SIAM J. IMAGING SCIENCES
Vol. 7, No. 4, pp. 2258–2300

On Infimal Convolution of TV-Type Functionals and Applications to Video and Image Reconstruction*

Martin Holler† and Karl Kunisch‡

Abstract. The infimal convolution of total (generalized) variation–type functionals and its application as regularization for video and image reconstruction is considered. The definition of this particular type of regularization functional is motivated by the need of suitably combining spatial and temporal regularity requirements for video processing. The proposed functional is defined in an infinite dimensional setting, and important analytical properties are established. As applications, the reconstruction of compressed video data and of noisy still images is considered. The resulting problem settings are posed in function space, and suitable numerical solution schemes are established. Experiments confirm a significant improvement compared to standard total variation–type methods, which originates from the introduction of spatio-temporal and spatial anisotropies.

Key words. video reconstruction, infimal convolution, total generalized variation, image reconstruction, temporal regularization, line enhancement

AMS subject classifications. 94A08, 49M29, 65F22

DOI. 10.1137/130948793

1. Introduction.

Motivated by the aim of defining a suitable spatio-temporal regularization for image sequences we consider the infimal convolution of an arbitrary number of modified total (generalized) variation functionals. We discuss analytical properties and propose a numerical realization for two concrete applications.

The combination of different regularization functionals by infimal convolution was mentioned early in [15] and is discussed in [31] in a more general, discrete setting, both motivated by combining first and higher order derivatives in order to reduce staircasing effects for still images. Given two functionals J_1 and J_2, their infimal convolution is defined by

$$(J_1 \Delta J_2)(u) = \inf_v (J_1(u - v) + J_2(v)).$$

In some settings, e.g., when both functionals are positive one-homogeneous and their infimal convolution is convex and lower semicontinuous, it is the convex lower semicontinuous envelope of the nonconvex function

$$u \mapsto \min\{J_1(u), J_2(u)\}.$$

*Received by the editors December 11, 2013; accepted for publication (in revised form) July 31, 2014; published electronically November 18, 2014. This work was supported by the Austrian Science Fund (FWF) special research grant SFB-F32, *Mathematical Optimization and Applications in Biomedical Sciences*.
http://www.siam.org/journals/siims/7-4/94879.html

†Institute of Mathematics and Scientific Computing, University of Graz, Graz A-8010, Austria (martin.holler@uni-graz.at).

‡Institute of Mathematics and Scientific Computing, University of Graz, Graz A-8010, Austria, and Radon Institute, Austrian Academy of Sciences, Wien 1030, Austria (karl.kunisch@uni-graz.at).

As a regularization functional, the latter favors reconstructions having either low J_1 or low J_2 value, while the former has the advantage that, besides being convex, it balances the contribution of J_1 and J_2 locally. More specifically, it provides an additive decomposition of u where one part fits the model enforced by J_1 and the other part fits the model enforced by J_2.

This property will be useful when applying the infimal convolution of total variation–type functionals for image sequence regularization, where we consider an image sequence to be a function defined on a three dimensional space-time domain. In such a setting, it allows us to suitably combine spatial and temporal regularity constraints. A combination of two or more total variation–type functionals can also be used advantageously in other situations beyond image sequences. As a second application we consider the improvement of the reconstruction of certain line structures in the still image setting.

Space-time regularization for videos is quite an open topic in mathematical image processing. The application of infimal convolution of total variation–type functionals in this context is discussed in section 4, where we also give a short overview on the available literature. An application to the still image setting is then addressed in section 5. Here the literature has already evolved further, as will be discussed in the same section.

In order to motivate the particular type of functional that we consider in this paper, let us discuss some aspects of regularization for image sequences: Applying regularization means enforcing a particular model on the obtained reconstruction. The model should be rich enough to cover a broad class of realistic applications, but also simple enough for a practical implementation. The *total generalized variation* (TGV) functional [11] has been proven to satisfy these demands and to allow a good reconstruction quality for still images. The underlying assumption of TGV regularization, namely piecewise smoothness, is a reasonable approximation for many realistic images. Extending TGV regularization to image sequences in a straightforward way, by applying it on functions defined on a three dimensional domain, we can expect to obtain a good reconstruction quality for image sequences that are piecewise smooth in *space and time*.

The question of whether such an approximation again allows a good visual reconstruction quality for a broad enough class of realistic image sequences remains open. Indeed, the perception of brightness variations in time is different from the perception of brightness variations in space. The human eye is, for example, very sensitive to brightness variations in time on a stable background. This indicates that what might be a reasonable approximation for still images is not necessarily a good approximation for videos.

In addition, the extension of TGV regularization to space time forces us to fix the scale of time with respect to space. Fixing this scale means deciding how much emphasis is put on regularity in space compared to regularity in time.

For simplicity, consider the *total variation* (TV) functional and assume that a given image sequence of k frames of size $n \times n$ is modeled as the discretization of a function defined on the space-time domain $(0,n)^2 \times (0,m)$. This means choosing a spatial stepsize of 1 and a temporal stepsize of m/k. Note that, even knowing that one frame step relates to a physical timestep of t seconds, we have no information on how to choose the ratio m/k. This ratio, however, weights the norm of the discretized gradient. With $\kappa = k/m$ and δ_x, δ_y, and δ_t being finite difference operators with a stepsize of 1 in the spatial and temporal directions, respectively,

we get for the discrete spatio-temporal gradient

$$|\nabla u(z)| = \sqrt{(\delta_x u(z))^2 + (\delta_y u(z))^2 + \kappa^2 (\delta_t u(z))^2}.$$

Using TV regularization, the weight κ thus defines how strongly temporal variations are punished.

By applying the infimal convolution of total variation–type functionals (ICTV), we can utilize the additional freedom of the parameter κ to benefit from the additional information of space-time correspondence in videos. We make the (arbitrary) choice that one pixel step corresponds to one frame step but fix a parameter $1 < \kappa$. Defining the norms

$$|x|_{\beta_1} = \sqrt{\kappa^2(x_1^2 + x_2^2) + x_3^2}, \quad |x|_{\beta_2} = \sqrt{x_1^2 + x_2^2 + \kappa^2 x_3^2},$$

we can use the functional

$$u \mapsto \min_v \left(\| |\nabla (u - v)|_{\beta_1} \|_1 + \| |\nabla v|_{\beta_2} \|_1 \right)$$

for spatio-temporal regularization. This means regularizing with TV by either using a large or a small space-time ratio. Using this concept, the additional temporal information of time correspondence of frames allows us to relax the piecewise smoothness assumption in *space and time* by assuming that the image sequence is locally piecewise smooth in *space or time*. In other words, the function $u \mapsto \| |\nabla u|_{\beta_1} \|_1$ enforces piecewise regularity in space and time, but it allows more model deviation in time than in space. The mapping $u \mapsto \| |\nabla u|_{\beta_2} \|_1$ acts the other way around by allowing more deviation in space than in time. The combination of the two functionals via infimal convolution locally emphasizes one of these assumptions. In practice, we can thus hope to get a good reconstruction of videos not only in areas that are piecewise smooth in space and time, but also in the two cases of either rapidly moving objects that are piecewise smooth in space or textured background regions, both of which do not satisfy the assumption of piecewise smoothness in space and time simultaneously.

Mathematically, the infimal convolution of a number of n TV functionals using norms $|\cdot|_{\beta_1}, \dots, |\cdot|_{\beta_n}$ on the gradient can be rigorously defined, for $u \in L^1_{\mathrm{loc}}(\Omega)$, by

$$\text{(1)} \quad \mathrm{ICTV}^n_\beta(u) = \sup \Bigg\{ \int_\Omega u \operatorname{div} p \,|\, p \in \mathcal{C}^1_c(\Omega, \mathbb{R}^d), \text{ such that } \operatorname{div} p = \operatorname{div} q_i, \\ \text{with } q_i \in \mathcal{C}^1_c(\Omega, \mathbb{R}^d), \|q_i\|_{\infty, \beta_i^*} \le 1, 1 \le i \le n \Bigg\},$$

where $\|q\|_{\infty,\beta_i^*} = \operatorname{ess\,sup}_{r\in\Omega} |q(r)|_{\beta_i^*}$ and $|\eta|_{\beta_i^*} := \sup_{\xi \in \mathbb{R}^d} (\eta, \xi) - \mathcal{I}_{\{|\xi|_{\beta_i} \le 1\}}(\xi)$ are the dual norms. Since we are actually interested in using the infimal convolution of TGV-type functionals of arbitrary order, this definition will appear in a more general setting later on.

We show in this paper that the infimal convolution of TGV functionals with arbitrary norms provides an analytically well-justified, flexible, convex regularization approach that allows applications not only for video but also for still image reconstruction. After stating some preliminary definitions in section 2, we will define the ICTV functional and derive interesting properties in section 3, and afterwards discuss two concrete applications in sections 4 and 5.

2. Preliminaries. For the sake of a streamlined presentation of the analysis of later sections, it will be convenient to summarize the necessary notation. For more detailed information on the presented concepts we refer the reader to the preliminaries section of [10].

Throughout this work, let $\Omega \subset \mathbb{R}^d$, $d \geq 1$, be a bounded Lipschitz domain. For any $p \in [1, \infty]$ we set $p' = p/(p-1)$ for $p > 1$ and $p' = \infty$ for $p = 1$.

Definition 2.1. *By* $\mathrm{Sym}^k(\mathbb{R}^d)$ *we denote the space of symmetric* k*-tensors, i.e., the space of multilinear mappings* $\xi : \mathbb{R}^d \times \cdots \times \mathbb{R}^d \to \mathbb{R}$ *such that* $\xi(x_1, \ldots, x_k) = \xi(x_{\pi(1)}, \ldots, x_{\pi(k)})$ *for any permutation* $\pi : \{1, \ldots, k\} \to \{1, \ldots, k\}$*. The standard inner product and norm on* $\mathrm{Sym}^k(\mathbb{R}^d)$ *are defined as*

$$(\xi, \xi) = |\xi|^2 = \sum_{p \in \{1,\ldots,d\}^k} \xi(e_{p_1}, \ldots, e_{p_k})^2,$$

where $e_i \in \mathbb{R}^d$ *denotes the* i*th standard basis vector.*

Mappings from Ω to $\mathrm{Sym}^k(\mathbb{R}^d)$ are called tensor fields, and the spaces

$$\mathcal{C}^n(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)),\ \mathcal{C}_c^\infty(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)),\ L^p(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)),$$

for $p \in [1, \infty]$, are defined in the usual way. The space of Radon measures $\mathcal{M}(\Omega, \mathrm{Sym}^k(\mathbb{R}^d))$ is defined by duality as

$$\mathcal{M}(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)) = \mathcal{C}_0(\Omega, \mathrm{Sym}^k(\mathbb{R}^d))^*,$$

where $\mathcal{C}_0(\Omega, \mathrm{Sym}^k(\mathbb{R}^d))$ denotes the closure of $\mathcal{C}_c(\Omega, \mathrm{Sym}^k(\mathbb{R}^d))$ with respect to the sup-norm. The norm $\|\cdot\|_{\mathcal{M}}$ on $\mathcal{M}(\Omega, \mathrm{Sym}^k(\mathbb{R}^d))$ is defined as the dual norm.

Definition 2.2. *For* $u \in \mathcal{C}^1(\Omega, \mathrm{Sym}^k(\mathbb{R}^d))$*, the derivative* $\nabla \otimes u$ *and divergence* $\operatorname{div} u$ *are again tensor fields defined by*

$$(\nabla \otimes u)(x)(x_1, \ldots, x_{k+1}) := \mathrm{D}u(x)(x_1)(x_2, \ldots, x_{k+1})$$

and

$$\operatorname{div} u(x)(x_1, \ldots, x_{k-1}) = \operatorname{tr}(\nabla \otimes u) := \sum_{i=1}^{d} (\nabla \otimes u)(x)(e_i, x_1, \ldots, x_{k-1}, e_i),$$

where $\mathrm{D}u$ *denotes the standard Fréchet derivative. For distributions* $u \in \mathcal{D}(\Omega, \mathrm{Sym}^k(\mathbb{R}^d))$*, the weak symmetrized derivative* $\mathcal{E}u$ *is defined in the weak sense by*

$$\langle \mathcal{E}u, \varphi \rangle = -\langle u, \operatorname{div} \varphi \rangle, \quad \varphi \in \mathcal{C}_c^\infty(\Omega, \mathrm{Sym}^{k+1}(\mathbb{R}^d)).$$

Note that the derivative of a symmetric tensor field is in general not symmetric, but the symmetrized derivative and the divergence are.

Since norms on finite dimensional vector spaces will play an important role later on, we introduce the following notation:

$\mathbb{S}$: To denote different norms on $\mathbb{R}^d$, we define $\mathbb{S}$ to be a set of symbols such that, for $\gamma \in \mathbb{S}$, $|\cdot|_\gamma$ denotes an arbitrary but fixed norm on $\mathbb{R}^d$.

$|\cdot|_\beta$: Given $n \in \mathbb{N}$, we use parameter matrices $\beta \in \mathbb{S}^{k_1} \times \cdots \times \mathbb{S}^{k_n}$ with orders $k_i \in \mathbb{N}$, $i = 1, \dots, n$, to denote a set of norms on symmetric tensor spaces, i.e.,

$$|\cdot|_{\beta_{i,j}} \text{ denotes a norm on } \operatorname{Sym}^{k_j - i}(\mathbb{R}^d)$$

for $i = 0, \dots, k_j - 1$, $j = 1, \dots, n$.

$|\cdot|_{\beta^*}$: Given $n \in \mathbb{N}$ and a norm parameter matrix $\beta \in \mathbb{S}^{k_1} \times \cdots \times \mathbb{S}^{k_n}$, we denote by β^* the norm parameter matrix corresponding to the dual norms, i.e.,

$$|\eta|_{\beta^*_{i,j}} = \sup_{\xi \in \operatorname{Sym}^{k_j - i}(\mathbb{R}^d)} (\eta, \xi) - \mathcal{I}_{\{\nu \,:\, |\nu|_{\beta_{i,j}} \le 1\}}(\xi)$$

for $\eta \in \operatorname{Sym}^{k_j - i}(\mathbb{R}^d)$, $i = 0, \dots, k_j - 1$, $j = 1, \dots, n$.

$\|\cdot\|_{p,\gamma}$: For a given norm parameter $\gamma \in \mathbb{S}$, $\|\cdot\|_{p,\gamma}$ denotes the p norm of a symmetric tensor field, i.e.,

$$\|\phi\|_{p,\gamma} = \| |\phi|_\gamma \|_p.$$

$\operatorname{md}(\beta)$: This denotes the maximal vector dimension of a parameter matrix β, i.e.,

$$\operatorname{md}(\beta) = \max\{k_i \,|\, i = 1, \dots, n\} \quad \text{if } \beta \in \mathbb{S}^{k_1} \times \cdots \times \mathbb{S}^{k_n}.$$

(n, β): By a *parameter tuple* (n, β) we denote a natural number n together with a norm parameter matrix $\beta \in \mathbb{S}^{k_1} \times \cdots \times \mathbb{S}^{k_n}$. With such a tuple, also the vector dimensions $k_1, \dots, k_n$, the maximal vector dimension $\operatorname{md}(\beta)$, and the dual norm parameter matrix β^* are implicitly introduced.

We point out that the purpose of the norm parameter matrices is to fix and name different, arbitrary norms finite dimensional vector spaces. The notation is, however, motivated by our applications later on, where we are in particular interested in weighted Euclidean norms, e.g.,

$$|\xi|_\gamma = \sqrt{\xi_1^2 + \xi_2^2 + \gamma \xi_3^2},$$

where $\xi \in \mathbb{R}^3$ and $\gamma > 0$ is a positive parameter.

Finally, we recall the space of functions of bounded variation and related functionals.

Definition 2.3. *The space* $\operatorname{BV}(\Omega)$ *is defined as the set of* $L^1(\Omega)$ *functions* u *such that*

$$\operatorname{TV}(u) = \sup \left\{ \int_\Omega u \operatorname{div} \varphi \, \mathrm{d}x \,|\, \varphi \in C_c^1(\Omega, \mathbb{R}^d),\ \|\varphi\|_\infty \le 1 \right\} < \infty.$$

A norm on $\operatorname{BV}(\Omega)$ *is given by* $\|u\|_{\operatorname{BV}} = \|u\|_{L^1} + \operatorname{TV}(u)$.

Remember that $u \in L^1(\Omega)$ is contained in $\operatorname{BV}(\Omega)$ if and only if its weak derivative, denoted by $\mathrm{D}u$, can be identified with a finite Radon measure, i.e., $\mathrm{D}u \in \mathcal{M}(\Omega, \mathbb{R}^d)$; see [1].

A generalization of the TV functional is the *total generalized variation* (TGV) functional, introduced in [11]. We slightly extend its definition by allowing arbitrary norms on the test functions.

Definition 2.4. *Denote with* $\beta \in \mathbb{S}^k$ *parameters for* k *arbitrary norms on* $\mathbb{R}^d$. *The* TGV

functional of order k is then defined as

$$\text{(2)}\quad \mathrm{TGV}^k_\beta(u) = \sup\left\{ \int_\Omega u \cdot \operatorname{div}^k \xi \,\mathrm{d}x \,\middle|\, \xi \in C^k_c(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)), \right.$$
$$\left. \|\operatorname{div}^i \xi\|_{\infty,\beta_i} \le 1,\ i = 0,\dots,k-1 \right\}.$$

Note that with $|\cdot|_{\beta_i} = \alpha_i^{-1}|\cdot|$ the definition of $\mathrm{TGV}^{\mathrm{k}}_\alpha$ coincides with that of [11]. In [10] it has been shown that TGV^k_β, with $|\cdot|_{\beta_i} = \alpha_i^{-1}|\cdot|$, can equivalently be written as

$$\mathrm{TGV}^k_\beta(u) = \min_{\substack{v_i \in \mathrm{BD}(\Omega,\mathrm{Sym}^i(\mathbb{R}^d)),\\ i=1,\dots,k,\\ v_k=0}} \alpha_{k-1}\|\mathrm{D}u - v_1\|_{\mathcal{M}} + \sum_{i=2}^{k} \alpha_{k-i}\|\mathcal{E}(v_{i-1}) - v_i\|_{\mathcal{M}},$$

where $\mathrm{BD}(\Omega, \mathrm{Sym}^i(\mathbb{R}^d))$ denotes the space of tensor fields of bounded deformation (see [10]). Also, the norm $\|u\|_{\mathrm{BGV}} = \|u\|_{L^1} + \mathrm{TGV}^k_\beta(u)$ is an equivalent norm on $\mathrm{BV}(\Omega)$.

