ defined by

$$(4.2) \qquad \Psi^*(v^*) = \sup_{v \in V} \left\{ \langle v, v^* \rangle_{V, V^*} - \Psi(v) \right\}.$$

Moreover, $(\bar{v}, \bar{q})$ is a solution pair for (4.1) *if and only if*

$$(4.3) \qquad\qquad\qquad \Lambda^* \bar{q} \in \partial\Psi(\bar{v}),$$

$$(4.4) \qquad\qquad\qquad -\bar{q} \in \partial\Phi(\Lambda\bar{v}).$$

4.1. Fenchel calculus.
We begin with the formal computation of the Fenchel-Legendre dual of (3.5). For this purpose, set $\Lambda := \nabla$,

$$\Psi(u) := \int_\Omega \lambda\,|Ku - z|\,dx \quad \text{and} \quad \Phi(\vec{q}) := \int_\Omega |\vec{q}|_{l^2}\,dx.$$

Thus, the conjugate of $\Psi(u)$ is defined by

$$(4.5) \qquad\qquad \Psi^*(u^*) = \sup_u \left\{ \langle u, u^* \rangle - \int_\Omega \lambda|Ku - z|dx \right\},$$

where $\langle u, v \rangle = \int_\Omega uv\,dx$. From this we obtain that

$$u^* - K^*\lambda\sigma(Ku - z) = 0,$$

where $\sigma(v) \in \partial|v|$. Here and below $\partial(\cdot)$ denotes the subdifferential of convex analysis. Hence Ψ^* has a maximum if $u^* = K^*\lambda\sigma(Ku - z)$. This implies $Ku^* = KK^*\lambda\sigma(Ku - z)$ and, if KK^* is invertible, $\lambda\sigma(Ku - z) = (KK^*)^{-1}Ku^*$. Therefore, we find

$$\Psi^*(u^*) = \langle u, K^*\lambda\sigma(Ku - z) \rangle - \int_\Omega \lambda|Ku - z|dx$$

$$= \int_{\{Ku \geq z\}} (\lambda Ku - \lambda Ku + \lambda z)\ dx + \int_{\{Ku < z\}} (-\lambda Ku + \lambda Ku - \lambda z)\ dx$$

$$= \int_\Omega z\lambda\sigma(Ku - z)\ dx$$

$$= \langle z, (KK^*)^{-1}Ku^* \rangle$$

together with the condition $|(KK^*)^{-1}Ku^*| \leq \lambda$ due to $(KK^*)^{-1}Ku^* = \lambda\sigma(Ku - z)$ and $\sigma(Ku - z) \in \partial|Ku - z|$. Thus, the conjugates of $\Psi(u)$ and $\Phi(\vec{q})$ are

$$\Psi^*(u^*) := \langle z, (KK^*)^{-1}Ku^* \rangle + I_{\{|(KK^*)^{-1}Ku^*| \leq \lambda\}}(u^*),$$

$$\Phi^*(\vec{q}^*) := I_{\{|\vec{w}(x)|_{l^2} \leq 1\}}(\vec{q}^*),$$

where I_S denotes the indicator function of S. For the latter we refer to [16].

According to the Fenchel duality theorem, for $u^* = \nabla^*\vec{p}^* = -\text{div}\vec{p}$, the dual problem of (3.5) is

$$(4.6) \qquad\qquad \begin{aligned} &\text{minimize } -\langle z, (KK^*)^{-1}K\,\text{div}\vec{p}^* \rangle \text{ over } \vec{p}^* \\ &\text{s.t. } |(KK^*)^{-1}K\,\text{div}\vec{p}^*| \leq \lambda, \\ &\qquad |\vec{p}^*|_{l^2} \leq 1. \end{aligned}$$

4.2. Fenchel predual.
The computations so far have been just formal. In this section we fix an appropriate function space setting so that we are able to apply Theorem 5 and rigorously establish (4.6) as the Fenchel pre-dual of (3.5). From now on we assume that

$$(4.7) \qquad\qquad\qquad (KK^*) \text{ is continuously invertible.}$$

We note that this assumption imposes a condition on the possible blurring operators. In our numerics we utilize a Gaussian convolution with a small window size (9-by-9) which satisfies (4.7).

Define $\Lambda := (KK^*)^{-1}K\text{div}$, $Y := L^2(\Omega)$, $V := H_0(\text{div})$, where

$$\mathbb{L}^2(\Omega) = L^2(\Omega) \times L^2(\Omega), \quad H_0(\text{div}) = \{\vec{v} \in \mathbb{L}^2(\Omega) :\ \text{div}\,\vec{v} \in L^2(\Omega),\ \vec{v}\cdot\vec{n} = 0 \text{ on } \partial\Omega\}, \quad \text{and}$$

$$(4.8) \qquad\qquad \begin{aligned} &\Phi : Y \to \mathbb{R}, \quad \Phi(v) = -\langle z, v \rangle + I_{\{v \in L^2(\Omega)\ :\ |v| \leq \lambda \text{ a.e. in } \Omega\}}(v), \\ &\Psi : V \to \mathbb{R}, \quad \Psi(\vec{p}) = I_{\{\vec{w} \in H_0(\text{div})\ :\ |\vec{w}|_{l^2} \leq 1 \text{ a.e. in } \Omega\}}(\vec{p}), \end{aligned}$$

where $\vec{n}$ denotes the outward unit normal to $\partial\Omega$. Consider the problem

$$(4.9) \qquad\qquad \text{minimize } \Phi(\Lambda\vec{p}) + \Psi(\vec{p}) \text{ over } \vec{p} \in H_0(\text{div}).$$

Note that Theorem 5 requires the continuity of Φ at a point $\Lambda \vec{p}_0$ where $\vec{p}_0 \in H_0(\mathrm{div})$. This, however, can not be guaranteed with such a setting due to the presence of the indicator function. Therefore, the following Moreau-Yosida type regularization of the indicator function in Φ is introduced:

(4.10)
$$\text{minimize} - \langle z, \Lambda \vec{p} \rangle + \tfrac{\mu}{2} \| (|\Lambda \vec{p}| - \lambda)^+ \|^2_{L^2(\Omega)} \ \text{ over } \ \vec{p} \in H_0(\mathrm{div})$$
$$\text{s.t. } |\vec{p}|_{l^2} \le 1 \ \text{ a.e. in } \Omega.$$

where $(v)^+ = \max(v, 0)$ in the pointwise a.e. sense and $\mu > 0$. Rather than the choices in (4.8), the following functions are considered for dualization:

(4.11)
$$\Phi_\mu : Y \to \mathbb{R}, \quad \Phi_\mu(v) = - \langle z, v \rangle + \tfrac{\mu}{2} \| (|v| - \lambda)^+ \|^2_{L^2(\Omega)},$$
$$\Psi : V \to \mathbb{R}, \quad \Psi(\vec{p}) = I_{\{\vec{w} \in H_0(\mathrm{div}) \, : \, |\vec{w}|_{l^2} \le 1 \text{ a.e. in } \Omega\}}(\vec{p}).$$

The convex conjugates of Φ_μ and Ψ are given by

(4.12)
$$\Phi_\mu^* : Y^* \to \mathbb{R}, \quad \Phi_\mu^*(v^*) = \sup_{v \in Y} \left\{ \langle v, v^* \rangle + \langle z, v \rangle - \tfrac{\mu}{2} \| (|v| - \lambda)^+ \|^2_{L^2(\Omega)} \right\},$$
$$\Psi^* : V^* \to \mathbb{R}, \quad \Psi^*(\vec{p}^*) = \sup_{\vec{p} \in V} \left\{ \langle \vec{p}, \vec{p}^* \rangle_{V, V^*} - I_{\{\vec{w} \in H_0(\mathrm{div}) \, : \, |\vec{w}|_{l^2} \le 1 \text{ a.e. in } \Omega\}}(\vec{p}) \right\}.$$

For Φ_μ^* we obtain

(4.13)
$$v^* + z - \mu \, \mathrm{sign}(v)(|v| - \lambda)^+ = 0.$$

For our further computations, we consider the following sets:

$$\Omega_0 = \{x \in \Omega : |v(x)| \le \lambda(x)\}, \quad \text{implying} \quad v^* + z = 0,$$
$$\Omega_+ = \{x \in \Omega : v(x) > \lambda(x)\}, \quad \text{implying} \quad v = \frac{v^* + z}{\mu} + \lambda,$$
$$\Omega_- = \{x \in \Omega : v(x) < -\lambda(x)\}, \quad \text{implying} \quad v = \frac{v^* + z}{\mu} - \lambda.$$

By splitting the domain Ω into the previous subsets, i.e. $\Omega = \Omega_0 \dot\cup \Omega_+ \dot\cup \Omega_-$, it can be shown that

(4.14)
$$\Phi_\mu^*(v^*) = \frac{1}{2\mu} \| v^* + z \|^2_{L^2(\Omega)} + \| \lambda(v^* + z) \|_{L^1(\Omega)}.$$

According to the results in [16], the conjugate Ψ^* is given by

(4.15)
$$\Psi^*(\vec{p}^*) = \sup_{\vec{p} \in S_1} \langle \vec{p}, \vec{p}^* \rangle,$$

where $S_1 = \{\vec{p} \in H_0(\mathrm{div}) \, : \, |\vec{p}|_{l^2} \le 1 \text{ a.e. in } \Omega\}$. Let us define the set

(4.16)
$$S_2 = \{\vec{p} \in C_0^1(\Omega) \times C_0^1(\Omega) \, : \, |\vec{p}|_{l^2} \le 1 \text{ a.e. in } \Omega\},$$

which is dense in the topology of $H_0(\mathrm{div})$ in S_1. Further, $(\mathcal{D}(\Omega))^2$ is dense in $H_0(\mathrm{div})$, where $\mathcal{D}(\Omega)$ is the space of test functions [2]. Let $\vec{p}$ be an arbitrary element of S_1 and let the sequence $\{\vec{p}_n\} \in (\mathcal{D}(\Omega))^2$ converge in $H_0(\mathrm{div})$ to $\vec{p}$. By $\mathcal{P}$ we denote the canonical projection in $H_0(\mathrm{div})$ onto the closed convex subset S_1. Since $\vec{p} \in S_1$, we have that

$$\| \vec{p} - \mathcal{P}\vec{p}_n \|_{H_0(\mathrm{div})} \le \| \vec{p} - \vec{p}_n \|_{H_0(\mathrm{div})} + \| \vec{p}_n - \mathcal{P}\vec{p}_n \|_{H_0(\mathrm{div})}$$
$$\le 2\| \vec{p} - \vec{p}_n \|_{H_0(\mathrm{div})} \to 0 \quad \text{as} \quad n \to \infty,$$

where $\| \vec{w} \|^2_{H_0(\mathrm{div})} = \| \vec{w} \|^2_{\mathbb{L}^2(\Omega)} + \| \mathrm{div}\ \vec{w} \|^2_{L^2(\Omega)}$. According to (4.15), for $v^* \in L^2(\Omega)$ and $\Lambda^* \in \mathcal{L}(L^2(\Omega), H_0(\mathrm{div})^*)$ we have

(4.17)
$$\Psi^*(\Lambda^* v^*) = \sup_{\vec{p} \in S_2} \langle v^*, -(KK^*)^{-1}K \mathrm{div}\vec{p} \rangle.$$

We define $u = K^*(KK^*)^{-1}v^*$, such that $\Lambda^* v^* = (-\mathrm{div})^* K^*(KK^*)^{-1}v^* = (-\mathrm{div})^* u$. Hence,

(4.18)
$$\Psi^*((-\mathrm{div})^* u) = \sup_{\vec{p} \in S_2} \langle u, -\mathrm{div}\vec{p} \rangle.$$

This function is finite if and only if $u \in BV(\Omega)$, i.e.

