A MULTI-SCALE VECTORIAL L^τ-TV FRAMEWORK FOR COLOR IMAGE RESTORATION

YIQIU DONG*, MICHAEL HINTERMÜLLER†, AND M. MONSERRAT RINCON-CAMACHO°

Abstract. A general multi-scale vectorial total variation model with spatially adapted regularization parameter for color image restoration is introduced in this paper. This total variation model contains an L^τ-data fidelity for any $\tau \in [1, 2]$. The use of a spatial dependent regularization parameter improves the reconstruction of features in the image as well as an adequate smoothing for the homogeneous parts. The automated adaptation of this regularization parameter is made according to local statistical characteristics of the noise which contaminates the image. The corresponding multiscale vectorial total variation model is solved by Fenchel-duality and inexact semismooth Newton techniques. Numerical results are presented for the cases $\tau = 1$ and $\tau = 2$ which reconstruct images contaminated with salt-and-pepper noise and Gaussian noise, respectively.

1. Introduction

In many applications, the deblurring and denoising of images are fundamental for subsequent image processing tasks, such as edge detection, segmentation, object recognition, etc. In this paper, we consider the problem of recovering color images degraded by cross-channel blurring and Gaussian or salt-and-pepper noise. For this purpose and without loss of generality, we assume a color image $\hat{\mathbf{u}}$ is a vector-valued function defined on a bounded and piecewise smooth open subset $\Omega \in \mathbb{R}^2$, i.e. $\hat{\mathbf{u}} : \Omega \to \mathbb{R}^M$, where M is the number of channels in the color model. Depending on the different noise types addressed above, the degraded form $\mathbf{z}$ of $\hat{\mathbf{u}}$ is obtained in case of

- *Gaussian noise* as

$$\mathbf{z} = K\hat{\mathbf{u}} + \mathbf{n},$$

where K is a linear and continuous cross-channel blurring operator from $L^2(\Omega; \mathbb{R}^M)$ to $L^2(\Omega; \mathbb{R}^M)$, i.e., $K \in \mathcal{L}(L^2(\Omega; \mathbb{R}^M))$, and $\mathbf{n}$ represents white Gaussian noise with zero mean and variance σ^2;

and in case of

- *salt-and-pepper noise* as

$$\mathbf{z} = \begin{cases} K\hat{\mathbf{u}}, & \text{with probability } 1 - r, \\ \mathbf{n}, & \text{with probability } r. \end{cases}$$

Key words and phrases. Multi-scale vectorial total variation, color image, L^2-data fidelity, L^1-data fidelity, Gaussian noise, salt-and-pepper noise.

This research was funded by the Austrian Ministry of Science and Research through START program Y305 "Interfaces and Free Boundaries" and the Austrian Science Fund FWF through SFB "Mathematical Optimization and Applications in Biomedical Sciences".

where the elements of $\mathbf{n}$ only attain the maximum or minimum of the intensity region and r denotes the noise ratio. For instance, if the (intensity) values of $\hat{\mathbf{u}}$ are confined to $[0,1]^M$, then $\mathbf{n}(x) \in \{0,1\}^M$ for $x \in \Omega$. Moreover, K is as above.

The problem of restoring $\hat{\mathbf{u}}$ from $\mathbf{z}$ with unknown $\mathbf{n}$ is known to be typically ill-posed [31] and it requires appropriate regularization for a successful solution. In order to preserve significant edges when restoring images, Rudin, Osher and Fatemi proposed total variation (TV) regularization [25] for removing Gaussian noise in gray-level images. In this approach (which we refer to as the L^2TV-model in what follows), the image u is recovered from given data z by solving the optimization problem

$$\min_{u \in BV(\Omega)} \int_\Omega |Du| + \frac{\lambda}{2} \int_\Omega |Ku - z|^2 dx, \tag{1.1}$$

where $BV(\Omega)$ denotes the space of functions of bounded variation (see, e.g., [31] for its definition) and $\lambda > 0$ induces the regularization parameter $1/\lambda$. Due to its ability to preserve edges, the L^2TV-model is widely accepted as a reliable tool in image restoration. Over the years, various research efforts have been devoted to studying, solving and extending the L^2TV-model; see, e.g., [12, 14, 20, 22, 25, 30] as well as the monograph [31] and the many references therein.

Whenever $\hat{u}$ is corrupted by salt-and-pepper noise, a non-smooth L^1-data-fidelity term proves to be more appropriate than the differentiable least-squares data fit in (1.1); see, e.g., [21, 22]. In this case, the associated minimization problem for recovering u from noisy data z becomes

$$\min_{u \in BV(\Omega)} \int_\Omega |Du| + \lambda \int_\Omega |Ku - z| dx, \tag{1.2}$$

which we call the L^1TV-model subsequently.

It is well-known that the choice of λ is crucial for the preservation of image details during the restoration process [7,26,27,32]. In addition, as images are typically comprised of multiple objects at different scales, variable values of λ localized at image features of different scales appear to be desirable to obtain better restoration results. For this reason, a multi-scale total variation model with a spatially varying choice of parameters was introduced in [1], and, in order to enhance image regions containing details while still sufficiently smoothing homogeneous features, two spatially dependent regularization parameter selections based on the L^2TV- and the L^1TV-model were proposed in [10] and [16], respectively. In [8], the scalar L^2TV-approach of [10] was extended to restore color images. Further literature addressing this issue can be found in these references.

In this paper, based on the analysis of the statistical characteristics of the underlying noise we unify the models for Gaussian noise and salt-and-pepper noise removal to obtain a general total variation model. Utilizing the vectorial total variation introduced in [4, 5], the resulting general model is employed for restoring degraded color images. Moreover, we introduce a multi-scale vectorial total variation with a spatially dependent regularization parameter. The proposed automated adjustment strategy of the regularization parameter is based on a local statistical estimator (LSE for short), and the resulting optimization problems are solved by a superlinearly convergent algorithm based on Fenchel-duality and inexact semismooth Newton techniques. We emphasize that our solver is capable of handling

the L^1TV-model and the L^2TV-model within one framework. Compared to [8] our L^τTV-framework requires appropriate generalizations of the statistical measures for computing the regularization parameter as well as the semismooth Newton method with respect to both theory and numerical implementation. The numerical results in this paper show that our method outperforms the vectorial total variation [4,5] with a scalar regularization parameter in both noise removal and detail preservation.

The outline of the rest of the paper is as follows. In Section 2 we introduce the unified multi-scale vectorial total variation model for Gaussian and salt-and-pepper noise and the primal-dual algorithm for solving the associated minimization problem. Section 3 addresses the LSE-based parameter selection for color images. A method for color image restoration combining the multi-scale representation and spatially adaptive parameter selection is proposed in Section 4. Section 5 gives numerical results to demonstrate the performance of the new method, and, finally, conclusions are drawn in Section 6.

Notation. Function spaces with bold face characters refer to spaces whose elements are vector fields. For instance, $\mathbf{L}^2(\Omega) = L^2(\Omega;\mathbb{R}^M)$ or $\mathbf{BV}(\Omega) = BV(\Omega;\mathbb{R}^M)$. By $\langle\cdot,\cdot\rangle$ we denote the L^2-inner product where we do not distinguish between scalar- or vector-valued functions. Moreover, for $w \in \mathbb{R}^M$ we use $|w|_\tau = (\sum_{i=1}^M |w_i|^\tau)^{1/\tau}$ for $1 \le \tau \le 2$, and $|w|_\infty = \max_{i=1,\dots,M} |w_i|$. Moreover, $|w|_c = (|w_1|,\dots,|w_M|)^\top$ and $|w|_c^\tau = (|w_1|^\tau,\dots,|w_M|^\tau)^\top$. For matrices $|\cdot|_F$ denotes the Frobenius-norm.

2. Multi-Scale Vectorial Total Variation

In our restoration approach we use the same regularization for both noise models, but we utilize different data fidelity terms. Based on statistical characteristics of the underlying noise, however, Gaussian or salt-and-pepper noise removal will be cast into one framework relying on constrained minimization.

