

HABILITATIONSSCHRIFT  
eingereicht bei der  
Fakultät für Mathematik und Informatik  
der Ruprecht – Karls – Universität Heidelberg

Vorgelegt von  
Dr. rer. nat. Frank Lenzen  
aus Brühl

Juli 2015



# Adaptive Variational Regularization Techniques in Image Processing and Computer Vision

by  
Frank Lenzen

July 2, 2015



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Variational Methods . . . . .	2
1.2	Regularization with TV . . . . .	2
1.3	General Framework for Adaptive TV Regularization . . . . .	3
1.4	Related Work . . . . .	4
1.5	From Data-Driven to Solution-Driven Adaptivity . . . . .	5
1.6	Contributions . . . . .	8
1.7	Organization . . . . .	8
<b>2</b>	<b>Introduction to TV Regularization</b>	<b>11</b>
2.1	Notation . . . . .	11
2.2	Image Denoising . . . . .	11
2.3	The ROF Functional . . . . .	12
2.4	Theory . . . . .	14
2.5	Numerics . . . . .	15
2.6	Drawbacks of TV Regularization . . . . .	17
2.7	Measure-theoretic interpretation of TV . . . . .	17
2.8	Variants of TV Regularization . . . . .	20
2.8.1	Anisotropic TV Regularization . . . . .	20
2.8.2	Vector-valued TV . . . . .	20
2.8.3	Higher-Order TV . . . . .	21
2.8.4	Nonlocal TV Regularization . . . . .	22
2.8.5	Structure-Texture Decomposition . . . . .	22
<b>3</b>	<b>Variational Methods in Banach Spaces</b>	<b>25</b>
3.1	Notation . . . . .	25
3.2	Overview . . . . .	25
3.3	Existence Theory . . . . .	26
3.4	Fenchel Duality, Sub-Differential and Sub-Gradient . . . . .	28
3.5	Constrained Convex Problems and Variational Inequalities . . . . .	29

<b>4</b>	<b>Convex Adaptive TV Regularization</b>	<b>33</b>
4.1	Notation and Basic Definitions . . . . .	33
4.1.1	Multi-Indices for Partial Derivatives . . . . .	33
4.1.2	Gradient and Divergence-Operator of Order 1 . . . . .	33
4.1.3	Higher Order Total Variation . . . . .	34
4.2	Motivation . . . . .	35
4.3	General Framework . . . . .	38
4.4	Theory . . . . .	41
4.5	Examples of Data Terms . . . . .	45
4.6	Detecting Image Structures for Adaptivity . . . . .	50
4.6.1	Edge Detection . . . . .	50
4.6.2	Corner Detection . . . . .	51
4.6.3	Detection of Slope Discontinuities . . . . .	54
4.7	Examples of Adaptive Regularization . . . . .	57
4.7.1	First-Order TV Regularization . . . . .	57
4.7.2	Anisotropic TV of First- and Second-Order . . . . .	59
<b>5</b>	<b>Non-Convex Optimization</b>	<b>63</b>
5.1	Motivation . . . . .	63
5.2	Problem Relaxation . . . . .	64
5.3	Casting Variational Problems into Gradient Flows . . . . .	66
<b>6</b>	<b>QVIs for Regularization</b>	<b>71</b>
6.1	Introduction . . . . .	71
6.2	Approach . . . . .	73
6.3	Theory for the Continuous Case . . . . .	77
6.3.1	Well-Posedness of Operator T . . . . .	78
6.3.2	Existence of a Fixed Point . . . . .	79
6.4	Theory for QVIs . . . . .	87
6.4.1	Problem Statement . . . . .	87
6.4.2	Existence . . . . .	89
6.4.3	Uniqueness . . . . .	90
6.4.4	An Analytic Example . . . . .	95
6.5	Data Terms for Image Restoration . . . . .	97
<b>7</b>	<b>Numerics</b>	<b>99</b>
7.1	Algorithms for Convex Problems . . . . .	99
7.1.1	Primal-Dual . . . . .	100
7.1.2	FISTA . . . . .	101
7.2	An Algorithm for Solving QVIP . . . . .	104
7.2.1	Proposed Algorithm . . . . .	104
7.2.2	Convergence . . . . .	105
7.2.3	Numerical Experiments . . . . .	109
7.3	A Newton-like Algorithm for Non-convex Optimization Problems . . . . .	110

7.3.1	Introduction . . . . .	110
7.3.2	Preliminaries . . . . .	110
7.3.3	Optimization . . . . .	112
<b>8</b>	<b>Applications</b>	<b>115</b>
8.1	Image Restoration . . . . .	115
8.1.1	Denoising . . . . .	117
8.1.2	Non-blind Deblurring . . . . .	126
8.1.3	Inpainting . . . . .	138
8.1.4	Image Evolution by Solution-driven Adaptivity . . . . .	146
8.1.5	Relation to Non-Convex Regularization . . . . .	149
8.1.6	Appendix: Parameter Selection . . . . .	150
8.2	Displacement Regularization . . . . .	154
8.3	Denoising Time-of-Flight Data . . . . .	163
8.3.1	Introduction . . . . .	163
8.3.2	Approach . . . . .	166
8.3.3	Results . . . . .	174
8.4	Fusion of ToF and Stereo Sensors . . . . .	175
8.4.1	Camera Setup and Available Data . . . . .	175
8.4.2	Approach . . . . .	177
8.4.3	Experiments . . . . .	181
8.5	Upsampling of Color Images . . . . .	187
8.5.1	Problem Description . . . . .	187
8.5.2	Interpolation Constraints . . . . .	187
8.5.3	Related work . . . . .	188
8.5.4	Approach . . . . .	189
8.5.5	Experiments . . . . .	190
8.6	Joint Estimation of Scene and Egomotion . . . . .	196
8.6.1	Task . . . . .	198
8.6.2	Related Work . . . . .	198
8.6.3	A Model for Optical Flow . . . . .	199
8.6.4	Approach . . . . .	201
8.6.5	Numerics . . . . .	206
8.6.6	Experiments . . . . .	206
<b>9</b>	<b>Conclusion</b>	<b>215</b>
	<b>Nomenclature</b>	<b>218</b>
	<b>Figure Acknowledgements</b>	<b>221</b>
	<b>Acknowledgements</b>	<b>223</b>
	<b>Bibliography</b>	<b>225</b>





# Chapter 1

## Introduction

Various applications in image processing and computer vision are concerned with the reconstruction or restoration of a signal from measured data. Examples are the tasks of denoising, deblurring and inpainting, segmentation and labeling, the optical flow problem and depth estimation from stereo.

From a mathematical point of view, these reconstruction tasks constitute inverse problems, which for the considered applications are usually ill-posed. Consequently, a direct inversion of the underlying operator is infeasible. To solve these inverse problems, variational methods have proven favorable due to the fact that introducing regularization leads to well-posed optimization problems. Besides ensuring well-posedness, regularization in addition allows to incorporate *a-priori knowledge (a prior)* on the unknown signal to be reconstructed. Such a prior has a strong impact on the reconstruction quality and therefore, the choice of a meaningful regularization approach is crucial.

Due to the fact that signals in image processing and computer vision typically contain discontinuities, regularization with total variation (TV) respecting this property has become standard in this field and serves as starting point for various generalizations. The focus of this thesis lies on *adaptive* total variation regularization, where the prior is changing locally and, in particular, adjusts to instance-specific structures. This adaptivity improves the reconstruction quality compared to non-adaptive regularization.

One major contribution of this thesis is to provide a general model for adaptive TV regularization. Our model also covers recent trends proposed in the literature, such as to higher-order and non-local regularization. Consequently, our model combines the benefits of advanced total variation approaches with the advantages of adaptivity.

A second major contribution of this thesis is to present a novel concept for *solution-driven* adaptivity, i.e. an adaptivity which is determined by the unknown signal itself. In contrast, most adaptive approaches in the literature consider a data-driven adaptivity, i.e. an adaptivity steered by the input data. Our approach has the advantage that a-priori knowledge can be modeled in a more distinct way, which in practical applications further improves the reconstruction quality. Solution-driven

adaptivity, however, raises a more involved mathematical optimization problem. In this work, we provide a rigorous mathematical theory, which in particular answers the question of existence and uniqueness.

Before going into more detail, we briefly give some background on variational methods and TV regularization.

## 1.1 Variational Methods

Let us consider an inverse problem, where one is interested in finding an unknown function  $u$  solving the equation

$$A(u) = f, \tag{1.1}$$

with  $A$  being some linear/non-linear operator and function  $f$  representing some measured data. Typically, the problem of finding  $u$  satisfying (1.1) is ill-posed in the sense that a solution might not exist, might not be unique, or might not depend continuously on the input data  $f$  [71].