$$(4.19) \qquad \Psi^*((-\mathrm{div})^* u) = \int_\Omega |Du| < \infty \ \text{ for } \ u \in BV(\Omega).$$

Since $u = K^*(KK^*)^{-1}v^*$, we infer $v^* = Ku$ and the function $\Phi_\mu^*(-v^*)$ becomes

$$(4.20) \qquad \Phi_\mu^*(-v^*) = \frac{1}{2\mu}\|Ku - z\|^2_{L^2(\Omega)} + \|\lambda(Ku - z)\|_{L^1(\Omega)}.$$

By (4.1), the dual problem of (4.10) is given by

$$(4.21) \qquad \text{minimize } \int_\Omega |Du| + \frac{1}{2\mu}\|Ku - z\|^2_{L^2(\Omega)} + \|\lambda(Ku - z)\|_{L^1(\Omega)} \ \text{ over } \ u \in BV(\Omega).$$

Moreover, according to (4.3), a primal-dual solution pair $(\bar{u}, \vec{\bar{p}})$ of the problems (4.10) and (4.21) satisfies

$$(4.22) \qquad \begin{aligned} -\mathrm{div}\vec{\bar{p}} &= \tfrac{1}{\mu}K^*(Ku - z) + K^*\lambda\,\sigma(K\bar{u} - z), \\ \langle(-\mathrm{div})^*\bar{u}, \vec{p} - \vec{\bar{p}}\rangle_{H_0(\mathrm{div})^*, H_0(\mathrm{div})} &\leq 0, \ \text{ for all } \ \vec{p} \in S_1. \end{aligned}$$

Notice the relation between (4.21) and the original problem (1.6). In fact, the penalization of the dual constraint $|\Lambda\vec{p}| \leq \lambda$ a.e. in Ω in (4.10) yields, after yet another dualization, a least-squares data fidelity term in $L^2(\Omega)$ weighted by $\frac{1}{2\mu}$ as it is the case in (4.21). Thus, the resulting problem has a combined L^1-L^2-fidelity term. Since we are interested in $\mu \to \infty$, the L^1-fidelity term increasingly dominates. It can even be shown that (4.21) converges to (1.6) as $\mu \to \infty$. In our numerics, we choose $\mu = 10^6$.

4.3. Reconstructabiliy, a first regularization.
It turns out that for a given dual solution $\vec{\bar{p}}$, the image intensity u, i.e. the primal solution, cannot be recovered in general from (4.22). As a remedy we introduce an appropriate regularization of the dual problem (4.10), which is more amenable to computations than the primal problem (4.21). In fact, we consider the following dual regularization:

$$(4.23) \qquad \begin{aligned} &\text{minimize } -\langle z, \Lambda\vec{p}^*\rangle + \tfrac{\mu}{2}\|\,(|\Lambda\vec{p}^*| - \lambda)^+\,\|^2_{L^2(\Omega)} + \tfrac{\beta}{2}\|\Lambda\vec{p}^*\|^2_{L^2(\Omega)} \text{ over } \vec{p}^* \in H_0(\mathrm{div}), \\ &\text{s.t. } |\vec{p}^*|_{l^2} \leq 1 \text{ a.e. in } \Omega, \end{aligned}$$

where $\beta > 0$ is a regularization parameter. As we shall see from dualization, the β-term results in a local smoothing of the L^1 fidelity term in (4.21). In order to see this, define

$$(4.24) \qquad \begin{aligned} \Phi_{\mu,\beta}(v) &:= -\langle z, v\rangle + \tfrac{\mu}{2}\|(|v| - \lambda)^+\|^2_{L^2(\Omega)} + \tfrac{\beta}{2}\|v\|^2_{L^2(\Omega)}, \\ \Psi(\vec{p}) &:= I_{\{\vec{w} \in H_0(\mathrm{div}) : |\vec{w}|_{l^2} \leq 1 \text{ a.e. in } \Omega\}}(\vec{p}^*). \end{aligned}$$

The dual of $\Phi_{\mu,\beta}(v)$ is given by $\Phi_{\mu,\beta}^*(v^*) = \sup_v \{\langle v, v^*\rangle - \Phi_{\mu,\beta}(v)\}$. Hence, we obtain

$$(4.25) \qquad v^* + z - \mu\,\sigma(v)(|v| - \lambda)^+ - \beta v = 0.$$

Splitting the domain Ω according to $\Omega = \Omega_0 \dot{\cup} \Omega_+ \dot{\cup} \Omega_-$ yields

$$\Omega_0 = \{x \in \Omega \ : \ |v(x)| \leq \lambda(x)\}, \text{ which implies } v = \frac{v^* + z}{\beta};$$

$$\Omega_+ = \{x \in \Omega \ : \ v(x) > \lambda(x)\}, \text{ which implies } v = \frac{v^* + z + \mu\lambda}{\beta + \mu};$$

$$\Omega_- = \{x \in \Omega \ : \ v(x) < -\lambda(x)\}, \text{ which implies } v = \frac{v^* + z - \mu\lambda}{\beta + \mu}.$$

From this we obtain that

$$(4.26)$$
$$\Phi_{\mu,\beta}^*(v^*)(x) = \begin{cases} \frac{1}{2\beta}|v^* + z|^2(x), & \text{if } |v^* + z|(x) \leq \lambda(x)\beta, \\ \frac{1}{2(\beta+\mu)}|v^* + z|^2(x) + \frac{\mu}{\mu+\beta}|\lambda(v^* + z)|(x) - \frac{\mu\beta}{2(\mu+\beta)}\lambda^2(x) & \text{if } |v^* + z|(x) > \lambda(x)\beta. \end{cases}$$

As $u = K^*(KK^*)^{-1}v^*$, we have $v^* = Ku$ and the function $\Phi^*_{\mu,\beta}(-v^*)$ becomes

(4.27)

$$\Phi^*_{\mu,\beta}(-Ku)(x) = \begin{cases} \frac{1}{2\beta}|Ku - z|^2(x), & \text{if } |Ku - z|(x) \le \lambda(x)\beta, \\ \frac{1}{2(\beta+\mu)}|Ku - z|^2(x) + \frac{\mu}{\mu+\beta}|\lambda(Ku - z)|(x) - \frac{\mu\beta}{2(\mu+\beta)}\lambda^2(x) & \text{if } |Ku - z|(x) > \lambda(x)\beta. \end{cases}$$

As before, we obtain

(4.28)
$$\Psi^*((-\mathrm{div})^*u) = \int_\Omega |Du| < \infty \ \text{ for } \ u \in BV(\Omega).$$

Thus, the dual problem of (4.23) is given by

(4.29)
$$\text{minimize } \int_\Omega |Du| + \int_\Omega \Phi^*_{\mu,\beta}(-Ku) \, dx \ \text{ over } \ u \in BV(\Omega).$$

4.4. Uniqueness of $\vec{p}$, a second regularization.
To ensure the uniqueness of the solution of the problem (4.23), a Tikhonov type regularization may be used. It is given by
(4.30)
$$\text{minimize } -\langle z, \Lambda\vec{p}\rangle + \tfrac{\mu}{2}\| (|\Lambda\vec{p}| - \lambda)^+ \|^2_{L^2(\Omega)} + \tfrac{\beta}{2}\|\Lambda\vec{p}\|^2_{L^2(\Omega)} + \tfrac{\gamma}{2}\|P_{\mathrm{div}}\vec{p}\|^2_{\mathbb{L}^2(\Omega)} \ \text{ over } \ \vec{p} \in H_0(\mathrm{div}),$$
$$\text{s.t. } |\vec{p}|_{l^2} \le 1 \text{ a.e. in } \Omega,$$

where, as before, $\Lambda = (KK^*)^{-1}K\mathrm{div}$. Here, P_{div} denotes the orthogonal projection in $\mathbb{L}^2(\Omega)$ onto $H_0(\mathrm{div}\, 0)$ with $H_0(\mathrm{div}\, 0) = \{\vec{v} \in H_0(\mathrm{div}) \ : \ \mathrm{div}\,\vec{v} = 0 \text{ a.e. in } \Omega\}$. We have

(4.31)
$$H_0(\mathrm{div}\, 0)^\perp = \{\vec{v} \in \mathrm{grad}\, H^1(\Omega) \ : \ \mathrm{div}\,\vec{v} \in L^2(\Omega), \ \vec{v} \cdot n = 0 \ \text{ on } \ \partial\Omega\},$$
$$H_0(\mathrm{div}) = H_0(\mathrm{div}\, 0)^\perp \oplus H_0(\mathrm{div}\, 0);$$

see [9] for details. We obtain the following result.

Theorem 6. *Let $\bar{\vec{p}} \in H_0(\mathrm{div})$ be the solution to (4.30) and assume that $\mathrm{Ker}_{\mathrm{div}}(K) := \{s \in \mathrm{div}(H_0(\mathrm{div})) \ : \ Ks = 0\} = \{0\}$. Then there exists $\bar{\vec{v}} \in H_0(\mathrm{div})^*$ such that*

(4.32)
$$-\Lambda^* z + \mu \, \Lambda^* \sigma(\Lambda\bar{\vec{p}})(|\Lambda\bar{\vec{p}}| - \lambda)^+ + \beta\Lambda^*\Lambda\bar{\vec{p}} + \gamma P_{\mathrm{div}}\bar{\vec{p}} + \bar{\vec{v}} = 0,$$
$$\langle \bar{\vec{v}}, \vec{p} - \bar{\vec{p}} \rangle_{H_0(\mathrm{div})^*, H_0(\mathrm{div})} \le 0 \ \text{ for all } \ \vec{p} \in H_0(\mathrm{div})$$

with $|\vec{p}|_{l^2} \le 1$ a.e. in Ω, and the solution $\bar{\vec{p}}$ is unique.

Proof. Due to (4.31), every $\vec{w} \in H_0(\mathrm{div})$ can be decomposed according to $\vec{w} = \vec{w}_1 + \vec{w}_2 \in H_0(\mathrm{div}\, 0)^\perp \oplus H_0(\mathrm{div}\, 0)$. Thus, the problem (4.30) can be expressed in the following way:

(4.33)
$$\text{minimize } \Psi(\vec{p}) + \Phi_1(\Lambda_1\vec{p}_1) + \Phi_2(\Lambda_2\vec{p}_2) \ \text{ over } \ \vec{p} \in H_0(\mathrm{div}),$$

where

(4.34)
$$\Psi(\vec{p}) := I_{\{\vec{w}\in H_0(\mathrm{div}):|\vec{w}|_{l^2}\le 1 \text{ a.e. in } \Omega\}}(\vec{p}),$$
$$\Phi_1(v) := -\langle z, v\rangle + \tfrac{\mu}{2}\|(|v| - \lambda)^+\|^2_{L^2(\Omega)} + \tfrac{\beta}{2}\|v\|^2_{L^2(\Omega)}.$$

and $\Lambda_1 = \Lambda$. As it was done in [16], we set

(4.35)
$$\Phi_2 : \mathbb{L}^2(\Omega) \to \mathbb{R}, \quad \Phi_2(\vec{p}) := \frac{\gamma}{2}\|\vec{p}\|^2_{\mathbb{L}^2(\Omega)}$$

and $\Lambda_2 \in \mathcal{L}(H_0(\mathrm{div}0), \mathbb{L}^2(\Omega))$ with Λ_2 the canonical injection. According to Remark 4.3 of Chapter III in [14], there exists $\bar{\vec{u}}_1 \in \mathbb{L}^2(\Omega)$ and $\bar{\vec{u}}_2 \in \mathbb{L}^2(\Omega)$ such that

(4.36)
$$-\bar{\vec{u}}_1 = -z + \mu\sigma(\Lambda_1\bar{\vec{p}}_1)(|\Lambda_1\bar{\vec{p}}_1| - \lambda)^+ + \beta\Lambda_1\bar{\vec{p}}_1 = -z + \mu\sigma(\Lambda\bar{\vec{p}})(|\Lambda\bar{\vec{p}}| - \lambda)^+ + \beta\Lambda\bar{\vec{p}}$$
$$-\bar{\vec{u}}_2 = \gamma\bar{\vec{p}}_2 = \gamma P_{\mathrm{div}}\bar{\vec{p}}$$
$$\langle \Lambda^*\bar{\vec{u}}_1 + \bar{\vec{u}}_2, \vec{p} - \bar{\vec{p}} \rangle_{H_0(\mathrm{div})^*, H_0(\mathrm{div})} \le 0$$

which yields (4.32) with $\bar{\vec{v}} = \Lambda^*\bar{\vec{u}}_1 + \bar{\vec{u}}_2$.