2.1. **A general total variation model.**

Before we introduce our general total variation model, we recall some basic statistical characteristics of Gaussian and salt-and-pepper noise. For this purpose consider the noise η as a random variable (r.v.) with f denoting its probability density function. For the description of the behavior of the random variable η several measures are defined [18, 24]. Some of the most important ones are

- the *mean* given by

$$E(\eta) = \begin{cases} \int_{-\infty}^{\infty} x_\eta f(x_\eta)\,dx_\eta & \text{if } \eta \text{ is a continuous r.v. (c.r.v.),} \\ \sum_{x_\eta \in \mathcal{U}} x_\eta f(x_\eta) & \text{if } \eta \text{ is a discrete r.v. (d.r.v.) in a universe } \mathcal{U} \subset \mathbb{R} \end{cases}$$

- the *variance* defined as

$$\mathrm{Var}(\eta) = E[(\eta - E(\eta))^2] = E(\eta^2) - E(\eta)^2 = \begin{cases} \int_{-\infty}^{\infty} x_\eta^2 f(x_\eta)\,dx_\eta - E(\eta)^2 & \text{if } \eta \text{ is a c.r.v.,} \\ \sum_{x_\eta \in \mathcal{U}} x_\eta^2 f(x_\eta) - E(\eta)^2 & \text{if } \eta \text{ is a d.r.v.} \end{cases}$$

- and the *expected absolute value*

$$\mathrm{EAV}(\eta) = E(|\eta|) = \begin{cases} \int_{-\infty}^{\infty} |x_\eta| f(x_\eta)\,dx_\eta & \text{if } \eta \text{ is a c.r.v.,} \\ \sum_{x_\eta \in \mathcal{U}} |x_\eta| f(x_\eta) & \text{if } \eta \text{ is a d.r.v.} \end{cases}$$

In the case of Gaussian noise considered in this paper, η is normally distributed with mean $E(\eta) = 0$ and variance $\mathrm{Var}(\eta) = \sigma^2$. The associated probability density function is given by

$$f(x_\eta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x_\eta^2}{2\sigma^2}\right),$$

and the expected absolute value is $\mathrm{EAV}(\eta) = \sqrt{\frac{2}{\pi}}\sigma$. Making use of the variance of Gaussian noise and $n = z - K\hat{u}$, the image restoration problem can be written as the constrained minimization problem

$$\min_{u\in BV(\Omega)} \int_\Omega |Du| \quad \text{subject to (s.t.)} \quad \int_\Omega |Ku - z|^2 dx \le \sigma^2|\Omega|, \tag{2.1}$$

which is equivalent to the unconstrained optimization problem (1.1) for an appropriate choice of $\lambda > 0$; see [6].

For salt-and-pepper noise, the value of the random variable η depends on u (here with some misuse of notation). In fact, suppose the intensity region is $[0, 1]$, then the conditional probability density function depending on u is given by

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

where r is the noise ratio. The conditional mean, variance and expected absolute value of the associated random variable η are readily obtained as

$$E(\eta \mid u) = \frac{r}{2}(1 - 2u), \quad \mathrm{Var}(\eta \mid u) = \frac{r}{2}(1 - 2u + 2u^2) - \frac{r^2}{4} + r^2u(1 - u), \quad EAV(\eta \mid u) = \frac{r}{2}.$$

Based on the expected absolute value, one may attempt to restore $\hat{u}$ by solving the following constrained minimization problem:

$$\min_{u\in BV(\Omega)} \int_\Omega |Du| \quad \text{s.t.} \quad \int_\Omega |Ku - z| dx \le \frac{r}{2}|\Omega|, \tag{2.2}$$

which is naturally linked to (1.2); see [16].

The formulations (2.1) and (2.2) can be condensed to obtain a combined model for removing Gaussian or salt-and-pepper noise:

$$\min_{u\in BV(\Omega)} \int_\Omega |Du| \quad \text{s.t.} \quad \frac{1}{|\Omega|} \int_\Omega \frac{1}{\tau} |Ku - z|^\tau \, dx \le \nu_\tau, \tag{2.3}$$

where $\tau = 2$ with $\nu_2 = \frac{\sigma^2}{2}$ for removing Gaussian noise, and $\tau = 1$ with $\nu_1 = \frac{r}{2}$ for removing salt-and-pepper noise. The above general total variation model corresponds to the unconstrained minimization problem

$$\min_{u\in BV(\Omega)} \int_\Omega |Du| + \frac{\lambda}{\tau} \int_\Omega |Ku - z|^\tau dx \tag{2.4}$$

with an appropriate choice of $\lambda > 0$. In what follows we call (2.4) the TV-model.

2.2. **Multi-scale vectorial total variation.** In [4,5] the L^2TV-model (1.1) was extended to recover color images based on vectorial total variation regularization. Transferring this regularization scheme to (2.4) we obtain the general vectorial total variation model (VTV-model)

$$\min_{\mathbf{u}\in\mathbf{BV}(\Omega)} \int_\Omega |D\mathbf{u}| + \frac{\lambda}{\tau}\int_\Omega |K\mathbf{u}-\mathbf{z}|_\tau^\tau dx, \tag{2.5}$$

with $\tau \in \{1,2\}$ based on the noise model. The space $\mathbf{BV}(\Omega)$ is the set of vector-valued functions $\mathbf{u}\in\mathbf{L}^1(\Omega)$ such that $\int_\Omega |D\mathbf{u}| < +\infty$, where the vectorial TV semi-norm $\int_\Omega |D\mathbf{u}|$ is defined as

$$\int_\Omega |D\mathbf{u}| = \sup\left\{ \int_\Omega \mathbf{u}\cdot \operatorname{div}\vec{\mathbf{v}}\; dx :\; \vec{\mathbf{v}}\in C_c^1(\Omega,\mathbb{R}^{M\times 2}), |\vec{\mathbf{v}}|_F \le 1 \text{ in } \Omega \right\}.$$

Note that the vectorial TV semi-norm is not the sum of the TV semi-norms of the respective channels but rather it couples the channels [5]. The space $\mathbf{BV}(\Omega)$ endowed with the norm

$$\|\mathbf{u}\|_{\mathbf{BV}(\Omega)} = \|\mathbf{u}\|_{\mathbf{L}^1(\Omega)} + \int_\Omega |D\mathbf{u}|$$

is a Banach space. Corresponding to (2.5), we may equivalently consider the constrained minimization problem

$$\min_{\mathbf{u}\in\mathbf{BV}(\Omega)} \int_\Omega |D\mathbf{u}| \quad \text{s.t. } \frac{1}{M|\Omega|}\int_\Omega \frac{1}{\tau}|K\mathbf{u}-\mathbf{z}|_\tau^\tau\; dx \le \nu_\tau,$$

with τ and ν_τ as in (2.3).

In the VTV-model, the regularization parameter λ controls the trade-off between a good fit of $\mathbf{z}$ and a smoothness requirement due to the vectorial total variation regularization. Since images are usually comprised of multiple objects at different scales, locally different λ depending on image features is desirable in order to obtain better restoration results. Roughly speaking, for small scale features, large λ leads to little smoothing so that details are usually preserved well. On the other hand, for large homogeneous features, small λ leads to smoothing so that noise is removed considerably. For this reason and based on (2.5), we consider a multi-scale vectorial total variation model (MVTV-model) for restoring the degraded color image $\mathbf{z}$ with Gaussian noise or salt-and-pepper noise:

$$\min_{\mathbf{u}\in\mathbf{BV}(\Omega)} \int_\Omega |D\mathbf{u}| + \frac{1}{\tau}\int_\Omega \lambda(x)|K\mathbf{u}-\mathbf{z}|_\tau^\tau\; dx, \tag{2.6}$$

where $\lambda\in L^\infty(\Omega)$ with $0<\epsilon_\lambda\le\lambda(x)\le\bar{\lambda}$ for almost all $x\in\Omega$, and

$$\lambda(x) = \int_\Omega w(x,y)\hat{\lambda}(y)\; dy \tag{2.7}$$

with $\hat{\lambda}\in L^1(\Omega)$. In addition, here w is the mean filter defined as

$$w(x,y) = \begin{cases} \frac{1}{\omega_\epsilon^2} & \text{if } |y-x|_\infty \le \frac{\omega}{2}, \\ \epsilon & \text{else,} \end{cases} \tag{2.8}$$

where $x \in \Omega$ is fixed, $\omega > 0$ sufficiently small is the essential width of the filter window, $0 < \epsilon \ll 1$ and ω_ϵ are such that $\int_\Omega \int_\Omega w(x,y)\, dydx = 1$. The corresponding locally constrained model is

$$\min_{\mathbf{u}\in\mathbf{BV}(\Omega)} \int_\Omega |D\mathbf{u}| \quad \text{s.t.} \ \frac{1}{M}\int_\Omega \frac{1}{\tau} w(x,y)|K\mathbf{u}-\mathbf{z}|_\tau^\tau(x)\, dx \le \nu_\tau. \tag{2.9}$$

The quantity $\hat{\lambda}$ in (2.7) can be interpreted as the Lagrange multiplier for the constraint in (2.9). Similar to [10, 16], the relation between (2.6) and (2.9) as well as existence and uniqueness can be studied.

2.3. **A unified Moreau-Yosida based primal-dual approach for the MVTV-model.** The unifying approach of this section enables us to study the TV-regularization with either L^1- or L^2-data-fidelity within one framework. In fact, one may even generalize to L^τ-data-fidelity for any $\tau \in [1,2]$; see (2.6). As a consequence, this allows us to introduce and analyze a unifying algorithmic framework for solving the respective image restoration problem as well.