Examples for inverse problems in image processing are the tasks of denoising, deblurring and inpainting (i.e. the restoration of missing data, see e.g. [36]), where  $u$  represents the undistorted image, and in computer vision the stereo [73] and the optical flow problem [18], where  $u$  is a vector field encoding displacements in the image plane.

The standard approach to solve inverse problems is to consider an optimization problem of the form

$$\min_{u \in X} \mathcal{F}(u), \quad \mathcal{F}(u) := \mathcal{S}(u) + \alpha \mathcal{R}(u), \tag{1.2}$$

where minimization over an appropriate function space  $X$  is considered.

The term  $\mathcal{S}(u)$  in (1.2) is the *data* or *fidelity* term penalizing the difference between  $A(u)$  and data  $f$ . The term  $\mathcal{R}(u)$  is the *regularization term*. Introducing a suitable regularization term, in particular, asserts that the minimization problem is well-defined. The *regularization parameter*  $\alpha > 0$  steers the coupling between both terms.

## 1.2 Regularization with TV

Regularization approaches commonly used in image processing and computer vision are based on the total variation semi-norm [5, 55],

$$\text{TV}(u) := \sup \left\{ \int u \operatorname{div} \varphi \, dx \mid \varphi \in C_c^1(\Omega, \mathbb{R}^2), \|\varphi(x)\|_2 \leq 1 \right\}, \tag{1.3}$$

and several variants of it. Total variation was first introduced as regularizer for the task of image denoising by Rudin, Osher and Fatemi [142]. One important reason for TV to become the standard regularization method in these fields is that

it allows a weak concept of differentiability (weaker than differentiability in Sobolev spaces), which also includes piecewise constant functions with finitely many jumps. Consequently, using

$$\mathcal{R}(u) := \alpha \text{TV}(u) \tag{1.4}$$

as a regularizer allows discontinuities of  $u$  up to a certain degree. We remark that for (weakly) differentiable functions, (1.4) simplifies to

$$\mathcal{R}(u) := \int_{\Omega} \alpha \|\nabla u(x)\|_2 \, d\mathcal{L}. \tag{1.5}$$

Total variation has advantageous theoretical properties. Firstly,  $\text{TV}(u)$  is convex, which eases both theoretical and numerical analysis of this approach. Secondly, the space of functions of bounded total variation provides sufficient conditions to show existence of a solution.

For the reasons mentioned above, TV serves as a well-suited starting point for various generalizations, for example in view of higher-order models, non-local operators or adaptive TV regularization.

### 1.3 General Framework for Adaptive TV Regularization

Based on the standard TV regularization (1.4), we introduce a general model of adaptive TV regularization of arbitrary order. To this end, we generalize (1.3) by introducing arbitrary local constraints  $\mathcal{D}^{loc}(x) \subset \mathbb{R}^m$  (cf. [53]) and an arbitrary differential operator  $D$ , i.e.

$$\mathcal{R}(u) := \sup \left\{ \int u D\varphi \, dx \mid \varphi \in C_c^\infty(\Omega, \mathbb{R}^m), \varphi(x) \in \mathcal{D}^{loc}(x) \right\}, \tag{1.6}$$

with the prerequisite that the sets  $\mathcal{D}^{loc}(x)$  are closed, non-empty and convex.

The above generalization allows us to

1. locally vary the regularization strength (size of  $\mathcal{D}^{loc}(x)$ ),
2. introduce an anisotropic (or directional) regularization (shape of  $\mathcal{D}^{loc}(x)$ ),
3. consider higher-order total variation (operator  $D$ ).

In particular, these adaptations enable us to promote the restoration of certain structures contained in  $u$  (e.g. edges, corners in image data), provided that we have some a-priori information about where to expect these image structures.

To model this a-priori information, we introduce an additional parameter  $v$  for the regularization term, i.e  $\mathcal{R}(u; v)$ . As an example, we consider the model of weighted total variation, where we set

$$\alpha(x; v) := \alpha_0 \frac{1}{1 + \kappa \|\nabla v(x)\|_2^2} \tag{1.7}$$

for some  $\alpha_0 > 0$ ,  $\kappa \geq 0$  and function  $v$  as introduced above. The function (1.7) was originally proposed by Perona and Malik [132] in the field of diffusion methods. Based on the function  $\alpha(x; v)$ , we choose local constraint sets  $\mathcal{D}^{loc}(x, v) = B_{\alpha(x; v)}(0)$ , where  $B_r(x)$  is the closed ball (w.r.t. the Euclidean norm) of radius  $r$  centered at  $x$ . We observe that with these sets and  $D = \text{div}$  the regularization term (1.6) for *differentiable* functions  $u$  simplifies to

$$\mathcal{R}(u; v) = \int_{\Omega} \alpha(x; v) \|\nabla u(x)\|_2 \, dx, \quad (1.8)$$

(cf. (1.5)), i.e. we have introduced a locally varying regularization parameter  $\alpha$ , which changes depending on the absolute value of the gradient of  $v$ .

With the generalized regularization term  $\mathcal{R}(u; v)$ , the considered optimization problem (1.2) becomes

$$\min_{u \in X} \mathcal{F}(u; v), \quad \mathcal{F}(u; v) := \mathcal{S}(u) + \mathcal{R}(u; v). \quad (1.9)$$

We discuss our generalized approach as well as different models of adaptivity in detail in Chapter 4.

## 1.4 Related Work

Since this thesis focuses on adaptive TV regularization, we concentrate on work related in this context.

We can roughly classify the proposed adaptive regularizers based on *first-order TV* into four classes. The first class consists of approaches which estimate the *scalar* regularization parameter  $\alpha$  [106, 128, 184]. The second, and probably the largest class of approaches, deals with a *spatially adaptive* regularization parameter  $\alpha(x)$  (e.g. (1.8) in the case of Sobolev functions) [43, 67, 76, 99, 100, 103, 109, 138, 141, 156, 158, 159, 170, 180]. In the related literature, those methods are referred to as *weighted TV*, *adaptive TV* (ATV), and *spatially adaptive TV* (SATV). In this thesis, we use the term *weighted TV*. The third class of approaches, see e.g. [25, 26, 42, 59, 185, 186], considers a spatially changing *exponent*  $q(x)$  for  $\|\nabla u(x)\|_2$  in (1.5), i.e.  $\int_{\Omega} \|\nabla u(x)\|^{q(x)} \, dx$ . Finally, there exists a class of TV regularizers, which penalize the image gradient depending on its *orientation*. We refer to such methods as *anisotropic* methods. Sometimes these methods are also termed *directional TV* regularization. Such anisotropic approaches have been proposed e.g. in [7, 14, 22, 54, 68, 99, 100, 103, 90, 91, 152, 154].

Following the recent development towards higher-order total variation approaches (see e.g. [28, 153]), also *adaptive higher-order TV* methods have been proposed [38, 80, 94, 98]. Moreover, we remark that non-local total variation [65, 84] by its nature is adaptive with respect to the data. Besides being applied in the field of image restoration, adaptive TV has been used to regularize a broad set of problems. Among these, we mention optical flow estimation [172, 178], de-interlacing [82],

compressed imaging [40], segmentation and labeling [93, 96, 125, 143, 183] and sensor fusion [118].

The approaches mentioned above use different models for determining meaningful structures in the data. We summarize the most common ones. In case of weighted TV, several approaches use model (1.7) with  $v$  being the pre-smoothed input data [43, 109, 170]. Other approaches, in particular several anisotropic models, build on the structure tensor [57] evaluated on either the input data  $f$ , a current estimate of  $u$ , or the unknown  $u$  itself [54, 68, 90, 91, 99, 100, 103, 141, 152, 154, 180]. Due to the use of the structure tensor, we observe a relationship to the anisotropic diffusion approach by Weickert [175].

We emphasize that a wide range of the approaches mentioned above, e.g. the weighted TV approaches [43, 76, 80, 109, 141, 170, 180], the anisotropic TV approaches [14, 22, 53, 99, 100, 103, 152, 154] as well as higher-order and non-local TV models [65, 84, 98], are covered by our general framework (1.6) by choosing appropriate local constraint set  $\mathcal{D}^{loc}$  and differential operators  $D$ .