In order to prove uniqueness of the solution, we consider the variational form of the first equation in (4.32) given by
(4.37)
$$-\langle z, \Lambda\vec{w}\rangle + \mu\,\left\langle \sigma(\Lambda\bar{\vec{p}})(|\Lambda\bar{\vec{p}}| - \lambda)^+, \Lambda\vec{w}\right\rangle + \beta\left\langle \Lambda\bar{\vec{p}}, \Lambda\vec{w}\right\rangle + \gamma\left\langle P_{\mathrm{div}}\bar{\vec{p}}, P_{\mathrm{div}}\vec{w}\right\rangle + \left\langle\bar{\vec{v}}, \vec{w}\right\rangle_{H_0(\mathrm{div})^*, H_0(\mathrm{div})} = 0$$

for all $\vec{w} \in H_0(\mathrm{div})$. Let us suppose that $(\vec{p}_i, \vec{v}_i) \in H_0(\mathrm{div}) \times H_0(\mathrm{div})^*$, $i = 1, 2$, are two solutions with associated multipliers satisfying (4.32), for $\delta\vec{p} = \vec{p}_2 - \vec{p}_1$, $\delta\vec{v} = \vec{v}_2 - \vec{v}_1$ we have

(4.38)
$$F(\vec{p}_1, \vec{p}_2, \vec{w}) + \beta\langle\Lambda\delta\vec{p}, \Lambda\vec{w}\rangle + \gamma\langle P_{\mathrm{div}}\delta\vec{p}, P_{\mathrm{div}}\vec{w}\rangle + \langle\delta\vec{v}, \vec{w}\rangle_{H_0(\mathrm{div})^*, H_0(\mathrm{div})} = 0,$$
$$\langle\delta\vec{v}, \delta\vec{p}\rangle_{H_0(\mathrm{div})^*, H_0(\mathrm{div})} \geq 0,$$

where

(4.39)
$$F(\vec{p}_1, \vec{p}_2, \vec{w}) = \mu\,\left\langle \sigma(\Lambda\vec{p}_2)(|\Lambda\vec{p}_2| - \lambda)^+ - \sigma(\Lambda\vec{p}_1)(|\Lambda\vec{p}_1| - \lambda)^+, \Lambda\vec{w}\right\rangle.$$

Replacing $\vec{w}$ by $\delta\vec{p}$, we have that

(4.40)
$$F(\vec{p}_1, \vec{p}_2, \delta\vec{p}) = \mu\,\left\langle \sigma(\Lambda\vec{p}_2)(|\Lambda\vec{p}_2| - \lambda)^+ - \sigma(\Lambda\vec{p}_1)(|\Lambda\vec{p}_1| - \lambda)^+, \Lambda\vec{p}_2 - \Lambda\vec{p}_1\right\rangle.$$

We define the following functions

(4.41)
$$A(x) := A(x, \vec{p}_1, \vec{p}_2) = \sigma(\Lambda\vec{p}_2(x))(|\Lambda\vec{p}_2(x)| - \lambda(x))^+ - \sigma(\Lambda\vec{p}_1(x))(|\Lambda\vec{p}_1(x)| - \lambda(x))^+,$$
$$B(x) := B(x, \vec{p}_1, \vec{p}_2) = \Lambda\vec{p}_2(x) - \Lambda\vec{p}_1(x)$$

and the following splitting of the domain $\Omega = \dot{\cup}_{i=0}^8 \Omega_i$ and its implications on $A(x)$ and $B(x)$:

(4.42)

i	Ω_i	$A(x)$	$B(x)$
0	$\{x \in \Omega\ :\ |\Lambda\vec{p}_2(x)| \leq \lambda(x), |\Lambda\vec{p}_1(x)| \leq \lambda(x)\}$	$= 0$	
1	$\{x \in \Omega\ :\ |\Lambda\vec{p}_2(x)| \leq \lambda(x), \Lambda\vec{p}_1(x) < -\lambda(x)\}$	$-\Lambda\vec{p}_1(x) - \lambda(x) \geq 0$	≥ 0
2	$\{x \in \Omega\ :\ |\Lambda\vec{p}_2(x)| \leq \lambda(x), \Lambda\vec{p}_1(x) > \lambda(x)\}$	$-\Lambda\vec{p}_1(x) + \lambda(x) \leq 0$	≤ 0
3	$\{x \in \Omega\ :\ \Lambda\vec{p}_2(x) < -\lambda(x), |\Lambda\vec{p}_1(x)| \leq \lambda(x)\}$	$\Lambda\vec{p}_2(x) + \lambda(x) \leq 0$	≤ 0
4	$\{x \in \Omega\ :\ \Lambda\vec{p}_2(x) < -\lambda(x), \Lambda\vec{p}_1(x) < -\lambda(x)\}$	$\Lambda\vec{p}_2(x) - \Lambda\vec{p}_1(x)$	$= A(x)$
5	$\{x \in \Omega\ :\ \Lambda\vec{p}_2(x) < -\lambda(x), \Lambda\vec{p}_1(x) > \lambda(x)\}$	$\Lambda\vec{p}_2(x) - \Lambda\vec{p}_1 + 2\lambda \leq 0$	≤ 0
6	$\{x \in \Omega\ :\ \Lambda\vec{p}_2(x) > \lambda(x), |\Lambda\vec{p}_1(x)| \leq \lambda(x)\}$	$\Lambda\vec{p}_2(x) - \lambda \geq 0$	≥ 0
7	$\{x \in \Omega\ :\ \Lambda\vec{p}_2(x) > \lambda(x), \Lambda\vec{p}_1(x) < -\lambda(x)\}$	$\Lambda\vec{p}_2(x) - \Lambda\vec{p}_1 - 2\lambda \geq 0$	≥ 0
8	$\{x \in \Omega\ :\ \Lambda\vec{p}_2(x) > \lambda(x), \Lambda\vec{p}_1(x) > \lambda(x)\}$	$\Lambda\vec{p}_2(x) - \Lambda\vec{p}_1$	$= A(x)$

Thus we infer that $F(\vec{p}_1, \vec{p}_2, \delta\vec{p}) \geq 0$. Replacing $\vec{w}$ by $\delta\vec{p}$ in equation (4.38), we obtain

(4.43)
$$F(\vec{p}_1, \vec{p}_2, \delta\vec{p}) + \beta\|\Lambda\delta\vec{p}\|_{L^2(\Omega)}^2 + \gamma\|P_{\mathrm{div}}\delta\vec{p}\|_{\mathbb{L}^2(\Omega)}^2 = -\langle\delta\vec{v}, \delta\vec{p}\rangle_{H_0(\mathrm{div})^*, H_0(\mathrm{div})}.$$

From (4.38) and (4.43) we obtain that

(4.44)
$$\|\Lambda\delta\vec{p}\|_{L^2(\Omega)} = 0 \ \text{ and } \ \|P_{\mathrm{div}}\delta\vec{p}\|_{\mathbb{L}^2(\Omega)} = 0.$$

Note that the invertibility of (KK^*) and our assumption that $\mathrm{Ker}_{\mathrm{div}}(K) = \{0\}$ yield $\|\Lambda\delta\vec{p}\|_{L^2(\Omega)} \geq c\|\mathrm{div}\delta\vec{p}\|_{L^2(\Omega)}$ for some constant $c > 0$. This and (4.44) imply

(4.45)
$$\|\mathrm{div}\,\delta\vec{p}\|_{L^2(\Omega)} = 0.$$

Hence, $\delta\vec{p} \in H_0(\mathrm{div}0)$ and therefore $P_{\mathrm{div}}\delta\vec{p} = \delta\vec{p}$. This result together with the second part of (4.44) implies that $\|\delta\vec{p}\|_{\mathbb{L}^2(\Omega)} = 0$, thus $\vec{p}_1 = \vec{p}_2$. □

4.5. **Alternative setting of the problem.** Here, we mention an alternative way of approaching the solution of (1.6) which is convenient for numerical purposes.

Instead of considering the problem (4.21) the following setting is proposed:
(4.46)
$$\text{minimize } \int_\Omega |\nabla u|_{l^2} dx + \frac{\alpha}{2}\int_\Omega |\nabla u|_{l^2}^2 dx + \frac{1}{2\mu}\int_\Omega |Ku - z|^2 dx + \int_\Omega \lambda|Ku - z|\,dx \ \text{ over } u \in H_0^1(\Omega).$$

where $0 < \alpha \ll 1/\|\lambda\|_{L^\infty(\Omega)}$. The model (4.46) is a close approximation of the problem (4.21). Via dual regularization, similar as in Section 4.3, the following problem is obtained

(4.47)
$$\text{minimize } \int_\Omega |\nabla u|_{l^2}\,dx + \frac{\alpha}{2}\int_\Omega |\nabla u|_{l^2}^2 dx + \int_\Omega \Phi_{\mu,\beta}^*(-Ku)\,dx \ \text{ over } u \in H_0^1(\Omega).$$

where $\Phi^*_{\mu,\beta}(-Ku)$ is given by (4.27). In the same way, if the (pre)dual of (4.47) is regularized by $\frac{\gamma}{2}\|\vec{p}\|^2_{\mathbb{L}^2(\Omega)}$ rather than by $\frac{\gamma}{2}\|P_{\text{div}}\vec{p}\|^2_{\mathbb{L}^2(\Omega)}$ like in Section 4.4, and the resulting problem is dualized, we obtain

$$(4.48) \qquad \text{minimize } \frac{\alpha}{2}\int_\Omega |\nabla u|^2_{l^2}\,dx + \int_\Omega \Phi^*_{\mu,\beta}(-Ku)\,dx + \int_\Omega \Psi^*_\gamma(\nabla u)\,dx \quad \text{over } u \in H^1_0(\Omega)$$

where Ψ^*_γ is given by

$$(4.49) \qquad \Psi^*_\gamma(\vec{w}) = \begin{cases} \frac{1}{2\gamma}|\vec{w}(x)|^2_{l^2}, & \text{if } |\vec{w}(x)|_{l^2} < \gamma, \\ |\vec{w}(x)|_{l^2} - \frac{\gamma}{2}, & \text{if } |\vec{w}(x)|_{l^2} \geq \gamma; \end{cases}$$

see [17] for details. According to the Fenchel theorem, the optimality conditions for the coupled solutions $(\bar{u}, \vec{\bar{p}})$ are

$$-\text{div}\vec{\bar{p}} = -\alpha\Delta\bar{u} + \frac{1}{\mu+\beta}K^*(K\bar{u}-z) + \frac{\mu}{\mu+\beta}K^*\frac{\lambda(K\bar{u}-z)}{\max\{\lambda\beta, |K\bar{u}-z|\}},$$

$$-\vec{\bar{p}} = \frac{\nabla\bar{u}}{\max\{\gamma, |\nabla\bar{u}|_{l^2}\}}.$$

Let $\vec{\bar{q}} = -\vec{\bar{p}}$ and $\bar{v} = \frac{\lambda(K\bar{u}-z)}{\max\{\lambda\beta, |K\bar{u}-z|\}}$, then the following system of equations is obtained

$$(4.50) \qquad -\max\{\beta\lambda, |K\bar{u}-z|\}\bar{v} + \lambda(K\bar{u}-z) = 0,$$

$$(4.51) \qquad \text{div}\vec{\bar{q}} + \alpha\Delta u - \frac{1}{\mu+\beta}K^*(Ku-z) - \frac{\mu}{\mu+\beta}K^*\bar{v} = 0,$$

$$(4.52) \qquad \max\{\gamma, |\nabla\bar{u}|_{l^2}\}\vec{\bar{q}} - \nabla\bar{u} = 0.$$

In what follows, we use (4.50)-(4.52) rather than the system associated with (4.30) and its dual.