3. The ICTV functional. With these preparations we can define the ICTV functional. Let (n,β) with $\beta \in \mathbb{S}^{k_1} \times \cdots \times \mathbb{S}^{k_n}$ be a parameter tuple. For $u \in L^1_{\mathrm{loc}}(\Omega)$, we define the extended real valued functional

$$\text{(3)}\quad \mathrm{ICTV}^n_\beta(u) = \sup\left\{ \int_\Omega u\phi \,\middle|\, \phi = \operatorname{div}^{k_i} q_i, \text{ with } q_i \in \mathcal{C}^{k_i}_c(\Omega, \mathrm{Sym}^{k_i}(\mathbb{R}^d)), \right.$$
$$\left. \|\operatorname{div}^l q_i\|_{\infty,\beta^*_{l,i}} \le 1,\ l = 0,\dots,k_i-1,\ i = 1,\dots,n \right\}.$$

Note that ICTV^n_β can equivalently be written as

$$\mathrm{ICTV}^n_\beta(u) = \sup_{\phi \in \bigcap_{i=1}^n U_i} \int_\Omega u\phi = \sup_{\phi \in C_c(\Omega)} \left(\int_\Omega u\phi - \sum_{i=1}^n \mathcal{I}_{U_i}(\phi) \right),$$

where

$$U_i = \{\operatorname{div}^{k_i} p \,|\, p \in C^{k_i}_c(\Omega, \mathrm{Sym}^{k_i}(\mathbb{R}^d)), \|\operatorname{div}^l p\|_{\infty,\beta^*_{l,i}} \le 1,\ l = 0,\dots,k_i-1\}.$$

By applying Fenchel duality formally one gets that

$$\text{(4)}\qquad \mathrm{ICTV}^n_\beta(u) = \inf_{\substack{v_0,\dots,v_n \in \mathrm{BV}(\Omega)\\ v_0 = u, v_n = 0}} \sum_{i=1}^n \mathrm{TGV}^{k_i}_{\beta_{\cdot,i}}(v_{i-1} - v_i).$$

Thus, n determines the number of TGV functionals we want to convolute and k_i, $i = 1,\dots,n$, their orders. In particular, for $n = 2$ and $k_1 = k_2 = 1$ we get formally

$$\mathrm{ICTV}^n_\beta(u) = \inf_{v \in \mathrm{BV}(\Omega)} \|\mathrm{D}(u - v)\|_{\mathcal{M},\beta_1} + \|\mathrm{D}v\|_{\mathcal{M},\beta_2}.$$

After providing some basic properties of ICTV^n_β we will show that this primal formulation is indeed equivalent.

3.1. Properties of ICTV. The following propositions sum up basic properties of the ICTV functional.

Proposition 3.1. *Let (n,β) be a given parameter tuple. Then the* ICTV^n_β *functional enjoys the following properties:*

1. *For any parameter tuple $(\tilde n,\tilde\beta)$ with $\mathrm{md}(\beta)\geq\mathrm{md}(\tilde\beta)$, there exists a constant $c>0$ such that, for all $u\in L^1_{\mathrm{loc}}(\Omega)$,*
$$c\,\mathrm{ICTV}^n_\beta(u)\leq\mathrm{ICTV}^{\tilde n}_{\tilde\beta}(u).$$
In particular, ICTV^n_β *is equivalent to* $\mathrm{ICTV}^{\tilde n}_{\tilde\beta}$ *for any parameter tuple $(\tilde n,\tilde\beta)$ with $\mathrm{md}(\beta)=\mathrm{md}(\tilde\beta)$, i.e., there exist constants $c_1,c_2>0$ such that, for all $u\in L^1_{\mathrm{loc}}(\Omega)$,*
$$c_1\,\mathrm{ICTV}^n_\beta(u)\leq\mathrm{ICTV}^{\tilde n}_{\tilde\beta}(u)\leq c_2\,\mathrm{ICTV}^n_\beta(u). \tag{5}$$
2. ICTV^n_β *is equivalent to* $\mathrm{TGV}^{\tilde k}_{\tilde\beta}$ *with $\tilde k=\mathrm{md}(\beta)$ and any parameter vector $\tilde\beta\in\mathbb{S}^{\tilde k}$.*
3. ICTV^n_β *is proper, convex, and lower semicontinuous with respect to weak L^1 convergence.*
4. ICTV^n_β *is a seminorm on the space* $\mathrm{BV}(\Omega)$.
5. *For $u\in L^1_{\mathrm{loc}}(\Omega)$, we have $\mathrm{ICTV}^n_\beta(u)=0$ if and only if u is a polynomial of degree less than* $\mathrm{md}(\beta)$.
6. *The norm $\|\cdot\|_{\mathrm{ICTV}^n_\beta}=\|\cdot\|_{L^1}+\mathrm{ICTV}^n_\beta(\cdot)$ is equivalent to $\|\cdot\|_{\mathrm{BV}}$ for any $k\in\mathbb{N}$ and any choice of norms β.*

Proof. We start by showing that, for an arbitrary parameter tuple $(\tilde n,\tilde\beta)$ with $\tilde\beta\in\mathbb{S}^{\tilde k_1}\times\cdots\times\mathbb{S}^{\tilde k_n}$ and $\mathrm{md}(\beta)\geq\mathrm{md}(\tilde\beta)$, there exists a constant c such that, for all $u\in L^1_{\mathrm{loc}}(\Omega)$,

$$c\,\mathrm{ICTV}^n_\beta(u)\leq\mathrm{ICTV}^n_{\tilde\beta}(u). \tag{6}$$

Fix $j\in\mathbb{N}$ such that $k_j=\mathrm{md}(\beta)$. By equivalence of norms, there exist constants $c_{l,i}>0$ such that
$$c_{l,i}|\eta|_{\tilde\beta^*_{l,i}}\leq|\eta|_{\beta^*_{l,j}}\quad\text{for all }\eta\in\mathrm{Sym}^{\tilde k_i-l}(\mathbb{R}^d),$$
and $l=0,\dots,\tilde k_i-1$, $i=1,\dots,\tilde n$. Define $c=\min_{l,i}\{c_{l,i}\}$. For any norm parameter matrix $\gamma\in\mathbb{S}^{l_1}\times\cdots\times\mathbb{S}^{l_m}$, introduce the set

$$U_{\gamma,i}(\Omega)=\{\mathrm{div}^{l_i}p\,:\,p\in\mathcal{C}^{l_i}_c(\Omega,\mathrm{Sym}^{l_i}(\mathbb{R}^d))\,:\,\|\mathrm{div}^l p\|_{\infty,\gamma^*_{l,i}}\leq 1,\,l=0,\dots,l_i-1\}.$$

Recall that

$$\mathrm{ICTV}^n_\gamma(u)=\sup_{\phi\in\bigcap_{i=1}^n U_{\gamma,i}(\Omega)}\int_\Omega u\,\phi \tag{7}$$

for any $u\in L^1_{\mathrm{loc}}(\Omega)$. Now take any $\phi\in\bigcap_{i=1}^n U_{\beta,i}$. Hence $\phi=\mathrm{div}^{k_j}p\in U_{\beta,j}(\Omega)$. Then, $c\phi\in U_{\tilde\beta,i}$ since $c\phi=\mathrm{div}^{\tilde k_i}(\mathrm{div}^{k_j-\tilde k_i}cp)$ and

$$|\mathrm{div}^l(\mathrm{div}^{k_j-\tilde k_i}cp)|_{\tilde\beta^*_{l,i}}\leq|\mathrm{div}^l(\mathrm{div}^{k_j-\mathrm{div}\,\tilde k_i})p|_{\beta^*_{l,j}}\leq 1$$

for any $0 \le l < k_i$. Hence, for any $u \in L^1_{\mathrm{loc}}(\Omega)$,

$$c\,\mathrm{ICTV}^n_\beta(u) = \sup_{\phi \in \bigcap_{i=1}^n U_{\beta,i}(\Omega)} \int_\Omega u(c\phi) \le \sup_{\tilde\phi \in \bigcap_{i=1}^n U_{\tilde\beta,i}(\Omega)} \int_\Omega u(\tilde\phi) = \mathrm{ICTV}^n_{\tilde\beta}(u)$$

and (6) holds.

The equivalence of ICTV^n_β to $\mathrm{ICTV}^{\tilde n}_{\tilde\beta}$ with $\mathrm{md}(\tilde\beta) = \mathrm{md}(\beta)$ follows by interchanging the roles of β and $\tilde\beta$. Also the equivalence to $\mathrm{TGV}^{\tilde k}_{\tilde\beta}$ with $\tilde k = \mathrm{md}(\beta)$ and $\tilde\beta \in \mathbb{S}^{\tilde k}$ is immediate once we notice that $\mathrm{TGV}^{\tilde k}_{\tilde\beta} = \mathrm{ICTV}^1_{\tilde\beta}$.

To verify claim 3, note that by (7), ICTV^n_β is given as pointwise supremum of a family of affine functions $u \mapsto \int_\Omega u \,\mathrm{div}\, p$ that are continuous with respect to weak L^1 topology. This yields convexity and lower semicontinuity of ICTV^n_β. Since ICTV^n_β is obviously proper, claim 3 follows.

It is also easy to see that $\mathrm{ICTV}^n_\beta(u)$ is positive-homogeneous since $\bigcap_{i=1}^n U_{\beta,i}$ is balanced and hence satisfies the seminorm properties.

The remaining assertions follow immediately from the equivalence of ICTV^n_β to $\mathrm{TGV}^{\mathrm{k}}_\alpha$ with $k = \mathrm{md}(\beta)$, $\alpha \in \mathbb{R}^k_+$, and the corresponding assertions on $\mathrm{TGV}^{\mathrm{k}}_\alpha$ as in [10]. ∎

Below, for any orthogonal matrix $O \in \mathbb{R}^{d\times d}$ and $\xi \in \mathrm{Sym}^k(\mathbb{R}^d), k \in \mathbb{N}$, we define the right multiplication $\xi O \in \mathrm{Sym}^k(\mathbb{R}^d)$ by

$$(\xi O)(a_1, \dots, a_k) = \xi(Oa_1, \dots, Oa_k).$$

Proposition 3.2. *Let again (n, β) be a given parameter tuple. ICTV^n_β is invariant under rotations that leave all norms $|\cdot|_{\beta^*_{l,i}}$ invariant, i.e., for any orthogonal matrix $O \in \mathbb{R}^{d\times d}$ satisfying $|\xi O^T|_{\beta^*_{l,i}} = |\xi|_{\beta^*_{l,i}}$ for all $\xi \in \mathrm{Sym}^{k_i-l}(\mathbb{R}^d)$ and $l = 0, \dots, k_i - 1$, $i = 1, \dots, n$, it follows that for any $u \in L^1_{\mathrm{loc}}(\Omega)$, $u \circ O \in L^1_{\mathrm{loc}}(O^{\mathrm{T}}\Omega)$ and*

$$\mathrm{ICTV}^n_\beta(u \circ O) = \mathrm{ICTV}^n_\beta(u).$$

Proof. For $i = 1, \dots, n$ we recall the definition of $U_{\beta,i}(\Omega)$ from the proof of the previous proposition. We shall also use $U_{\beta,i}(O^T\Omega)$, where O satisfies the assumptions of the proposition. The assertion will be verified by showing that

$$\sup_{\phi \in \bigcap_{i=1}^n U_{\beta,i}(O^T\Omega)} \int_{O^T\Omega} u(Ox)\phi(x)\,dx = \sup_{\psi \in \bigcap_{i=1}^n U_{\beta,i}(\Omega)} \int_\Omega u(x)\psi(x)\,dx. \tag{8}$$

First, let $\phi \in \bigcap_{i=1}^n U_{\beta,i}(O^T\Omega)$ so that for each $i = 1, \dots, n$ we have $\phi = \mathrm{div}^{k_i} p_i \in U_{\beta,i}(O^T\Omega)$ for some $p_i \in \mathcal{C}^{k_i}_c(O^T\Omega, \mathrm{Sym}^{k_i}(\mathbb{R}^d))$. From (A.4) in [11] it follows that

$$p_i \in \mathcal{C}^{k_i}_c(O^T\Omega, \mathrm{Sym}^{k_i}(\mathbb{R}^d)) \Leftrightarrow \tilde p_i := (p_i \circ O^T)O^T \in \mathcal{C}^{k_i}_c(\Omega, \mathrm{Sym}^{k_i}(\mathbb{R}^d))$$

and

$$\mathrm{div}^l \tilde p_i = ((\mathrm{div}^l p_i) \circ O^T)O^T \text{ for } l = 0, \dots, k_i.$$

Consequently by the assumed invariance of norms

$$\begin{aligned}\|\operatorname{div}^l \tilde{p}_i\|_{\infty,\beta^*_{l,i}} &= \|(\operatorname{div}^l p_i \circ O^T)O^T\|_{\infty,\beta^*_{l,i}} \\ &= \|\operatorname{div}^l p_i \circ O^T\|_{\infty,\beta^*_{l,i}} = \|\operatorname{div}^l p_i\|_{\infty,\beta^*_{l,i}},\end{aligned} \tag{9}$$

where the ∞ norm is taken over Ω in the first three expressions and over $O^T\Omega$ in the last one. It follows that $\operatorname{div}^{k_i} \tilde{p}_i \in U_{\beta,i}(\Omega)$. Also, they coincide since $\phi = \operatorname{div}^{k_i} p$ for all $i = 1, \dots, n$, and hence $\psi := \operatorname{div}^{k_i} \tilde{p}_i \in \bigcap_{i=1}^n U_{\beta,i}(\Omega)$. These arguments can be reversed: For $\psi = \operatorname{div}^{k_i} q_i \in U_{\beta,i}(\Omega)$ with $q_i \in \mathcal{C}_c^{k_i}(\Omega, \operatorname{Sym}^{k_i}(\mathbb{R}^d))$ we have $p_i = (q_i \circ O)O \in \mathcal{C}_c^{k_i}(O^T\Omega, \operatorname{Sym}^{k_i}(\mathbb{R}^d))$ and $\phi = \operatorname{div}^{k_i} p_i \in U_{\beta,i}(O^T\Omega)$.

Finally for each $i = 1, \dots, n$ we find

$$\begin{aligned}\int_{O^T\Omega} u(Ox)\phi(x)\,dx &= \int_{O^T\Omega} u(Ox) \operatorname{div}^{k_i} p_i(x)\,dx = \int_\Omega u(x) \operatorname{div}^{k_i}(p_i(O^T x))\,dx \\ &= \int_\Omega u(x)(\operatorname{div}^{k_i} \tilde{p}_i)(x)\,dx = \int_\Omega u(x)\psi(x)\,dx.\end{aligned}$$

Together with (9) this implies (8). ∎

The assumption on the relationship between O and $|\cdot|_{\beta^*_{l,i}}$ in Proposition 3.2 is satisfied, for instance, if the weighted norm is given by

$$|\xi|_{\beta^*_{l,i}} = \left(\sum_{p \in \{1,\dots,d\}^{k_i - l}} \xi(Be_1, \dots, Be_{p_{k_i - l}})^2 \right)^{\frac{1}{2}},$$

with $B \in \mathbb{R}^{d\times d}$ a positive definite matrix which commutes with O. In fact, since the operation $\xi \to \xi O^T$ is an orthogonal transformation on $\operatorname{Sym}(\mathbb{R}^d)^{k_i - l}$ [11], we have

$$\begin{aligned}|\xi|^2_{\beta^*_{l,i}} &= \textstyle\sum_{p \in \{1,\dots,d\}^{k_i - l}} \xi(Be_1, \dots, Be_{p_{k_i - l}})^2 \\ &= \textstyle\sum_{p \in \{1,\dots,d\}^{k_i - l}} \xi(BO^T e_1, \dots, BO^T e_{p_{k_i - l}})^2 \\ &= \textstyle\sum_{p \in \{1,\dots,d\}^{k_i - l}} \xi(O^T B e_1, \dots, O^T B e_{p_{k_i - l}})^2 = |\xi O^T|^2_{\beta^*_{l,i}}.\end{aligned}$$

This situation will be relevant for our treatment of image sequences, where the temporal and spatial coordinates are uncoupled and the spatial coordinates are subject to rotation.

The next result shows that the dual formulation of ICTV^n_β as in (4) is indeed equivalent and that, for any $u \in \mathrm{BV}(\Omega)$, the infimum in (4) is actually achieved. The proof of this statement will be elaborated in the appendix.

Proposition 3.3. *For $n \in \mathbb{N}$ and $\beta \in \mathbb{S}^{k_1} \times \cdots \times \mathbb{S}^{k_n}$, let k_i and $\beta_{\cdot,i}$, $i \in \{1, \dots, n\}$, be the order and parameter vectors for the $\mathrm{TGV}^{k_i}_{\beta_{\cdot,i}}$ functionals, respectively. Then, for any $u \in L^{d'}(\Omega)$,*

$$\mathrm{ICTV}^n_\beta(u) = \min_{\substack{v_i \in L^{d'}(\Omega),\\ 1 \le i < n,\\ v_0 = u,\, v_n = 0}} \left(\sum_{i=1}^n \mathrm{TGV}^{k_i}_{\beta_{\cdot,i}}(v_{i-1} - v_i) \right). \tag{10}$$

Remark 3.1. *By the embedding* $\mathrm{BV}(\Omega) \subset L^{d'}(\Omega)$ *it follows that for* $u \in L^1_{\mathrm{loc}}(\Omega) \setminus L^{d'}(\Omega)$ *both* ICTV^n_β *and the right-hand side of* (10) *equal infinity. Thus the right-hand side of* (10) *is an equivalent definition for* ICTV^n_β *for any* $u \in L^1_{\mathrm{loc}}(\Omega)$.

Remark 3.2. *Also note that when two norms* $\beta_{\cdot,i}, \beta_{\cdot,j}$ *are chosen equal in the definition of* ICTV^n_β *given in* (3), *then the corresponding* L^∞ *restrictions simplify to one restriction and hence the* ICTV^n_β *functional reduces to* $\mathrm{ICTV}^{n-1}_{\tilde{\beta}}$.

Remark 3.3. *For the sake of illustration let us consider a special case with* $k = 2$, $d = 2$ *and two norms on* $\mathbb{R}^2$ *given by* $|x|_{\beta_i} = (\sum_{j=1}^2 \beta_i^j x_j^2)^{\frac{1}{2}}$, *with* $\beta_i^j > 0$, *for* $i = 1, 2$; $j = 1, 2$. *By Proposition* 3.3 *we have*

$$\mathrm{ICTV}^2_\beta(u) = \min_{v \in BV(\Omega)} \|\mathrm{D}(u-v)\|_{\mathcal{M},\beta_1} + \|\mathrm{D}v\|_{\mathcal{M},\beta_2}. \tag{11}$$

If $\beta_1 = \beta_2$, *then the two restrictions* $\|q_i\|_{\infty,\beta_i^*} \le 1$ *in the definition of* $\mathrm{ICTV}^2_\beta(u)$ *according to* (3) *reduce to just one constraint. Equivalently, since* $\|\mathrm{D}(u-v)\|_{\mathcal{M},\beta_1} + \|\mathrm{D}v\|_{\mathcal{M},\beta_1} \ge \|\mathrm{D}u\|_{\mathcal{M},\beta_1}$, *we observe that the minimum on the right-hand side of* (11) *is attained at* $v = \alpha u$ *for any* $\alpha \in [0, 1]$. *As soon as* $\beta_1 \neq \beta_2$ *the coefficients* β_i^j *influence the contributions along the* x_j*-directions in such a way that the effect of* $|\partial_{x_1} \cdot|_{\mathcal{M}}$ *is emphasized if* β_i^1 *is large relative to* β_i^2. *The contribution of* $\|\mathrm{D}(u-v)\|_{\mathcal{M},\beta_1}$ *is balanced with that of* $\|\mathrm{D}v\|_{\mathcal{M},\beta_2}$. *This balancing takes place locally. Therefore, as a regularization term, the effect of* ICTV^2_β *is different from that of*

$$\alpha\|\mathrm{D}u\|_{\mathcal{M},\beta_1} + (1-\alpha)\|\mathrm{D}u\|_{\mathcal{M},\beta_2} \tag{12}$$

with $\alpha \in [0, 1]$ *since* (12) *acts globally while* (11) *acts locally.*

3.2. ICTV as regularization.

Let $p \in [1, \infty]$. We consider

$$\min_{u \in L^p(\Omega)} j(u) + \mathrm{ICTV}^n_\beta(u), \tag{P}$$

where $j : L^p(\Omega) \to \overline{\mathbb{R}}_+$ is convex. We additionally make the following assumption:

$$\text{(A)} \quad \begin{cases} \text{Any sequence } (u_n)_{n=1}^\infty \text{ in } L^p(\Omega), \text{ for which } \mathrm{ICTV}^n_\beta(u_n) + j(u_n) \text{ is bounded,} \\ \text{admits a subsequence } L^1\text{-weakly converging to some } u \in L^p(\Omega) \\ \text{such that } j(u) \le \liminf_{i\to\infty} j(u_{n_i}). \end{cases}$$

Then we have the following result.