## 1.5 From Data-Driven to Solution-Driven Adaptivity

In the following, we give an overview of the proposed solution-driven strategy. To this end, let us focus on the function  $v$  determining the adaptivity of  $\mathcal{R}(u; v)$ . Without loss of generality we make the assumption that the function  $v$  is of the same kind as  $f$  or  $u$  (e.g. that  $u, v, f$  represent images).

We distinguish between two different cases for the choice of  $v$ . The first one is *data-driven adaptivity*, where we choose  $v = f$ , i.e. structures in  $f$  (e.g. edges, corners, slopes) are used to steer the adaptivity. In the second case, which is *solution-driven adaptivity*, we choose  $v = u$ , i.e. structures in the unknown  $u$  determine the adaptivity.

The latter case has the advantage that the adaptivity is determined directly by the reconstructed function and thus, is less affected by distortions present in the input data. Consequently, no pre-processing for removing these distortions is required. We will see in the discussion of various applications, that, in general, *solution-driven adaptivity* provides a *higher reconstruction quality* than data-driven adaptivity. These benefits come along with a more complex setting from a mathematical point of view. We therefore provide *rigorous mathematical theory* for this new ansatz.

Before turning to the theory for solution-driven adaptivity, let us first present theoretical and practical properties of both approaches in more detail.

### Data-Driven Adaptivity

With the choice  $v = f$ , the local constraint sets  $\mathcal{D}^{loc}(x; v)$  in (1.6) become fixed. Then, provided that these sets are convex, a convex minimization problem is obtained. Theory on existence and uniqueness of minimizers, which we present in Chapter 4, therefore is analogously to the non-adaptive case.

The convexity of  $\mathcal{F}(u)$  and the availability of efficient numerical solvers might be the reason why data-driven adaptivity is the most common form of adaptivity used in the literature. From a practical point of view, however, the data-driven approaches face the problem that  $f$  might be corrupted e.g. by noise, and important image structures might be missing in  $f = A(u)$  in case of  $A \neq \text{Id}$ .

We present the general framework of data-driven adaptive TV regularization in detail in Chapter 4 along with specific models of adaptivity, that allow the restoration of prominent image structures in the data such as image edges, corners and intensity slopes.

### Solution-Driven Adaptivity

Let us first consider the *direct* approach, which addresses the minimization problem

$$\min_{u \in X} \mathcal{F}(u; u). \quad (1.10)$$

We remark that convexity of problem (1.10) is in general not guaranteed. For the models of adaptivity we study in this thesis (cf. Chapter 4), we will see that the *direct* solution-driven approach (1.10) would always result in a *non-convex* optimization problem. An example is choosing  $v = u$  in (1.7). A consequence of the non-convexity is, that the direct approach leads to problems, when it comes to showing existence and uniqueness and providing convergent numerical methods.

To circumvent the difficulties with non-convex minimization problems, we can think of relaxing the problem. By relaxation, the minimization problem becomes convex, while the set of minimizers in general is enlarged. We will describe a class of relaxation methods in Chapter 5. For the case of adaptive total variation regularization, however, it turns out that relaxation is not a suitable strategy, since the optimization of the relaxed functional is an open issue.

We therefore choose a different strategy, in which we pose the problem of solution-driven adaptive regularization in terms of a *fixed point approach* as detailed next. As a starting point, we consider an arbitrary  $u_0 \in X$  and the mapping

$$u_0 \mapsto T(u_0) := \arg \min_{u \in X} \mathcal{S}(u) + \mathcal{R}(u; u_0). \quad (1.11)$$

Having  $u_0$  fixed, existence theory from the data-driven case applies and thus, operator  $T$  is well-defined. Moreover, assuming that the data term  $\mathcal{S}(u)$  is strictly convex,  $T$  is single-valued (a property which is required to define adaptivity in a meaningful way).

We then search for a fixed point  $u^*$  of operator  $T$ . Please note that  $u^*$  is the solution of

$$\min_{u \in X} \mathcal{S}(u) + \mathcal{R}(u; u^*), \quad (1.12)$$

i.e. the adaptivity of  $\mathcal{R}$  is determined by  $u^*$ . Thus, our ansatz provides a solution-driven adaptivity.

It has to be remarked that the problem (1.10) and the fixed point problem are not equivalent in general. However, we observe that the fixed point approach, depending on the mapping  $v \mapsto \mathcal{R}(\cdot; v)$ , mimics a non-convex behavior (cf. Section 8.1).

Considering related work, we find that only a few *solution-driven* regularization approaches have been proposed [7, 54, 90, 138]. The approaches proposed by Lefkimiatis et al. [90] based on the structure tensor and Åström et al. [7] based on the gradient energy tensor consider solution-driven adaptivity in a direct way. In both cases, the authors have shown convexity of their regularizers. Both approaches, however, face the problem that they are only defined on function spaces of sufficient regularity, i.e. Sobolev spaces, while the generalization to the function space  $BV$  as a suitable space for proving existence of a minimizer is unclear.

The methods by Estellers et al. [54] and Rojas et al. [138] propose an iterative update of the adaptivity. Therefore, these can be regarded as the most related ones to our work. However, theory on the existence of fixed points for these approaches is not provided. In contrast, for the approach presented here, we provide theory for fixed points in a suitable function space.

### Theory for Solution-Driven Adaptivity

For the operator  $T$  introduced above, we provide existence theory for fixed points both in the continuous and discrete setting, cf. Sections 6.3 and 6.4, respectively.

In the *continuous* setting, we build on our previous work [101]. The main ingredient for proving existence is the Himmelberg fixed point theorem for compact, weakly sequentially upper semi-continuous operators [3, 74].

In the *discrete* setting, we observe that the problem of finding a fixed point of operator  $T$  is equivalent to solving the quasi-variational inequality problem (QVIP)

$$\begin{aligned} \text{find } p^* \in \mathcal{D}(p^*) \text{ such that} \\ \langle G(p^*), p - p^* \rangle \geq 0 \text{ for all } p \in \mathcal{D}(p^*), \end{aligned} \tag{1.13}$$

where  $G$  comprises the gradient of the dual objective of  $u \mapsto \mathcal{F}(u; u^*)$  and  $\mathcal{D}$  is a constraint set related to the local constraint sets  $\mathcal{D}^{loc}(x, u^*)$ , but now depending on the unknown  $p^*$ . For details, we refer to Chapter 6.

We show that under sufficient conditions a solution  $p^*$  of (1.13) exists, from which existence of a fixed point  $u^*$  of  $T$  follows. Here, we can make use of existence theory for QVIPs from the literature, see e.g. [37, 122, 127].

In view of *uniqueness*, QVIP (1.13) exhibits difficulties due to missing strong monotonicity of the operator  $G$ . Consequently, theory from the literature (see e.g. [122]) on the uniqueness of  $p^*$  is not applicable. However, we can close this gap and show uniqueness of  $u^*$  under sufficient assumptions. Both existence and uniqueness theory is described in detail in Section 6.4. This extends our previous work [99, 100, 103].

Our theoretical results show that our fixed point strategy is not just an ad hoc approach, but is supported by rigorous mathematical theory.

Finally, we remark that our concept of solution-driven adaptivity can be applied to various existing data-driven approaches. In particular, the adaptive methods in [14, 22, 43, 53, 76, 80, 109, 141, 152, 154, 170, 180] can directly be transferred into solution-driven methods by our approach. In these cases, the theory presented in Chapter 6 applies.

## 1.6 Contributions

We briefly summarize the contributions of our work:

1. We provide a general model for adaptive TV regularization of arbitrary order, which covers various data-driven approaches proposed in the literature. For the general model, we provide theory on existence and uniqueness.
2. We propose a novel concept of solution-driven adaptivity, which is supported by rigorous mathematical theory both in the continuous and discrete setting. Our concept can be directly applied to various data-driven approaches existing in the literature.
3. We discuss several models of adaptivity, which can be used to preserve prominent image structures such as edges, corners and slopes when reconstructing the unknown function  $u$ . Each model can be applied in a data-driven or a solution-driven way.
4. We provide a numerical method for solving the solution-driven regularization approaches. This algorithm is supported by convergence theory.
5. We present various applications in the field of image processing and computer vision in which adaptive regularization is used. These applications clearly demonstrate that adaptive regularization is preferable over non-adaptive regularization. Moreover, we show applications where the solution-driven variants clearly outperform the data-driven variants.