We start by studying the Fenchel-Legendre pre-dual of a close approximation of (2.6). For this purpose we define

$$K_\lambda \mathbf{u} := \lambda^{1/\tau} K\mathbf{u}, \quad \mathbf{z}_\lambda := \lambda^{1/\tau}\mathbf{z} \quad \text{and} \quad \mathcal{F}(\mathbf{u}) := \frac{1}{\tau}\int_\Omega |K_\lambda \mathbf{u} - \mathbf{z}_\lambda|_\tau^\tau\, dx.$$

From now on we invoke the assumption that

$$(K_\lambda K_\lambda^*) \text{ is invertible;} \tag{A}$$

see, e.g., [9] for examples of blurring operators satisfying this assumption. In the denoising case, i.e. $K = \mathrm{id}$, this assumptions is obviously satisfied.

Applying the Fenchel calculus [11] we find that the convex conjugate of $\mathcal{F}$ is given by

$$\mathcal{F}^*(\mathbf{u}^*) = \langle (K_\lambda K_\lambda^*)^{-1} K_\lambda \mathbf{u}^*, \mathbf{z}_\lambda\rangle + \mathcal{J}^*((K_\lambda K_\lambda^*)^{-1} K_\lambda \mathbf{u}^*),$$

where

$$\mathcal{J}^*(\mathbf{v}) = \begin{cases} \frac{1}{\varsigma}\int_\Omega |\mathbf{v}|_\varsigma^\varsigma\, dx, & \text{if } 1 < \tau \le 2, \\ I_{\{\mathbf{w}\in\mathbf{L}^2(\Omega):|\mathbf{w}|_\infty \le 1 \text{ a.e. in } \Omega\}}(\mathbf{v}), & \text{if } \tau = 1, \end{cases}$$

where $\tau^{-1} + \varsigma^{-1} = 1$, with $\varsigma = \infty$ for $\tau = 1$, and $I_S(\cdot)$ denotes the indicator function of the set S. Next we define the problem

$$\begin{aligned} \min_{\vec{\mathbf{p}}\in\mathbf{H}_0(\mathrm{div})} \quad & \langle (K_\lambda K_\lambda^*)^{-1} K_\lambda \operatorname{div}\vec{\mathbf{p}}, \mathbf{z}_\lambda\rangle + \mathcal{J}^*((K_\lambda K_\lambda^*)^{-1} K_\lambda \operatorname{div}\vec{\mathbf{p}}), \\ \text{s. t.} \quad & |\vec{\mathbf{p}}|_F \le 1 \text{ a.e. in } \Omega, \end{aligned} \tag{2.10}$$

where $\mathbf{H}_0(\mathrm{div}) := \{\vec{\mathbf{v}} \in L^2(\Omega, \mathbb{R}^{M\times 2}) : \operatorname{div}\vec{\mathbf{v}} \in \mathbf{L}^2(\Omega),\ \vec{\mathbf{v}}_i \cdot \eta = 0 \text{ for } i = 1,\dots,M\}$. Here η denotes the outward unit normal to $\partial\Omega$ and $\vec{\mathbf{v}}_i$ refers to the i-th row of $\vec{\mathbf{v}}$. Assuming $(K_\lambda K_\lambda^*)^{-1} K_\lambda \in \mathcal{L}(\mathbf{L}^\varsigma(\Omega))$, note that continuity of $\mathcal{J}^*(\cdot)$ can only be guaranteed on $\mathbf{H}_0^\varsigma(\mathrm{div}) := \{\vec{\mathbf{v}} \in \mathbf{H}_0(\mathrm{div}) : \operatorname{div}\vec{\mathbf{v}} \in \mathbf{L}^\varsigma(\Omega)\}$, with $\varsigma = \infty$ for $\tau = 1$. As such a continuity property is essential for the Fenchel calculus and also in view of an efficient numerical approach, we replace $\mathcal{J}^*$ by its Moreau-Yosida regularization [19, 33]

$$\mathcal{J}_\mu^*(\mathbf{v}) := \inf_{\mathbf{w}\in\mathbf{L}^2(\Omega)} \left\{ \mathcal{J}^*(\mathbf{w}) + \frac{\mu}{2}\|\mathbf{w}-\mathbf{v}\|_{\mathbf{L}^2}^2 \right\} \tag{2.11}$$

with a positive parameter $\mu > 0$. Then, the resulting problem is given by

$$\begin{aligned} \min_{\vec{\mathbf{p}} \in \mathbf{H}_0(\mathrm{div})} \quad & \langle (K_\lambda K_\lambda^*)^{-1} K_\lambda \operatorname{div} \vec{\mathbf{p}}, \mathbf{z}_\lambda \rangle + \mathcal{J}_\mu^*((K_\lambda K_\lambda^*)^{-1} K_\lambda \operatorname{div} \vec{\mathbf{p}}), \\ \text{s. t.} \quad & |\vec{\mathbf{p}}|_F \leq 1 \text{ a.e. in } \Omega, \end{aligned} \tag{2.12}$$

Note that $\mathcal{J}_\mu^*(\mathbf{v}) \leq \frac{\mu}{2} \|\mathbf{v}\|_{\mathbf{L}^2}^2$ and $\mathcal{J}_\mu^*(\mathbf{v}) \to \mathcal{J}^*(\mathbf{v})$ as $\mu \to \infty$ for all $\mathbf{v} \in \mathbf{L}^2(\Omega)$. Moreover, well-known results [3] yield that $\mathcal{J}_\mu^*$ is continuously differentiable for every $0 < \mu < +\infty$. In what follows, the choices $\tau = 1$ and $\tau = 2$ will play an important role. Hence, let us briefly study $\mathcal{J}_\mu^*$ in these cases.

- *Case* $\tau = 2$. One readily finds that

$$\mathcal{J}_\mu^*(\mathbf{v}) = \frac{\mu}{2(1+\mu)} \int_\Omega |\mathbf{v}|_2^2 \, dx.$$

  For $\lambda(x) \equiv \lambda_0$ with $\lambda_0 > 0$, the problem (2.12) corresponds to a slightly rescaled version of the Fenchel pre-dual of the classical TV-problem; see [15]. The re-scaling factor is $\mu/(1+\mu)$ and may be easily incorporated in the choice of λ_0 to obtain equivalence between (2.12) and the pre-dual of [15]. Moreover, since (2.12) is interesting only for large values of μ, we have $\mu/(1+\mu) \approx 1$ and find that (2.12) closely approximates the Fenchel pre-dual established in [15] even without re-scaling λ.
- *Case* $\tau = 1$. The optimal $\mathbf{v}_\mu$ in (2.11) satisfies pointwise a.e.

$$0 \in \partial I_{\{\mathbf{w} \in \mathbf{L}^2(\Omega) : |\mathbf{w}|_\infty \leq 1 \text{ a.e. in } \Omega\}}(\mathbf{v}_\mu) + \mu(\mathbf{v}_\mu - \mathbf{v}).$$

  Obviously, $\mathbf{v}_\mu \in \{\mathbf{w} \in \mathbf{L}^2(\Omega) : |\mathbf{w}|_\infty \leq 1 \text{ a.e. in } \Omega\}$; otherwise $\mathcal{J}^*(\mathbf{v}_\mu) = +\infty$. Thus,

$$\mathcal{J}_\mu^*(\mathbf{v}) = \frac{\mu}{2} \int_\Omega \max(0, |\mathbf{v}|_\infty - 1)^2 \, dx.$$

Now we study the effect of the Moreau-Yosida regularization, i.e., replacing (2.10) by (2.12), on the primal problem. By the Fenchel-Legendre calculus [11], the convex conjugates of

$$\begin{aligned} \Psi_\mu^*(\mathbf{v}^*) &:= \mathcal{J}_\mu^*(\mathbf{v}^*) - \langle \mathbf{v}^*, \mathbf{z}_\lambda \rangle, \\ \Phi^*(\vec{\mathbf{p}}^*) &:= I_{\{\vec{\mathbf{w}} \in \mathbf{H}_0(\mathrm{div}) : |\vec{\mathbf{w}}|_F \leq 1 \text{ a.e. in } \Omega\}}(\vec{\mathbf{p}}^*) \end{aligned}$$

are obtained as

$$\begin{aligned} \Psi_\mu(\mathbf{v}) &:= \mathcal{J}_\mu(\mathbf{v} + \mathbf{z}_\lambda), \\ \Phi(\vec{\mathbf{p}}) &:= \sup_{\vec{\mathbf{q}} \in S_1} \langle \vec{\mathbf{q}}, \vec{\mathbf{p}} \rangle, \end{aligned}$$

where $|\vec{\mathbf{p}}|_F \leq 1$ a.e. in Ω and $S_1 = \{\vec{\mathbf{q}} \in \mathbf{H}_0(\mathrm{div}) : |\vec{\mathbf{q}}|_F \leq 1 \text{ a.e. in } \Omega\}$. For a further analysis of Φ let $S_2 = \{\vec{\mathbf{q}} \in C_c^1(\Omega, \mathbb{R}^{M \times 2}) : |\vec{\mathbf{q}}|_F \leq 1 \text{ in } \Omega\}$, which is dense in S_1 with respect to the topology of $\mathbf{H}_0(\mathrm{div})$. Extending the results in [15] to vector fields, we get that $\tilde{\Phi}((-\operatorname{div})^* \mathbf{u}) := \sup_{\vec{\mathbf{q}} \in S_2} \langle \mathbf{u}, -\operatorname{div} \vec{\mathbf{q}} \rangle$ is finite if and only if $\mathbf{u} \in \mathbf{BV}(\Omega)$, that is,

$$\tilde{\Phi}((-\operatorname{div})^* \mathbf{u}) = \int_\Omega |D\mathbf{u}| < \infty \quad \text{for} \quad \mathbf{u} \in \mathbf{BV}(\Omega). \tag{2.13}$$