Proposition 3.4. *With* (A) *holding, there exists at least one solution to* (P).

Of course, assumption (A) is constructed in such a way that existence of a solution to (P) follows by properties of the ICTV functional and standard arguments. It will be shown in the subsequent corollaries that this assumption is convenient to cover all generic settings we are interested in.

Proof. For the sake of completeness, we provide a short proof. Take $(u_n)_n$ to be a minimizing sequence for (P). Obviously, $\mathrm{ICTV}^n_\beta + j$ is bounded below by zero. If $\mathrm{ICTV}^n_\beta + j$ does not admit a finite value, any $u \in L^p(\Omega)$ will be a minimizer. In the other case, we can assume that the sequence $\{\mathrm{ICTV}^n_\beta(u_n) + j(u_n)\}$ is bounded. By assumption (A) there exist $\overline{u} \in L^p(\Omega)$

and a subsequence $(u_{n_i})_i$ such that $u_{n_i} \to \overline{u}$ weakly in $L^1(\Omega)$ and $j(\overline{u}) \leq \liminf_{i\to\infty} j(u_{n_i})$. By lower semicontinuity of ICTV^n_β with respect to weak L^1 convergence we conclude that

$$\begin{aligned} j(\overline{u}) + \mathrm{ICTV}^n_\beta(\overline{u}) &\leq \liminf_{i\to\infty} j(u_{n_i}) + \liminf_{i\to\infty} \mathrm{ICTV}^n_\beta(u_{n_i}) \\ &\leq \liminf_{i\to\infty} \left(j(u_{n_i}) + \mathrm{ICTV}^n_\beta(u_{n_i})\right) = \inf_{u\in L^p(\Omega)} j(u) + \mathrm{ICTV}^n_\beta(u). \end{aligned}$$

Hence $\overline{u}$ is a solution to (P). ∎

A first consequence is the following.

Corollary 3.1. *In the problem setting* (P)*, suppose that* $j(u) := \frac{1}{p}\|u - u_0\|_p^p$ *in case* $p < \infty$ *or* $j(u) = \|u - u_0\|_\infty$ *in case* $p = \infty$ *for given* $u_0 \in L^p(\Omega)$*. Then assumption* (A) *holds true and thus there exists a solution to* (P)*.*

Proof. In general, $\|\cdot\|_p^p$ is weakly lower semicontinuous with respect to L^q convergence for any $p \leq q$.

If $p \in (1, \infty)$, assumption (A) holds trivially since any sequence $(u_n)_n$ in $L^p(\Omega)$ for which $(\|u_n - u_0\|_p^p)_n$ is bounded admits a subsequence L^p-weakly converging to some $u \in L^p(\Omega)$.

If $p = 1$, boundedness of $\|u_n\|_1 + \mathrm{ICTV}^n_\beta(u_n)$ yields, by compact embedding of $\mathrm{BV}(\Omega)$ into $L^{d'}(\Omega)$, existence of a subsequence $(u_{n_i})_i$ weakly converging in $L^{d'}(\Omega)$ to some $u \in L^{d'}(\Omega)$ and thus assumption (A) again holds true (replace d' by any $r \in (1, \infty)$ if $d = 1$).

If $p = \infty$, boundedness of $\|u_n\|_\infty$ yields existence of subsequence $(u_{n_i})_i$ weakly-* converging to some $u \in L^\infty(\Omega)$. Since $L^\infty(\Omega) \subset L^1(\Omega) \subset (L^\infty(\Omega))^*$, $(u_{n_i})_i$ converges to u also weakly in $L^1(\Omega)$, and by weak-* lower semicontinuity of $\|\cdot\|_\infty$ [14, *Proposition* 3.13], assumption (A) holds. ∎

Similarly, one can show the following.

Corollary 3.2. *In the problem setting* (P)*, with* $p \in [1, \infty)$*, suppose that* $j(u)$ *is lower semicontinuous and coercive in* $L^p(\Omega)$*, i.e.,* $j(u) \to \infty$ *for* $\|u\|_p \to \infty$*. Then, assumption* (A) *holds true and there exists a solution to* (P)*.*

In fact, the properties of $j(u) = \frac{1}{p}\|u - u_0\|_p^p$ used in the proof of Corollary 3.1 are exactly lower semicontinuity and coercivity in $L^p(\Omega)$, and hence the result can be shown as above. Of course, the result also holds for $p = \infty$ assuming that j is weak-* lower semicontinuous.

For another important choice of j, existence of a solution to (P) follows from Proposition 3.4, even if j itself is not coercive.

Corollary 3.3. *In the problem setting* (P)*, suppose that*

$$j(u) := \frac{1}{q}\|Ku - u_0\|_Y^q \quad \text{for} \quad K \in \mathcal{L}(L^p(\Omega), Y),$$

where $p, q \in [1, \infty)$*,* $p \leq d'$*, and* Y *is a normed space. Then there exists a solution to* (P)*.*

Proof. We partly follow [10, Corollary 4.3]: Define $M = \ker(K) \cap \mathcal{P}^{k-1}(\Omega) \subset L^p(\Omega)$. Since M is a subspace of the finite dimensional space $\mathcal{P}^{k-1}(\Omega)$, we can use Lemma 6.1 from the appendix to obtain that

$$M^\perp_{k,p} := \{u \in L^p(\Omega) \,|\, (u, q) = 0 \text{ for all } q \in M\}$$

is closed as a subspace of $L^p(\Omega)$ and the existence of a continuous linear projection $R_{k,p} : L^p(\Omega) \to M \subset L^p(\Omega)$ with $\ker(R_{k,p}) = M^\perp_{k,p}$.

Since any $u \in L^p(\Omega)$ admits a unique decomposition $u = u - R_{k,p}(u) + R_{k,p}(u)$ and both j and ICTV_β^n are zero on M, it is sufficient to find a minimizer in $M_{k,p}^\perp \subset L^p(\Omega)$. Consider the convex function $F(u) = j(u) + \mathcal{I}_{M_{k,p}^\perp}(u)$. Any sequence $(u_n)_n$ in $L^p(\Omega)$ such that $F(u_n) + \mathrm{ICTV}_\beta^n(u_n)$ is bounded satisfies $R_{k,p}(u_n) = 0$. In order to bound $(u_n)_n$ we decompose it as $u_n - P(u_n) + P(u_n)$, $P : L^p(\Omega) \to \mathcal{P}^{k-1}(\Omega)$ being a projection to the polynomials. As a consequence of a Poincaré-type inequality for $\mathrm{TGV}_\alpha^{\mathrm{k}}$ [10, Proposition 3.11] and equivalence of TGV_α^n to ICTV_β^n as in Proposition 3.1, we get that $\|u_n - P(u_n)\|_{d'} \le C\,\mathrm{ICTV}_\beta^n(u_n)$, with $C > 0$, and thus is bounded. (Note that by embedding, $(u_n)_n$ indeed is in $L^{d'}(\Omega)$.) For $(P(u_n))_n$ we note that K is injective on the finite dimensional space $\mathcal{P}^{k-1}(\Omega) \cap M_{k,p}^\perp$ to get

$$\|P(u_n)\|_{d'} \le C_1 \|K(P(u_n))\|_Y \le C_2 \left(\|Ku_n - u_0\|_Y + \|K\| \|u_n - P(u_n)\|_{d'} + \|u_0\|_Y\right)$$

with $C_1, C_2 > 0$ and hence boundedness of $\|P(u_n)\|_{d'}$.

Since $p \le d'$ and $p < \infty$, we can thus choose a subsequence $(u_{n_i})_i$ weakly converging to u in L^p. By weak-weak continuity of K and weak lower semicontinuity of any norm it follows that $F(u) \le \liminf_{i\to\infty} F(u_{n_i})$. Thus assumption (A) holds true for $F + \mathrm{ICTV}_\beta^n$. Since we already know that any minimizer of $F + \mathrm{ICTV}_\beta^n$ also minimizes $j + \mathrm{ICTV}_\beta^n$, existence of a solution follows. ∎

4. Application to image sequence reconstruction. As a first application we deal with the reconstruction of corrupted image sequences. More specifically, we consider the decompression of videos where each frame has been compressed individually using JPEG compression. This setting serves as a first test situation for the application of ICTV regularization for image sequence reconstruction. It has the particular advantage that there is no parameter necessary to weight between data fidelity and regularization, and thus the comparison of different methods is simpler. Similar applications that are also captured with our framework are image sequence denoising or deblurring.

A future goal could be the application of ICTV regularization to MPEG decompression, which combines framewise JPEG decompression and motion compensation. However, due to the more involved coding and, as a consequence, the more involved data constraints, we consider only framewise JPEG encoding (MJPEG) in this paper, where the focus lies on regularization rather than data modeling.

The mathematical field of image sequence reconstruction is by far not as evolved as still image reconstruction. While there is rich literature on problems such as optical flow computations, publications that propose variational models for image sequence reconstruction are, to the best of the knowledge of the authors, relatively rare.

We refer the reader to [4, 22] for a short overview on image sequence reconstruction. A crucial point is the incorporation of the additional time dimension. Methods for image sequence reconstruction can thus be very well classified by their approach to resolving this issue.

A first approach would of course be to ignore the time correspondence of frames and regularize each frame in space only. This, however, ignores both the fact that time correlation provides important information for spatial regularization of each frame and that the observed visual reconstruction quality is heavily influenced by the transitions between the individual frames. Simple numerical experiments later on will support these claims.

A second, more suitable approach is to explicitly handle time correlation between frames by coupling them via PDEs. Methods using optical flow or transport models follow this direction. Regularization can then be carried out either in space only, since the PDE deals with time correspondence, or, in case of optical flow models, additionally in the direction of the optical flow. We refer the reader to [17, 28, 6, 26] for optical flow based methods that are applied for image restoration. Drawbacks of such methods are that they are typically limited to particular situations and have a nonconvex nature. With standard optical flow based methods, one either has to apply prior optical flow computations which can be error-prone due to noisy data or has to solve the nonconvex problem of obtaining the optical flow and the reconstructed image sequence simultaneously.

Another possibility, which is also considered in the present paper, is to deal with image sequences as functions defined on the space-time domain and apply suitable regularization techniques for those functions. This raises questions concerning scaling of the time dimension with respect to the space dimension (does one pixel step correspond to one timestep?) and how to regularize in time. It was already noted in [4], for example, that piecewise constancy in time is not a well-suited model for realistic image sequences. Thus, TV regularization does not naturally transfer to the situation where an additional time dimension appears. This will also be confirmed by our numerical experiments. In [22], the authors tried to resolve such issues by separating the fore- and background of an image sequence and regularizing the static background.

A fore- and background separation is also the aim of increasingly popular approaches of low-rank and sparse decomposition, where an additive separation of images sequences into a low-rank component and a sparse component is achieved by penalizing the nuclear norm and the L^1 norm of a matrix containing the individual frames as column vectors. The usage of such techniques for regularization of course requires some additional modifications [21].

Using the ICTV functional, we are able to propose a more flexible regularization functional for image sequence reconstruction. In particular, a combination of different norms on the space-time gradient deals with the problem of space-time scaling and allows for regularization that automatically locally adapts to different situations.

Usage of the ICTV functional for image sequence regularization is motivated by two important observations: The first one is that static background regions should be reconstructed as such since flickering effects on static regions are very well visible by the human observer. The second one is that the additional information provided by the temporal correspondence of frames allows one to loosen some regularity assumptions: It is sufficient to focus on piecewise regularity either in space or in time; regularity for the other direction can be relaxed.

As explained in the introduction to this paper, these observations can be incorporated by using the infimal convolution of two TV-type functionals with different norms. Choosing $1 < \kappa$ and defining the norms

$$|x|_{\beta_1} = \sqrt{\kappa^2(x_1^2 + x_2^2) + x_3^2}, \quad |x|_{\beta_2} = \sqrt{x_1^2 + x_2^2 + \kappa^2 x_3^2} \tag{13}$$

for $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, we use the functional

$$u \mapsto \min_v \left(\| |\nabla(u - v)|_{\beta_1} \| + \| |\nabla v|_{\beta_2} \| \right)$$

for regularization. This will separate the image sequence u into two image sequences, $u - v$ and v, one with little illumination change in space and the other one with little illumination change in time. The formulation as infimal convolution balances the contribution from each of the two sequences.

Based on the results of section 3 we can now rigorously state a resulting minimization problem for MJPEG decompression with ICTV regularization in the following subsection.

4.1. The ICTV regularized MJPEG decompression model. The modeling of data fidelity for MJPEG decompression is very similar to that for JPEG decompression. We refer the reader to [8] for a detailed description of data modeling for individual JPEG images and give only a brief explanation in the following: As already mentioned, MJPEG compression means applying JPEG compression to each frame of the image sequence. JPEG compression is a lossy image compression standard whose main part is a block-cosine transformation of the image, followed by a quantization and rounding to integer of each resulting block of coefficients. As a consequence, given a JPEG compressed image file, one can obtain a set of integer coefficients $(d_n)_n$ as quantized values of a block-cosine transform of the image. Knowledge of the quantization table thus allows one to obtain maximal error bounds for each of the coefficients. Denoting now the block-cosine operator by BDCT, we can obtain closed, bounded intervals $(J_n)_n$ from a given, compressed JPEG file, such that each possible source image for the compression process must be contained in the data set U_D, where

$$U_D = \{u \,|\, (\mathrm{BDCT}(u))_n \in J_n \text{ for all } n\}.$$

In an infinite dimensional setting for MJPEG compressed image sequences, given a domain $\Omega = (0, 8k) \times (0, 8l) \times (0, T)$ and $q = 3/2$, we define the operator

$$A : L^q(\Omega) \to \left\{ (\lambda_i(t))_{i\in\mathbb{N}} \,|\, \sup_i \int_0^T |\lambda_i(t)|^q \,\mathrm{d}t < \infty \right\} \quad (Au)_i(t) = \int_{(0,8k)\times(0,8l)} c_i(x)u(x,t)\,\mathrm{d}x,$$

where the $(c_i)_{i\in\mathbb{N}}$ correspond to the blockwise cosine basis as given in [8]. Note that A is indeed well defined and linear. Now given the closed, bounded intervals $(J_n(t))_{n\in\mathbb{N}}$, $t \in (0, T)$, the data set can be defined as

$$U_D = \{u \in L^q(\Omega) \,|\, (Au)_n(t) \in J_n(t) \text{ for all } n \in \mathbb{N} \text{ a.e. } t \in (0, T)\}.$$

Given $1 < \kappa$ and defining the norms $|\cdot|_{\beta_1}$, $|\cdot|_{\beta_2}$ as in (13), we can define the ICTV^2_β functional for $u \in L^q(\Omega)$ as

$$\mathrm{ICTV}^2_\beta(u) = \sup\Bigg\{ \int_\Omega u \,\mathrm{div}\, p \,|\, p \in \mathcal{C}^1_c(\Omega, \mathbb{R}^d), \text{ such that } \mathrm{div}\, p = \mathrm{div}\, q_0 = \mathrm{div}\, q_1, \text{ with } q_0, q_1 \in \mathcal{C}^1_c(\Omega, \mathbb{R}^d), \| |q_0|_{\beta_1^*} \|_\infty \le 1, \| |q_1|_{\beta_2^*} \|_\infty \le 1 \Bigg\}. \tag{14}$$

According to Proposition 3.3, an equivalent definition is given by

$$\mathrm{ICTV}^2_\beta(u) = \min_{v\in L^q(\Omega)} \mathrm{TV}_{\beta_1}(u - v) + \mathrm{TV}_{\beta_2}(v),$$

where TV_{β_1} and TV_{β_2} denote the standard TV functional with the Euclidean norm replaced by $|\cdot|_{\beta_1}$ and $|\cdot|_{\beta_2}$, respectively. The minimization problem for MJPEG decompression is then given as

$$\min_{u\in L^q(\Omega)} \mathrm{ICTV}^2_{\beta}(u) + \mathcal{I}_{U_D}(u), \tag{15}$$

where $\mathcal{I}_{U_D}$ denotes the convex indicator function of the set U_D, i.e., $\mathcal{I}_{U_D}(u) = 0$ if $u \in U_D$ and infinity else. Using the particular definition of the cosine basis and the data intervals, it follows by direct calculation that, for any sequence $(u_n)_n \in U_D$, the integral mean $\overline{u}_n = \int_\Omega u_n$ is bounded. Hence coercivity of (15) follows from the Poincaré-type inequality for TV and equivalence of TV and ICTV^2_{β}, and by applying Proposition 3.4 existence of a solution follows.

Remark 4.1. *We end this subsection with a comment on the choice of the gradient norms* (13) *for the definition of the* ICTV *functional in the context of video regularization. As explained, the different weighting of the coordinates is motivated by the aim of separating areas with little motion and areas with moving objects. However, one could also think of using weighted L^1-type product norms, i.e.,*

$$|x|_{\tilde{\beta}_1} = \sqrt{\kappa^2(x_1^2 + x_2^2)} + |x_3|, \quad |x|_{\tilde{\beta}_2} = \sqrt{x_1^2 + x_2^2} + \kappa|x_3|.$$

In the situation of still image regularization, rotational invariance is a strong argument for using the Euclidean norm on the gradient, but this argument does not apply to our setting since there is no obvious interpretation of rotational invariance in space and time. When applying an L^1-type product norm on the space-time gradient of the image sequence, where the third coordinate contains the time derivative, rotational invariance in space would still hold. A counterargument against this choice of norm is that it leads to a decoupling of space and time derivatives, i.e., the functional gives equal cost whether variations in space and time are correlated or not. But this is not desirable since we expect space-time variations to be correlated, i.e., in areas of one frame with constant brightness values, also a variation in time is less likely to appear in the next frame. Also, in areas of a frame with strong brightness variations, high time variations appear once some movement occurs. Noise, on the other hand, is typically not space and time correlated and should be penalized more strictly, which is an argument for coupling the derivatives.

Coupling the space and time gradient with an L^∞ norm would be another possibility but has the disadvantage that, once a high derivative in one direction occurs, noise in the other direction will not be penalized.

4.2. Numerics of ICTV based MJPEG decompression. In this subsection we discuss the discretization and numerical solution of (15). For discretization we use a finite difference scheme with forward differences. Given the image sequence dimensions $N \times M \times T \in \mathbb{N}^3$, we define the space of discrete images as $U := \mathbb{R}^{N\times M\times T}$ and the space for discrete gradient information as $V := U \times U \times U$. Note that the space dimensions N, M are assumed to be multiples of 8 due to the 8×8-block-cosine transform as part of the JPEG standard.