## 1.7 Organization

Let us give a short overview of the organization of this thesis: We start with an introduction to TV regularization in Chapter 2. Chapter 3 provides background on variational methods and convex optimization. In Chapter 4, we focus on data-driven adaptive TV regularization, which results in convex minimization problems. We discuss our general framework for adaptivity and provide existence and uniqueness results. Moreover, we discuss models to steer adaptivity and various examples of adaptive TV regularization. Chapter 5 recalls relaxation approaches in  $W^{1,p}(\Omega)$  and  $BV(\Omega)$ , which in certain cases can serve as an alternative to tackle non-convex minimization problems. In Chapter 6, we present our fixed point strategy for solution-driven adaptive regularization. This strategy enables us to transfer the data-driven



adaptive approaches from Chapter 4 into solution-driven approaches. We stress that in the context of adaptive regularization, this strategy is preferable over relaxation methods. In Chapter 7, we consider numerical approaches for convex, non-convex, and fixed point problems. In particular, we present a new algorithm for solving QVIPs with non-strongly monotone operators (Section 7.2) and show convergence. Moreover, we provide a Newton-like algorithm for solving non-convex optimization problems on Lie groups (Section 7.3). Finally, in Chapter 8, we discuss applications using the framework presented in Chapters 4 to 6 and the numerical methods from Chapter 7. The considered applications comprise image restoration (denoising, deblurring, inpainting), displacement regularization (i.e. correcting errors in the location of sampling points), denoising and fusion of time-of-flight depth maps, upsampling of color images and, estimation of scene and egomotion from image sequences in an automotive environment. We conclude this thesis with a discussion in Chapter 9.



# Chapter 2

## Introduction to Total Variation Regularization

### 2.1 Notation

Throughout this chapter, let  $\Omega \subset \mathbb{R}^2$  be an open bounded domain with Lipschitz-boundary. The integral of some  $f$  over  $\Omega$  is denoted by  $\int_{\Omega} f \, d\mathcal{L}$ , where we omit the dependency of  $f$  on  $(x, y)$ . The (weak) derivatives of  $f$  are denoted by  $\partial_x f$  and  $\partial_y f$ .

### 2.2 Image Denoising

In the process of recording image data, it is inevitable that the recording is distorted by random errors, which commonly is referred to as *noise*. This process is typically described in a discrete setting. When the signal is recorded at pre-defined sampling points  $x_i \in \Omega$ ,  $i = 1, \dots, n$ , the original signal  $u(x_i)$  is distorted by a random process. For the moment, we concentrate on grayscale images and the model of an *additive* distortion, such that the recorded data  $f(x_i) \in \mathbb{R}$  are given as

$$f(x_i) := u(x_i) + v(x_i), \quad (2.1)$$

where  $v(x_i)$  are the realizations of some random variables. We define  $u_i := u(x_i)$  and analogously for  $v, f$ . (Please note that we denote the resulting vectors also by  $u, v, f$ . In the following, the respective setting will be clear from the context.)

Another often considered noise model is the model of multiplicative noise (see e.g. [106, 155] and the reference therein), which via the log-function can be transformed into an additive model.

The standard model for additive noise is, that the random process is an independent and identically distributed (i.i.d.) Gaussian process with zero mean and standard deviation  $\sigma > 0$ , i.e.

$$v(x_i) \sim \mathcal{N}(0, \sigma) \quad (2.2)$$

independent from the location  $x_i$ . In order to retrieve a noise-free approximation  $\bar{u}$  of the recorded data  $f$ , variational methods motivated by the statistical approach of maximum-a-posteriori (MAP) estimation [23, 63] can be used:

$$\bar{u} := \arg \min_{u \in \mathbb{R}^n} \frac{1}{2} \|u - f\|_2^2 + \mathcal{R}(u), \quad (2.3)$$

where the first term is motivated by the Gaussian noise model and  $\mathcal{R} : \mathbb{R}^n \rightarrow \mathbb{R}^{\geq 0}$  is a functional obtained from an a-priori assumption on the distribution of the signal  $u$ . We refer to [23, 97] for further details.

The above ansatz can also be considered in a continuous setting, where  $u, v$  and  $f = u+v$  are assumed to be functions in  $L^2(\Omega)$ . The variational approach introduced above then reads

$$\bar{u} := \arg \min_{u \in X} \frac{1}{2} \int_{\Omega} (u - f)^2 d\mathcal{L} + \mathcal{R}(u), \quad (2.4)$$

where  $X \subseteq L^2(\Omega)$  is some suitable function space on which  $\mathcal{R}(u) : X \rightarrow \mathbb{R}_{\geq 0}$  is well-defined. We remark that in the continuous setting a mathematically rigorous definition of the noise process generating  $v$  is non-trivial. However, describing such a model is out of the scope of this thesis.

The problem of denoising can also be studied from the viewpoint of the theory of *inverse problems* [148], where the process of generating noisy data  $f$  from given  $u$  is referred to as the *forward* and the problem of estimation  $u$  from given  $f$  as the *inverse* problem.

In this view, we refer to the first term in (2.4) as the *data term*, for which we introduce the notation

$$\mathcal{S}(u) := \frac{1}{2} \int_{\Omega} (u - f)^2 d\mathcal{L} = \frac{1}{2} \|u - f\|_{L^2}^2. \quad (2.5)$$

The term  $\mathcal{R}$  in (2.4) is referred to as the *regularization term*.

## 2.3 The ROF Functional

For the task of image denoising, Rudin, Osher and Fatemi [142] proposed in 1992 a variational approach where the total variation (TV) semi-norm is used as regularization term. The functional to be minimized, which is commonly referred to as the *ROF functional*, reads follows: for given  $f \in L^2(\Omega)$ ,  $\alpha > 0$  let

$$\mathcal{F}(u) := \frac{1}{2} \|u - f\|_{L^2}^2 + \alpha \text{TV}(u), \quad (2.6)$$

where

$$\text{TV}(u) := \sup \left\{ \int_{\Omega} u \operatorname{div} \varphi \mid \varphi \in C_c^1(\Omega, \mathbb{R}^d), \|\varphi(x)\|_2 \leq 1 \right\} \quad (2.7)$$

is the total variation semi-norm [55, 148]. We recall that (2.6) takes the classical form of a variation approach to solve the inverse problem (cf. Chapter 1) of retrieving  $u$  from  $f$ , where  $\mathcal{S}(u) := \frac{1}{2}\|u - f\|_{L^2}^2$  is the data term and  $\alpha \text{TV}(u)$  serves as regularization term  $\mathcal{R}(u)$ . Both terms are coupled by parameter  $\alpha$ , which is referred to as the *regularization strength* or *regularization parameter*. Since the introduction of the ROF functional TV has become very popular as regularization term also for other inverse problems in image processing and computer vision.

Let us discuss the properties of the functional in (2.6). First, we determine an appropriate function space on which this functional well defined. We start with the regularization term  $\alpha \text{TV}(u)$  in (2.6). Please note that in the definition of  $\text{TV}(u)$  in (2.7) the integral is well-defined for functions  $u \in L^1(\Omega)$ ; however, the supremum might take the value  $+\infty$ . Let us therefore consider the subset (subspace) of  $L^1(\Omega)$ , for which  $\text{TV}(u)$  is finite:

**Definition 2.1.** [Space  $BV(\Omega)$ ] Let  $u \in L^1(\Omega)$  and let the total variation of  $u$ ,  $\text{TV}(u)$  be defined as in (2.7). We call the space of functions

$$BV(\Omega) := \{u \in L^1(\Omega) \mid \text{TV}(u) < \infty\} \quad (2.8)$$

the space of functions of bounded variation.

Equipped with the norm  $\|u\|_{BV} := \|u\|_{L^1} + \text{TV}(u)$  the space  $BV(\Omega)$  becomes a Banach space.

Now, let us turn to the data term in (2.6). Please note that in the case  $\Omega \subset \mathbb{R}^2$ , on which we focus here,  $L^2(\Omega)$  is continuously embedded in  $BV(\Omega)$  and thus (2.6) is finite on  $BV(\Omega)$ . Consequently, the appropriate space to minimize (2.6) over is  $BV(\Omega)$ :

$$\min_{u \in BV(\Omega)} \mathcal{F}(u) := \min_{u \in L^2} \frac{1}{2}\|u - f\|_{L^2}^2 + \alpha \text{TV}(u). \quad (2.9)$$

One might think of further restricting the functional to the space  $W^{1,1}(\Omega) \subset BV(\Omega)$ , on which

$$\text{TV}(u) = \int_{\Omega} \|\nabla u\|_2 \, d\mathcal{L} \quad (2.10)$$

(see also Section 2.7 below). However, this function space does not have the required properties to show existence of a minimizer.