Notice that by adding the term $\alpha\int_\Omega |\nabla u|^2_{l^2}\,dx$ with $\alpha > 0$ we lift the solution of (4.21) into $H^1_0(\Omega)$. This has the mathematical advantage of yielding a problem in a reflexive function space and makes the dualization process convenient, and it has the numerical advantage of allowing us to reduce the first order system to a system in u (as compared to a system in the vector field $\vec{p}$). It can be shown that (4.46) converges to (4.21) as $\alpha \to 0$ and to (1.6) for $\alpha \to 0$ and $\mu \to \infty$.

4.6. **Primal-dual algorithm.** Note that the system (4.50)-(4.52) is non-smooth, i.e. not necessarily Fréchet-differentiable. However, relying on generalized differentiation, it turns out that the discrete version of this system can be solved efficiently by a semismooth Newton method.

For this purpose let $u^h \in \mathbb{R}^M$, $p^h \in \mathbb{R}^{2M}$, $\lambda^h \in \mathbb{R}^M$ denote the discrete image intensity, dual variable and spatially dependent λ, respectively, for some $M \in \mathbb{N}$ which depends on the image size $m \times m$. Further, let $z^h \in \mathbb{R}^M$ denote the discrete data vector, $\nabla^h \in \mathbb{R}^{2M\times M}$ the discrete gradient operator, $\Delta^h \in \mathbb{R}^{M\times M}$ the discrete Laplace operator, and $K^h \in \mathbb{R}^{M\times M}$ the discrete blurring operator. Here the mappings $|\cdot|$, $\max\{\cdot,\cdot\}$ and $\text{sign}(\cdot)$ are understood for vectors in a componentwise sense. We use the mapping $[|\cdot|] : \mathbb{R}^{2M} \to \mathbb{R}^{2M}$ with $[|v^h|]_i = |(v^h_i, v^h_{i+M})^T|_{l^2}$ for $i \in \{1,\ldots,M\}$, and $e^h \in \mathbb{R}^{2M}$ is the vector of all ones. The discrete version of (4.50)-(4.52) is given by

$$(4.53) \qquad -\max\{\beta\lambda^h, |K^h u^h - z^h|\}v^h + \lambda^h(K^h u^h - z^h) = 0,$$

$$(4.54) \qquad -(\nabla^h)^T q^h + \alpha\Delta^h u^h - \frac{1}{\mu+\beta}(K^h)^T(K^h u^h - z^h) - \frac{\mu}{\mu+\beta}(K^h)^T v^h = 0,$$

$$(4.55) \qquad \max\{\gamma e^h, [|\nabla^h u^h|]\}q^h - \nabla^h u^h = 0.$$

For the generalized linearization of (4.53)-(4.55), which is required for the semismooth Newton solver, we use the following element of the generalized Jacobian of $\max : \mathbb{R}^M \to \mathbb{R}^M$, the diagonal matrix $G_{\max} \in \mathbb{R}^{M\times M}$ with

$$(4.56) \qquad (G_{\max}(w))_{ii} := \begin{cases} 1 & \text{if } w_i \geq 0, \\ 0 & \text{if } w_i < 0 \end{cases} \quad \text{for } 1 \leq i \leq 1;$$

compare [8].

Hence, in every step of our Newton method for solving (4.53)-(4.55) the following system needs to be solved:

$$(4.57) \quad \begin{pmatrix} A_k^h & -D(m_{\beta_k}) & 0 \\ -\frac{1}{\mu+\beta}K^TK + \alpha\Delta^h & -\frac{\mu}{\mu+\beta}(K^h)^T & -(\nabla^h)^T \\ B_k^h & 0 & D(m_{\gamma_k}) \end{pmatrix} \begin{pmatrix} \delta_u \\ \delta_v \\ \delta_q \end{pmatrix} = \begin{pmatrix} -F_1^k \\ -F_2^k \\ -F_3^k \end{pmatrix}$$

where

$$A_k^h = [D(\lambda_k^h) - D(v_k^h)\chi_{\mathcal{A}_{\beta_k}}D(\text{sign}(K^hu_k^h - z^h))]K^h := \Gamma_k K^h,$$
$$B_k^h = [-I_{2M} + D(q_k^h)\chi_{\mathcal{A}_{\gamma_k}}D(m_{\gamma_k})^{-1}N^h(\nabla^hu_k^h)]\nabla^h := -C_k\nabla^h,$$
$$F_1^k := D(\lambda_k^h)(K^hu_k^h - z^h) - D(m_{\beta_k})v_k^h,$$
$$F_2^k := -(\nabla^h)^Tq_k^h + \alpha\Delta^hu_k^h - \frac{1}{\mu+\beta}(K^h)^T(K^hu_k^h - z^h) - \frac{\mu}{\mu+\beta}(K^h)^Tv_k^h,$$
$$F_3^k := -\nabla^hu_k^h + D(m_{\gamma_k})q_k^h,$$

$I_{2M} \in \mathbb{R}^{2M\times2M}$ is the identity matrix, $D(v)$ is a diagonal matrix with the vector v in the main diagonal, $m_{\beta_k} = \max\{\beta\lambda_k^h, |K^hu_k^h - z^h|\}$, $m_{\gamma_k} = \max\{\gamma e^h, [|\nabla^hu_k^h|]\}$,

$$\chi_{\mathcal{A}_{\beta_k}} = D(t_{\beta_k}) \quad \text{with} \quad (t_{\beta_k})_i = \begin{cases} 0, & \text{if } (m_{\beta_k})_i = \beta(\lambda_k^h)_i, \\ 1, & \text{else;} \end{cases}$$

$$\chi_{\mathcal{A}_{\gamma_k}} = D(t_{\gamma_k}) \quad \text{with} \quad (t_{\gamma_k})_i = \begin{cases} 0, & \text{if } (m_{\gamma_k})_i = \gamma, \\ 1, & \text{else;} \end{cases}$$

$$N^h(v) = \begin{pmatrix} D(v_x) & D(v_y) \\ D(v_x) & D(v_y) \end{pmatrix} \quad \text{with} \quad v = (v_x, v_y)^T \in \mathbb{R}^{2M}.$$

The diagonal matrices $D(m_{\beta_k})$ and $D(m_{\gamma_k})$ are invertible. Therefore, δ_v and δ_q are obtained from the first and third equation in (4.57), respectively, and substituted into the second equation. The resulting equation for δ_u is written as

$$(4.58) \quad H_k\delta_u = f_k,$$

where the matrix H_k and the right-hand side f_k are defined by

$$H_k := \frac{1}{\mu+\beta}(K^h)^TK^h - \alpha\Delta^h + \frac{\mu}{\mu+\beta}(K^h)^TD(m_{\beta_k})^{-1}\Gamma_kK^h + (\nabla^h)^TD(m_{\gamma_k})^{-1}C_k\nabla^h,$$
$$f_k := F_2^k - \frac{\mu}{\mu+\beta}(K^h)^TD(m_{\beta_k})^{-1}F_1^k + (\nabla^h)^TD(m_{\gamma_k})^{-1}F_3^k.$$

First note that the matrix H_k is in general not symmetric, because C_k is not. In [17] it was shown that the matrix C_k at the solution $(u_k^h, v_k^h, q_k^h) = (\bar{u}, \bar{v}, \bar{q})$ is positive definite whenever

$$(C1) \quad [|q_k^h|]_i \le 1, \quad \text{for } i = 1, \ldots, 2M.$$

Another important condition is

$$(C2) \quad (|v_k^h|)_j \le (\lambda_k^h)_j, \quad \text{for } j = 1, \ldots, M,$$

which yields feasibility of the dual variable. In fact, when both (C1) and (C2) are satisfied, the following result holds true.

Lemma 7. *Suppose the conditions* (C1) *and* (C2) *hold true and* $\alpha > 0$*, then for all* $k \in \mathbb{N}$*, the matrix* H_k *is positive definite.*

For the proof of the Lemma 7 we refer to [17].

In case the conditions (C1) and (C2) are not satisfied, q_k^h and v_k^h are projected onto their respective feasible set. In fact, if (C1) is not satisfied, $((q_k^h)_i, (q_k^h)_{i+m})$ is replaced by $\max\{1, [|q_k^h|]_i\}^{-1}$ $((q_k^h)_i, (q_k^h)_{i+m})$. Analogously, if (C2) is violated at some index j, $(v_k^h)_j$ is replaced by $(\lambda_k^h)_j$ $\max\{(\lambda_k^h)_j, (|v_k^h|)_j\}^{-1}(v_k^h)_j$. Thus, the modified system matrix, denoted by H_k^+, is positive definite; see [11]. Summarizing the above discussion, our semismooth Newton solver is as follows.

Semismooth Newton for step 2 of the SA-TV-algorithm.

1: Initialize $(u_0^h, p_0^h) \in \mathbb{R}^M \times \mathbb{R}^{2M}$ and set $k := 0$.

2: Estimate the active sets, i.e., determine $\chi_{\mathcal{A}_{\beta_{k+1}}} \in \mathbb{R}^{M \times M}$, $\chi_{\mathcal{A}_{\gamma_{k+1}}} \in \mathbb{R}^{2M \times 2M}$.

3: If $(C1)$ and $(C2)$ are not satisfied, then compute H_k^+; otherwise $H_k^+ := H_k$.

4: Solve $H_k^+ \delta_u = f_k^h$ for δ_u, and let δ_{u_k} denote the solution.

5: Use δ_{u_k} to compute δ_{p_k}.

6: Update $u_{k+1}^h := u_k^h + \delta_{u_k}$, $p_{k+1}^h := p_k^h + \delta_{p_k}$.

7: Stop; or set $k := k + 1$ and return to step 2.

We point out that instead of $\alpha > 0$, one may use $H_k^+ + \varepsilon_k I_M$ with $\alpha = 0$, $\varepsilon_k > 0$ and $\varepsilon_k \downarrow 0$ as $k \to \infty$ to obtain a positive definite system matrix. For $\epsilon_k = \epsilon > 0$ for all k, this regularization corresponds to replacing $\alpha \int_\Omega |\nabla u|_{l^2}^2 dx$ by $\epsilon \int_\Omega |u|_{l^2}^2 dx$ in (4.46).

The locally superlinear convergence of the above algorithm follows from standard theory; see [16, 17].

5. SPATIAL ADAPTATION BY LOCAL EXPECTED ABSOLUTE VALUE ESTIMATORS

In the case were the local expected absolute value ν of the noise is at our disposal, as it was the case in Section 3, the parameter λ may be chosen depending on the local constraint

$$(5.1) \qquad\qquad S(u) \le \nu \quad \text{a.e. in } \Omega.$$

We are interested in a reconstructed image such that its expected absolute value of the residual is close to the expected absolute value of the noise in both the detail regions and the homogeneous parts. Hence, we introduce a local expected absolute value estimator ($LEAVE$) for an automated adaptive choice of λ. The adjustment rule makes use of the constraint (5.1).