Norms and inner products on these spaces are given as

$$\|u\|_U^2 = (u,u)_U = \sum_{i,j,k} u_{i,j,k}^2, \quad \|v\|_V^2 = (v,v)_V = \sum_{i,j,k} (v^1_{i,j,k})^2 + (v^2_{i,j,k})^2 + (v^3_{i,j,k})^2$$

for $u \in U$, $v = (v^1, v^2, v^3) \in V$. Using the forward differences

$$(16)\qquad \begin{aligned}(\delta_{x+}u)_{i,j,k} &= \begin{cases}(u_{i+1,j,k} - u_{i,j,k}) & \text{if } \ 0 \le i < N-1,\\ 0 & \text{if } \ i = N-1,\end{cases}\\ (\delta_{y+}u)_{i,j,k} &= \begin{cases}(u_{i,j+1,k} - u_{i,j,k}) & \text{if } \ 0 \le j < M-1,\\ 0 & \text{if } \ j = M-1,\end{cases}\\ (\delta_{t+}u)_{i,j,k} &= \begin{cases}(u_{i,j,k+1} - u_{i,j,k}) & \text{if } \ 0 \le k < T-1,\\ 0 & \text{if } \ k = T-1,\end{cases}\end{aligned}$$

the spatio-temporal gradient ∇ is defined as

$$(\nabla u) = (\delta_{x+}u, \delta_{y+}u, \delta_{t+}u)^{\mathrm{T}}.$$

The discrete, blockwise component operator $C : U \to U$ is defined locally, for each 8×8 block $(z_{i,j})_{0 \le i,j \le 7}$ of each frame, as

$$(17)\qquad (Cz)_{p,q} = c_p c_q \sum_{n,m=0}^{7} z_{n,m} \cos\left(\frac{\pi(2n+1)p}{16}\right) \cos\left(\frac{\pi(2m+1)q}{16}\right)$$

for $0 \le p, q \le 7$ and

$$c_s = \begin{cases}\frac{1}{\sqrt{8}} & \text{if } s = 0,\\ \frac{1}{2} & \text{if } 1 \le s \le 7.\end{cases}$$

For each JPEG compressed frame of the MJPEG compressed image sequence, we can obtain a set of maximal error intervals for the coefficients of its block-cosine transform, and thus we suppose the closed, bounded intervals $(J_{i,j,k})_{0 \le i,j,k < N,M,T}$ to be given. The set of possible source images can then be defined as

$$(18)\qquad U_D = \{u \in U \,|\, (Cu)_{i,j,k} \in J_{i,j,k} \text{ for all } 0 \le i, j, k < N, M, T\}.$$

Setting $|\cdot|_{\beta_1}$, $|\cdot|_{\beta_2}$ to be the norms of (13) with $1 < \kappa$, the discrete version of the ICTV^2_β functional is given as

$$(19)\qquad \mathrm{ICTV}^2_\beta(u) = \min_{v \in U} \| |\nabla(u - v)|_{\beta_1} \|_1 + \| |\nabla(v)|_{\beta_2} \|_1,$$

with $\|\cdot\|_1$ the discrete L^1 norm. Note that, as can be shown by working on the subspace $\ker(\nabla)^\perp$ of the finite dimensional Hilbert space U similarly as in Corollary 3.3, the minimum in the definition of the discrete ICTV functional indeed exists.

In the following, we will use an equivalent reformulation of the ICTV functional with weighted gradients instead of different norms, which turns out to be more convenient for the implementation. We define the weighted discrete gradients

$$(20)\qquad \nabla_1 u = (\kappa(\delta_{x+}u), \kappa(\delta_{y+}u), \delta_{t+}u), \quad \nabla_2 u = (\delta_{x+}u, \delta_{y+}u, \kappa(\delta_{t+}u))$$

such that

$$\mathrm{ICTV}^2_\beta(u) = \min_{v \in U} \|\nabla_1(u - v)\|_1 + \|\nabla_2 v\|_1.$$

We solve the discrete minimization problem by applying the globally convergent primal dual algorithm of [16] to an equivalent saddle point formulation. Equivalence follows from the following proposition, which can be shown by standard arguments from convex analysis.

Proposition 4.1. *With U_D and ∇_1, ∇_2 as in (18) and (20), respectively, and $\mathcal{I}_{U_D}$ the convex indicator function of the set U_D, there exists a solution to the primal problem*

$$\inf_{u,v\in U} \|\nabla_1(u-v)\|_1 + \|\nabla_2 v\|_1 + \mathcal{I}_{U_D}(u), \tag{21}$$

the dual problem

$$\sup_{p,q\in V} -\mathcal{I}_{\|\cdot\|_\infty\le 1}(p) - \mathcal{I}_{\|\cdot\|_\infty\le 1}(q) - \sup_{w_1\in U_D} (-\operatorname{div}_1 p, w_1)_U - \mathcal{I}_{\{0\}}(\operatorname{div}_1 p - \operatorname{div}_2 q), \tag{22}$$

and the saddle point problem

$$\inf_{u,v\in U}\sup_{p,q\in V} \left(\begin{pmatrix} \nabla_1 & -\nabla_1 \\ 0 & \nabla_2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} p \\ q \end{pmatrix} \right)_{V\times V} - \mathcal{I}_{\|\cdot\|_\infty\le 1}(p) - \mathcal{I}_{\|\cdot\|_\infty\le 1}(q) + \mathcal{I}_{U_D}(u). \tag{23}$$

Further, the infimum of the primal and the supremum of the dual problem coincide and the problems are equivalent in the sense that (u,v,p,q) solves (23) if and only if (u,v) solves (21) and (p,q) solves (22).

In the above proposition, the operators $\operatorname{div}_1, \operatorname{div}_2 : V \to U$ are the negative adjoints of the discrete gradient operators ∇_1, ∇_2 and are given by

$$\operatorname{div}_1 p = \kappa(\delta_{x-}p^1 + \delta_{y-}p^2) + \delta_{t-}p^3, \quad \operatorname{div}_2 p = \delta_{x-}p^1 + \delta_{y-}p^2 + \kappa(\delta_{t-}p^3),$$

where

$$(\delta_{x-}z)_{i,j,k} = \begin{cases} -z_{i-1,j,k} & \text{if} \quad i = N-1, \\ (z_{i,j,k} - z_{i-1,j,k}) & \text{if} \quad 0 < i < N-1, \\ z_{i,j,k} & \text{if} \quad i = 0, \end{cases} \tag{24}$$

and similarly for the y,t coordinates. The expression $\mathcal{I}_{\|\cdot\|\le 1}$ is an abbreviation for the convex indicator function of the set $\{q \in V \,|\, \|q\| \le 1\}$.

The implementation of our solution algorithm is described in Algorithm 1. The operants proj_1 and $\operatorname{proj}_{U_D}$ given there denote projections to the sets $\{p \in V \,|\, \|p\|_\infty \le 1\}$ and U_D, respectively. Note that these projections can be evaluated easily by pointwise projections and evaluation of the operators C, C^* for $\operatorname{proj}_{U_D}$.

The stepsizes σ, τ are chosen adaptively, as proposed in [20], i.e., for $\theta, \delta \in (0,1)$,

$$\sigma_{n+1}\tau_{n+1} = S_K(\sigma_n,\tau_n) = \begin{cases} \frac{\delta\|x_n - x_{n-1}\|_{U^2}^2}{\|K(x_n - x_{n-1})\|_{V^2}^2} & \text{if } \theta\sigma_n\tau_n \ge \frac{\|x_n - x_{n-1}\|_{U^2}^2}{\|K(x_n - x_{n-1})\|_{V^2}^2}, \\ \theta\sigma_n\tau_n & \text{if } \sigma_n\tau_n \ge \frac{\|x_n - x_{n-1}\|_{U^2}^2}{\|K(x_n - x_{n-1})\|_{V^2}^2} > \theta\sigma_n\tau_n, \\ \sigma_n\tau_n & \text{if } \sigma_n\tau_n < \frac{\|x_n - x_{n-1}\|_{U^2}^2}{\|K(x_n - x_{n-1})\|_{V^2}^2}, \end{cases} \tag{25}$$

with

$$K = \begin{pmatrix} \nabla_1 & -\nabla_1 \\ 0 & \nabla_2 \end{pmatrix} \tag{26}$$

Algorithm 1. Scheme of implementation for JPEG decompression.

function ICTV-MJPEG(J_{comp})
  $(J_n)_n \leftarrow$ Decoding of MJPEG-Object J_{comp}
  $u \leftarrow u_0$ (Standard decompression)
  $v \leftarrow 0, \overline{u} \leftarrow u, \overline{v} \leftarrow 0, p \leftarrow 0, q \leftarrow 0$
  choose $\sigma, \tau > 0$ arbitrary
  repeat
    $p \leftarrow \text{proj}_1\,(p + \sigma\nabla_1(\overline{u} - \overline{v}))$
    $q \leftarrow \text{proj}_1(q + \sigma\nabla_2\overline{v})$
    $u_+ \leftarrow u - \tau(-\operatorname{div}_1 p)$
    $v_+ \leftarrow v - \tau(\operatorname{div}_1 p - \operatorname{div}_2 q)$
    $u_+ \leftarrow \text{proj}_{U_D}(u_+)$
    $\overline{u} \leftarrow (2u_+ - u)$, $\overline{v} \leftarrow (2v_+ - v)$
    $u \leftarrow u_+$, $v \leftarrow v_+$
    $\sigma\tau \leftarrow S_K(\sigma, \tau)$
  until Stopping criterion fulfilled
  return u_+
end function

and $x_n = (u_n, v_n)$ the primal iterates of the algorithm. As can be seen directly by checking the convergence proof given in [16], for this adaptive stepsize choice convergence of the primal dual algorithm can still be guaranteed even if the standard condition $\sigma\tau\|K\|^2 < 1$ might be violated.

As stopping criterion we can use a primal dual gap.

Proposition 4.2. *Let $(\hat{u}, \hat{v})$ be a solution to (21) and $(x_n, y_n) = ((u_n, v_n), (p_n, q_n))$ be the iterates of Algorithm 1. For $\gamma > 1$, define*

$$\begin{aligned}\mathcal{G}(x_n, y_n) = {} & \|\nabla_1(u_n - v_n)\|_1 + \|\nabla_2 v_n\|_1 \\ & + \sup_{w \in U_D} (w, \operatorname{div}_1 p_n)_U + \gamma\|v_n\|_2 \|\operatorname{div}_1 p_n - \operatorname{div}_2 q_n\|_2.\end{aligned} \tag{27}$$

Then $\mathcal{G}(x_n, y_n)$ converges to zero as $n \to \infty$ and

$$\mathcal{G}(x_n, y_n) \geq \|\nabla_1(u_n - v_n)\|_1 + \|\nabla_2 v_n\|_1 - \text{ICTV}^n_\beta(\hat{u}) \geq 0$$

whenever $\gamma\|v_n\|_2 \geq \|\hat{v}\|_2$.

Proof. First note that since the iterates (x_n, y_n) are known to converge to an optimal solution of the primal and dual problems (21) and (22), respectively, and all quantities in the definition of $\mathcal{G}$ are continuous with respect to the iterates, convergence of $\mathcal{G}(x_n, y_n)$ to zero follows.

Take now $(\hat{p}, \hat{q})$ a solution of the dual problem (22). For convenience we define

$$F(p, q) = \|p\|_1 + \|q\|_1$$

and note that

$$F^*(p, q) = \mathcal{I}_{\|\cdot\|_\infty \leq 1}(p) + \mathcal{I}_{\|\cdot\|_\infty \leq 1}(q).$$

Due to the projections involved in Algorithm 1, $F^*(p_n, q_n) = \mathcal{I}_{U_D}(u_n) = 0$ for all iterates. We can estimate, with K as in (26),

$$\mathcal{G}(x_n, y_n) \geq F(K(x_n)) + \sup_{\substack{w_0, w_1 \in U \\ \|w_1\|_2 \leq \gamma \|v_n\|_2}} ((w_0, w_1), -K^*(p_n, q_n))_{U\times U} + F^*(p_n, q_n) - \mathcal{I}_{U_D}(w_0).$$

Now if $\gamma\|v_n\|_2 \geq \|\hat{v}\|_2$, and since $(\hat{u}, \hat{v}, \hat{p}, \hat{q})$ solves the saddle point problem (23), we get

$$\begin{aligned}
\mathcal{G}(x_n, y_n) &\geq F(K(x_n)) + ((\hat{u}, \hat{v}), -K^*(p_n, q_n))_{U\times U} + F^*(p_n, q_n) - \mathcal{I}_{U_D}(\hat{u}) \\
&\geq F(K(x_n)) - \left[\sup_{p,q \in V} ((\hat{u}, \hat{v}), K^*(p, q))_{U\times U} - F^*(p, q) + \mathcal{I}_{U_D}(\hat{u}) \right] \\
&= F(K(x_n)) - F(K(\hat{x})) \geq 0
\end{aligned}$$

and the assertion follows. ∎

Note that, due to orthogonality of the basis transformation operator, the supremum in the definition of the duality gap (27) can be easily calculated as

$$\begin{aligned}
\sup_{w \in U_D} (w, \operatorname{div}_1 p_n) = &\sum_{\substack{i,j,k \\ (A(\operatorname{div}_1 p_n))_{i,j,j} < 0}} (A(\operatorname{div}_1 p_n))_{i,j,k} l_{i,j,k} \\
&+ \sum_{\substack{i,j,k \\ (A(\operatorname{div}_1 p_n))_{i,j,j} \geq 0}} (A(\operatorname{div}_0 p_n))_{i,j,k} r_{i,j,k},
\end{aligned}$$

where $(l_{i,j,k})$ and $(r_{i,j,k})$ are such that

$$J_{i,j,k} = [l_{i,j,k}, r_{i,j,k}]\,.$$

For the numerical experiments we will use the normalized primal dual gap

$$\tilde{\mathcal{G}}(x_n, y_n) = \mathcal{G}(x_n, y_n)/(NMT) \tag{28}$$

as stopping criterion. This is done to make the primal dual gap independent of the image size and to estimate the average pixelwise error

$$[|\nabla_1(u_n - v_n)| + |\nabla_2 v_n|]_i - [\nabla_1(\hat{u} - \hat{v})| + |\nabla_2 \hat{v}|]_i.$$

4.3. Numerical experiments for MJPEG decompression. In this subsection, numerical experiments for the decompression of MJPEG compressed image sequences are presented. We start with an evaluation of some straightforward methods on a test image sequence. The one that gives the best results will then be compared to ICTV regularization. The following approaches are tested:

$\mathrm{TV}_{\mathrm{st}}$	Total variation regularization in space and time,
$\mathrm{TV}_{\mathrm{fl}}$	Total variation regularization in the direction of a precomputed optical flow,
$\mathrm{TGV}_{\mathrm{fr}}$	Second order total generalized variation regularization in space only,
$\mathrm{TGV}_{\mathrm{st}}$	Second order TGV regularization in space and time.

These functionals are combined with the data fidelity term $\mathcal{I}_{U_D}$, where U_D is given in (18), in order to reconstruct MJPEG compressed videos. While we have already discussed arguments against the first three types of regularization, the qualitative behavior of TGV regularization in space and time is a priori not clear. With TV regularization in space and time we expect flickering of moving objects as a consequence of a temporal staircasing effect. Thus the use of TGV regularization in space and time may give an improvement.

TV regularization along the lines of the optical flow means that we penalize the space gradient in a standard way and use directional time derivatives in the direction of a precomputed optical flow. The optical flow is computed using the method presented in [36], where the implementation for computing the flow has been obtained from the MATLAB package provided in [19].

The discretization for these methods is again done by finite differences and is very similar to the one presented in subsection 4.2. The spatial stepsize is fixed to 1 for all methods. The temporal stepsize for TGV regularization in space and time is chosen to be 2, meaning that illumination changes in space are twice as "cheap" as illumination changes in time. The ratio of the parameters for TGV is fixed to $\alpha_0/\alpha_1 = \sqrt{2}$, as this has delivered good results for still images in previous works [9]. For the optical flow based method we choose a temporal stepsize of 0.1, which has been proposed in [36] for inpainting and restoration and puts much more emphasis on regularity along the optical flow. We tried a large range of different temporal stepsizes and observed that the proposed choices lead to the best results in terms of observed visual image quality.

All four methods were implemented by the authors using the general framework of the primal dual algorithm of [16]. For this first experiment, we use a fixed iteration number of 5000 as stopping criterion. A modified primal dual gap, similar to (27), has been used to ensure a proper implementation and optimality for all implementations.

As it is already very difficult to define a good measure of visual image quality for still images, it seems almost impossible to obtain such a measure for image sequences. We thus visualize quality improvements by showing frames and plotting time graphs for the image sequences. Also, all tested image sequences and obtained results are available as MATLAB data files at one of the authors' webpage [35].

At first, Figure 1 shows the original and standard compressed/decompressed version of frame 21 of a 50-frame test image sequence. The movement of the objects is indicated with blue arrows on the original frame. As one can see, the standard decompressed frame suffers from heavy JPEG artifacts.

Figure 2 then shows the improved decompression of the same test image sequence using the regularizations $\mathrm{TV}_{\mathrm{st}}$, $\mathrm{TV}_{\mathrm{fl}}$, $\mathrm{TGV}_{\mathrm{fr}}$, $\mathrm{TGV}_{\mathrm{st}}$. The left-hand side shows again frame 21 with one pixel line marked in red. For the right-hand side, we have moved this red line with the underlying object, such that it always marks the same region of the object, and plotted the development of the brightness values at this line in time. With a perfect reconstruction, the graphs on the right-hand side should thus be constant in time direction for all 35 shown frames. This allows us to evaluate the flickering of this region, which is an eye-catching artifact when watching the video. It can be observed on the surface plots that the methods $\mathrm{TV}_{\mathrm{st}}$, $\mathrm{TV}_{\mathrm{fl}}$, $\mathrm{TGV}_{\mathrm{fr}}$ suffer heavily from this artifact. In particular the method $\mathrm{TGV}_{\mathrm{fr}}$, which leads to a good reconstruction quality for each individual frame, shows the importance of such surface plots

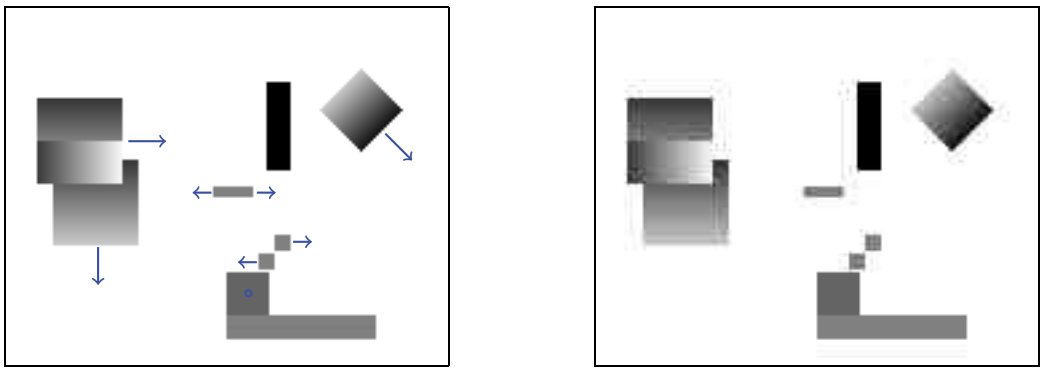

Figure 1. *Frame* 21 *of the original and standard compression/decompression of a test image sequence. The blue arrows indicate the moving direction of the objects, and the object with the circle continuously gets darker.*

to validate reconstruction quality. In contrast to that, $\mathrm{TGV}_{\mathrm{st}}$ yields a similar framewise image quality, but with the surface plot being almost constant in time. This shows an improved visual reconstruction quality for this method. For the TV based reconstructions, one can also observe staircasing in space when looking at the linear part of the surfaces. Motivated by this first experiment, we now use the method $\mathrm{TGV}_{\mathrm{st}}$ for comparison to ICTV regularization.