The ROF functional has the properties of being non-negative, proper, i.e.  $\mathcal{F}(u) < \infty$  for at least one element  $u \in BV(\Omega)$  and strictly convex. The non-negativity follows from the fact that both data term and regularization term are non-negative, where  $\text{TV}(u) \geq 0$  follows from the fact that  $p \equiv 0$  belongs to the set of functions considered on the right hand side of (2.7). Properness follows e.g. by choosing  $u \equiv 0$ . Finally, since  $\text{TV}(u)$  is the supremum over convex functionals, it is also convex. The strict convexity of  $\mathcal{S}(u)$  then implies the strict convexity of the ROF functional.

It remains to discuss the issue of existence of a unique minimizer. This will be part of the next section.

## 2.4 Theory

We briefly recall the main steps to prove the existence of a minimizer of

$$\min_{u \in BV(\Omega)} \mathcal{F}(u) := \min_{u \in BV(\Omega)} \frac{1}{2} \|u - f\|_{L^2}^2 + \alpha \text{TV}(u). \quad (2.11)$$

Uniqueness then follows from the strict convexity of  $\mathcal{F}(u)$ .

The strategy to prove existence is as follows (see e.g. [49]). We consider a minimizing sequence  $(u^{(k)})_k$ , i.e.  $\mathcal{F}(u^{(k)}) \rightarrow \inf_{u \in BV(\Omega)} \mathcal{F}(u)$ . After showing boundedness of this sequence in the norm  $\|\cdot\|_{BV}$ , a subsequence, also denoted by  $(u^{(k)})_k$  has to be found, which converges to some  $\bar{u} \in BV(\Omega)$  in a weak sense. Here, we require the property of  $BV(\Omega)$ , that bounded sets are pre-compact with respect to this weak convergence. (This requirement is the reason why it does not suffice to consider minimization of  $\mathcal{F}$  over  $W^{1,1}(\Omega)$ .) Finally, it has to be shown that  $\bar{u}$  is a minimizer of  $\mathcal{F}(u)$ :

$$\mathcal{F}(\bar{u}) \stackrel{!}{=} \inf_k \mathcal{F}(u^{(k)}) = \inf_{u \in BV(\Omega)} \mathcal{F}(u), \quad (2.12)$$

where the second equality holds due to the fact that we chose a minimizing sequence.

We address each step in a little more detail:

**Boundedness** Since  $\mathcal{F}(0) = \frac{1}{2} \|f\|_{L^2}^2 =: C$ , we can assume without loss of generality that  $\mathcal{F}(u^{(k)}) \leq C$  for all  $k \in \mathbb{N}$ . For arbitrary  $u$  we find from

$$\|u\|_{L^2} \leq \|u - f\|_{L^2} + \|f\|_{L^2} \leq \sqrt{2\mathcal{F}(u)} + \sqrt{2C} \quad (2.13)$$

and

$$\text{TV}(u) \leq \frac{1}{\alpha} \mathcal{F}(u), \quad (2.14)$$

that

$$\|u\|_{BV(\Omega)} \leq \sqrt{2\mathcal{F}(u)} + \sqrt{2C} + \frac{1}{\alpha} \mathcal{F}(u) \quad (2.15)$$

(coercivity of  $\mathcal{F}(u)$ ). It follows, that

$$\|u^{(k)}\|_{BV(\Omega)} \leq 2\sqrt{2C} + \frac{C}{\alpha}. \quad (2.16)$$

**Pre-compactness of bounded sets** We consider the following weak convergence in  $BV(\Omega)$ :

**Definition 2.2.** [Weak\*-convergence] A sequence  $(u^{(k)})_k$  weakly\* converges to  $u \in BV(\Omega)$ , denoted by  $u^{(k)} \rightharpoonup u$ , if and only if  $\|u^{(k)} - u\|_{L^1} \rightarrow 0$  and  $\sup_k \text{TV}(u^{(k)}) < \infty$ .

Furthermore, we consider the weak topology induced by this weak convergence. It can be shown that each bounded set in  $BV$  is pre-compact with respect to this weak topology [55].

Since  $(u^{(k)})_k$  is bounded, there exists some  $\bar{u} \in BV(\Omega)$ , such that  $u^{(k)} \overset{*}{\rightharpoonup} \bar{u}$ . This implies also that  $\bar{u} \in L^2(\Omega)$  and  $u^{(k)} \rightharpoonup \bar{u}$  in  $L^2(\Omega)$ .

**Minimizing property of  $\bar{u}$**  We have to show that

$$\mathcal{F}(\bar{u}) \leq \inf_k \mathcal{F}(u^{(k)}). \quad (2.17)$$

Firstly, we observe that  $\mathcal{S}(u)$  is weakly lower semi-continuous in  $L^2(\Omega)$ , i.e.

$$\mathcal{S}(\bar{u}) \leq \lim_{u^k \rightharpoonup \bar{u}} \mathcal{S}(u^k) \quad (2.18)$$

for sequences  $(u^k)_k$  weakly converging to  $u$  in  $L^2(\Omega)$ , see Lemma 2.3 below. Secondly, the TV functional is weakly\* lower semi-continuous (see e.g. [55, Theorem 1 in Section 5.2.]), i.e.

$$\text{TV}(\bar{u}) \leq \liminf_{u^{(k)} \overset{*}{\rightharpoonup} \bar{u}} \text{TV}(u^{(k)}). \quad (2.19)$$

Combining (2.18) and (2.19) shows (2.17). In our case, where  $(u^{(k)})_k$  is a minimizing sequence, the right hand side of (2.17) equals  $\inf_{u \in BV(\Omega)} \mathcal{F}(u)$ . Thus, it follows immediately that  $\bar{u}$  is a minimizer of  $\mathcal{F}(u)$ .

**Lemma 2.3.** *Consider  $f \in L^2(\Omega)$  fixed. The functional  $\mathcal{S}(u) := \frac{1}{2} \|u - f\|_{L^2}^2$  defined on  $L^2(\Omega)$  is weakly lower semi-continuous.*

*Proof.* We show the lower semi-continuity of  $\mathcal{S}$  in  $L^2(\Omega)$ , which, since  $\mathcal{S}$  is convex, induces its weak lower semi-continuity by Theorem 1.2 in [49]. To this end, let  $(u^{(k)})_k, \bar{u} \in L^2(\Omega)$ , such that  $u^{(k)} \rightharpoonup \bar{u}$ . In particular,  $(u^{(k)})_k$  is bounded in  $L^2(\Omega)$ . Then,

$$\begin{aligned} \mathcal{S}(\bar{u}) - \mathcal{S}(u^{(k)}) &= \frac{1}{2} \int_{\Omega} (\bar{u} - f)^2 - (u^{(k)} - f)^2 \, d\mathcal{L} \\ &\leq \frac{1}{2} \int_{\Omega} |(u^{(k)} - \bar{u})(u^{(k)} + \bar{u} - 2f)| \, d\mathcal{L} \\ &\leq \frac{1}{2} \|u^{(k)} - \bar{u}\|_{L^2} \|u^{(k)} + \bar{u} - 2f\|_{L^2} \\ &\leq \frac{1}{2} \|u^{(k)} - \bar{u}\|_{L^2} (\|u^{(k)}\|_{L^2} + \|\bar{u}\|_{L^2} + 2\|f\|_{L^2}). \end{aligned} \quad (2.20)$$

Since  $(\|u^{(k)}\|_{L^2} + \|\bar{u}\|_{L^2} + 2\|f\|_{L^2})$  is bounded and  $\|u^{(k)} - \bar{u}\|_{L^2} \rightarrow 0$ , it follows that

$$\liminf_{k \rightarrow \infty} \mathcal{S}(\bar{u}) - \mathcal{S}(u^{(k)}) \leq 0. \quad (2.21)$$

□

## 2.5 Numerics

Concerning the numerical optimization of  $\mathcal{F}$ , one particular issue is the non-smoothness of the regularization term. Some algorithms proposed in the literature circumvent this problem by smoothing the regularization term, i.e. they consider

$$\text{TV}_{\varepsilon}(u) := \int_{\Omega} \sqrt{(\partial_x u)^2 + (\partial_y u)^2 + \varepsilon^2} \, d\mathcal{L} \quad (2.22)$$

---

**Algorithm 1:** A primal-dual algorithm for minimizing the ROF functional.

---

**Input:**  $N \geq 1$ , data  $f$ , initial guess  $u^{(0)} = f$ ,  $\tilde{u}^{(0)} = u$ ,  $p^{(0)} \in \mathbb{R}^{2n}$

**Output:**  $u = u^{(N)}$

Let  $\tau, \sigma > 0$  such that  $\tau\sigma \leq \frac{1}{\|L\|^2}$

**begin**

<p><b>for</b> <math>i = 1, \dots, N</math> <b>do</b></p>	<p><math>p^{(i)} \leftarrow \Pi_{\mathcal{D}}(p^{(i-1)} + L^\top \tilde{u}^{i-1})</math>      // update dual variable</p> <p><math>u^{(i)} \leftarrow \frac{1}{1+\tau}(u^{(i)} - \tau L p^{(i)} + \tau f)</math>      // update primal variable</p> <p><math>\tilde{u}^{(i)} \leftarrow 2u^{(i)} - u^{(i-1)}</math>      // extrapolation step</p>
--	--

---

for some  $\varepsilon > 0$ , and solve the Euler-Lagrange equation for the smoothed minimization problem in a semi-implicit scheme. In order to avoid a blurring of the result, a small  $\varepsilon$  has to be used.