According to the Fenchel duality theorem [11] we have that the dual of (2.12) is given by

$$\min_{\mathbf{u}\in\mathbf{BV}(\Omega)} \int_\Omega |D\mathbf{u}| + \mathcal{J}_\mu(K_\lambda \mathbf{u} - \mathbf{z}_\lambda), \tag{2.14}$$

where

$$\mathcal{J}_\mu(\mathbf{v}) = \frac{1}{2\mu}\int_\Omega |\mathbf{v}|_2^2 \, dx + \frac{1}{\tau}\int_\Omega |\mathbf{v}|_\tau^\tau \, dx. \tag{2.15}$$

This proves the following result.

Theorem 1. *For $\mu > 0$ and $\tau \in [1,2]$ the dual problem of (2.12) is given by*

$$\min_{\mathbf{u}\in\mathbf{BV}(\Omega)} \int_\Omega |D\mathbf{u}| + \frac{1}{2\mu}\int_\Omega |K_\lambda \mathbf{u} - \mathbf{z}_\lambda|_2^2 \, dx + \frac{1}{\tau}\int_\Omega |K_\lambda \mathbf{u} - \mathbf{z}_\lambda|_\tau^\tau \, dx \tag{2.16}$$

From this general result we come back to the two cases $\tau = 1$ and $\tau = 2$, which we focus on in our numerics.

- *Case $\tau = 2$.* The convex conjugate of $\mathcal{J}_\mu^*$ is readily obtained as

  $$\mathcal{J}_\mu(\mathbf{v}) = \frac{1+\mu}{2\mu}\int_\Omega |\mathbf{v}|_2^2 dx$$

  such that (2.16) becomes

  $$\min_{\mathbf{u}\in\mathbf{BV}(\Omega)} \int_\Omega |D\mathbf{u}| + \frac{1+\mu}{2\mu}\int_\Omega |K_\lambda \mathbf{u} - \mathbf{z}_\lambda|_2^2 \, dx. \tag{2.17}$$

  Thus, (2.17) is just a re-scaled version of the original TV-problem. As noted above, only large μ-values are of interest such that $(1+\mu)/\mu \approx 1$.
- *Case $\tau = 1$.* Similarly to [16] one finds that

  $$\mathcal{J}_\mu(\mathbf{v}) = \frac{1}{2\mu}\int_\Omega |\mathbf{v}|_2^2 \, dx + \int_\Omega |\mathbf{v}|_1 \, dx.$$

  Thus, (2.16) becomes

  $$\min_{\mathbf{u}\in\mathbf{BV}(\Omega)} \int_\Omega |D\mathbf{u}| + \frac{1}{2\mu}\int_\Omega |K_\lambda \mathbf{u} - \mathbf{z}_\lambda|_2^2 \, dx + \int_\Omega |K_\lambda \mathbf{u} - \mathbf{z}_\lambda|_1 \, dx. \tag{2.18}$$

The problem (2.12) is appealing as, in contrast to (2.18), it consists of minimizing a smooth objective subject to pointwise constraints. For numerical stability reasons, the only difficulty to overcome is the non-uniqueness of the solution of the general dual problem (2.12). Following [17] we first study the case $\tau = 2$ and generalize afterwards. In fact, as a remedy for the non-uniqueness we propose the following regularization, which we write here for the primal problem:

$$\min_{\mathbf{u}\in\mathbf{H}_0^1(\Omega)} \frac{\varrho}{2}\int_\Omega |\nabla \mathbf{u}|_F^2 \, dx + \int_\Omega \Phi_\gamma(\nabla \mathbf{u}) \, dx + \frac{1+\mu}{2\mu}\int_\Omega |K_\lambda \mathbf{u} - \mathbf{z}_\lambda|_2^2 \, dx, \tag{2.19}$$

where

$$\Phi_\gamma(\vec{\mathbf{w}})(x) = \begin{cases} |\vec{\mathbf{w}}(x)|_F - \frac{\gamma}{2}, & \text{if } |\vec{\mathbf{w}}(x)|_F \geq \gamma, \\ \frac{1}{2\gamma}|\vec{\mathbf{w}}(x)|_F^2, & \text{if } |\vec{\mathbf{w}}(x)|_F < \gamma. \end{cases} \tag{2.20}$$

Note that the regularization acts in two ways: First it increases the regularity of u by adding the ϱ-term, where we assume $0 < \varrho \ll \epsilon_\lambda \leq \lambda(x)$ for a.e. $x \in \Omega$, and secondly it smoothes the TV-regularization locally. For the latter study Φ_γ. The effect is a unique solution of the associated Fenchel dual which is given by

$$\begin{aligned} \min_{\vec{\mathbf{p}} \in L^2(\Omega, \mathbb{R}^{M\times 2})} \quad & \tfrac{1}{2}\| \operatorname{div} \vec{\mathbf{p}} + \tfrac{1+\mu}{\mu} K_\lambda^* \mathbf{z}_\lambda \|^2_{\mathbf{H}^{-1}} + \tfrac{\gamma}{2} \textstyle\int_\Omega |\vec{\mathbf{p}}|_F^2 \, dx - \tfrac{1+\mu}{2\mu} \int_\Omega |\mathbf{z}_\lambda|_2^2 dx \\ \text{s.t.} \quad & |\vec{\mathbf{p}}|_F \leq 1 \text{ a.e. in } \Omega. \end{aligned} \tag{2.21}$$

where $\|\mathbf{v}\|^2_{\mathbf{H}^{-1}} = \langle H_{\varrho,\mu,K_\lambda} \mathbf{v}, \mathbf{v} \rangle_{\mathbf{H}^1_0, \mathbf{H}^{-1}}$ for $\mathbf{v} \in \mathbf{H}^{-1}(\Omega)$ with $H_{\varrho,\mu,K_\lambda} = (\frac{1+\mu}{\mu} K_\lambda^* K_\lambda - \varrho\Delta)^{-1}$. Note that in this case the involved function spaces are reflexive. Hence the dual problem of (2.21) is equal to the primal problem (2.19). Although the explicit computation of the dual might be tedious, the approach is readily generalized. In fact, for $\tau \in [1, 2]$ we consider

$$\min_{\mathbf{u} \in \mathbf{H}^1_0(\Omega)} \frac{\varrho}{2} \int_\Omega |\nabla \mathbf{u}|_F^2 \, dx + \int_\Omega \Phi_\gamma(\nabla \mathbf{u}) \, dx + \mathcal{J}_\mu(K_\lambda \mathbf{u} - \mathbf{z}_\lambda). \tag{2.22}$$

The following first-order optimality conditions characterize a primal-dual solution pair $(\bar{\mathbf{u}}, \bar{\vec{\mathbf{p}}})$:

$$-\varrho\Delta\bar{\mathbf{u}} - \operatorname{div} \bar{\vec{\mathbf{p}}} + \frac{1}{\mu} K_\lambda^*(K_\lambda \bar{\mathbf{u}} - \mathbf{z}_\lambda) + K_\lambda^*(|K_\lambda \bar{\mathbf{u}} - \mathbf{z}_\lambda|_c^{\tau-1} \star \sigma(K_\lambda \bar{\mathbf{u}} - \mathbf{z}_\lambda)) = 0, \tag{2.23a}$$

$$\max(\gamma, |\nabla \bar{\mathbf{u}}|_F) \bar{\vec{\mathbf{p}}} - \nabla \bar{\mathbf{u}} = 0, \tag{2.23b}$$

where $\sigma(\mathbf{w}) \in \partial |\mathbf{w}|_1 \in \mathbb{R}^M$ for $\mathbf{w} \in \mathbb{R}^M$, and $\star$ denotes the Hadamard product of vectors in $\mathbb{R}^n$, i.e., $\mathbf{v} \star \mathbf{w} = (\mathbf{v}_1 \mathbf{w}_1, \ldots, \mathbf{v}_n \mathbf{w}_n)^\top$.