The next evaluations are carried out on realistic images sequences. As a first example, we use a section of the *juggler* image sequence from the Middlebury optical flow test dataset [25]. Frame 5 of the original and MJPEG compressed image sequence are shown in Figure 3. This 8-frame image sequence contains both fast and complex movements of the balls and the hands of the juggler and is therefore a challenging test scenario. As stopping criterion for the subsequent experiments we require a modified primal dual gap (see (27) for its definition in case of the ICTV functional) to be below the threshold of 0.5. The resulting iteration numbers are given in the figures. We point out that, for the implementation of ICTV regularization, we fixed the ratio between the primal and dual stepsize to 0.2^2 since this accelerated convergence significantly.

Figure 4 shows test results for different regularization approaches. On the left we show frame 5 of the reconstructed image sequences together with a line marked in red, while on the right we plot the temporal development of the brightness values along this line. Since the line is contained entirely in the background during the whole sequence, this graph should ideally again be constant in the temporal direction.

The first line shows the result obtained with $\mathrm{TGV}_{\mathrm{st}}$ regularization. While the visual quality of the individual frame is quite good, the surface plot shows a flickering of the background region in time, which is again an eye-catching artifact when watching the image sequence. The flickering can be explained by low penalization of brightness change in time due to a stepsize of 2 and the originally texture-type background region. This causes the model to focus on spatial regularity rather than temporal constancy.

To resolve this, an obvious solution would be to put more emphasis on time regularity, for example by choosing a timestep of 0.2. As can be observed on the second line of Figure 4, this indeed keeps background regions constant but leads to strong motion artifacts in the

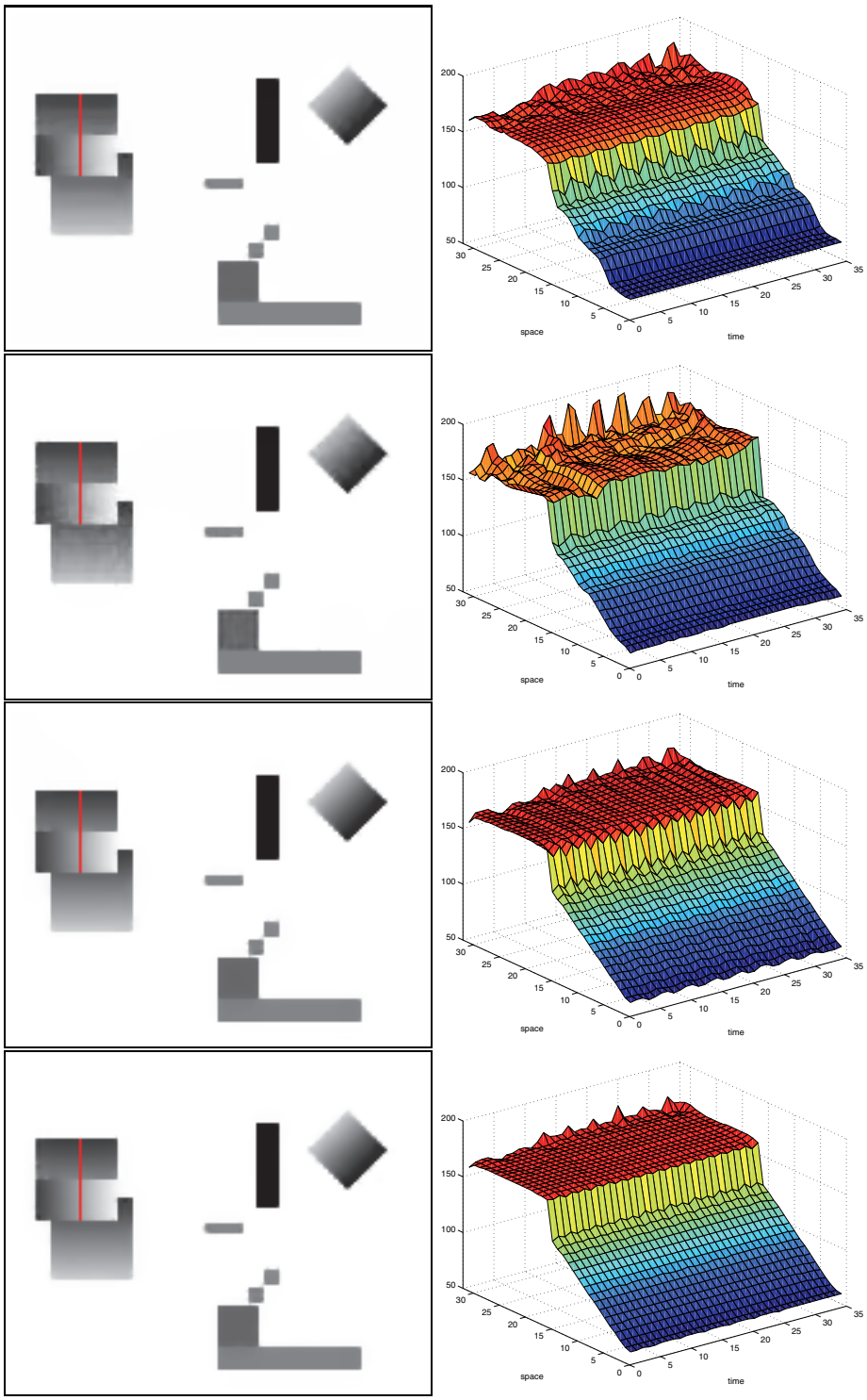

Figure 2. *Comparison of different regularization techniques. From top to bottom: TV in space/time, TV along the optical flow, TGV in space only, TGV in space/time. Left: Frame* 21 *of the compressed test image sequence of Figure* 2 *with marked region of one object. Right: Development of pixel values of this object region in time.*

Figure 3. *Frame 5 of the original and standard compressed/decompressed juggler image sequence.*

frame. All intermediate timestep choices suffer from the same problem of balancing between background flickering and motion artifacts.

In contrast to that, the result with ICTV regularization, as can be seen in the third line of Figure 4, avoids both motion artifacts and background flickering. Drawing on the experience with TGV regularization we choose $\kappa = 5$ in the definition of the norms (13), which corresponds to a space/time step relation of 1/5 and 5 for the two different functionals. We want to point out, however, that, since for simplicity we use only TV functionals for the infimal convolution, some spatial staircasing can be observed in the individual frames. This can be overcome by using the infimal convolution of second order TGV functionals.

Solving the minimization problem for ICTV reconstruction provides not only the reconstructed image sequence but also a separation of this sequence into two components possessing either little spatial or little temporal brightness changes. As can be seen in Figure 5, in case of the juggler image sequence, this yields a separation of the image sequence into the fast moving hands and balls on the one hand and the slowly moving person together with the stable background on the other. The success of this separation depends on the parameter κ and the type of object and movement.

As a last experiment we again compare TGV regularization in space and time for different timesteps with ICTV regularization on the Minicooper sequence, also obtained as a section of a video from the Middlebury optical flow test dataset [25]. In this case, a much larger region of the image sequence is moving. Figure 6 shows the original image sequence and the standard decompression of an MJPEG compression version. Figure 7 then shows the result of using the three different types of regularization, again with a surface plot of a background line marked in red. The behavior is very similar to that in the juggler experiment. TGV in space and time suffers from either motion artifacts or background flickering, depending on the timestep choice, and ICTV regularization is able to reconstruct each frame with fewer motion

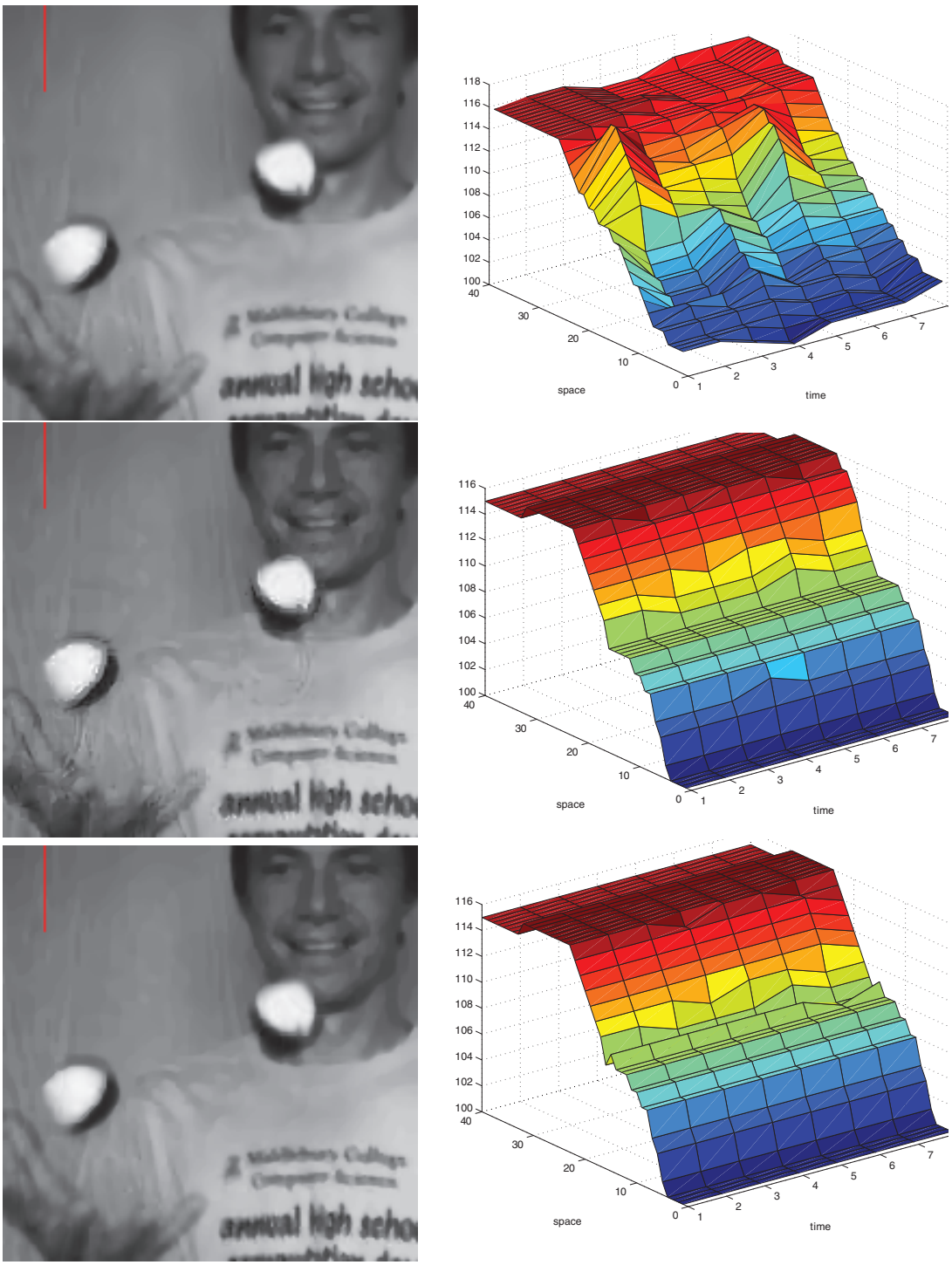

Figure 4. *Comparison of different regularization functionals for the juggler image sequence. From top to bottom:* TGV *regularization in space and time using temporal stepsizes of* 2 *and* 0.2*, and ICTV regularization in space and time. On the left, frame* 5 *of the image sequence is shown with line of the background region marked in red; on the right, the temporal development of this line is plotted. Iteration numbers to reach the stopping criterion, from top to bottom:* 250, 2102, 2546.

Figure 5. *Decomposition of the juggler image sequence, frame* 5.

Figure 6. *Frame* 5 *of the original and standard compressed/decompressed Minicooper image sequence.*

artifacts and keeps the background constant.

5. Application to still image reconstruction. Infimal convolution of TV-type functionals can also be used to introduce anisotropies in still image reconstruction. Combining the standard total (generalized) variation functional, where the Euclidean norm is used to penalize the gradient, with TV-type functionals whose gradient norm unit balls are ellipses, allows a separation of images into isotropic and anisotropic components. Indeed, the anisotropic com-

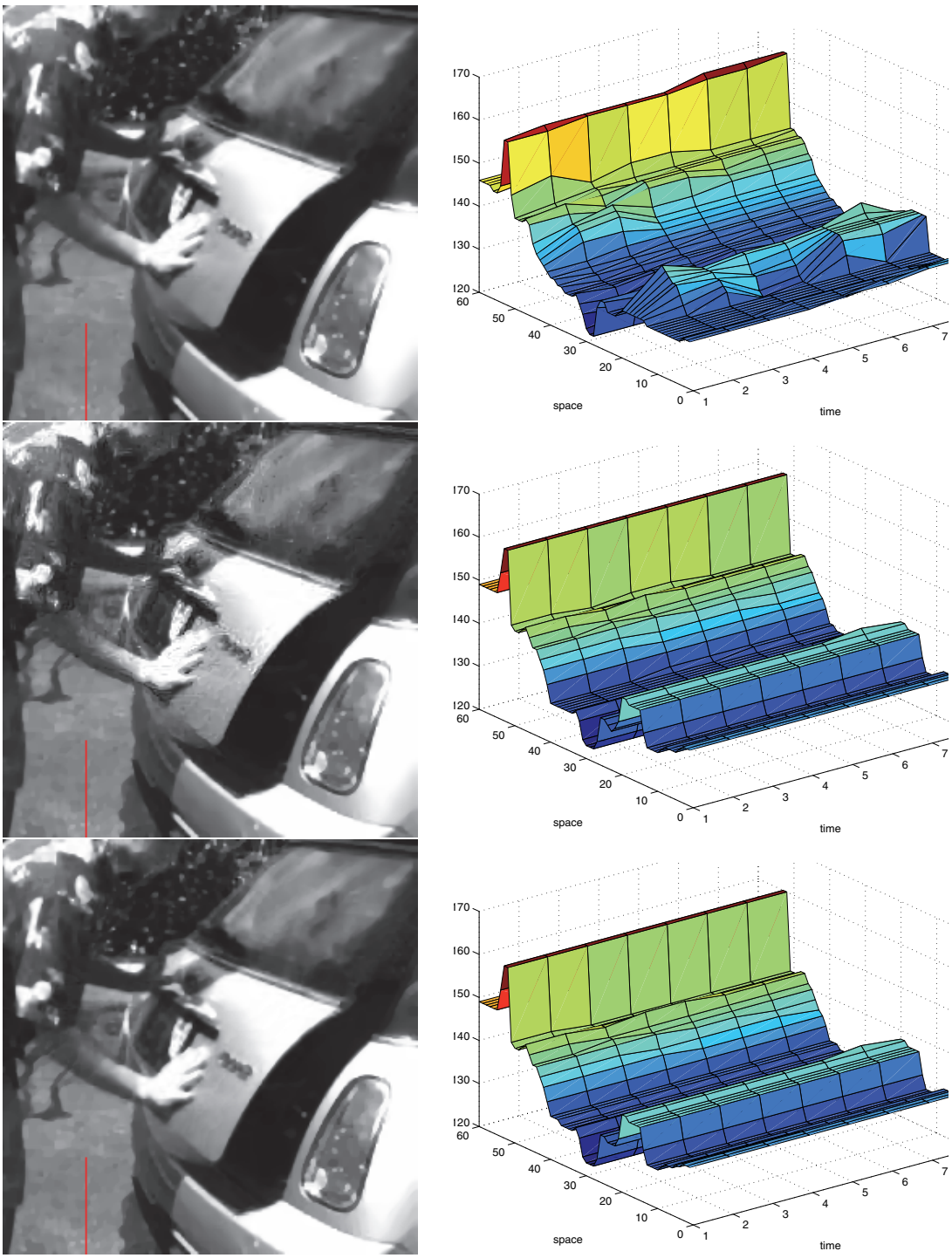

Figure 7. *Comparison of different regularization functionals for the Minicooper image sequence. From top to bottom:* TGV *regularization in space and time using temporal stepsizes of* 2 *and* 0.2*, and ICTV regularization in space and time. On the left, frame* 5 *of the image sequence is shown with line of the background region marked in red; on the right, the temporal development of this line is plotted. Iteration numbers to reach the stopping criterion, from top to bottom:* 542, 5416, 6621.

ponent will contain line structures whose direction is determined by the major axes of these ellipses. An immediate application of such a functional is the enhancement or separation of lines pointing in a predefined direction. But also, by combining a standard TV functional and four TV-type functionals with the major axis of the ellipses pointing in the directions 0, $\pi/4$, $\pi/2$, $3\pi/4$, we can hope for separation of cartoon and line structures.

As a particular case of the analytic framework described in section 3, we suggest the following functional for still image regularization:

$$\mathrm{ICTV}^5_\beta(u) = \min_{v_1,\dots,v_4} \mathrm{TGV}^2_\alpha(u - v_1) + \sum_{\substack{i=1 \\ v_5=0}}^{4} \mathrm{TV}_{\beta_{i+1}}(v_i - v_{i+1}),$$

where TGV^2_α is the standard TGV functional of second order using Euclidean norms and the $\mathrm{TV}_{\beta_{i+1}}$ functionals use gradient norms whose unit balls are ellipses with major axis pointing in direction $(i-1)\pi/4$, $i = 1, \dots, 4$.

In order to support the choice of using four different anisotropic components, we provide a numerical example for this setting in advance. Figure 8 shows the result of applying standard TV and ICTV^5_β denoising on a test image, corrupted by salt and pepper noise on 10 percent of the pixels. For information on the numerical methods and the choice of ICTV^5_β related parameters we refer the reader to subsequent sections. Data fidelity has been ensured by using L^1 discrepancy, and the regularization parameter has been chosen optimal in terms of *peak signal to noise ratio* (PSNR). Figure 8 shows that with ICTV^5_β we achieve a significant improvement compared to TV, which is also reflected in a much higher PSNR value. In addition, the figure shows two of the anisotropic components v_2, v_3, obtained with the ICTV^5_β reconstruction. The norm unit balls of these components are ellipses pointing in directions $\pi/4$ and $\pi/2$. As one can see, the functional allows a certain flexibility regarding these directions, and we can thus hope that four anisotropic components are sufficient to capture line structures of different orientations. Nevertheless, ICTV^5_β regularization will always favor line structures that are exactly orthogonal to that of the major axis, and thus we cannot expect rotational invariance.

Application of this kind of regularization yields a separation of the image into a piecewise smooth and a texture component. There have been various attempts in this direction in the past years. One approach, which has been taken in [5, 33, 27], originates in the work of Meyer [24] and decomposes the image into a cartoon part with low total variation and a texture/noise part capturing oscillatory components. For the texture part, the original idea was to minimize the norm

$$\|v\|_G = \inf\{\|g\|_\infty \,|\, v = \operatorname{div} g\}$$

over a suitable space. This approach works very well for the separation of cartoon and texture of noise-free images or the denoising of cartoon-type images. Its application to general denoising problems seems, however, limited by the fact that both texture and noise possess a low G norm.