An alternative, which became popular in the recent years, is to make use of the formulation of the total variation via the supremum of linear functions (cf. (2.7)) and to tackle the saddle point problem

$$\min_{u \in BV(\Omega)} \frac{1}{2} \|u - f\|_{L^2}^2 + \sup_{p \in C^1(\Omega; \mathbb{R}^2), \|p(x)\|_2 \leq \alpha} \int_{\Omega} \operatorname{div} p u \, d\mathcal{L}. \quad (2.23)$$

In the discrete setting (2.23) turns into

$$\min_{u \in \mathbb{R}^n} \frac{1}{2} \|u - f\|_2^2 + \sup_{p \in \mathcal{D}} (Lp)^\top u, \quad (2.24)$$

where  $u, f \in \mathbb{R}^n$  are vector representations of the functions  $u$  and  $f$  on a regular grid (e.g. ordered lexicographically),  $L : \mathbb{R}^{2n} \rightarrow \mathbb{R}^n$  is a discretization of the divergence operator and

$$\mathcal{D} := \{p = (p^x, p^y) \in \mathbb{R}^{2n} \mid \forall i = 1, \dots, n : \sqrt{(p_i^x)^2 + (p_i^y)^2} \leq \alpha\}. \quad (2.25)$$

Since  $\mathcal{D}$  is a closed and convex set, the projection  $\Pi_{\mathcal{D}}(p)$  of  $p \in \mathbb{R}^{2n}$  onto this set is well-defined.

**Remark 2.4.** For the discretization we choose left-sided finite differences together with homogeneous Dirichlet boundary conditions, such that  $G^\top := -L^\top$  is a discretization of the gradient with right-sided finite differences and homogeneous Neumann boundary conditions. Moreover, note that

$$\sup_{p \in \mathcal{D}} (Lp)^\top u = \sup_{p \in \mathcal{D}} p^\top Gu \quad (2.26)$$

as long as the set  $\mathcal{D}$  is point-symmetric with respect to 0.



We briefly recall one basic algorithm, see Algorithm 1, to numerically solve the saddle point problem (2.24). This algorithm is taken from [36].

The algorithm consist of iterated updates of the dual variable  $p$  and the primal variable  $u$ , combined with an extrapolation step. The update of  $p$  involves the projection  $\Pi_{\mathcal{D}}(p)$  of  $p = (p_x, p_y) \in \mathbb{R}^{2n}$  onto the convex set  $\mathcal{D}$  in (2.25). This projection can be performed for each index  $i$  independently by scaling  $(p_i^x, p_i^y)$  by a factor  $\min(\alpha/\|(p_i^x, p_i^y)\|, 1)$ . Moreover, we remark that  $\|L\|^2 \leq 8$  (cf. e.g. [34]). This bound is sharp in the sense that it is attained in the limit for increasing image size.

A convergence proof for Algorithm 1 can be found in [36]. We will consider more sophisticated algorithms, i.e. algorithms which are faster and have a-posteriori error estimates, in Chapter 7.

## 2.6 Drawbacks of TV Regularization

There exist three well-known drawbacks of using TV for regularization . The first one is the loss of contrast, which is caused by the fact that any kind of slope or jump in the data is penalized, which from an optimization point of view favors  $u < f$  at local maxima of  $f$  and  $u > f$  at local minima of  $f$ . We refer to Fig. 2.1 for an example.

The second drawback is the so called stair-casing effect. To understand this effect, we consider a one-dimensional example. Let  $f : [0, 1] \rightarrow \mathbb{R}$  be continuous and monotonically increasing. The total variation of  $f$  then is given by  $\max_{x \in [0, 1]} f(x) - \min_{x \in [0, 1]} f(x) = f(1) - f(0)$  independent from its values in the interior of  $[0, 1]$ . The solution  $u$  to the ROF functional with this particular  $f$  – away from the boundary – will stick to  $f$  since this minimizes the data term among all monotonic functions. At the boundary we observe the already mentioned loss of contrast. In a similar manner, for any noisy data  $f$ , the solution  $u$  of the ROF functional will preserve monotonic structures induced by the noise. This in particular becomes obvious in regions where the original noise-free data is affine, see Fig. 2.2, bottom row.

The third drawback is the over-smoothing of image textures. This effect is caused by the fact that textures usually have a high total variation, which is significantly reduced when TV is used as a regularization term. We refer to Fig. 2.3 for an illustration. To overcome this shortcoming, non-local TV variants as well as structure/texture decomposition have been proposed. These will be discussed below.

## 2.7 Measure-theoretic interpretation of TV

In the case that  $u \in W^{1,1}(\Omega)$ , we find by the Gauss Green formula (cf. [148, Corollary 9.32]) that

$$\text{TV}(u) = \int_{\Omega} \|\nabla u\| \, d\mathcal{L}, \quad (2.27)$$

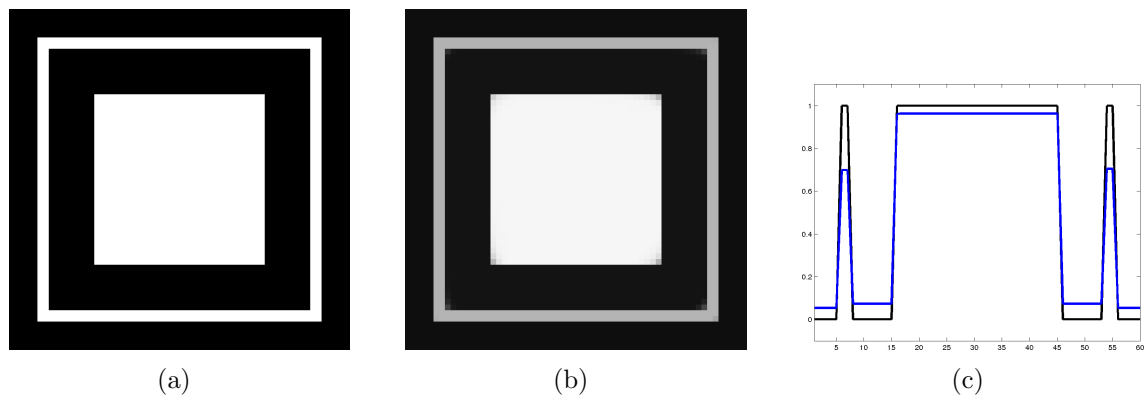


Figure 2.1: Drawbacks of TV regularization – loss of contrast. We illustrate the loss of contrast occurring with TV regularization. The input data (a) are piecewise constant without noise. The smoothed image (b) obtained from minimizing the ROF functional resembles this piecewise structure, but with a loss of contrast. This can in particular be observed when comparing the profiles (c) of input (black line) and result (blue line) along e.g. a horizontal line through the image center.