Remark 2. *The case $\tau = 1$ is still problematic, as we are only able to characterize $(K_\lambda \mathbf{u})_i(x)$ as greater than or less than $(\mathbf{z}_\lambda)_i(x)$ whenever $|((K_\lambda K_\lambda^*)^{-1} K_\lambda \operatorname{div} \vec{\mathbf{p}})_i(x)| = 1$ for $i \in \{1, \ldots, M\}$. Analogously to [9, 16], for $\tau = 1$ we propose the following modification of* (2.22)

$$\min_{\mathbf{u} \in \mathbf{H}^1_0(\Omega)} \frac{\varrho}{2} \int_\Omega |\nabla \mathbf{u}|_F^2 \, dx + \int_\Omega \Phi_\gamma(\nabla \mathbf{u}) \, dx + \int_\Omega \Psi_{\mu,\beta}(K_\lambda \mathbf{u} - \mathbf{z}_\lambda) \, dx, \tag{2.24}$$

where

$$\Psi_{\mu,\beta}(\mathbf{v}_i)(x) = \begin{cases} \frac{1}{2\beta}(\mathbf{v}_i(x))^2, & \text{if } |\mathbf{v}_i(x)| \leq \beta, \\ \frac{1}{2(\beta+\mu)}(\mathbf{v}_i(x))^2 + \frac{\mu}{\mu+\beta}|\mathbf{v}_i(x)| - \frac{\mu\beta}{2(\mu+\beta)}, & \text{if } |\mathbf{v}_i(x)| > \beta, \end{cases} \tag{2.25}$$

where $i \in \{1, \ldots, M\}$ corresponds to each channel. Hence, (2.23) *modifies to*

$$-\varrho\Delta\bar{\mathbf{u}} - \operatorname{div} \bar{\vec{\mathbf{p}}} + \frac{1}{\mu+\beta} K_\lambda^*(K_\lambda \bar{\mathbf{u}} - \mathbf{z}_\lambda) + \frac{\mu}{\mu+\beta} K_\lambda^* \bar{\mathbf{v}} = 0, \tag{2.26a}$$

$$\max(\gamma, |\nabla \bar{\mathbf{u}}|_F) \bar{\vec{\mathbf{p}}} - \nabla \bar{\mathbf{u}} = 0, \tag{2.26b}$$

$$\max(\boldsymbol{\beta}, |K_\lambda \bar{\mathbf{u}} - \mathbf{z}_\lambda|) \star \bar{\mathbf{v}} - (K_\lambda \bar{\mathbf{u}} - \mathbf{z}_\lambda) = 0, \tag{2.26c}$$

where $\boldsymbol{\beta} = (\beta, \ldots, \beta)^\top \in \mathbb{R}^M$, and $\max(\mathbf{v}, \mathbf{w}) = (\max(\mathbf{v}_1, \mathbf{w}_1), \ldots, \max(\mathbf{v}_n, \mathbf{w}_n))^\top \in \mathbb{R}^n$ for the vectors $\mathbf{v}$, $\mathbf{w} \in \mathbb{R}^n$.

Note that the systems (2.23) and (2.26) are non-smooth, i.e. not necessarily Fréchet-differentiable. But the discrete version of these systems can be solved by semismooth Newton methods, respectively, see, e.g., [9, 10, 16, 17]. Equipped with a line search strategy, it can be shown that the generalized Newton solver converges globally, i.e. regardless of its initialization, and locally at a superlinear rate. The proof is analogous to the one in [17].

3. Spatially Dependent Regularization Parameter Selection

Since the capability of multi-scale vectorial total variation is mainly limited by the selection of the parameter λ, in this section we discuss a way to choose λ in the MVTV-model.

3.1. **Local statistical estimator for noise.** Suppose the variance σ^2 of Gaussian noise or the noise ratio r of salt-and-pepper noise is known; in practice estimates of σ^2 and r can be obtained by following, e.g., the techniques in [2, 13]. In order to enhance color image details while preserving homogenous regions, we search for a reconstruction where some statistical characteristics of the residual $K\mathbf{u} - \mathbf{z}$ are closer to those characteristics of the noise in both the detail regions and the homogeneous parts. For this purpose, in view of (2.9) and temporarily assuming $\epsilon = 0$ in (2.8) we are interested in spatially adapted $\lambda(x)$ to satisfy the constraint

$$\frac{1}{M|\Omega^\omega|}\int_{\Omega^\omega}\frac{1}{\tau}|K\mathbf{u}-\mathbf{z}|_\tau^\tau\,dx \le \nu_\tau \tag{3.1}$$

locally where $\Omega^\omega \subset \Omega$ is specified below. Here, we introduce a local statistical estimator (LSE for short) for an automated adaptive choice of λ.

In the rest of this section, only discrete versions of images are considered. The residual image is denoted by $\mathbf{r}^h = \mathbf{z}^h - K^h\mathbf{u}^h$, where $\mathbf{r}^h$, $\mathbf{z}^h$, $K^h\mathbf{u} \in \mathbb{R}^{m\times n\times M}$, $m \times n$ corresponds to the image size, $\mathbf{u}^h \in \mathbb{R}^{m\times n\times M}$ is the restored image from the minimization problem (2.6) with $\lambda > 0$. In order to satisfy the constraint (3.1), λ has to be chosen on whether Ω^ω is located in homogeneous or detail regions of the image. If we use a relatively small parameter λ, the residual $\mathbf{r}^h$ will include noise as well as details. Then, based on the satisfaction of the constraint (3.1), the distribution of details in the image is reflected.

Let $\Omega^\omega_{i,j}$ 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.,

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

where $\lfloor\cdot\rfloor$ means rounding to the nearest integer towards zero. Then we apply the mean filter with window size ω to the residual image $\mathbf{r}^h$ as follows:

$$\mathrm{LSE}^{\omega,\tau}_{i,j} = \frac{1}{\tau\omega^2 M}\sum_{l=1}^{M}\sum_{(s,t)\in\Omega^\omega_{i,j}}\left|\mathbf{r}^h_{s,t,l}\right|^\tau = \frac{1}{\tau\omega^2 M}\sum_{l=1}^{M}\sum_{(s,t)\in\Omega^\omega_{i,j}}\left|\mathbf{z}^h_{s,t,l} - (K^h\mathbf{u}^h)_{s,t,l}\right|^\tau.$$

In general, when there is only noise left in the residual, the constraint (3.1) is likely to be satisfied, i.e., $\mathrm{LSE}^{\omega,\tau}_{i,j} \le \nu_\tau$; otherwise we consider $\mathrm{LSE}^{\omega,\tau}$ to have a large value, which is usually due to details.

3.2. **Selection of the parameter λ.** According to its definition, $\mathrm{LSE}_{i,j}^{\omega,\tau}$ reflects the statistical characteristic of the residual image in a given window $\Omega_{i,j}^{\omega}$. Ideally, the residual contains noise only. Hence, whenever $\mathrm{LSE}_{i,j}^{\omega,\tau} \leq \nu_\tau$, it is assumed that in the window the residual primarily consists of noise; otherwise, i.e. if $\mathrm{LSE}_{i,j}^{\omega,\tau} > \nu_\tau$, significant image details are left in the residual. Therefore, λ needs to be increased in this region in order to preserve the details in the reconstruction.

For adapting λ algorithmically we proceed as follows. Initially we set λ to a same small constant to obtain an over-smoothed restored image. Given $\hat{\lambda}_k$ (which yields λ_k) and an associated reconstruction $\mathbf{u}_k$, based on whether $\mathrm{LSE}^{\omega,\tau,k} \leq \nu_\tau$ or not, we generate an improved regularization parameter as follows:

$$(\hat{\lambda}_{k+1})_{i,j} = \zeta \cdot \min\left((\hat{\lambda}_k)_{i,j} + \rho \max\left(\left(\mathrm{LSE}_{i,j}^{\omega,\tau,k}\right)^{\frac{1}{\tau}} - (\nu_\tau)^{\frac{1}{\tau}}, 0\right), L\right), \tag{3.2a}$$

$$(\lambda_{k+1})_{i,j} = \frac{1}{\omega^2} \sum_{(s,t)\in\Omega_{i,j}^{\omega}} \left(\hat{\lambda}_{k+1}\right)_{s,t}, \tag{3.2b}$$

where $\zeta \geq 1$, $\rho > 0$, $\mathrm{LSE}^{\omega,\tau,k}$ is obtained from $\mathbf{u}_k$, and L is a large positive value to ensure bounded $\hat{\lambda}_k$. In our numerics we set the parameter $\rho = \rho_k = ||\hat{\lambda}_k||_\infty/\nu_\tau$ in order to keep the new $\hat{\lambda}_{k+1}$ at the same scale as $\hat{\lambda}_k$, $L = 1000$, $\zeta = 2$ for Gaussian noise removal, and we choose $\zeta = 1.1$ for salt-and-pepper noise removal since in this case the model is much more sensitive to changes in λ; see the analysis in [7] for the latter. In addition, we typically set $\omega = 17$. In Section 5, however, we study the influence of the window size on the restoration results.