Another very interesting approach is to introduce two transform operators, such as (wavelet) frames, which are well suited to approximate piecewise smooth and texture structures, respectively. One can then decompose images by penalizing the ℓ^1 norm of the coefficients needed

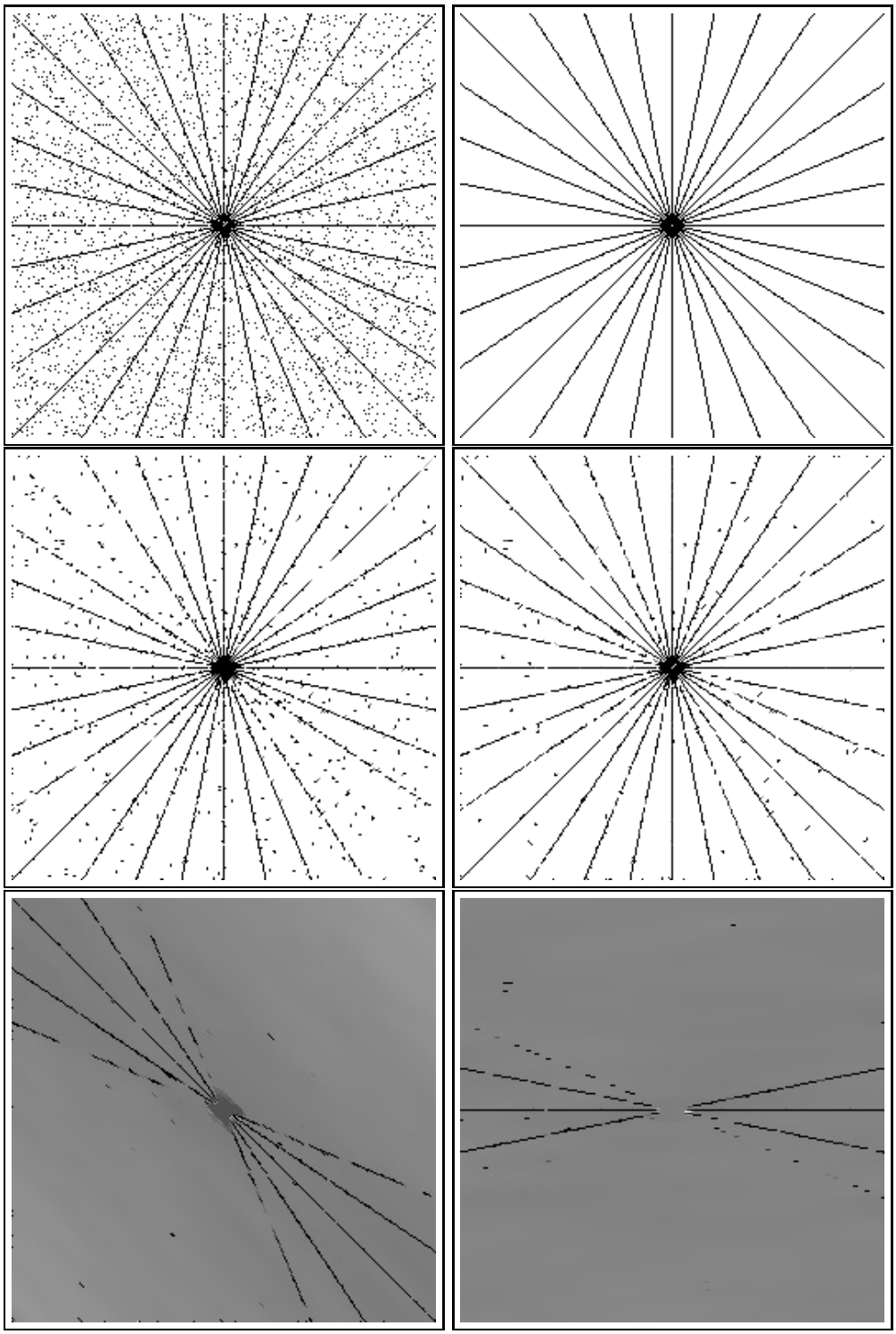

Figure 8. *Test image. Top: Noisy and original image. Middle:* TV *(left) and* ICTV^5_β *(right) based reconstruction. Bottom: Two anisotropic components obtained with the* ICTV^5_β *reconstruction. PSNR values are* 16.58 *(*TV*) and* 19.52 *(*ICTV^3_γ*).*

for both transforms to approximate the original image. We refer the reader to [23] for an overview and the analysis of such models and to [32] for numerical experiments.

Recently, in [30], the ideas of low-rank and sparse decomposition have been applied for still image regularization in a discrete setting. Again, the images get decomposed into a cartoon and texture part, where the cartoon part is expected to have low total variation. For the texture part, [30] suggests decomposing the image into small, nonoverlapping blocks, reordering them as vectors, and trying to keep the rank of a matrix, whose columns consist of all such vectors, low. This is done by minimizing the nuclear norm of this matrix. The numerical experiments of [30] suggest that this approach works very well for cartoon-texture decomposition. The main disadvantage seems to be the difficulty in selecting the block size and position for the matrix decomposition. The size of the chosen blocks has to somehow match the scale of oscillations one wants to recover, and one might expect problems when their position has a bad fit to texture-cartoon boundaries.

While above we focus on methods utilizing a separation in cartoon and texture parts, there are of course alternative approaches which allow a good reconstruction of textured regions, which include nonlocal methods [18]. In the context of enhancing line structures, we also refer the reader to the works [12, 13], where a convex relaxation of Euler's elastica functional is applied, and to [3], where the aim is to reconstruct point or line structures as low dimensional objects.

Despite being quite simple, our method works very well, in particular, for line structures. The formulation as infimal convolution of TV-type functionals allows a nice embedding of this approach into a general convex model in function space setting. Even though the numerical implementation does not pose any additional difficulties compared to standard TV regularization, we are able to obtain a surprisingly high improvement with respect to TV regularization in terms of visual image quality and error measures. Disadvantages are of course the bias of this regularization approach towards certain directions and the fact that we cannot expect to recover texture that is not decomposed of different line structures. However, for a certain, not too small class of images, the proposed approach works very well.

5.1. ICTV based image denoising. For image denoising, we consider the following problem setting:

$$\min_{u\in L^1(\Omega)} \mathrm{ICTV}^5_\beta(u) + \lambda\|u-f\|_1, \tag{29}$$

where $f \in L^1(\Omega)$ is given, and $\Omega \subset \mathbb{R}^2$ is again a bounded Lipschitz domain. For the norm parameter $\beta \in \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$, we define the norms

$$|\cdot|_{\beta_{0,1}} = \alpha_0^{-1}|\cdot|, \quad |\cdot|_{\beta_{1,1}} = \alpha_1^{-1}|\cdot|,$$

and, for $\xi \in \mathbb{R}^2$,

$$|\xi|_{\beta_{0,i}} = \mu|O_i \cdot \xi|_\kappa$$

with

$$O_i = \begin{pmatrix} \cos\phi_i & -\sin\phi_i \\ \sin\phi_i & \cos\phi_i \end{pmatrix}, \quad |\xi|_\kappa = \sqrt{(\kappa_1\xi_1)^2 + (\kappa_2\xi_2)^2},$$

$i = 2, \dots, 5$. Note that the O_i are simple rotation matrices and α_0, α_1 are the standard parameters for the TGV functional of second order. The scalars μ and κ_1, κ_2 are weights for the anisotropic part and the directional derivatives, respectively. To reduce the amount of parameters, we further normalize the integral mean of the norm $|\cdot|_\kappa$ over all possible directions to one; i.e., we require

$$1 = \frac{1}{2\pi} \int_0^{2\pi} \sqrt{(\kappa_1 \cos\phi)^2 + (\kappa_2 \sin\phi)^2}\, \mathrm{d}\phi.$$

This corresponds to choosing κ_1, κ_2 such that the ellipses with κ_1, κ_2 as semiaxes have perimeter 2π, and hence we obtain the second weight parameter κ_2 as a function of the first one. With this we expect the anisotropic TV functionals to be at the same scale as the isotropic one, independent of the ratio κ_1/κ_2. Thus our functional needs two parameters, μ as a weight of all anisotropic components together and κ_1 to control the amount of anisotropy corresponding to each direction.

For this setting, the results of section 3 apply; in particular, existence of a solution follows by Corollary 3.1.

5.2. Numerics for ICTV based denoising. For discretization and numerical solution of the ICTV denoising problem (29) we use a framework similar to that in section 4.2 for image sequences.

Defining the space of discrete images to be $U = \mathbb{R}^{N \times M}$ and $V = U \times U$, the discretized version of (29) is

$$\min_{\substack{u_0,\dots,u_5 \in U \\ u_5 = 0,\, v \in V}} \alpha_1 \|\nabla(u_0 - u_1) - v\|_1 + \alpha_0 \|\mathcal{E} v\|_1 + \mu \sum_{i=1}^{4} \|\nabla_i (u_i - u_{i+1})\|_1 + \lambda \|u_0 - f\|_1. \tag{30}$$

The gradient operators are given as

$$\nabla u = (\delta_{x+} u, \delta_{y+} u)^T, \quad \mathcal{E} v = (\delta_{x-} v_1, \delta_{y-} v_2, (\delta_{y-} v_1 + \delta_{x-} v_2)/2)^T,$$

$$\nabla_1 u = (\kappa_1 \delta_{x+} u, \kappa_2 \delta_{y+} u)^T, \quad \nabla_2 u = (\kappa_2 \delta_{x+} u, \kappa_1 \delta_{y+} u)^T,$$

$$\nabla_3 u = \begin{pmatrix} \kappa_1 \cos(\pi/4) & -\kappa_1 \sin(\pi/4) \\ \kappa_2 \sin(\pi/4) & \kappa_2 \cos(\pi/4) \end{pmatrix} (\delta_{x-} u, \delta_{y+} u)^T,$$

$$\nabla_4 u = \begin{pmatrix} \kappa_1 \cos(-\pi/4) & -\kappa_1 \sin(-\pi/4) \\ \kappa_2 \sin(-\pi/4) & \kappa_2 \cos(-\pi/4) \end{pmatrix} (\delta_{x+} u, \delta_{y+} u)^T,$$

where $\delta_{x+}, \delta_{y+}, \delta_{x-}, \delta_{y-}$ are forward and backward finite differences in the horizontal (x) and vertical (y) directions, respectively. Note that for the operator ∇_3 we use backward instead of forward differences for the horizontal derivative. This leads to an improvement in the model approximation since rays are in practice better resolved by finite difference gradients whose search directions both point to the same side of the ray.

Abusing notation, the norm $\|\cdot\|_1$ denotes the discrete L^1 norm for $\mathbb{R}$, $\mathbb{R}^2$, and $\mathbb{R}^3$ valued functions, where we use the Euclidean norm on the vector components in the $\mathbb{R}^2$ valued case and the norm $|\xi| = \sqrt{\xi_1^2 + \xi_2^2 + 2\xi_3^2}$ in the $\mathbb{R}^3$ valued case. The factor two in front of the

third component is to compensate for the symmetrization in the definition of $\mathcal{E}$; see [11, 9] for details.

To simplify notation, we introduce the space $W = U^3$ and denote by $F : V \times W \times V^4 \to \mathbb{R}$, $K : U \times V \times U^4 \to V \times W \times V^4$,

$$F(x^1, \dots, x^6) = \alpha_1 \|x^1\|_1 + \alpha_0 \|x^2\|_1 + \mu \sum_{i=3}^{6} \|x^i\|_1, \tag{31}$$

$$K = \begin{pmatrix} \nabla & -\mathrm{Id}_V & \nabla & 0 & 0 & 0 \\ 0 & \mathcal{E} & 0 & 0 & 0 & 0 \\ 0 & 0 & \nabla_1 & \nabla_1 & 0 & 0 \\ 0 & 0 & 0 & \nabla_2 & \nabla_2 & 0 \\ 0 & 0 & 0 & 0 & \nabla_3 & \nabla_3 \\ 0 & 0 & 0 & 0 & 0 & \nabla_4 \end{pmatrix}. \tag{32}$$

The discrete denoising problem (30) can then be rewritten as

$$\min_{x=(x_1,\dots,x_6)} F(Kx) + \lambda \|x_1 - f\|_1.$$

Note that, as in the discretization for the MJPEG decompression problem, we have included the weighting of the norms in the linear operators rather than the norms themselves. For the numerical solution, we again use a variant of the primal dual algorithm presented in [16] applied to an equivalent saddle point problem. The implementation is given in Algorithm 2, where we define

$$\mathrm{div} = -\nabla^*, \quad \mathrm{div}_{\mathcal{E}} = -\mathcal{E}^*,$$

$$\mathrm{div}_i = -\nabla_i^*, \quad i = 1, \dots, 4.$$

The operants proj_∞, proj_{L^1} can be evaluated pointwise and are given as

$$(\mathrm{proj}_\infty(y))^1 = \mathcal{P}_{\{\|y^1\|_\infty \le \alpha_1\}}(y^1), \quad (\mathrm{proj}_\infty(y))^2 = \mathcal{P}_{\{\|y^2\|_\infty \le \alpha_0\}}(y^2),$$

$$(\mathrm{proj}_\infty(y))^i = \mathcal{P}_{\{\|y^i\|_\infty \le \mu\}}(y^i), \quad i = 3, \dots, 6,$$

and

$$(\mathrm{proj}_{L^1}(x))^1_{i,j} = \begin{cases} x^1_{i,j} - \tau\lambda & \text{if } x^1_{i,j} - f_{i,j} > \tau\lambda, \\ x^1_{i,j} + \tau\lambda & \text{if } x^1_{i,j} - f_{i,j} < -\tau\lambda, \\ f_{i,j} & \text{else,} \end{cases} \quad i, j = 1, \dots, N, M,$$

$$(\mathrm{proj}_{L^1}(x))^i = x^i, \quad i = 2, \dots, 6,$$

where $\mathcal{P}_M$ denotes the projection on the set M. The operator S_K again realizes an adaptive stepsize choice similar to (25).

As stopping criterion we use a normalization of a partial primal dual gap

$$\tilde{\mathcal{G}}(x_n, y_n) = \mathcal{G}(x_n, y_n)/(NM),$$

Algorithm 2. Scheme of implementation for L^1 denoising.

function ICTV-$L^1(f)$
  $x^1 \leftarrow f$
  $x^i \leftarrow 0$, $i = 2, \dots, 6$,
  $\overline{x}^i \leftarrow 0$, $y^i \leftarrow 0$, $i = 1, \dots, 6$,
  choose $\sigma, \tau > 0$ arbitrary
  repeat
    $y \leftarrow \operatorname{proj}_\infty (y + \sigma K \overline{x})$
    $x_+ \leftarrow \operatorname{proj}_{L^1}(x - \tau K^* y)$
    $\overline{x} \leftarrow (2x_+ - x)$
    $x \leftarrow x_+$, $y \leftarrow y_+$
    $\sigma\tau \leftarrow S_K(\sigma, \tau)$
  until Stopping criterion fulfilled
  return x^1
end function

where $\mathcal{G}$ is defined by

$$\begin{aligned}\mathcal{G}(x_n, y_n) = F(Kx_n) &+ \lambda\|x_n^1 - f\|_1 + (-\operatorname{div}(\operatorname{div}_{\mathcal{E}} \tilde{y}_n^2), u_0)_U \\ &+ \gamma\|x_n^2\|_2\| \operatorname{div}(\operatorname{div}_{\mathcal{E}} \tilde{y}_n^2) - \operatorname{div}_1 y_3\|_2 \\ &+ \sum_{i=4}^{6} \gamma\|x_n^i\|_2\| \operatorname{div}_{i-3} y_{i-1} - \operatorname{div}_{i-2} y_i\|_2.\end{aligned} \tag{33}$$

It can be shown by techniques similar to those in the proof of Proposition 4.2 that

$$\mathcal{G}(x_n, y_n) \to 0 \text{ as } n \to \infty$$

and that

$$\mathcal{G}(x_n, y_n) \geq F(Kx_n) + \lambda\|x_1 - f\|_1 - (F(K\hat{x}_n) + \lambda\|\hat{x}_1 - f\|_1)$$

whenever $\gamma\|x_n^i\|_2 \geq \|\hat{x}^i\|_2$ for $i = 3, \dots, 6$.

5.3. Numerical results. In this section we compare numerical results of ICTV_β^5 denoising with results of standard TV and ICTV_γ^3 denoising, where for ICTV_γ^3 denoising we convolute three TV functionals using only the horizontal and vertical directions as anisotropic components, i.e.,

$$\mathrm{ICTV}_\gamma^3(u) = \min_{u_1, u_2} \|\nabla(u - u_1)\|_1 + \mu\left(\|\nabla_1(u_1 - u_2)\|_1 + \|\nabla_2 u_2\|_2\right).$$

For TV and ICTV_γ^3 based denoising we implemented a primal dual algorithm similar to Algorithm 2. We carry out experiments on three different images which have been corrupted by salt and pepper noise on 25 percent of the pixels. Parameters affecting the image quality are for all test cases fixed as follows:

- $\lambda = 1.4$: Chosen optimal for TV based denoising of the lighthouse image (see Figure 9) in terms of visual image quality.

- $\kappa_1 = 3/2$, $\mu = 1.6$: Chosen optimal for ICTV^3_γ based denoising of the lighthouse image (see Figure 9) in terms of visual image quality.
- $\alpha_0/\alpha_1 = \sqrt{2}$. The ratio of the TGV^2_α parameters for the isotropic part of ICTV^5_β is chosen in a standard way as suggested, for example, in [9].

Note that, while the optimal choice of the regularization parameter λ of course depends on the noise level, the parameters κ_1, μ, α_0/α_1 reflect the expected structure of realistic images and can be fixed noise-level independently.

Parameters affecting convergence speed are for all test cases fixed as follows:

- TV: $\sigma = (\sqrt{1/8})0.05$, $\tau = (\sqrt{1/8})/0.05$.
- ICTV^3_γ: Adaptive stepsize choice similar to (25), ratio σ/τ fixed to 0.05^2.
- ICTV^5_β: Adaptive stepsize choice similar to (25), ratio σ/τ fixed to 0.03^2.

As stopping criterion, we use a normalized, modified primal dual gap similar to (33) for all three implementations. The resulting iteration numbers are given in the captions of the figures. Concerning computation time, we remark that our nonoptimized MATLAB code including all computations for the stopping rules needs about 4.9 times (ICTV^3_γ), respectively, 10.7 times (ICTV^5_γ), as long as with TV per iteration.

We point out that the unequal stepsize choice for σ and τ resulted in a significantly accelerated convergence of the methods, in particular also for the standard TV-L^1 case: For denoising the lighthouse image, $\mathrm{TV}_k - \mathrm{TV}_{\mathrm{opt}} < 100$ was satisfied after $k = 80$ iterations for a σ-τ ratio of 0.05^2, and $k = 1680$ iterations for a ratio of 1. Similarly, after 2000 iterations, the modified duality gap with ratio 0.05^2 was at $1.6 \cdot 10^{-3}$, while with ratio 1 it was at $1.9 \cdot 10^{-1}$.