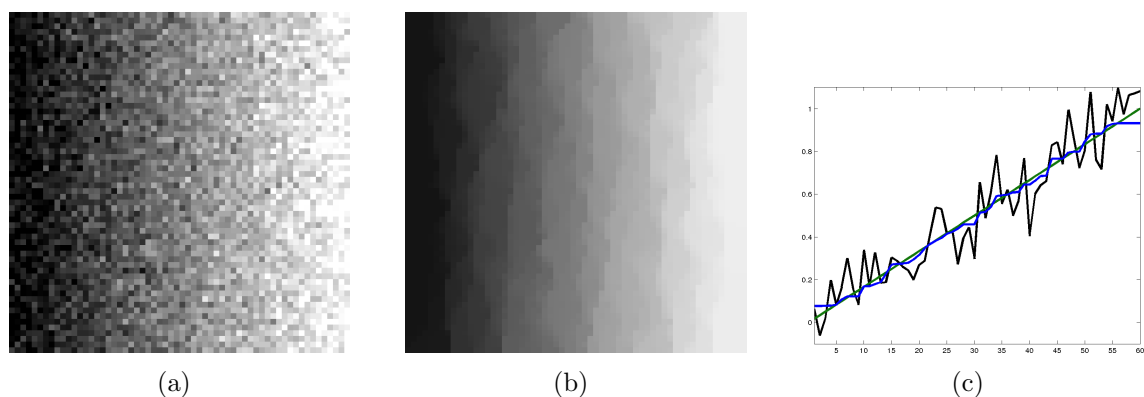


Figure 2.2: Drawbacks of TV regularization – stair-casing. This example illustrates the stair-casing effect of TV regularization. To this end we consider affine noise-free data distorted by additive Gaussian noise (a). The result (b) of minimizing the ROF functional shows the well-known stair-casing effect of TV regularization. The original affine image is reconstructed as a monotonic, but not affine structure. Investigating a horizontal cross-section (c), we also observe this stair-casing effect (blue line) together with a loss of contrast at the image borders (increased difference between the result and the original noise-free data (green line)).

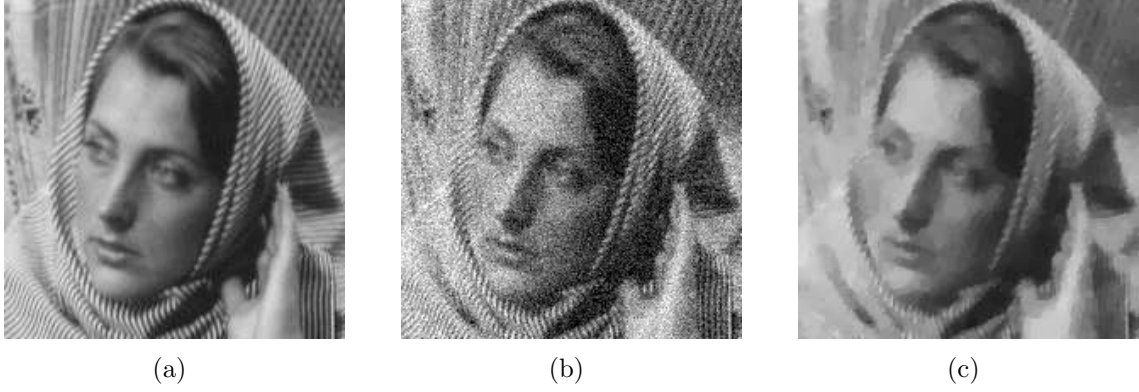


Figure 2.3: Drawbacks of TV regularization – over-smoothing of textured regions. We illustrate this effect by means of the face region of the *Barbara* test image (a). To smooth a noisy version (b) we minimize the ROF functional in (2.6). The result, depicted in (c) shows the over-smoothing of the stripes of Barbara’s scarf.

i.e. the total variation is linked to the weak derivative of  $u$ . In the more general case  $u \in BV(\Omega)$ , it can be shown that the weak derivatives of  $u$  are Radon measures.

To this end, we consider *signed* Radon measures  $\nu$ . Such a signed measure is obtained by combining an unsigned Radon measure  $\mu$  with a signed density  $\sigma$  (see [55, Section 1.3] for details on  $\sigma$ ) by setting

$$\nu(K) := \int_K \sigma \, d\mu \quad (2.28)$$

for all compact sets  $K \subseteq \mathbb{R}^n$ . Borrowing the notation from [55], we shortly write  $\nu = \mu \llcorner \sigma$ . Then, it can be shown that there exists a signed Radon measure  $Du = |Du| \llcorner \sigma$  on  $\Omega$ , such that for every  $p \in C_c^1(\Omega; \mathbb{R}^d)$

$$\int_{\Omega} u \operatorname{div} p \, d\mathcal{L} = - \int_{\Omega} p \, dDu = - \int_{\Omega} p \sigma \, d|Du|, \quad (2.29)$$

see e.g. [55, Section 5.1]. We can interpret the measure  $|Du| \llcorner \sigma_i$ , where  $\sigma_i$  for  $i = 1, 2$  is the  $i$ -th component of  $\sigma$ , as weak directional derivative of  $u$ , since

$$\int_{\Omega} u \partial_i p \, d\mathcal{L} = - \int_{\Omega} p \sigma_i \, d|Du| \quad (2.30)$$

for every  $p \in C_c^1(\Omega, \mathbb{R})$ .

The signed measure  $Du$  can be decomposed into a part  $Du_{ac}$ , which is absolutely continuous with respect to the  $d$ -dimensional Lebesgue measure  $\mathcal{L}^d$  and a part  $Du_s$  singular to  $\mathcal{L}^d$  [55]. An analogous decomposition exists for the positive measure  $|Du|$ . (We require the above decomposition in particular for our considerations in Chapter 5).

## 2.8 Variants of TV Regularization

In the following, we consider a generalized data model, i.e. we assume that  $f = M(u) + v$  with some operator  $M : X \rightarrow L^p(\Omega)$ , data  $f \in L^p(\Omega)$ ,  $p = 1, 2$  and a noise component  $v$ . The data term then becomes  $\mathcal{S}(u) := \frac{1}{2} \|M(u) - f\|_{L^p}^p$ .

We now consider some important variants of TV regularization. Hereby, we concentrate on non-adaptive variants while adaptive TV regularization will be part of the next chapter.

### 2.8.1 Anisotropic TV Regularization

Besides the TV-norm as defined in (2.7), one can think of anisotropic versions of TV. Here, we refer to anisotropy as the property of the regularization term to penalize the gradient  $\nabla u$  of  $u$  with a regularization strength depending on the orientation of  $\nabla u$ . One often used variant is

$$\text{TV}_a(u) = \sup \left\{ \int_{\Omega} u(\partial_x \phi_1) \, d\mathcal{L} \mid \phi_1 \in C_c^\infty(\Omega), |\phi_1(x)| \leq 1 \right\} \quad (2.31)$$

$$+ \sup \left\{ \int_{\Omega} u(\partial_y \phi_2) \, d\mathcal{L} \mid \phi_2 \in C_c^\infty(\Omega), |\phi_2(x)| \leq 1 \right\}. \quad (2.32)$$

This semi-norm shows the strongest penalization for gradients in diagonal direction and weakest for horizontal/vertically oriented gradients (comparing gradients of equal norm). As a consequence, when  $\text{TV}_a$  is used as a regularization term, the reconstruction of horizontal and vertical edges is favored. When in the literature the term *anisotropic TV* is used, then often  $\text{TV}_a$  is meant. However, we will consider various examples of anisotropic TV models in Chapters 4 and 6 and refer to all these models as *anisotropic TV* models. In particular, most of these models will be adaptive to the image content in order to improve the preservation/restoration of certain image structures.

### 2.8.2 Vector-valued TV

The definition of the TV-semi-norm extents naturally to vector-valued function: For  $u : \Omega \rightarrow \mathbb{R}^m$ ,  $u = (u_1, \dots, u_m)$ , let

$$\text{TV}(u) := \sup \left\{ \int_{\Omega} \sum_{i=1}^m u_i(\partial_x p_{i,1} + \partial_y p_{i,2}) \, d\mathcal{L} \mid p \in C_c^1(\Omega, \mathbb{R}^{m \times 2}), \|p(x)\|_2 \leq 1 \right\}, \quad (2.33)$$

which for functions  $u \in W^{1,1}(\Omega, \mathbb{R}^m)$  becomes  $\text{TV}(u) = \int_{\Omega} \|\nabla u\|_F \, d\mathcal{L}$ , where  $\|\cdot\|_F$  denotes the Frobenius norm. An alternative TV model considers in (2.33) the local constraint  $\|(p_{i,1}(x), p_{i,2}(x))\|_2 \leq 1$  instead of  $\|p(x)\|_2 \leq 1$ , which results in  $\text{TV}(u) = \sum_{i=1}^m \text{TV}(u_i)$ .

Depending on the application, it might be suitable to consider other variants of vector-valued TV. In particular for color images several models have been proposed in the literature, see e.g. [24, 29, 39], in order to cope with the specific human perception of edges in colored pictures.

### 2.8.3 Higher-Order TV

A recent trend is to consider total variation of higher-order, see e.g. [28, 35, 94, 130, 129, 147, 153]. In the following, we concentrate on regularization terms which include both first- and second-order total variation semi-norms. For the sake of brevity, we refer to such a regularization as *second-order TV*.