4. Our Method

In [28, 29] an image decomposition method (TNV-algorithm) was proposed, which captures different features in the image by varying the regularization parameter in (1.1). The TNV-algorithm relies on a scalar regularization parameter, increasing this parameter and appropriately feeding back the previous scales allows to recover small scale features. However, there is a point where the noise is also recovered as the parameter gets larger. Hence, if the latter is the case, the method should refrain from further increasing the regularization parameter and stop. In the following algorithm, we intertwine the idea of [28, 29] with the MVTV-model (2.6) and combine it with the spatially adapted regularization parameter selection.

Algorithm.

1: Initialize $\mathbf{u}_0^h = 0 \in \mathbb{R}^{m\times n\times M}$, $\vec{\mathbf{p}}_0^h = 0 \in \mathbb{R}^{m\times n\times M\times 2}$, $\lambda_0 = \tilde{\lambda}_0 \mathbf{1}_{mn} \in \mathbb{R}^{m\times n}$, $\mathbf{w}_0^h = \mathbf{z}^h$ and $k = 0$.

2: Utilizing the Moreau-Yosida regularization approach of Section 2.2 and 2.3, find $\tilde{\mathbf{u}}_k^h$ by solving the discrete version of the following problem:

$$\tilde{\mathbf{u}}_k \in \arg\min_{\check{\mathbf{u}}\in\mathbf{BV}(\Omega)} \int_\Omega |D\check{\mathbf{u}}| + \frac{1}{\tau}\int_\Omega \lambda_k |K\check{\mathbf{u}} - \mathbf{w}_k|_\tau^\tau dx$$

(a)

(b)

FIGURE 1. Original images: (a) "Barbara", (b) "Goldhill".

3: Update $\mathbf{u}_{k+1}^h = \mathbf{u}_k^h + \tilde{\mathbf{u}}_k^h$ and $\mathbf{w}_{k+1}^h = \mathbf{z}^h - K^h \mathbf{u}_{k+1}^h$.

4: Based on $\mathbf{u}_{k+1}^h$, update

$$(\hat{\lambda}_{k+1})_{i,j} = \zeta \cdot \min\left((\hat{\lambda}_k)_{i,j} + \rho \max\left(\left(\mathrm{LSE}_{i,j}^{\omega,\tau,k} \right)^{\frac{1}{\tau}} - (\nu_\tau)^{\frac{1}{\tau}}, 0 \right), L \right),$$

$$(\lambda_{k+1})_{i,j} = \frac{1}{\omega^2} \sum_{(s,t)\in\Omega_{i,j}^{\omega}} \left(\hat{\lambda}_{k+1} \right)_{s,t}.$$

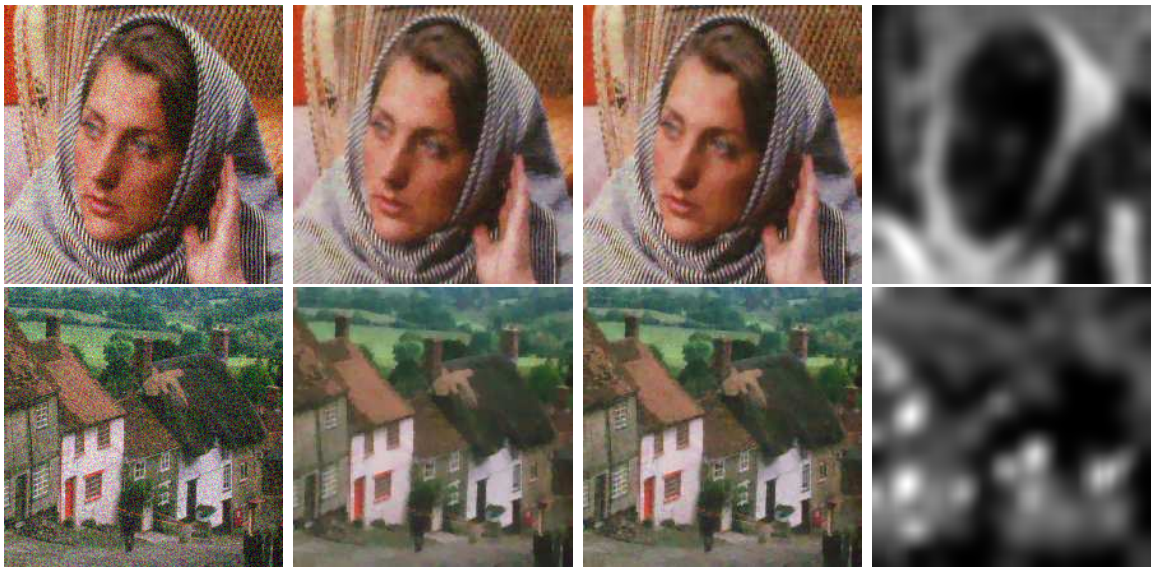
5: Stop; or set $k := k+1$ and go to step 2.

In this algorithm, $\mathbf{1}_{mn} \in \mathbb{R}^{m\times n}$ denotes the matrix with all elements equal to 1. In our numerical practice, initially we set λ to a relatively small positive constant $\tilde{\lambda}_0$. For Gaussian noise removal, we always use $\tilde{\lambda}_0 = 2.5$. However, for removing salt-and-pepper noise the model is more sensitive to the selection of $\tilde{\lambda}_0$, and hence, our choice will depend on the noise ratio. Similar to the Bregman iteration proposed in [23], we stop the iterative procedure as soon as the constraint (3.1) is satisfied globally, i.e. $\Omega^\omega := \Omega$. For Gaussian noise removal with a uniform initial $\tilde{\lambda}_0$, our algorithm always satisfies the stopping condition after three iterations. However, for removing salt-and-pepper noise the number of iteration steps depends on the choice of $\tilde{\lambda}_0$. The respective iteration numbers are listed together with the numerical results in the next section.

5. Numerical Results

In this section we discuss numerical results to study the behavior of our method and compare it with the vectorial total variation method, which is used to restore color images by solving the VTV-model with a scalar λ. For illustrations, the results for the 220-by-220 RGB color images (i.e., $M = 3$) "Barbara" and "Goldhill", are presented in Figure 1. In all of our experiments the image intensity range is scaled to $[0, 1]$.

5.1. **Color image denoising.** Here, we concentrate on image denoising, i.e., K is the identity operator. The degraded images contain Gaussian white noise with $\sigma = 0.1$, see Figure 2(a), or salt-and-pepper noise with the noise ratio $r = 0.4$, see Figure 3(a). For the VTV-model operating with a scalar λ, we adjust λ until the visually best results were obtained. In our method, we solve the MVTV-model with spatially dependent λ. Concerning



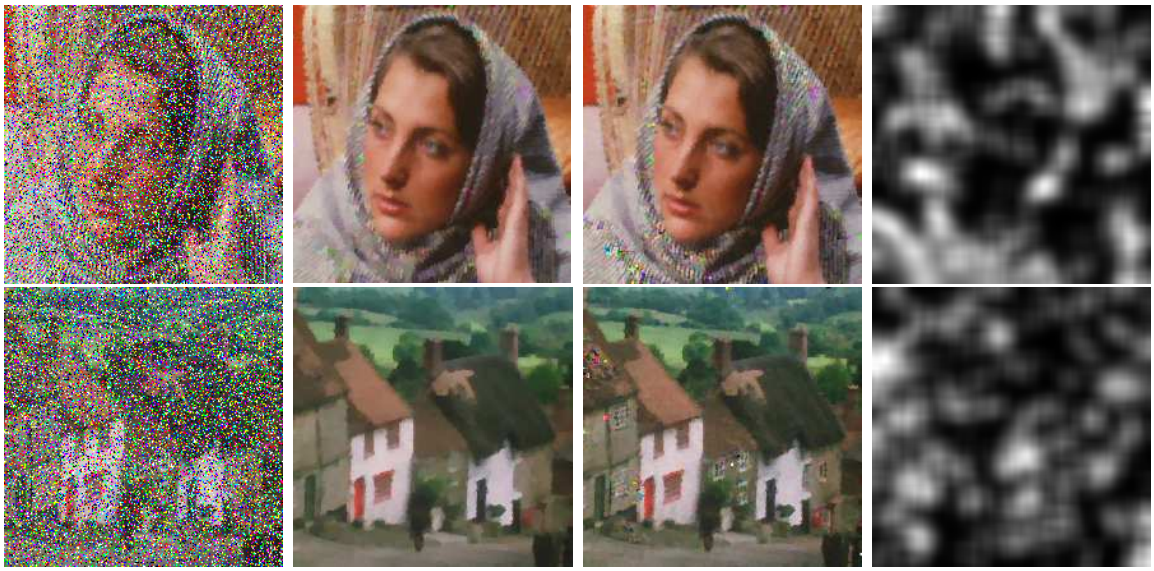
(a) (b) (c) (d)