As a first experiment we compare TV denoising and ICTV^3_γ denoising of the lighthouse test image. Figure 9 shows the noisy and original image on the top row, its TV denoised version and an error plot in the middle row, and the ICTV^3_γ denoised version with error plot on the bottom row. The aim of this comparison is to show that a significant improvement can be achieved even with using just standard TV for the isotropic component together with two anisotropic components. As can be seen in the images and the error plots, in particular the fine details of the fence are much better reconstructed using ICTV^3_γ. But also the wall of the lighthouse appears more realistic with ICTV^3_γ than with TV since line structure is less frayed. In terms of PSNR and structural similarity (SSIM) [34] a significant improvement is achieved: While the PSNR and SSIM values of the TV reconstruction are 23.56 and 0.6891, respectively, the ICTV^3_γ reconstruction achieves 26.39 and 0.7899, respectively.

In Figure 10 we show the decomposition of the lighthouse image obtained with ICTV^3_γ denoising. The left image shows the isotropic component, while the right image shows the addition of the two anisotropic components. As can be seen, the fence is almost entirely contained in the anisotropic components, which explains the improvement achieved with ICTV^3_γ in this part of the image.

In the next experiment we compare TV and ICTV^3_γ based denoising with ICTV^5_β based denoising. Remember that for ICTV^5_β we use the TGV functional of second order for the isotropic component. The reconstruction of a section of the standard Lena test image can be seen in Figure 11. The top row shows the noisy image together with the ICTV^3_γ based reconstruction. The middle row depicts the TV based reconstruction (left) and the ICTV^5_β based reconstruction (right). The areas indicated by the red squares are magnified below.

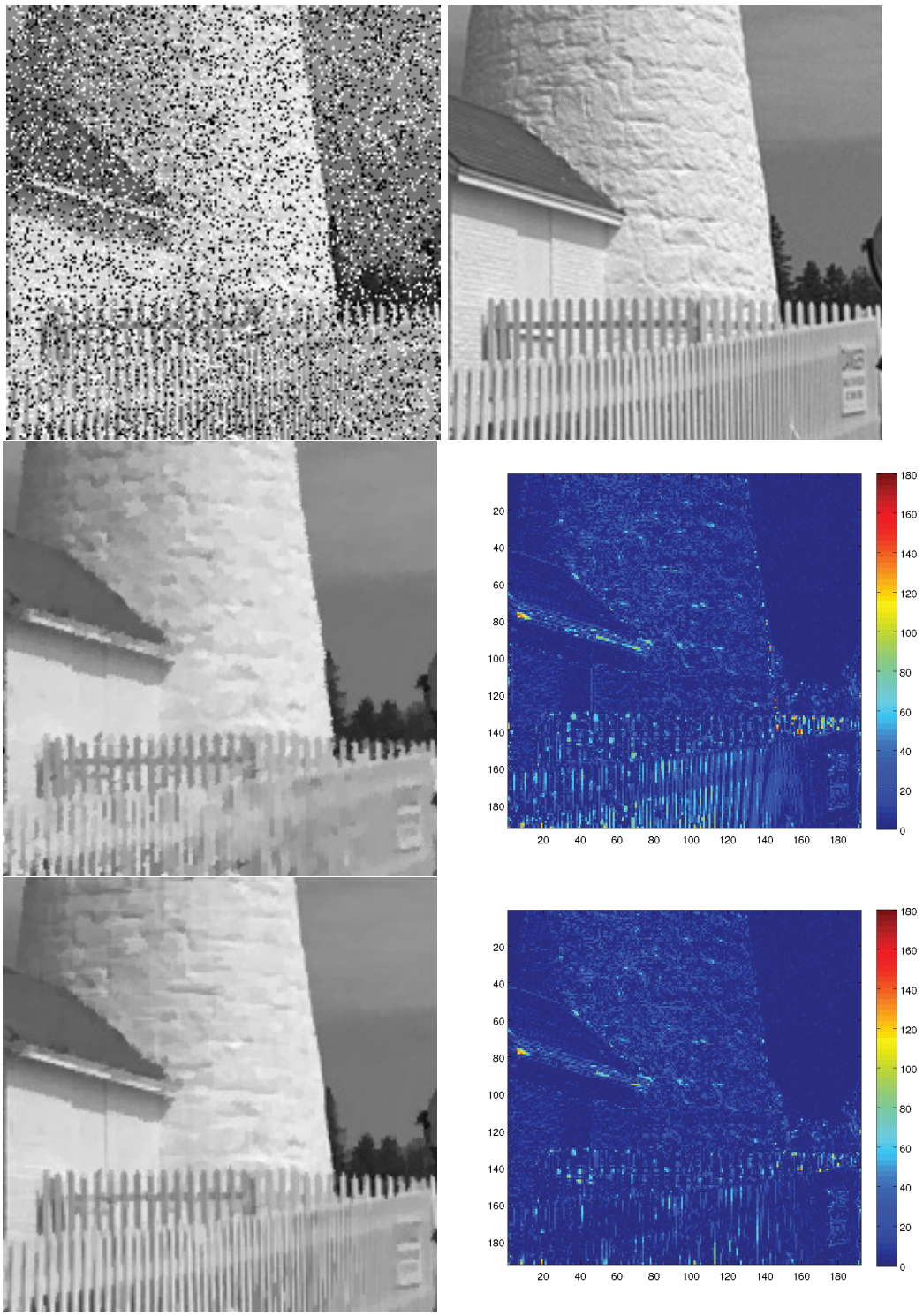

Figure 9. *Lighthouse test image. Top: Noisy and original image. Middle:* TV *based reconstruction and error plot. Bottom:* ICTV^3_γ *based reconstruction and error plot. PSNR values are* 23.60 *(*TV*) and* 26.39 *(*ICTV^3_γ*). Iteration numbers to reach the stopping criterion:* 1067 *(TV),* 3020 *(*ICTV^3_γ*).*

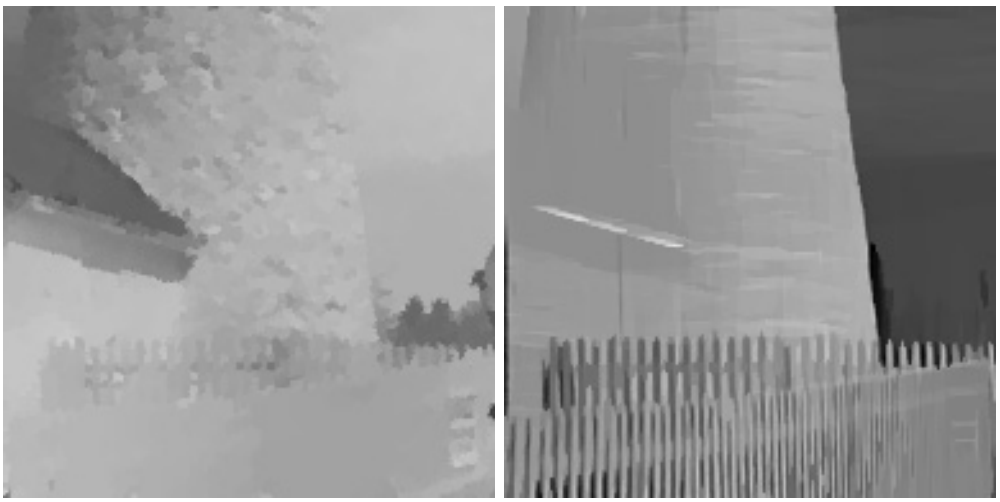

Figure 10. *Isotropic and anisotropic component of* ICTV_γ^3 *based reconstruction of the noisy lighthouse image.*

One can see that the top left boundary of the hat and the line in the background are better resolved with ICTV_γ^3 than with TV. The result with the ICTV_β^5 functional again yields an improvement even with respect to ICTV_γ^3. The brim of the hat appears much more realistic, and fine structure of the head is also much better resolved. This can in particular be seen in the magnified squares on the bottom. Also the face appears more natural since, as a result of using TGV_α^2 for the isotropic component, it suffers from fewer staircasing effects. PSNR values of the reconstructions are 30.22 (TV), 30.62 (ICTV_γ^3), 31.58 (ICTV_β^5), and SSIM values are given as 0.8683 (TV), 0.8749 (ICTV_γ^3), 0.8954 (ICTV_β^5). This in particular indicates a significant improvement also by using ICTV_β^5 instead ICTV_γ^3.

At last we compare TV and ICTV_β^5 denoising for the Barbara test image. The top row of Figure 12 shows the noisy image and a summation of the anisotropic components using ICTV_β^5 denoising. The middle row shows the result of TV (left) and ICTV_β^5 (right) denoising, and the bottom row again shows the corresponding close ups. It can be seen in the summation of the anisotropic components on the top right that many line structures are captured by the ICTV_β^5 functional, which are then resolved much better in the reconstruction. We also see, however, that curved lines are not resolved very well and that the reconstruction of lines still depends on the direction. In particular, lines on the left leg are not captured. We observed in experiments that adding additional directions further improves the result, but the drawback of favoring only particular, predefined directions of course remains.

We can conclude from these experiments that the ICTV_β^5 functional allows a surprisingly strong improvement with respect to standard TV denoising. While remaining in a very similar, convex problem setting, it is possible to capture certain structures much better and achieve a much higher signal to noise ratio. In general, the results are superior to standard TV denoising in all our experiments, while the level of improvement of course depends on the particular images.

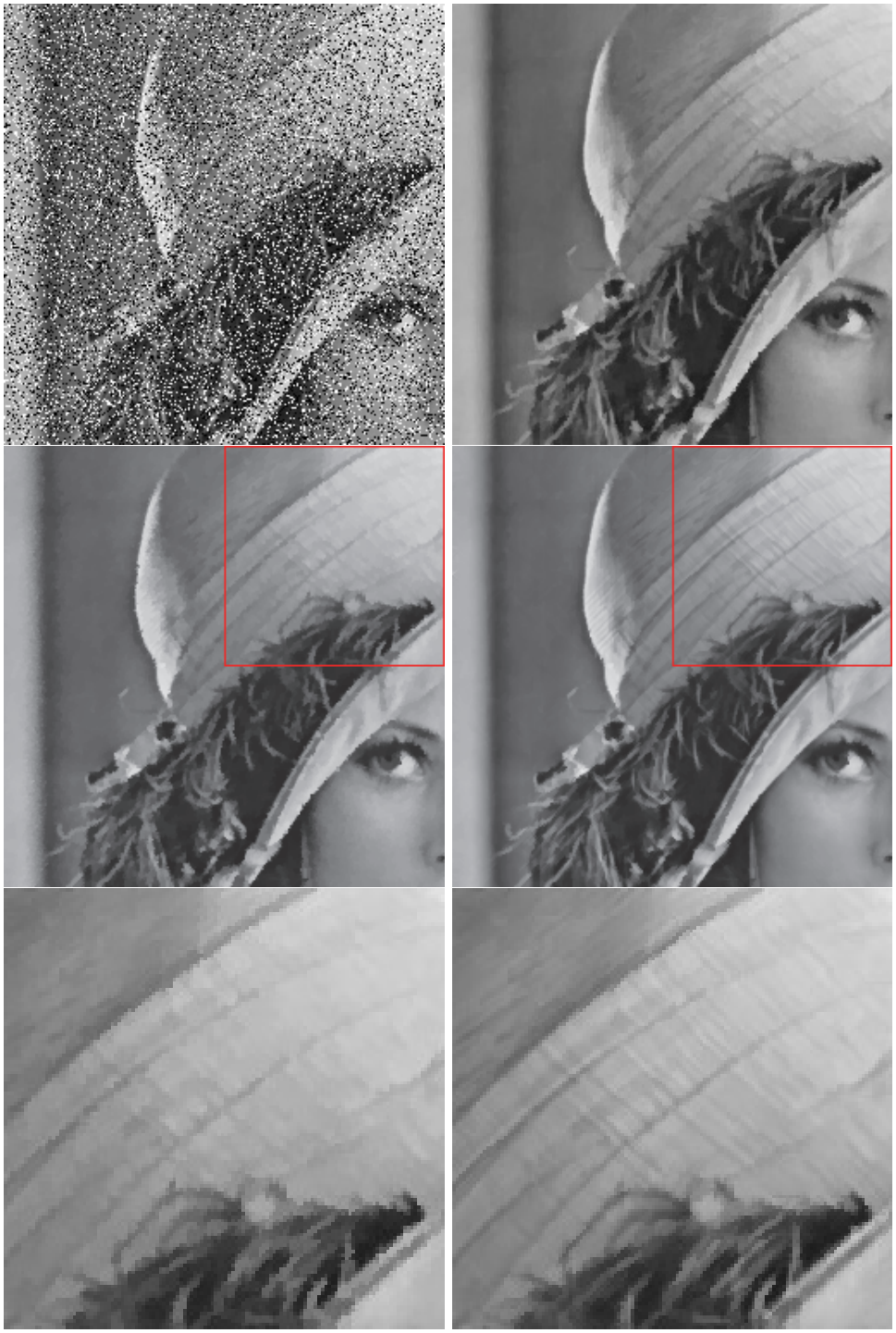

Figure 11. *Lena test image. Top: Noisy image (left) and* ICTV^3_γ *based reconstruction. Middle:* TV *based reconstruction (left) and* ICTV^5_β *based reconstruction with marked regions. Bottom: Magnification of marked regions of* TV *(left) and* ICTV^5_β *reconstructions. PSNR values are* 30.22 *(*TV*),* 30.62 *(*ICTV^3_γ*), and* 31.58 *(*ICTV^5_β*). Iteration numbers to reach the stopping criterion:* 561 *(TV),* 2100 *(*ICTV^3_γ*),* 4468 *(*ICTV^5_β*).*

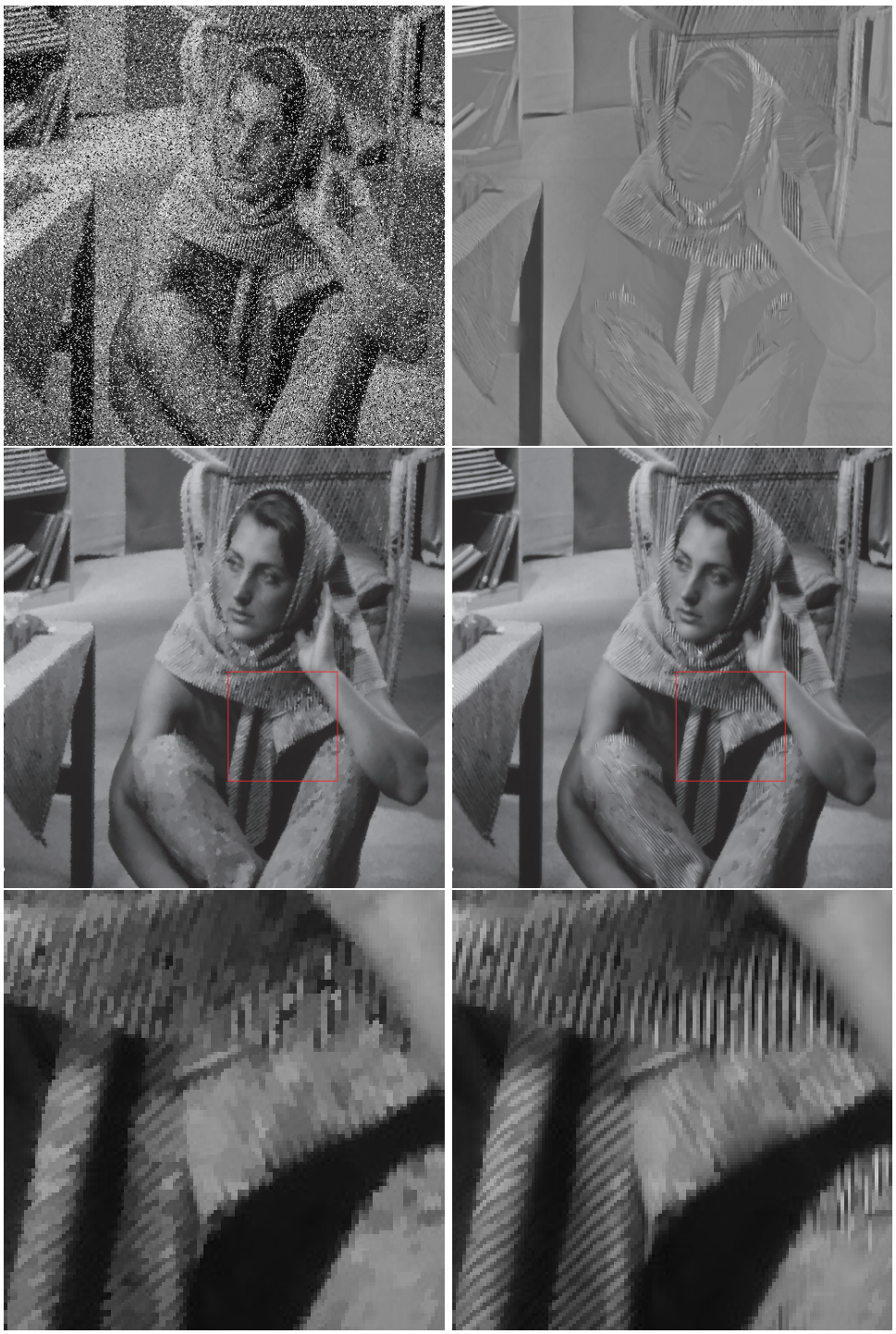

Figure 12. *Barbara test image. Top: Noisy image (left) and anisotropic components of* $\mathrm{ICTV}_{\beta}^{5}$ *based reconstruction. Middle:* TV *based reconstruction (left) and* $\mathrm{ICTV}_{\beta}^{5}$ *based reconstruction with marked regions. Bottom: Magnification of marked regions of* TV *(left) and* $\mathrm{ICTV}_{\beta}^{5}$ *reconstructions. PSNR values are* 24.94 *(TV),* 25.96 *(*$\mathrm{ICTV}_{\beta}^{5}$*); SSIM values are* 0.7660 *(TV) and* 0.8145 *(*$\mathrm{ICTV}_{\beta}^{5}$*). Iteration numbers to reach the stopping criterion:* 1508 *(TV),* 4468 *(*$\mathrm{ICTV}_{\beta}^{5}$*).*

6. Appendix. We will now elaborate the proof of Proposition 3.3 claiming that the dual formulation of ICTV^n_β as in (4) is indeed equivalent and that, for any $u \in \mathrm{BV}(\Omega)$, the infimum in (4) is actually achieved.

For this purpose, we first establish for $L^p(\Omega)$, $p \in [1,\infty]$, what would be an orthogonal decomposition of $L^2(\Omega)$ as $L^2(\Omega) = \mathcal{P}^{k-1}(\Omega) \oplus \mathcal{P}^{k-1}(\Omega)^\perp$. However, in a general Banach space setting, such a decomposition is somewhat more involved. For the reader's convenience, we provide an abstract functional analytic lemma from which the desired decomposition will follow.

For a normed space X with dual X^* and $M \subset X$, $N \subset X^*$ define the annihilators

$$M^\perp = \{f \in X^* \,|\, \langle f, x\rangle_{X^*,X} = 0 \text{ for all } x \in M\},$$

$$N^\perp = \{x \in X \,|\, \langle f, x\rangle_{X^*,X} = 0 \text{ for all } f \in N\}.$$

Note that the definition of the annihilator depends on whether the set is contained in X or X^*.

Lemma 6.1. *Let X be a Banach space and $P \subset X$ be a finite dimensional subspace. Further suppose that $F : P \to X^*$ is a linear mapping such that $\langle F(p), p\rangle_{X^*,X} = 0$ implies $p = 0$. Then, there exists a continuous, linear projection $R : X \to P \subset X$ such that*

$$\ker(R) = F(P)^\perp$$

and every $x \in X$ can be uniquely decomposed as $x = x - R(x) + R(x) \in F(P)^\perp + P$.