5.1. **Local expected absolute value estimator.** In the following, only discrete terms are used. The discrete image domain Ω^h contains $M = m \times m$ pixels. The discrete image residual is denoted by $r^h = K^h u^h - z^h$ with $r^h, z^h, u^h \in \mathbb{R}^M$ and $K^h \in \mathbb{R}^{M \times M}$. For convenience, for the remainder of this section, r^h, z^h and $K^h u^h$ are reshaped as $m \times m$ matrices. In the discrete version of the mean filter defined in (3.3), we use $\epsilon = 0$. Let Ω^ω denote the set of pixel-coordinates in a ω-by-ω window centered at (i, j) (with a symmetric extension at the boundary), i.e.

$$\Omega_{i,j}^\omega = \left\{ (s + i, t + j) \mid -\left\lfloor \frac{\omega}{2} \right\rfloor \le s, t \le \left\lfloor \frac{\omega}{2} \right\rfloor \right\},$$

where $\lfloor \cdot \rfloor$ denotes rounding to the nearest integer towards zero. The mean filter is applied to the absolute value of the residual image to obtain

$$(5.2) \qquad LEAVE_{i,j}^\omega = \frac{1}{\omega^2} \sum_{(s,t) \in \Omega_{i,j}^\omega} \left| (K^h u^h)_{s,t} - z_{s,t}^h \right| = \frac{1}{\omega^2} \sum_{(s,t) \in \Omega_{i,j}^\omega} \left| r_{s,t}^h \right|.$$

Based on the current estimate λ^h and the pertinent reconstruction u^h, the $LEAVE$ is a statistical measure which allows us to decide on the amount of details contained in the window around (i, j). In Figure 1 the $LEAVE^{11}$ is shown for images corrupted by salt-and-pepper noise and random-valued impulse noise, respectively. The $LEAVE^{11}$ is large (indicated in light gray) in image regions which contain fine scale details. One also observes that for fixed contrast, $LEAVE^{11}$ is the larger the finer (smaller) the scale is.

5.2. **Selection of λ in case of salt-and-pepper noise.** For this type of noise, the expected absolute value is $\nu = \frac{r}{2}$, where r is the noise ratio. As we observe in Figure 1 (b), the use of a small regularization parameter in the L^1-TV model results in an over-smoothing of the image and larger values of the $LEAVE^{11}$ in details regions, as we see in Figure 1 (c). This shows that the constraint (5.1) may have been violated. Thus, if the $LEAVE_{i,j}^\omega$ at a pixel (i, j) is larger than ν, the current value of λ^h at this pixel needs to be increased; otherwise it must not to be

(a)   (b)   (c)

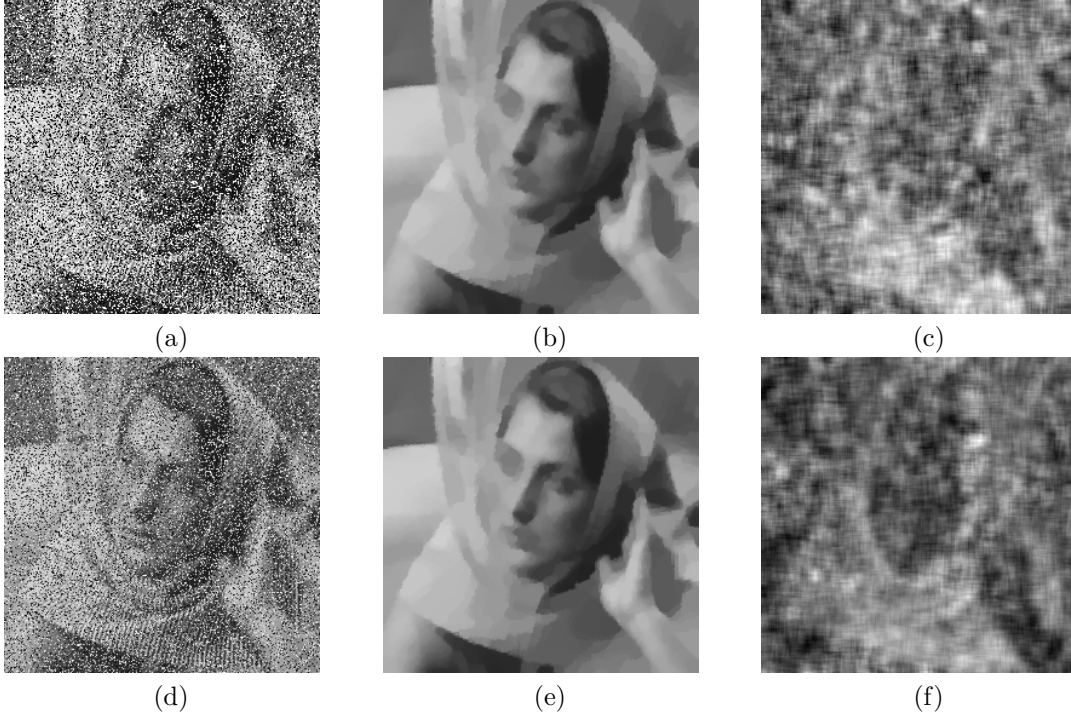(d)   (e)   (f)

FIGURE 1.   Local expected absolute estimator; **Salt and pepper noise**, (a) Noisy image, (b) Reconstructed image for too small regularization parameter λ, (c) $LEAVE^{11}$. **Random-valued impulse noise**, (d) Noisy image, (e) Reconstructed image for too small regularization parameter λ, (f) $LEAVE^{11}$.

increased in comparison to other pixels. Hence, for a given $\tilde{\lambda}_k^{\circ,h}$ (which yields λ_k^h), associated with a reconstructed image u_k^h, the following update of the regularization parameter λ^h is proposed:

$$(5.3) \qquad (\tilde{\lambda}_{k+1}^{\circ,h})_{i,j} = \eta \min \left((\tilde{\lambda}_k^{\circ,h})_{i,j} + \tau \left((LEAVE_k^\omega)_{i,j} - \nu \right)^+ , L \right),$$

$$(5.4) \qquad (\lambda_{k+1}^h)_{i,j} = \frac{1}{\omega^2} \sum_{(s,t)\in\Omega_{i,j}^\omega} (\tilde{\lambda}_{k+1}^{\circ,h})_{s,t}.$$

Note that (5.3)-(5.4) (for $\eta = 1$) are inspired by (3.14) and (3.15) together with the result (3.17). Here, $LEAVE_k^\omega$ is obtained from u_k, and L is a large positive value to ensure that $\tilde{\lambda}_k^{\circ,h}$ stays bounded. The parameter τ is set as $\tau = \tau_k = \|\tilde{\lambda}_k^{\circ,h}\|_{l^\infty}/\nu$ in order to keep the new $\tilde{\lambda}_{k+1}^{\circ,h}$ at the same scale as $\tilde{\lambda}_k^{\circ,h}$. The choice of $\eta > 1$ homogeneously increases $\tilde{\lambda}_k^{\circ,h}$ for improving the reconstruction and is motivated and explained in more detail in Section 6.

5.3. **Empirical selection of λ in case of random-valued impulse noise.** Similarly to the salt-and-pepper case, we observe in Figure 1 (e) that a small regularization parameter in the L^1-TV model results in an over-smoothing of the image and comparatively large values of $LEAVE^{11}$ in detail regions. Hence, in these regions an increase of λ is supposed to improve the restoration result. In our numerical test we found that an update rule for λ like (5.3) with a scalar $\nu \in \left[\frac{r}{4}, \frac{r}{2}\right]$ is improved by remembering that according to (2.13) ν actually depends on u. Rather than considering $S(u) \leq \nu(u)$ a.e. in Ω, which results in a quite nonlinear problem, we choose a reference value for u and compute an approximated expected absolute value which then is a function rather than a scalar. In our discrete setting, the approach works as follows. For a given $\tilde{\lambda}_k^{\circ,h}$ (which yields λ_k^h) associated with a reconstructed image u_k^h, a current expected absolute value

is computed, i.e.

$$(\nu_k^\omega)_{i,j} = \frac{1}{\omega^2} \sum_{(s,t)\in\Omega_{i,j}^\omega} r((K^h u_k^h)_{s,t}^2 - (K^h u_k^h)_{s,t} + \frac{1}{2}).$$

Then the update of the regularization parameter λ^h is as follows:

(5.5) $$(\tilde{\lambda}_{k+1}^{\circ,h})_{i,j} = \eta \min \left((\tilde{\lambda}_k^{\circ,h})_{i,j} + \tau \left((LEAVE_k^\omega)_{i,j} - (\nu_k^\omega)_{i,j}\right)^+, L \right),$$

(5.6) $$(\lambda_{k+1}^h)_{i,j} = \frac{1}{\omega^2} \sum_{(s,t)\in\Omega_{i,j}^\omega} (\tilde{\lambda}_{k+1}^{\circ,h})_{s,t}.$$

The parameters η, L have the same motivation as before. We note that we experimented with using different window sizes for λ_k^h, $LEAVE_k$ on one hand and ν_k on the other hand. We found that using equal window sizes produces the best robust results. In order to keep the new $\tilde{\lambda}_{k+1}^{\circ,h}$ at the same scale as $\tilde{\lambda}_k^{\circ,h}$, here τ is set as $\tau = \tau_k = 2\|\tilde{\lambda}_k^{\circ,h}\|_{l\infty}/r$. This setting is inspired by the fact that $\nu \in \left[\frac{r}{4}, \frac{r}{2}\right]$.

6. SPATIALLY ADAPTED TV-ALGORITHM

Based on the local expected absolute value estimator of Section 5, the reconstruction of the image by the primal-dual method presented in Section 4 is improved by modifying iteratively the regularization parameter. Hence, the following algorithm is proposed:

SA-TV-Algorithm (A).

1: Initialize $u_0^h \in \mathbb{R}^M$, $p_0^h \in \mathbb{R}^{2M}$, $\lambda_0^h \in \mathbb{R}_{++}^M$ and set $k = 0$.

2: Solve the discrete version of the problem

$$u_k \in \arg \min_{u\in BV(\Omega)} \int_\Omega |Du| + \int_\Omega \lambda_k |Ku - z| dx$$

by means of the primal-dual method proposed in Section 4. The discrete version of the solution is denoted by u_k^h.

3: Based on u_k^h, update

$$\tilde{\lambda}_{k+1}^{\circ,h} = \min \left(\tilde{\lambda}_k^{\circ,h} + \tau \left(LEAVE_k^\omega - \nu_k^\omega\right)^+, L \right),$$

$$(\lambda_{k+1}^h)_{i,j} = \frac{1}{\omega^2} \sum_{(s,t)\in\Omega_{i,j}^\omega} (\tilde{\lambda}_{k+1}^{\circ,h})_{s,t},$$

where

$$(\nu_k^\omega)_{i,j} = \begin{cases} \frac{r}{2} & \text{for salt-and-pepper noise,} \\ \frac{1}{\omega^2} \sum_{(s,t)\in\Omega_{i,j}^\omega} r((K^h u_k^h)_{s,t}^2 - (K^h u_k^h)_{s,t} + \frac{1}{2}) & \text{for random-valued impulse noise.} \end{cases}$$

4: Stop, or set $k := k + 1$ and return to step 2.