FIGURE 2. Results of removing Gaussian noise for the images "Barbara" (the 1st row) and "Goldhill" (the 2nd row): (a) Noisy images, (b) From VTV-model, (c) From MVTV-model, (d) Final λ for MVTV-model.

the choice of the parameters in (2.19) and (2.24), the guideline is to use rather small β, ϱ, γ and large μ in order to obtain an approximate MVTV-model. In our numerical experiments we choose $\mu = 10^6$, $\varrho = 0$ (the ϱ-regularization serves primarily function space purposes and is left out in the discrete setting), $\beta = 0.001$ and $\gamma = 0.01$. In addition, λ controls the trade-off between a good data fit and a smoothness requirement due to the total variation regularization. In [7] it was shown that the scalar TV-model for $\tau = 1$ is more sensitive with respect to λ than for $\tau = 2$. In fact, for $\tau = 1$, for every scale there exists a threshold for λ such that the features of this scale and smaller scales disappear. For $\tau = 2$, on the other hand, a scale dependent "fading-away"-effect is observed as λ is decreased. In our tests, in order to remove salt-and-pepper noise with 40% noise ratio, we initialize $\tilde{\lambda}_0 = 0.7$. In this case, our algorithm satisfies the stopping condition after four iterations. Moreover, we suggest that large $\tilde{\lambda}_0$ is used for low noise levels in order to preserve details whereas small $\tilde{\lambda}_0$ is used in case of high noise levels.

The results are shown in Figure 2 and 3. We can see that compared with the results from the VTV-model our method suppresses the noise successfully while preserving more details. In this respect observe the background and the scarf of "Barbara" and the window and roof regions in "Goldhill". In addition, we also show the final values of the parameter λ obtained by our choice rule (3.2). We find that in detail regions λ is usually large in order to preserve the details, and it is small in the homogeneous regions to remove noise.

In order to test our method for different values of the window size ω, Figure 4 shows the restored images with $\omega = 7, 17, 27$. For Gaussian noise removal we observe a remarkable



(a) (b) (c) (d)

FIGURE 3. Results of removing salt-and-pepper noise for the images "Barbara" (the 1st row) and "Goldhill" (the 2nd row): (a) Noisy images, (b) From VTV-model, (c) From MVTV-model, (d) Final λ for MVTV-model.

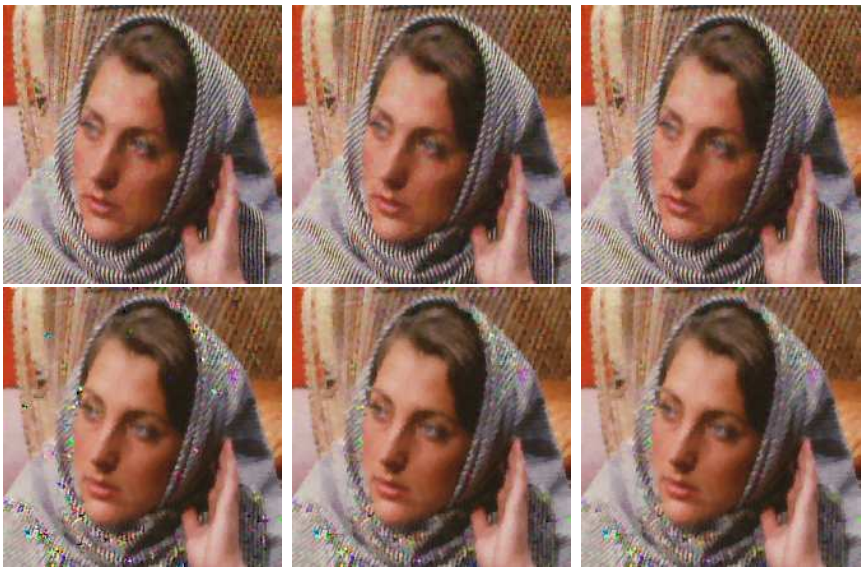
stability with respect to ω. However, for salt-and-pepper noise removal larger window size leads to better results. We clearly observe more noise remaining in the restored image for small ω. For $\tau = 1$, this is mainly due to the sensitivity of the MVTV-model with respect to the regularization parameter. Moreover, the reduced number of random variables contained in a small window increases the statistical error in the regularization parameter update (3.2).

5.2. **Color image deblurring and denoising.** In this section, we consider the restoration of noisy blurred images. The blurring operator K is a cross-channel blurring operator with the kernel

$$\begin{bmatrix} K_{rr} & K_{rg} & K_{rb} \\ K_{gr} & K_{gg} & K_{gb} \\ K_{br} & K_{bg} & K_{bb} \end{bmatrix} = \begin{bmatrix} 0.8 \cdot (M, 7, 135) & 0.1 \cdot (A, 7) & 0.1 \cdot (G, 9, 7) \\ 0.1 \cdot (A, 9) & 0.8 \cdot (G, 5, 1) & 0.1 \cdot (M, 7, 90) \\ 0.1 \cdot (G, 7, 5) & 0.1 \cdot (M, 7, 45) & 0.8 \cdot (A, 5) \end{bmatrix},$$

where (A, r) denotes the average blur with window size r, (G, r, σ) denotes Gaussian blur with window size r and standard deviation σ, (M, l, θ) denotes motion blur with length l and angle θ, and (r, g, b) are the three channels in the RGB color model. Further we have Gaussian white noise with $\sigma = 0.02$ or salt-and-pepper noise with $r = 0.6$.

Figures 5 and 6 show the degraded images and the results by minimizing the VTV-model and the MVTV-model by our method, respectively. In order to remove the salt-and-pepper noise with 60% noise ratio, we set $\tilde{\lambda}^0 = 1.8$ in our method. The selection of the other parameters is as in Section 5.1, and the number of iteration steps is 5. In contrast to the results obtained by the VTV-model which uses the scalar λ, our method preserves more



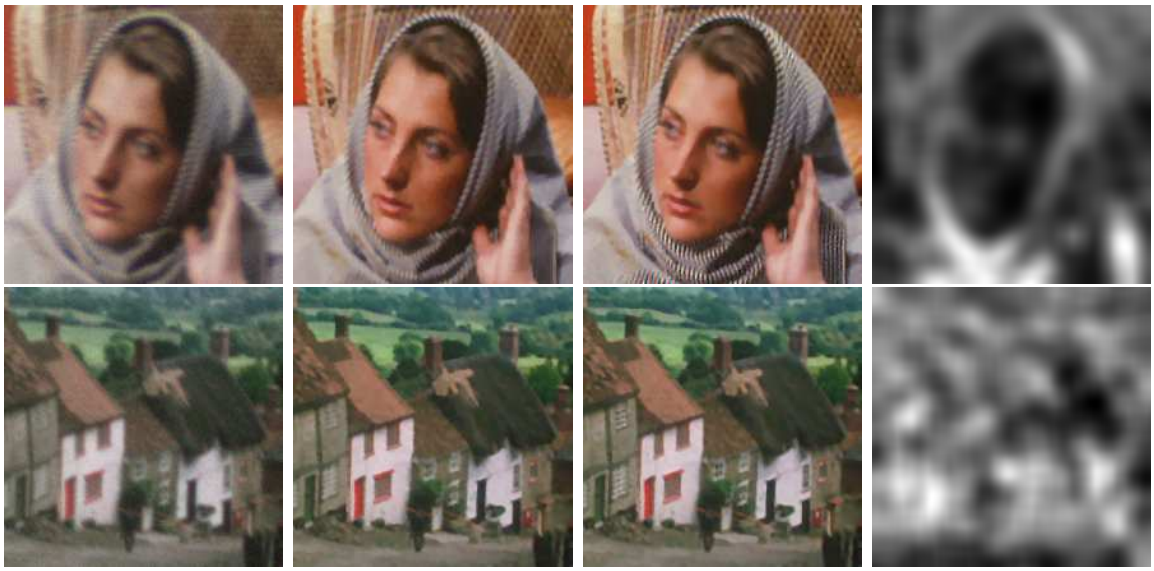
(a) (b) (c)

FIGURE 4. Restored images by our method with different ω for Gaussian noise removal (the 1st row) and salt-and-pepper noise removal (the 2nd row): (a) $\omega = 7$, (b) $\omega = 17$, (c) $\omega = 27$.

details; see, e.g., the background of "Barbara" and the window and roof regions of "Goldhill". Furthermore, for restoring blurred images with Gaussian noise our method is still able to distinguish most of the detail regions properly.