Proof. First remember [29, Theorem 4.7] that $(F(P)^\perp)^\perp = \overline{F(P)}^*$, where $\overline{F(P)}^*$ denotes the closure of $F(P) \subset X^*$ with respect to the weak-* topology. By $F(P)$ being a finite dimensional subspace of X^*, it is closed with respect to weak-* convergence [14, Remark 10, p. 64], and hence we get

$$F(P) = (F(P)^\perp)^\perp.$$

To obtain the decomposition of X as claimed, note that $A := P + F(P)^\perp$ is a closed subspace of X as the sum of a finite dimensional and a closed subspace [14, Theorem 11.4]. Also $P \cap F(P)^\perp = \{0\}$ since $0 = \langle F(p), p\rangle_{X^*,X}$ implies $p = 0$.

To show that $A = X$, take any $f \in X^*$ such that

$$\langle f, p + q\rangle_{X^*,X} = 0 \quad \text{for all} \quad p \in P, q \in F(P)^\perp \subset X.$$

Choosing $p = 0$, this implies that $f \in (F(P)^\perp)^\perp = F(P)$. Thus there exists $p_0 \in P$ with $F(p_0) = f$ and choosing $q = 0$ in the equation above implies that $\langle F(p_0), p_0\rangle_{X^*,X} = 0$ and thus $f = 0$.

Hence $A = \overline{A} = X$. Thus the linear mapping

$$T : P \times F(P)^\perp \to X, \quad T(p,q) := p + q$$

is bijective and the unique decomposition follows. Equipping $P \times F(P)^\perp$ with the norm $\|(p,q)\|_{P\times F(P)^\perp} := \|p\|_X + \|q\|_X$, T is also continuous and, as a consequence of the open mapping theorem [14, Corollary 2.7], it possesses a continuous inverse. Thus the linear projection given by $R(x) = p$ for $x = p + q \in P + F(P)^\perp$ is continuous since

$$\|p\|_X \le \|p\|_X + \|q\|_X = \|T^{-1}(x)\|_{P\times F(P)^\perp} \le C\|x\|_X$$

and the assertion follows. ∎

Lemma 6.1 implies the following corollary, which will be relevant.

Corollary 6.1. *With $k \in \mathbb{N}$ and $\mathcal{P}^{k-1}(\Omega)$ the space of polynomials of degree less than or equal to $k-1$, for each $p \in [1,\infty]$, there exists a linear, continuous projection $R_{k,p} : L^p(\Omega) \to \mathcal{P}^{k-1}(\Omega) \subset L^p(\Omega)$ such that $\ker(R_{k,p}) = P_{k,p}^{\perp}$ with*

$$P_{k,p}^{\perp} := \{u \in L^p(\Omega) \,|\, (u,p) = 0 \text{ for all } p \in \mathcal{P}^{k-1}(\Omega)\}.$$

Proof. We apply Lemma 6.1 with $X = L^p(\Omega)$ and $P = \mathcal{P}^{k-1}(\Omega)$.

If $p < \infty$, we can define the mapping F of Lemma 6.1 to be the identity from $\mathcal{P}^{k-1}(\Omega) \to L^{p'}(\Omega)$. The duality pairing $\langle \cdot, \cdot \rangle_{X^*,X}$ then coincides with the standard pairing $(\cdot,\cdot)$ between $L^{p'}$ and $L^p(\Omega)$, and hence the assertion follows.

If $p = \infty$, we can use the canonical injection $J : L^1(\Omega) \to (L^\infty(\Omega))^*$ and define F to be the restriction of J to P as subset of $L^1(\Omega)$. Then, for all $x \in X = L^\infty(\Omega)$ and $f \in J(L^1(\Omega)) \subset X^*$, the duality pairing $\langle f, x \rangle_{X^*,X}$ again coincides with the standard pairing (x,f) between L^p and $L^{p'}(\Omega)$ and the assertion follows. ∎

The important part of the previous corollary is that the exponent p is also allowed to take the values 1 and ∞. Without these two cases, the decomposition would have followed more easily as a result of reflexivity of $L^p(\Omega)$.

We are now able to state one of our main results, from which the primal representation of the ICTV functional will follow.

Proposition 6.1. *Let $k \in \mathbb{N}$, $\beta \in \mathbb{S}^k$ be a parameter vector for the TGV_β^k functional, and let $\psi : L^d(\Omega) \to \overline{\mathbb{R}}$ be a convex, lower semicontinuous function such that*

$$0 \in \mathrm{dom}(\psi) \subset P_{k,d}^{\perp},$$

with $P_{k,d}^{\perp}$ as in in Corollary 6.1. Denote

$$U = \{\mathrm{div}^k \phi \in C_c^\infty(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)) \,|\, \|\mathrm{div}^i \phi\|_{\infty,\beta_i} \le 1,\ 0 \le i \le k-1\} \subset L^d(\Omega)$$

such that $\mathcal{I}_U^* = \mathrm{TGV}_\beta^k$ in $L^{d'}(\Omega)$. Then, for any $u \in L^{d'}(\Omega)$,

$$(\mathcal{I}_U + \psi)^*(u) = \min_{v \in L^{d'}(\Omega)} \mathrm{TGV}_\beta^k(u-v) + \psi^*(v) \quad \text{in } L^{d'}(\Omega). \tag{34}$$

Proof. First note that it is sufficient to show the assertion with U replaced by its closure in $L^d(\Omega)$, which is denoted by $\overline{U}$.

By [2, Theorem 1.1], the claimed assertion is true if we can show that $\bigcup_{\lambda \ge 0} \lambda(\overline{U} - \mathrm{dom}(\psi))$ is a closed vector space. We shall verify that

$$\bigcup_{\lambda \ge 0} \lambda(\overline{U} - \mathrm{dom}(\psi)) = P_{k,d}^{\perp}, \tag{35}$$

which is a closed subspace of $L^d(\Omega)$.

This holds true if we can find $\epsilon > 0$ such that

$$B_\epsilon(0) \cap P_{k,d}^{\perp} \subset \overline{U}$$

as a subset in $L^d(\Omega)$. Indeed, to show the inclusion $\supset$ in (35), take $u \in P_{k,d}^{\perp}$. Then, with $\lambda := \epsilon^{-1} 2\|u\|_d$, $\tilde{u} := u/\lambda \in B_\epsilon(0)$ and since $0 \in \mathrm{dom}(\psi)$, it follows that $u = \lambda(\tilde{u} - 0) \in \bigcup_{\lambda \geq 0} \lambda(\overline{U} - \mathrm{dom}(\psi))$. The inclusion $\subset$ in (35) is satisfied trivially since both $\overline{U}$ and $\mathrm{dom}(\psi)$ are contained in $P_{k,d}^{\perp}$.

Now by the Poincare-type inequality for TGV_β^k there exists a constant $C > 0$ such that, with $R_{k,d'} : L^{d'}(\Omega) \to \mathcal{P}^{k-1}(\Omega)$ the continuous, linear projection of Corollary 6.1, for all $v \in L^{d'}(\Omega)$,

$$\|v - R_{k,d'}(v)\|_{L^{d'}} \leq C\, \mathrm{TGV}_\beta^k(v).$$

Choose $\epsilon < \frac{1}{C}$, and take $g \in B_\epsilon(0) \cap P_{k,d}^{\perp}$ arbitrary. Now for $v \in L^{d'}(\Omega)$

$$(v, g) = (v - R_{k,d'}(v), g) + (R_{k,d'}(v), g) \leq \|v - R_{k,d'}(v)\|_{d'} \|g\|_d \leq \mathrm{TGV}_\beta^k(v)$$

and hence

$$\mathcal{I}_{\overline{U}}(g) = \mathcal{I}_{\overline{U}}^{**}(g) = \sup_{v \in L^{d'}(\Omega)} (v, g) - \mathrm{TGV}_\beta^k(v) \leq 0,$$

which implies $g \in \overline{U}$, from which the assertion follows. ∎

The techniques of the above proof can also be applied to investigate the solvability of the equation $\mathrm{div}^k \phi = g$ with ϕ suitably defined. For that purpose, we define

$$\|\phi\|_{d,\mathrm{div}^k}^d := \sum_{i=0}^{k} \|\mathrm{div}^i \phi\|_d^d$$

and the space $W_0^d(\mathrm{div}^k; \Omega, \mathrm{Sym}^k(\mathbb{R}^d))$ as the closure of $C_c^\infty(\Omega, \mathrm{Sym}^k(\mathbb{R}^d))$ with respect to $\|\cdot\|_{d,\mathrm{div}^k}$ (see [20, section 2.3]) as norm. We then get the following result.

Corollary 6.2. *There exists a constant $C > 0$ such that, for any $g \in L^d(\Omega) \cap P_{k,d}^{\perp}$, there exists $\phi \in W_0^d(\mathrm{div}^k; \Omega, \mathrm{Sym}^k(\mathbb{R}^d))$ with $\|\mathrm{div}^i \phi\|_\infty < \infty$ for $i = 0, \ldots, k-1$ such that*

$$\mathrm{div}^k \phi = g \quad \text{and} \quad \max_{i=0,\ldots,k-1} \{\|\mathrm{div}^i \phi\|_\infty\} \leq C\|g\|_d. \tag{36}$$

Proof. With U defined as

$$U = \{\mathrm{div}^k \phi \in C_c^\infty(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)) \,|\, \|\mathrm{div}^i \phi\|_\infty \leq 1,\ 0 \leq i \leq k-1\},$$

as in the proof of Proposition 6.1 we can choose $\epsilon > 0$ such that $B_\epsilon(0) \cap P_{k,d}^{\perp} \subset \overline{U}$. Note also that

$$\overline{U} = \left\{\mathrm{div}^k \phi \,|\, \phi \in W_0^d(\mathrm{div}^k; \Omega, \mathrm{Sym}^k(\mathbb{R}^d)), \|\mathrm{div}^i \phi\|_\infty \leq 1,\ i = 0, \ldots, k-1\right\},$$

which can be shown by weak sequential compactness arguments (see [20, Proposition 4.3]). Thus, for $g \in L^d(\Omega) \cap P_{k,d}^{\perp}$ there exists $\phi \in \overline{U}$ such that $\mathrm{div}^k \phi = \epsilon g/(2\|g\|_d)$. Consequently, $\tilde{\phi} := \epsilon^{-1} 2\|g\|_d \phi$ satisfies

$$\mathrm{div}^k \tilde{\phi} = g \quad \text{and} \quad \|\mathrm{div}^i \tilde{\phi}\|_\infty \leq 2\epsilon^{-1}\|g\|_d \quad \text{for } i = 0, \ldots, k-1. \qquad ∎$$

This result can be put into context with a result in [7], which says that for $g \in L^d(\Omega)$ with $\int_\Omega g = 0$ there exists $\phi \in C^0(\overline{\Omega}) \cap W_0^{1,d}(\Omega)$ and a constant C, independent of g, such that

$$\operatorname{div} \phi = g \quad \text{and} \quad \|\phi\|_\infty + \|\phi\|_{W^{1,d}} \leq C\|g\|_d.$$

For the case $k = 1$, this can be used to obtain the results of Proposition 6.1. For arbitrary k, an alternative proof of Proposition 6.1 would have been to modify the arguments of [7] by defining div^k as unbounded operator and arguing with the closed range theorem. For our purposes, the proof presented above is more direct.

We are now able to obtain the equivalent definition of ICTV as stated in Proposition 3.3.

Proposition. *For $n \in \mathbb{N}$ and $\beta \in \mathbb{S}^{k_1} \times \cdots \times \mathbb{S}^{k_n}$, let k_i and $\beta_{\cdot,i}$, $i \in \{1, \dots, n\}$, be the order and parameter vectors for the $\mathrm{TGV}^{k_i}_{\beta_{\cdot,i}}$ functionals, respectively. Then, for any $u \in L^{d'}(\Omega)$,*

$$\mathrm{ICTV}^n_\beta(u) = \min_{\substack{v_i \in L^{d'}(\Omega),\\ 1 \leq i < n,\\ v_0 = u,\, v_n = 0}} \left(\sum_{i=1}^n \mathrm{TGV}^{k_i}_{\beta_{\cdot,i}}(v_{i-1} - v_i) \right). \tag{37}$$

Proof. Denote by U_i the sets such that $\mathcal{I}^*_{U_i} = \mathrm{TGV}^{k_i}_{\beta_{\cdot,i}}$ in $L^{d'}(\Omega)$. Then, by definition,

$$\left(\sum_{i=1}^n \mathcal{I}_{U_i} \right)^* (u) = \mathrm{ICTV}^n_\beta(u).$$

Choose i_1 such that $k_{i_1} = \min\{k_1, \dots, k_n\}$. Then,

$$0 \in \operatorname{dom} \left(\sum_{\substack{i=1 \\ i \neq i_1}}^n \mathcal{I}_{U_i} \right) \subset P^\perp_{k_{i_1},d}$$

and we can apply Proposition 6.1 to get

$$\left(\sum_{i=1}^n \mathcal{I}_{U_i} \right)^* (u) = \min_{v_1 \in L^{d'}(\Omega)} \mathrm{TGV}^{k_{i_1}}_{\beta_{\cdot,i_1}}(u - v_1) + \left(\sum_{\substack{i=1 \\ i \neq i_1}}^n \mathcal{I}_{U_i} \right)^* (v_1).$$

Proceeding inductively, the assertion follows after finitely many steps. ∎

REFERENCES

[1] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford University Press, New York, 2000.

[2] H. Attouch and H. Brezis, *Duality for the sum of convex functions in general Banach spaces*, Aspects Math. Appl., 34 (1986), pp. 125–133.

[3] G. Aubert, L. Blanc-Féraud, and D. Graziani, *Analysis of a new variational model to restore point-like and curve-like singularities in imaging*, Appl. Math. Optim., 67 (2013), pp. 73–96.

[4] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing*, Springer, New York, 2006.
[5] J.-F. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle, *Image decomposition into a bounded variation component and an oscillating component*, J. Math. Imaging Vision, 22 (2005), pp. 71–88.
[6] L. Bar, B. Berkels, M. Rumpf, and G. Sapiro, *A variational framework for simultaneous motion estimation and restoration of motion-blurred video*, in Proceedings of the 11th IEEE International Conference on Computer Vision (ICCV 2007), 2007, pp. 1–8.
[7] J. Bourgain and H. Brezis, *On the equation* $\operatorname{div} y = f$ *and application to control phases*, J. Amer. Math. Soc., 16 (2002), pp. 393–426.
[8] K. Bredies and M. Holler, *A total variation–based JPEG decompression model*, SIAM J. Imaging Sci., 5 (2012), pp. 366–393.
[9] K. Bredies and M. Holler, *Artifact-free decompression and zooming of JPEG compressed images with total generalized variation*, in Computer Vision, Imaging and Computer Graphics: Theory and Application, Comm. Comput. Inform. Sci. 359, Springer, Berlin, 2013, pp. 242–258.
[10] K. Bredies and M. Holler, *Regularization of linear inverse problems with total generalized variation*, J. Inverse Ill-Posed Probl., 2014, doi:10.1515/jip-2013-0068.
[11] K. Bredies, K. Kunisch, and T. Pock, *Total generalized variation*, SIAM J. Imaging Sci., 3 (2010), pp. 492–526.
[12] K. Bredies, T. Pock, and B. Wirth, *Convex relaxation of a class of vertex penalizing functionals*, J. Math. Imaging Vision, 47 (2013), pp. 278–302.
[13] K. Bredies, T. Pock, and B. Wirth, *A Convex, Lower Semi-continuous Approximation of Euler's Elastica Energy*, preprint, 2013.
[14] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, New York, 2010.
[15] A. Chambolle and P.-L. Lions, *Image recovery via total variation minimization and related problems*, Numer. Math., 76 (1997), pp. 167–188.
[16] A. Chambolle and T. Pock, *A first-order primal-dual algorithm for convex problems with applications to imaging*, J. Math. Imaging Vision, 40 (2011), pp. 120–145.
[17] J. P. Cocquerez, L. Chanas, and J. Blanc-Talon, *Simultaneous inpainting and motion estimation of highly degraded video-sequences*, in Scandinavian Conference on Image Analysis, Lecture Notes in Comput. Sci. 2749, Springer, Berlin, 2003, pp. 523–530.
[18] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, *Image denoising with block-matching and* 3*D filtering*, in Electronic Imaging'06, Proc. SPIE 6064, no. 6064A-30, 2006.
[19] GPU4Vision Project, http://gpu4vision.icg.tugraz.at/.
[20] M. Holler, *Higher Order Regularization for Model Based Data Decompression*, Ph.D. thesis, University of Graz, Graz, Austria, 2013; available online from http://www.uni-graz.at/~hollerm/.
[21] H. Ji, S. Huang, Z. Shen, and Y. Xu, *Robust video restoration by joint sparse and low rank matrix approximation*, SIAM J. Imaging Sci., 4 (2011), pp. 1122–1142.
[22] P. Kornprobst, R. Deriche, and G. Aubert, *Image sequence analysis via partial differential equations*, J. Math. Imaging Vision, 11 (1999), pp. 5–26.
[23] G. Kutyniok, *Clustered sparsity and separation of cartoon and texture*, SIAM J. Imaging Sci., 6 (2013), pp. 848–874.
[24] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, AMS, Providence, RI, 2002.
[25] Middlebury optical flow dataset, vision.middlebury.edu/flow/.
[26] T. Nir, R. Kimmel, and A. Bruckstein, *Variational Approach for Joint Optic-Flow Computation and Video Restoration*, Technical report, Department of Computer Science, Technion - Israel Institute of Technology, Haifa, Israel, 2005.
[27] S. Osher, A. Solé, and L. Vese, *Image decomposition and restoration using total variation minimization and the* H^{-1} *norm*, Multiscale Model. Simul., 1 (2003), pp. 349–370.
[28] T. Preusser, M. Droske, C. S. Garbe, A. Telea, and M. Rumpf, *A phase field method for joint denoising, edge detection, and motion estimation in image sequence processing*, SIAM J. Appl. Math., 68 (2007), pp. 599–618.
[29] W. Rudin, *Functional Analysis*, McGraw–Hill, New York, 1973.

[30] H. Schaeffer and S. Osher, *A low patch-rank interpretation of texture*, SIAM J. Imaging Sci., 6 (2013), pp. 226–262.
[31] S. Setzer, G. Steidl, and T. Teuber, *Infimal convolution regularizations with discrete l1-type functionals*, Commun. Math. Sci., 9 (2011), pp. 797–827.
[32] J.-L. Starck, M. Elad, and D. L. Donoho, *Image decomposition via the combination of sparse representations and a variational approach*, IEEE Trans. Image Process., 14 (2005), pp. 1570–1582.
[33] L. Vese and S. Osher, *Modeling textures with total variation minimization and oscillating patterns in image processing*, J. Sci. Comput., 19 (2002), pp. 553–572.
[34] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, *Image quality assessment: From error visibility to structural similarity*, IEEE Trans. Image Process., 13 (2004), pp. 600–612.
[35] Webpage of Martin Holler. http://www.uni-graz.at/~hollerm/.
[36] M. Werlberger, T. Pock, M. Unger, and H. Bischof, *Optical flow guided TV-L1 video interpolation and restoration*, in Energy Minimization Methods in Computer Vision and Pattern Recognition, Springer, Berlin, 2011, pp. 273–286.