As it will be shown in the numerical results, see Figure 3, this algorithm exhibits a slow convergence. For this reason, we propose to accompany the update by an image decomposition method inspired by [13] in order to accelerate the restoration of the image. In [13], an image is represented as the sum of "atoms" w_k, where every w_k reveals finer scales than the previous w_{k-1}. In [13], each w_k comes from an L^2-TV image restoration. The resulting algorithm for dyadic scales (i.e. $\lambda_j = 2^j \lambda_0$) is as follows:

i) Choose $\lambda_0 > 0$ and compute

(6.1) $$u_0 = \arg \min_{u\in BV(\Omega)} \int_\Omega |Du| + \frac{\lambda_0}{2} \int_\Omega (Ku - z)^2 dx$$

ii) For $j = 0, 1, 2, \ldots$ set $\lambda_j = 2^j \lambda_0$ and $v_j = z - Ku_j$. Then compute

(6.2) $$w_j = \arg \min_{u\in BV(\Omega)} \int_\Omega |Du| + \frac{\lambda_{j+1}}{2} \int_\Omega (Ku - v_j)^2 dx, \quad u_{j+1} = u_j + w_j.$$

The dual of the semi-norm $\int_\Omega |Du|$ is

(6.3)
$$\|u\|_\star = \sup_{\phi \in BV(\Omega), \int_\Omega |D\phi| \neq 0} \frac{\int_\Omega u\phi \, dx}{\int_\Omega |D\phi|}.$$

For $\hat{u}_0 := u_0$, one has $\sum_{j=0}^k Kw_j \to z$ as $k \to \infty$ in the following weak sense:

(6.4)
$$\|K^*(z - \sum_{j=0}^k Kw_j)\|_\star = \frac{1}{2^k \lambda_0}.$$

We transfer this idea to the L^1-TV case and incorporate it into our algorithm in order to accelerate the adjustment of the regularization parameter. This is done in steps 2 and 3 of the following algorithm. In step 4 we observe that the parameter $\eta > 1$ is used instead of the factor 2 in step (ii).

SA-TV-Algorithm (B).

    1: Initialize $u_0^h \in \mathbb{R}^M$, $p_0^h \in \mathbb{R}^{2M}$, $\lambda_0^h \in \mathbb{R}_{++}^M$, $v_0^h = z^h$ and set $k = 0$.

    2: Solve the discrete version of the problem

$$w_k \in \arg \min_{v \in BV(\Omega)} \int_\Omega |Dv| + \int_\Omega \lambda_k |Kv - v_k| dx$$

        by means of the primal-dual method proposed in Section 4. The discrete version of the solution is denoted by w_k^h.

    3: Update $u_{k+1}^h = u_k^h + w_k^h$, $v_{k+1}^h = z^h - K^h u_{k+1}^h$.

    4: Based on u_{k+1}^h, update

$$\tilde{\lambda}_{k+1}^{\circ,h} = \eta \min \left(\tilde{\lambda}_k^{\circ,h} + \tau \left(LEAVE_k^\omega - \nu_k^\omega\right)^+, L\right),$$

$$(\lambda_{k+1}^h)_{i,j} = \frac{1}{\omega^2} \sum_{(s,t) \in \Omega_{i,j}^\omega} (\tilde{\lambda}_{k+1}^{\circ,h})_{s,t},$$

        where

$$(\nu_k^\omega)_{i,j} = \begin{cases} \frac{r}{2} & \text{for salt-and-pepper noise,} \\ \frac{1}{\omega^2} \sum_{(s,t) \in \Omega_{i,j}^\omega} r((K^h u_k^h)_{s,t}^2 - (K^h u_k^h)_{s,t} + \frac{1}{2}) & \text{for random-valued impulse noise.} \end{cases}$$

    5: Stop, or set $k := k + 1$ and return to step 2.

    The parameter λ^h is initialized by a relatively small constant, i.e. $\lambda_0^h = (\bar{\lambda}^0, \bar{\lambda}^0, \cdots, \bar{\lambda}^0)^\top$ with $\bar{\lambda}^0 > 0$ small.

    We note that a full theoretical justification of the transfer of the image decomposition method of [13] to the L^1-TV context is out of the scope of the present paper. Some first steps in this direction can be found, e.g., in [23, 26, 27] and further studies are the subject of our outgoing investigations.

7. Numerical results

In this section different experiments assessing the quality of algorithm (B) are presented. The images called "cameraman" and "man" shown in the first row of Figure 2 and their blurred version in the second row of Figure 2 are corrupted by random-valued impulse noise and salt-and-pepper noise. The corrupted images are then restored by means of the L^1-TV model. More precisely, the algorithm presented in [11] is used where the regularization parameter is a scalar. This scalar is close to be optimal, in the sense that after many trials, this scalar gives the best restoration. Then, the images are also restored by using the model presented in this paper and both results are compared.

The results for denoising are discussed in detail in section 7.2 and the ones for the simultaneous deblurring and denoising in section 7.3. The next section addresses a performance comparison of algorithms (A) and (B).

FIGURE 2.   (a) Cameraman, (b) Man, (c) Cameraman with Gaussian convolution
with a 9-by-9 window, (d) Man with Gaussian convolution with a 9-by-9 window.

7.1. **General remarks.** The performance of the method is compared quantitatively by means of
the mean absolute error (MAE) and the peak signal-to-noise ratio (PSNR) [3]. Note that since the
PSNR is an L^2-based measure, it has to be taken with care. Therefore we also propose to consider
the MAE which is an L^1-based measure. The smaller MAE becomes, the better the reconstruction
results are.

In order to compare the performance of algorithms (A) and (B), we use them to denoise "cam-
eraman" corrupted by salt-and-pepper noise with noise ratio $r = 0.5$ (see Figure 5 (a)) and
"cameraman" corrupted by random-valued impulse noise with noise ratio $r = 0.4$ (see Figure 6
(a)). These results are shown in Figure 3. The images (a) and (e) are obtained after 20 iterations
of the algorithm (A). The images in (c) and (g) are obtained after 5 iterations of algorithm (B).
Note that the cost per iteration for each of the algorithms is approximately the same. We observe
that the convergence of algorithm (A) is significantly slower than the one of algorithm (B), and
that smaller details are better preserved by using algorithm (B) rather than algorithm (A). Thus,
in all other examples, algorithm (B) is used to restore the images.

In Figure 4, one can observe that the restoration results are stable with respect to the window
size, i.e. there is no considerable difference between the results, although a larger window is more
suitable in order to reduce the influence of noise clusters, as one can find upon comparing (a) and
(b) or (c) and (d) and (e) or (f) in Figure 4. In the results and unless otherwise specified, the
window size used to compute the $LEAVE^\omega$ explained in Section 5 is set to $\omega = 21$.

The parameter η for updating the regularization parameter in step 4 of the algorithm (B) in
Section 6 is set to 1.1. As the L^1-TV model is rather sensitive to the regularization parameter
(this coincides with the observations in [6]), a moderate acceleration of the $\hat{\lambda}$-update due to η is
appropriate. This is the reason why we choose η rather close to 1. The parameters μ, β, γ are set
to 10^6, 10^{-4} and 10^{-5}, respectively.

The quality of images corrupted by blurring and noise is worse than the quality of those corrupted by noise only. The reconstructed images, however, are usually better in the case of deblurring and denoising rather than denoising only. This is due to the characteristics of the impulse noise (random-valued or salt-and-pepper). Since the restoration method is concerned with detecting the noisy pixels, this task becomes easier when the original image is first "smoothed" i.e., a contrast reduction due to blurring, and then corrupted by noise. As a result, noisy pixels are better distinguished from other ones. This helps the restoration method to produce better results.

The blurring used to corrupt the images in Figures 7 and 8 (a) and (e) comes from a Gaussian convolution with a 9×9 window and a standard deviation of 1. The resulting matrix $K^h \in \mathbb{R}^{M \times M}$ is ill-conditioned, but invertible. It therefore satisfies the discrete versions of our assumptions on K; see (4.7), Theorem 6, and Proposition 1.

Step 2 of algorithm (B) presented in Section 6 applies the primal-dual algorithm presented at the end of Section 4. At most it needs 21 iterations until successful termination, i.e. until the residual $res_i = (|F_1^i|_{l^2}^2 + |F_2^i|_{l^2}^2 + |F_3^i|_{l^2}^2)/(|F_1^0|_{l^2}^2 + |F_2^0|_{l^2}^2 + |F_3^0|_{l^2}^2)$ in (4.57) drops below a given tolerance. The linear system in step 4 of the primal-dual algorithm in the denoising case is solved by a direct solver, but in the deblurring and denoising case, it is solved by an iterative solver (biconjugate gradient stabilized - BICGSTAB) which requires at most 20 iterations at each call. In the spirit of an inexact Newton approach, we stop the BICGSTAB iteration as soon as $|f_k^h - H_k^+ \delta u_k^\ell|_{l^2}/|f_k^h|_{l^2} \le 0.5$. Here, δu_k^ℓ denotes the direction in iteration ℓ of the BICGSTAB algorithm in the kth step of our semismooth Newton method. With this setting our Newton solver always converged to the desired accuracy.

The iterations of algorithm (B) of Section 6 are stopped when the expected absolute value estimator for the whole domain Ω^h, defined as

$$(7.1) \qquad EAVE_k^\Omega := \tfrac{1}{|\Omega^h|} \sum_{(s,t)\in\Omega^h} \left| (K^h u^h)_{s,t} - z_{s,t}^h \right|$$

is less or equal to the corresponding expected absolute value ν_k^Ω, given by
(7.2)

$$\nu_k^\Omega := \begin{cases} \frac{r}{2} & \text{for salt-and-pepper noise,} \\ \frac{1}{|\Omega^h|} \sum_{(s,t)\in\Omega^h} r((K^h u_k^h)_{s,t}^2 - (K^h u_k^h)_{s,t} + \frac{1}{2}) & \text{for random-valued impulse noise,} \end{cases}$$

where $|\Omega^h| = M = m \times m$ pixels, i.e. size of the image.

In Table 1 we illustrate these stopping conditions. In fact, we consider the reconstruction of "cameraman" corrupted by salt-and-pepper noise with ratio $r = 0.5$ as shown in Figure 5 (c) and the reconstruction of "man" corrupted by random-valued impulse noise with $r = 0.4$ shown in Figure 6 (g). In this table, we show the number of necessary inner iterations (i) of the semismooth Newton solver and the residual (res_i) upon termination of the inner iteration. The stopping condition $EAVE_k^\Omega \le \nu_k^\Omega$ is also contained in Table 1 when comparing the columns for $EAVE_k^\Omega$ and ν_k^Ω for the respective image. Further, the number of the necessary outer iterations (k) of algorithm (B) is shown.

Cameraman: Salt-and-pepper, Fig. 5 (c)					**Man:** Random-valued impulse noise, Fig. 6 (g)				
k	i	res_i	$EAVE_k^\Omega$	ν_k^Ω	k	i	res_i	$EAVE_k^\Omega$	ν_k^Ω
1	14	9.63e-007	0.2647	0.25	1	16	3.13e-007	0.1580	0.1305
2	17	9.64e-007	0.2600	0.25	2	13	3.57e-007	0.1526	0.1315
3	15	2.98e-007	0.2575	0.25	3	17	2.76e-007	0.1461	0.1321
4	18	8.18e-007	0.2539	0.25	4	21	4.66e-007	0.1374	0.1323
5	15	6.10e-007	0.2504	0.25	5	21	2.45e-005	0.1288	0.1324
6	20	7.77e-007	0.2447	0.25					

TABLE 1.   Outer iterations (k), inner iterations (i), residual (res_i), expected absolute value estimator ($EAVE_k^\Omega$), expected absolute value (ν_k^Ω).

7.2. **Image denoising.** In Figure 5, our results in the case where the image is corrupted by salt-and-pepper noise are shown. The noise corrupting "cameraman" has a noise ratio of $r = 0.5$, whereas the one corrupting the "man" has $r = 0.4$. The results obtained by a scalar regularization parameter only can be found in (b) and (f) and the results obtained with a spatially adapted regularization parameter in (c) and (g) for "cameraman" and "man", respectively. Note that in (c) and (g) small features are better recovered as compared to (b) and (f). Also, homogeneous regions appear smoother in (c) and (g) rather than in (b) and (f). For instance, in "cameraman" we observe that the center of the tripod is "sharper" in (c) than in (b). Also noise spots in the sky appear in (b) and not in (c). Moreover, the edges are better defined in (c). In (b) the edges seem to be more irregular. In the case of "man", the feathers in the hat exhibit more details in (g) rather than in (f). The background is also more homogeneous in (g) than in (f). The improvement of the images is due to the different values of the regularization parameter λ over the image domain which can be observed in (d) and (h). In (d) and (h) the values of the function λ are presented in a gray scale. Light gray regions refer to large values of λ, whereas dark gray belongs to zones where λ is rather small. Note that in (d) λ is large in the tripod area and in (h) λ is large in the regions of the feathers. Both image details are of rather small scale. The measure MAE shows that the restoration of "cameraman" in (c) is better than in (b) and that the restoration of "man" is better in (g) than in (f). However the PSNR is a little bit smaller in (c) than in (b) and it is also smaller in (g) than in (f). The algorithm (B) in Section 6 requires 6 iterations for "cameraman" and 5 for "man".