6. Conclusion

A multi-scale vectorial total variation model with a unified data-fidelity term $\|K\mathbf{u} - \mathbf{z}\|^{\tau}_{\mathbf{L}^{\tau}(\Omega)}$ and spatially dependent regularization parameter λ for color image restoration is proposed in this paper. We focus on the cases $\tau = 1$ and $\tau = 2$ suitable for salt-and-pepper noise and Gaussian noise, respectively. The discrete system of the MVTV-model are solved by Fenchel-duality and semismooth Newton techniques. It was shown that with a local statistical estimator to update λ iteratively, the new method can restore the degraded images efficiently while preserving details. Our method is able to properly reconstruct blurred images containing Gaussian noise. In case of salt-and-pepper noise, due to the high noise ratio and the significant gap between noise values and the noise-free pixel values, the detail regions are more difficult to detected. In the latter context, potential future research may focus on alternative statistical measures or alternative ways to select λ. In fact, we explored vector-valued choices of λ in the RGB-context, where each component function of λ corresponds to one color channel. In some cases we indeed obtained improved reconstructions. This approach, however, raises the question of an information balance when choosing λ.



(a) (b) (c) (d)

FIGURE 5. Results of restoring the blurred images with Gaussian noise: "Barbara" (the 1st row) and "Goldhill" (the 2nd row): (a) Noisy images, (b) From VTV-model, (c) From MVTV-model, (d) Final λ for MVTV-model.

REFERENCES

[1] A. Almansa, C. Ballester, V. Caselles, and G. Haro. A TV based restoration model with local constraints. *J. Sci. Comput.*, 34(3):209–236, 2008.

[2] H.C. Andrews and B.R. Hunt. *Digital Image Restoration*. Prentice Hall, 1977.

[3] Jean-Pierre Aubin and Ivar Ekeland. *Applied nonlinear analysis*. Dover Publications Inc., Mineola, NY, 2006. Reprint of the 1984 original.

[4] P. Blomgren and T. Chan. Color tv: Total variation methods for restoration of vector-valued images. *IEEE Transactions on Image Processing*, 7:304–309, 1998.

[5] X. Bresson and T.F. Chan. Fast dual minimization of the vectorial total variation norm and applications to color image processing. *Inverse Problems and Imaging*, 2(4):455–484, 2008.

[6] A. Chambolle and P-L. Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76:167–188, 1997.

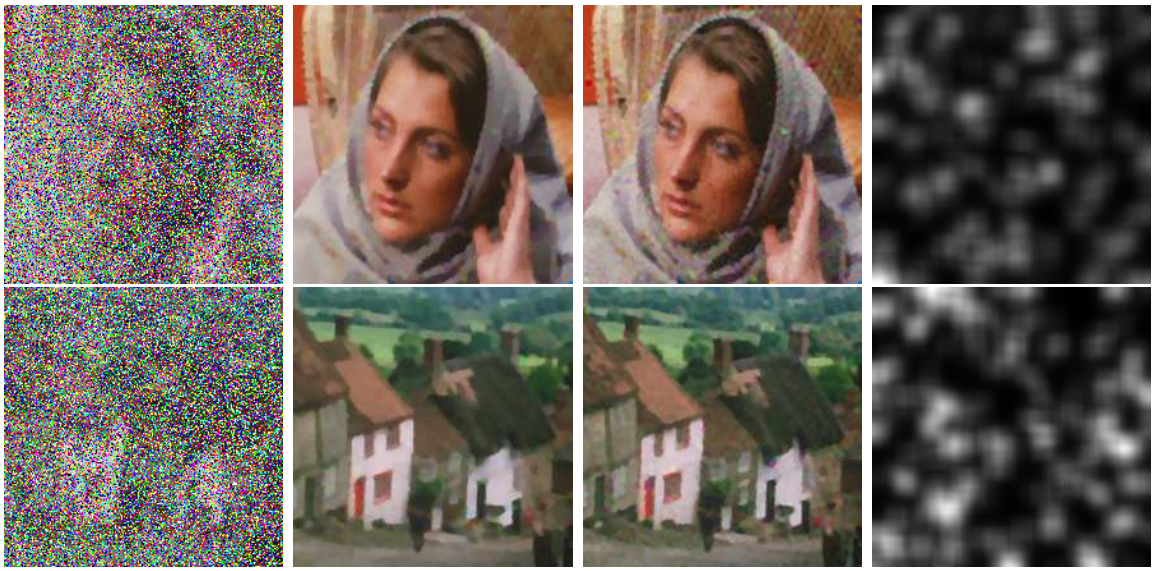
[7] T. F. Chan and S. Esedoḡlu. Aspects of total variation regularized L^1 function approximation. *SIAM J. Appl. Math.*, 65:1817–1837, 2005.

[8] Y.Q. Dong and M. Hintermüller. Multi-scale vectorial total variation with automated regularization parameter selection for color image restoration. In *Proc. SSVM'2009*, volume Lecture Notes in Computer Science, pages 271–281. Springer, 2009.

[9] Y.Q. Dong, M. Hintermüller, and M. Neri. A primal-dual method for L^1 TV image denoising. To appear in SIAM Journal on Imaging Sciences. URL: http://www.uni-graz.at/imawww/ifb/IFB_Report27_HiDoNe09.pdf.

[10] Y.Q. Dong, M. Hintermüller, and M. Rincon-Camacho. Automated parameter selection in a multi-scale total variation model. IFB-Report No. 22 (11/2008), Institute of Mathematics and Scientific Computing, University of Graz.

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



(a) (b) (c) (d)

FIGURE 6. Results of restoring the blurred images with salt-and-pepper noise: "Barbara" (the 1st row) and "Goldhill" (the 2nd row): (a) Noisy images, (b) From VTV-model, (c) From MVTV-model, (d) Final λ for MVTV-model.

[12] H. Fu, M. K. Ng, M. Nikolova, and J. L. Barlow. Efficient minimization methods of mixed $\ell 2 - \ell 1$ and $\ell 1 - \ell 1$ norms for image restoration. *SIAM J. Sci. Comput.*, 27:1881–1902, 2006.

[13] N.P. Galatsanos and A.K Ketsaggelos. Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation. *IEEE Trans. Image Process.*, 1:322–336, 1992.

[14] G. Gilboa, N. Sochen, and Y.Y. Zeevi. Texture preserving variational denoising using an adaptive fidelity term. In *in Proceeding of the IEEE Workshop on Variational, Geometric and Level Set Methods in Computer Vision*, pages 137–144, Nice, France, 2003.

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

[16] M. Hintermüller and M. Rincon-Camacho. L^1-TV model with spatially adapted regularization parameter. Preprint of the Department of Mathematics and Scientific Computing of the University of Graz, Austria.

[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] A. Mood. *Introduction to the Theory of Statistics*. McGraw Hill, 1974.

[19] J. Moreau. Proximité et dualité dans un espace hilbertien. *Bulletin de la Société Mathématique de France*, 93:273–299, 1965.

[20] D. Mumford and J. Shah. Optimal approximation by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.*, 17:577–685, 1989.

[21] M. Nikolova. Minimizers of cost-functions involving nonsmooth data-fidelity terms. Application to the processing of outliers. *SIAM Journal on Numerical Analysis*, 40:965–994, 2002.

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

[23] 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.

[24] A. Papoulis. *Probability, Random Variables, Stochastic Processes*. McGraw Hill, 1991.
[25] L.I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
[26] D. Strong, J.-F. Aujol, and T. Chan. Scale recognition, regularization parameter selection, and Meyer's G norm in total variation regularization. Technical report, UCLA, 2005.
[27] D. Strong and T. Chan. Spatially and scale adaptive total variation based regularization and anisotropic diffusion in image processing. Technical report, UCLA, 1996.
[28] E. Tadmor, S. Nezzar, and L. Vese. A multiscale image representation using hierarchical (BV, L^2) decompositions. *Multiscale Model. Simul.*, 2:554–579, 2004.
[29] E. Tadmor, S. Nezzar, and L. Vese. Multiscale hierarchical decomposition of images with applications to deblurring, denoising and segmentation. *Comm. Math. Sci.*, 6:1–26, 2008.
[30] A. Tikhonov and V. Arsenin. *Solutions of Ill-Posed Problems*. Winston and Sons, Washington, D.C, 1977.
[31] C.R. Vogel. *Computational Methods for Inverse Problems*. Frontiers Appl. Math. 23, SIAM Philadelphia, 2002.
[32] W. Yin, D. Goldfarb, and S. Osher. The total variation regularized L^1 model for multiscale decomposition. *SIAM Journal on Multiscale Modeling and Simulation*, 6:190–211, 2006.
[33] K. Yosida. *Functional analysis*. Springer Verlag, 1964.

* Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, A-8010 Graz, Austria.
E-mail address: yiqiu.dong@uni-graz.at

† Department of Mathematics, Humboldt-University of Berlin, Unter den Linden 6, 10099 Berlin, Germany.
E-mail address: hint@math.hu-berlin.de

° Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, A-8010 Graz, Austria.
E-mail address: maria.rincon-camacho@uni-graz.at