In Figure 6, images are corrupted by random-valued impulse noise with a noise ratio of $r = 0.4$. The results obtained by a scalar regularization parameter only can be found in (b) and (f) and the results obtained with a spatially adapted regularization parameter in (c) and (g) for "cameraman" and "man", respectively. The middle of the tripod and the details in the buildings in "cameraman" are better defined in (c) than in (b). In (d) we observe that the value of λ is bigger in these regions. A similar effect is found in "man" in the region of the feathers in the hat. Thus, the details are better recovered in (g) than in (f) due to the spatial adaptation of λ yielding the result in (g). The algorithm (B) in Section 6 requires 5 iterations for "cameraman" and 4 for "man". The measure MAE is better in (c) and (g) than in (b) and (f). The PSNR is better in (c) than in (b) although it is smaller in (g) than in (f). The algorithm (B) of Section 6 requires 5 iterations in both cases.

7.3. **Image deblurring.** In Figure 7 we observe the reconstruction of blurred images corrupted by salt-and-pepper noise. We find that the reconstructions for spatially adapted regularization are slightly better; compare (c) vs. (b) and (g) vs. (f). In this case we note that the difference in λ between details and homogeneous regions appears smaller in the deblurring rather than in the denoising case. The algorithm (B) of Section 6 required 5 iterations for "cameraman" and 6 for "man".

Further, in Figure 8 we observe a slightly better recovery of the details in (c) than in (b) and in (g) than (f). In (d) we observe that the values of λ are larger in regions with small details and edges. This behavior is less pronounced for "man" as can be seen in (h). The restoration of "cameraman" needs 5 iterations and of "man" 6 iterations.

We observe that the images restored by the algorithm requiring a scalar regularization parameter yield slightly better results of the MAE and the PSNR than the restoration given by the proposed algorithm (B) although these measures are close. As we mentioned before, the original image is first "smoothed" due to blurring, and then corrupted by noise. As a consequence, small details may get lost due to this process. Subsequently, the update of the spatially adapted parameter has not the same impact as it has in the case of images corrupted by noise only.

8. Conclusions

The L^1-TV method is modified by replacing the scalar regularization parameter λ by a function. A suitable choice of such a function is related to an equivalent problem with pointwise constraints. Moreover, statistical characteristics of the noise help to locally adjust λ. For this purpose, a local expected absolute value estimator is introduced for the parameter update. The resulting

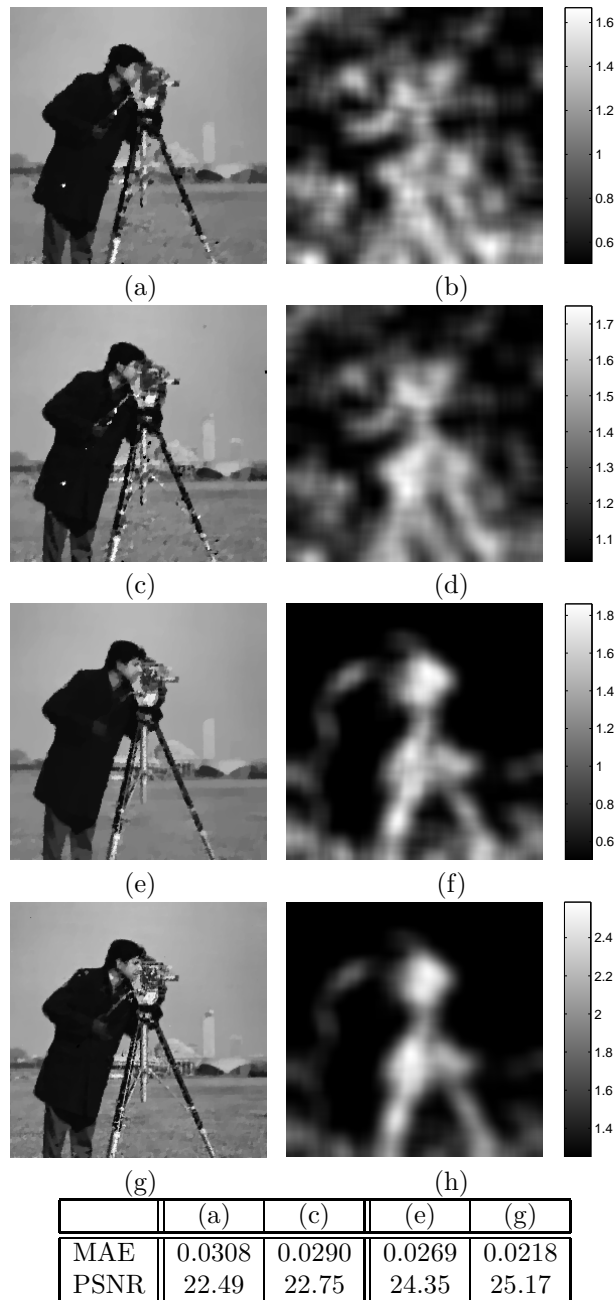|		(a)	(c)	(e)	(g)
MAE		0.0308	0.0290	0.0269	0.0218
PSNR		22.49	22.75	24.35	25.17

FIGURE 3.   **Salt-and-pepper noise:** (a) Restored image using algorithm (A), (b) Resulting λ, (c) Restored image using algorithm (B), (d) Resulting λ. **Random-valued impulse noise:** (e) Restored image using algorithm (A), (f) Resulting λ, (g) Restored image using algorithm (B), (h) Resulting λ.

spatially adapted regularization parameter λ is better suited in a L^1-TV model where small details are desired to be preserved. The two non-differentiable terms in the L^1-TV model (i.e., the TV-seminorm and the L^1 fidelity term) are appropriately treated by adequate regularization and Fenchel duality. The overall method combines an automated λ-adjustment scheme with an inexact semismooth Newton solver for the L^1-TV subproblems. The resulting method outperforms the restoration due to a scalar regularization parameter in the case of images contaminated by noise only. In the case of blurred and noisy images the reconstruction performed by a model using a
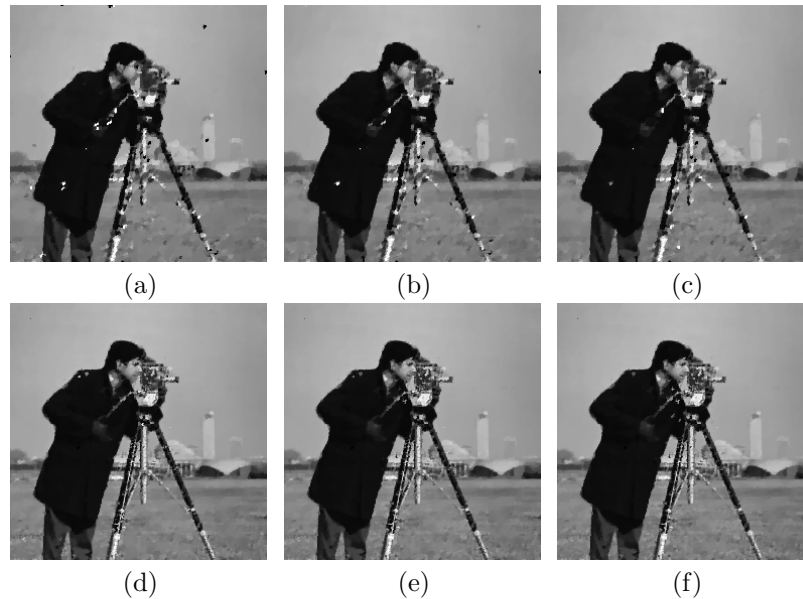
FIGURE 4. **Restoration of Cameraman with different windows size**. *Salt-and-pepper* (a) $\omega = 11$, (b) $\omega = 21$, (c) $\omega = 31$. *Random-valued impulse noise:* (d) $\omega = 11$, (e) $\omega = 21$, (f) $\omega = 31$.

scalar parameter is already very accurate for a large range of scalars, i.e. a precise tuning of the regularization is not necessary to find an appropriate reconstruction. This is due to the fact that details are lost in the blurring process. Thus, blurred details tend to remain in the residual image and do not contribute to a significant update of the spatially adapted regularization parameter. We also point out that the spatially adapted regularization parameter choice is fully automatic.

REFERENCES

[1] R. Acar and C.R. Vogel. Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Prolems*, 10:1217–1229, 1994.
[2] H. W. Alt. *Lineare Funktionalanalysis*. Springer, 2006.
[3] A. Bovik. *Handbook of Image and Video Processing*. Academic Press, 2000.
[4] A. Chambolle. An algorithm for total variation minimization and application. *Journal of Mathematical Imaging and Vision*, 20:89–97, 2004.
[5] A. Chambolle and P-L. Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76:167–188, 1997.
[6] T. Chan and S. Esedoglu. Aspects of total variation regularized l^1 function approximation. *SIAM Journal on Applied Mathematics*, 65(5):1817–1837, 2005.
[7] Q. Chang and I-L. Chern. Acceleration methods for total variation-based image denoising. *SIAM J. Applied Mathematics*, 25:982–994, 2003.
[8] F.H. Clarke. *Optimization and Nonsmooth Analysis*. Wiley Interscience, 1983.
[9] R Dautray and J.L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology*, volume 3. Springer-Verlag, 2000.
[10] D. C. Dobson and C. R. Vogel. Convergence of an iterative method for total variation denoising. *SIAM J. Numer. Anal.*, 34:1779–1791, 1997.
[11] Y. Dong, M. Hintermüller, and M. Neri. An efficient primal-dual method for l1-tv image restoration. *SIAM Journal on Image Sciences*, 2(4):1168–1189, 2009.
[12] Y. Dong, M. Hintermüller, and M.M. Rincon-Camacho. Automated regularization parameter selection in a multi-scale total variation model for image restoration. *IFB-Report No. 22 (11/2008), Institute of Mathematics and Scientific Computing, University of Graz*, 2008.
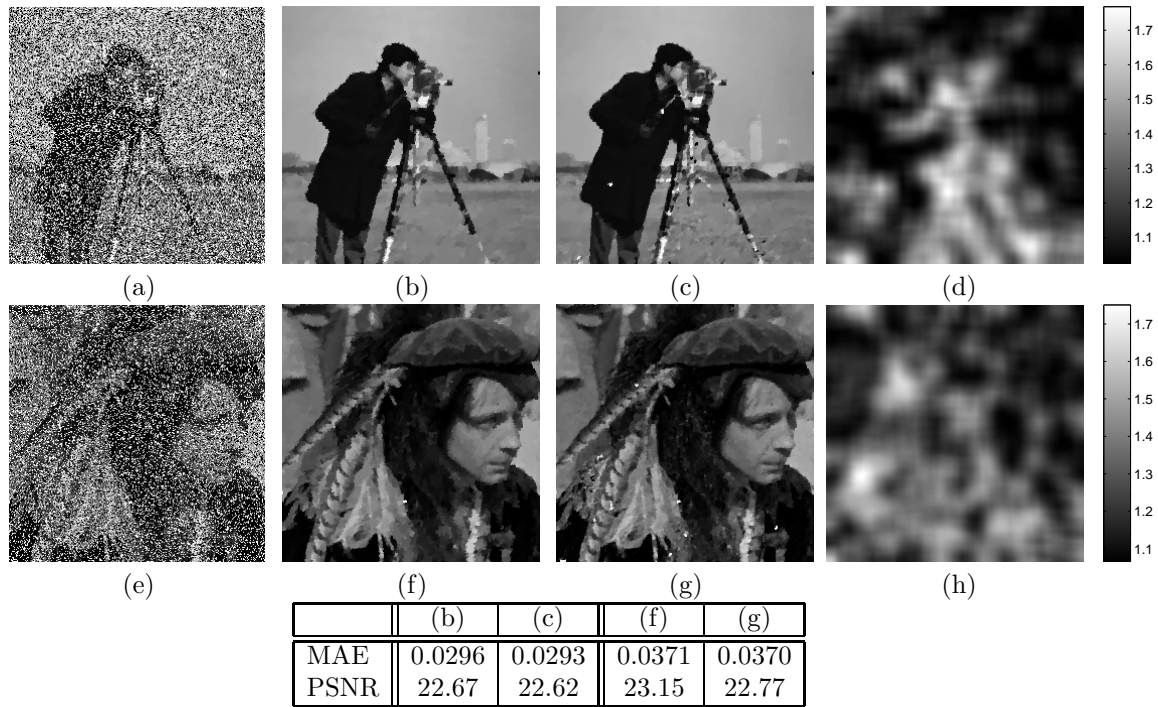
	(b)	(c)	(f)	(g)
MAE	0.0296	0.0293	0.0371	0.0370
PSNR	22.67	22.62	23.15	22.77

FIGURE 5. **Cameraman:** (a) Image with salt and pepper noise $r = 0.5$. (b) Denoising with L^1-TV model, $\lambda = 1.2$. (c) Denoising with the present method. (d) Resulting λ. **Man:** (e) Image with salt-and-pepper noise $r = 0.4$. (f) Denosing with L^1-TV model, $\lambda = 1.3$. (g) Denoising with the present method, (h) Resulting λ.

[13] E. Tadmor, S. Nezzar and L. Vese. A multiscale image representation using hierarchical (BV, L^2) decompositions. *Multiscale Model. Simul.*, 2:554–579, 2004.

[14] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems*. Classics Appl. Math. 28, SIAM, 1999.

[15] E. Giusti. *Minimal Surfaces and Functions of Bounded Variation*. Birkhäuser, 1984.

[16] M. Hintermüller and K. Kunisch. Total bounded variation regularization as bilaterally constrained optimization problem. *SIAM J. Appl. Math.*, 64:1311–1333, 2004.

[17] M. Hintermüller and G. Stadler. An infeasible primal-dual algorithm for total bounded variation-based inf-convolution-type image restoration. *SIAM Journal on Scientific Computing*, 28(1):1–23, 2006.

[18] L. I. Rudin, S. Osher and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[19] M. Nikolova. Minimizers of cost-functions involving nonsmooth data-fidelity term. aplication to the processing of outliers. *SIAM Journal on Numerical Analysis*, 40(3):965–994, 2002.

[20] M. Nikolova. A variational approach to remove outliers and impulse noise. *Journal of Mathematical Imaging and Vision*, 20:99–120, 2004.

[21] L. Rudin. MTV-multiscale total variation principle for a pde-based solution to nonsmooth ill-posed problem. Technical report, Cognitech, Inc. Talk presented at the Workshop on Mathematical Methods in Computer Vision, University of Minnesota, 1995.

[22] S. Osher, M. Burger, D. Goldfarb, J. Xu and W. Yin. An iterative regularization method for total variation-based image restoration. *SIAM Multiscale Model. and Simu.*, 4:460–489, 2005.

[23] O. Scherzer, W. Yin, and S. Osher. Slope and g-set characterization of set-valued functions and applications to non-differentiable optimization problems. *Commun. Math. Sci.*, 3(4):479–492, 2005.

[24] D. Strong and T. Chan. Spatially and scale adaptive total variation based regularization and anisotropic diffusion in image processing. Technical report, UCLA, 1996.

[25] D. Strong and T. Chan. Edge-preserving and scale-dependent properties of total variation regularization. *Inverse Problems*, 19:165–187, 2003.

[26] W. Yin, D. Goldfarb, and S. Osher. Image cartoon-texture decomposition and feature selection using the total variation regularized l^1 functional. *Springer-Verlag Berlin Heidelberg*, 3752:73–84, 2005.

[27] W. Yin, D. Goldfarb, and S. Osher. The total variation regularized l^1 model for multiscale decomposition. *Multiscale Model. Simul.*, 6(1):190–211, 2007.

	(b)	(c)	(f)	(g)
MAE	0.0226	0.0216	0.0476	0.0462
PSNR	24.84	25.21	22.36	22.18

FIGURE 6. **Cameraman:** (a) Image with random-valued impulse noise $r = 0.4$. (b) Denosing with L^1-TV model, $\lambda = 1.3$. (c) Denoising with the present method. (d) Resulting λ. **Man:** (e) Image with random-valued impulse noise $r = 0.4$. (f) Denosing with L^1-TV model, $\lambda = 1.3$. (g) Denoising with the present method. (h) Resulting λ.

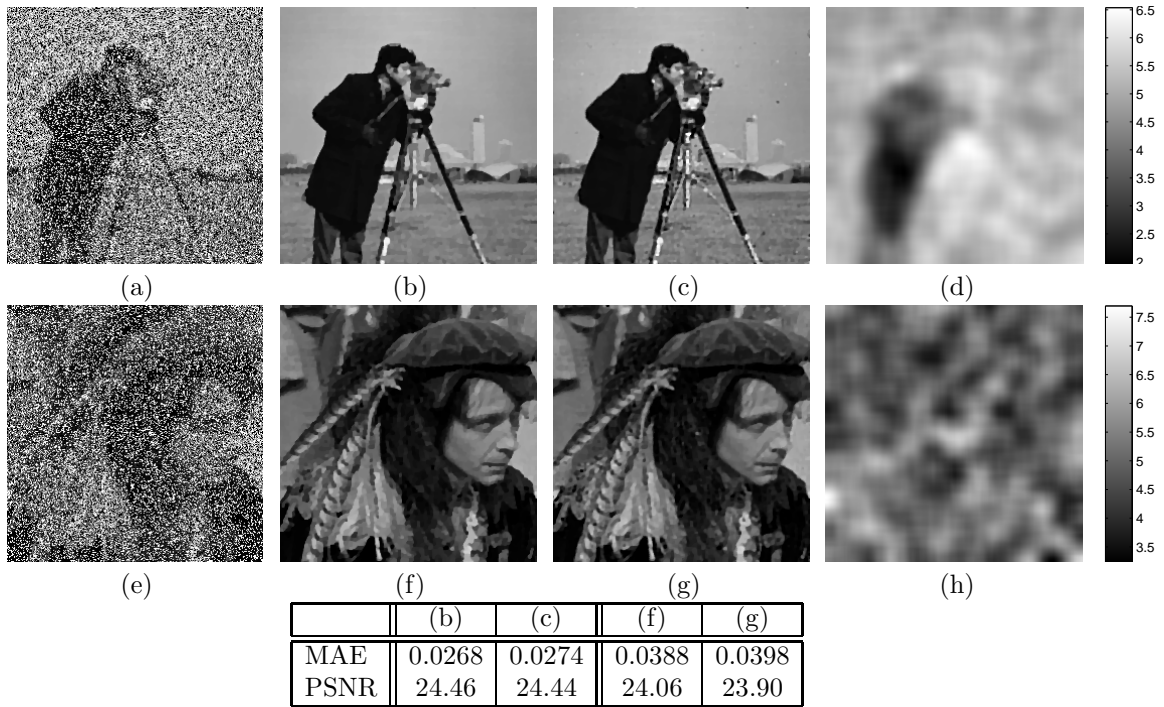|	(b)	(c)	(f)	(g)
MAE	0.0268	0.0274	0.0388	0.0398
PSNR	24.46	24.44	24.06	23.90

FIGURE 7. **Cameraman:** (a) Image with blur and salt-and-pepper noise $r = 0.5$. (b) Denosing and deblurring with L^1-TV model, $\lambda = 3$. (c) Denoising and deblurring with the present method. (d) Resulting λ. **Man:** (e) Image with blur and salt and pepper noise $r = 0.5$. (f) Denosing and deblurring with L^1-TV model, $\lambda = 3$. (g) Denoising and deblurring with the present method. (h) Resulting λ.

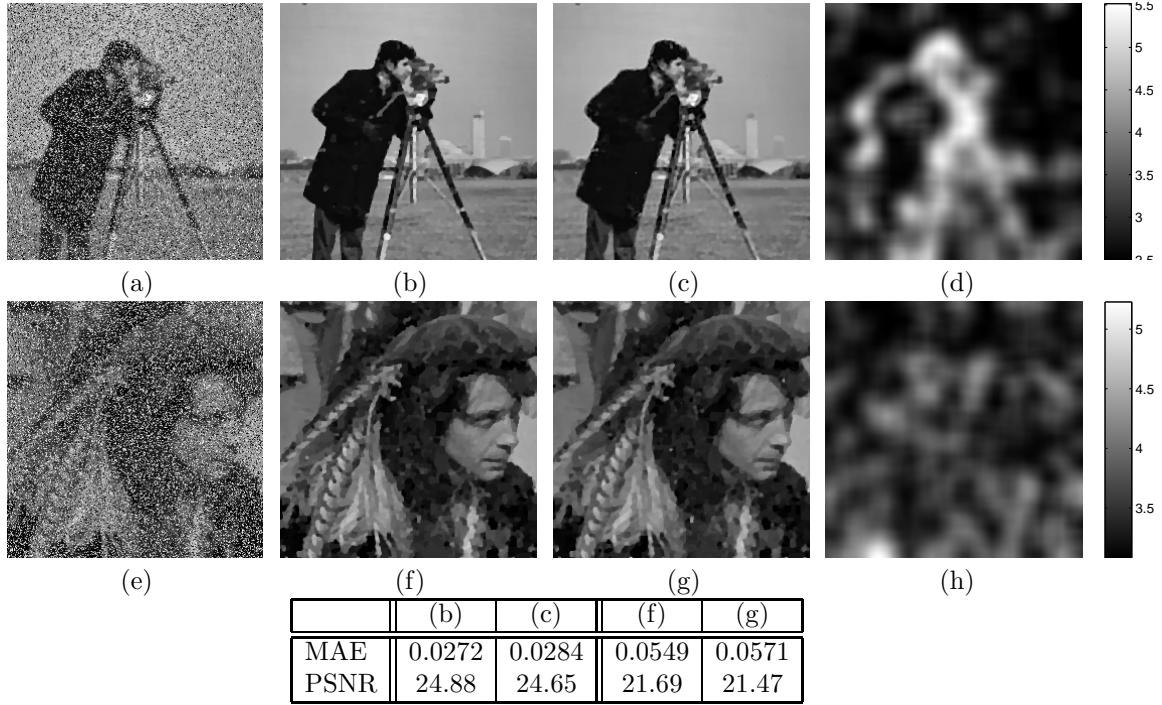|	(b)	(c)	(f)	(g)
MAE	0.0272	0.0284	0.0549	0.0571
PSNR	24.88	24.65	21.69	21.47

FIGURE 8. **Cameraman:** (a) Image with blur and random-valued impulse noise $r = 0.4$. (b) Denosing and Deblurring with L^1-TV model, $\lambda = 3$. (c) Denoising and deblurring with the present method. (d) Resulting λ. **Man:** (e) Image with blur and random-valued impulse noise $r = 0.5$. (f) Denosing and deblurring with L^1-TV model, $\lambda = 3$. (g) Denoising and deblurring with the present method. (h) Resulting λ.