The straightforward model is based on the following definition:

**Definition 2.5** ( $BV^2(\Omega)$ ). *Let  $u \in W^{1,1}(\Omega)$ . We define the space  $BV^2(\Omega)$  as the space of functions  $u \in W^{1,1}(\Omega)$ , for which  $\partial^x u, \partial^y u \in BV(\Omega)$ . We define the second-order total variation as*

$$\text{TV}^2(\Omega) = \sup \left\{ \int_{\Omega} u \operatorname{div}^2 \varphi \mid p \in C_c^1(\Omega, \mathbb{R}^4), \varphi(x) \in B_1(0) \right\}, \quad (2.34)$$

where

$$\operatorname{div}^2 \varphi = \partial_{xx} \varphi_1 + \partial_{xy} \varphi_2 + \partial_{yx} \varphi_2 + \partial_{yy} \varphi_2 \quad (2.35)$$

for  $p = (p_1, \dots, p_n)^\top \in C_c^1(\Omega, \mathbb{R}^4)$ .

Equipped with the norm  $\|f\|_{BV^2} := \|f\|_{W^{1,1}} + \text{TV}^2(u)$  the space  $BV^2(\Omega)$  becomes a Banach space. We remark that there is an alternative definition of  $\text{TV}^2$  via Radon measures, cf. [148, 9.63 and 9.81].

Based on this definition, a second-order TV regularization approach for denoising is

$$\min_{u \in L^2(\Omega) \cap BV^2(\Omega)} \frac{1}{2} \|u - f\|_{L^2}^2 + \alpha \text{TV}(u) + \beta \text{TV}^2(u). \quad (2.36)$$

with  $\alpha, \beta \geq 0$ , cf. [147]. We note that in order to obtain a coercive functional, it is necessary to include both first- and second-order TV, i.e.  $\alpha > 0$  should be assumed.

Besides this straightforward model, other TV variants of second order have been proposed in the literature. One frequently used approach is the *Total General Variation* (TGV) proposed by Bredies, Pock and Kunisch [28], which generalizes total variation up to an arbitrary order  $p$ . In the case  $p = 2$  TGV is defined as

$$\begin{aligned} \text{TGV}^2(u) = \\ \sup \left\{ \int_{\Omega} u \widetilde{\operatorname{div}}^2 \varphi \, dx \mid \varphi \in C_c^k(\Omega, \operatorname{Sym}^2(\mathbb{R}^2)), \|\varphi\|_{\infty} \leq \alpha_0; \|\widetilde{\operatorname{div}} \varphi\|_{\infty} \leq \alpha_1 \right\}, \end{aligned} \quad (2.37)$$

where  $\operatorname{Sym}^2(\mathbb{R}^2)$  is the space of symmetric matrices (2-tensors) and

$$(\widetilde{\operatorname{div}} \varphi)_i := \sum_{j=1}^2 \partial^j \varphi_{ij}, \quad \widetilde{\operatorname{div}}^2 \varphi := \sum_{i=1}^2 \left( \partial^i \partial^i \xi_{ii} + \sum_{i < j} 2 \partial^i \partial^j \varphi \right). \quad (2.38)$$

For details on Total Generalized Variation, including higher order models, we refer to [28].

The third model we are to mention here is based on infimal convolution (IC) [35, 153]. The corresponding optimization problem for denoising reads

$$\min_{u \in BV(\Omega), v \in BV^2(\Omega)} \frac{1}{2} \|u + v - f\|^2 + \alpha \int_{\Omega} |\nabla u| \, d\mathcal{L} + \beta \text{TV}(\nabla v). \quad (2.39)$$

where  $\text{TV}(\nabla v)$  is the total variation of the vector-valued function  $\nabla v \in W^{1,1}(\Omega)$ .

Please observe that the TV models considered in this section always comprise both first- and second-order TV terms. For the compactness of presentation, we refer to these models as *second-order* TV in the sense that the regularization is *up to* the this order.

### 2.8.4 Nonlocal TV Regularization

One alternative to avoid the drawback of TV regularization to over-smooth image textures are non-local variants of total variation, see e.g. [64, 65, 84]. The core idea behind non-local TV regularization is, that for each pair of pixels  $i, j$  a weight  $w_{ij}$  between 0 and 1 is assigned, which describes how closely related the image content at both pixels is (with a higher value indicating a closer relation). One possibility to derive  $w_{ij}$  is by comparing image patches centered at positions  $i$  and  $j$  and to utilize a (normalized) patch distance measure.

The two most often used non-local variants of TV are given by

$$\text{TV}_{\text{non-local}}(u) := \sum_{i,j} w_{ij} |u_i - u_j| \quad (2.40)$$

and

$$\text{TV}_{\text{non-local}}(u) := \sum_i \sqrt{\sum_j w_{ij}^2 |u_i - u_j|^2}, \quad (2.41)$$

respectively. A corresponding continuous representation can be obtained by replacing indices  $i$  and  $j$  by image coordinates  $x$  and  $y$  and the sums in (2.40) and (2.41) by integrals over these coordinates. In the same way, non-local gradient and divergence operators can be defined. In practical applications, it is quite costly from a numerical point of view to consider all possible pairs  $w_{ij}$ . Therefore, often only pixel pairs up to a maximal distance are considered.

An approach, which combines non-local with higher-order regularization has been proposed in [95].

### 2.8.5 Structure-Texture Decomposition

A second alternative to avoid over-smooth image textures is to introduce a decomposition of a given image  $f$  into a structure (cartoon) component  $u$ , a texture

component  $v$  and a noise component  $\eta$ . To this end, a variational approach of the form

$$\min_{u \in BV(\Omega), v \in V} \frac{1}{p} \|u + v - f\|_{L^p}^p + \text{TV}(u) + \mathcal{R}_{\text{texture}}(v) \quad (2.42)$$

can be considered, where the structural component of  $u$  is penalized by the TV semi-norm, and for the textural component a suitable function space  $v$  and a suitable regularization term  $\mathcal{R}_{\text{texture}}$  have to be chosen.

In the literature, several regularization terms have been proposed for this task. We exemplarily recall two regularizers based on the G-norm proposed by Meyer[115],

$$\|v\|_G := \inf \left\{ \left\| \sqrt{g_1^2(x) + g_2^2(x)} \right\|_{L^\infty} \mid g = (g_1, g_2), g_i \in L^\infty(\Omega), v = \text{div } g \right\}, \quad (2.43)$$

which is defined for functions belonging to the space

$$G := \{v \mid \exists g = (g_1, g_2), g_i \in L^\infty(\Omega) \text{ such that } v = \text{div } g\}. \quad (2.44)$$

The first example is to use  $\mathcal{R}_{\text{texture}}(v) = \|v\|_G$ . For the second example, proposed in [8, 9], we define

$$G_\lambda := \{v \in G \mid \|v\|_G \leq \lambda\} \quad (2.45)$$

for  $\lambda > 0$  and set

$$\mathcal{R}_{\text{texture}}(v) = \begin{cases} 0 & \text{if } v \in G_\lambda, \\ +\infty & \text{else.} \end{cases} \quad (2.46)$$

We refer to [10] for a detailed discussion on these and related decomposition models.





# Chapter 3

## Variational Methods in Banach Spaces

### 3.1 Notation

In the remainder of this thesis,  $\Omega \subset \mathbb{R}^d$  is considered to be a  $d$ -dimensional open, bounded domain with Lipschitz boundary. In some sections we restrict ourselves to the special case  $d = 2$ , which is then explicitly stated.

Moreover, by  $X$  we denote a Banach space with norm  $\|\cdot\|_X$ . Let  $X^*$  denote the space dual to  $X$ . In  $X$ , we consider a weaker form of convergence (in a weak topology) and mathematical concepts based on this form of convergence, e.g. sequential lower semi-continuity of functions and sequential pre-compactness (the detailed definition will be given below). For the compactness of notation, when referring to such concepts, we omit in the following the term 'sequential'.

In the following sections, we will consider functionals on  $X$ , which are allowed to take extended real values. To this end, we define  $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$  with the additional conventions that  $+\infty \leq +\infty$ ,  $+\infty + \infty = +\infty$ ,  $t \cdot (+\infty) = +\infty$  for  $t > 0$ .

### 3.2 Overview

In this chapter we consider the following general minimization problem

$$\min_{u \in \mathcal{K}} \mathcal{F}(u), \quad (3.1)$$

where  $\mathcal{F}(u) : X \rightarrow \overline{\mathbb{R}}$  and  $\mathcal{K} \subseteq X$  is closed and convex. (The case  $\mathcal{K} = X$  is explicitly included.) On  $\mathcal{F}$  we assume that it can be split into two summands  $\mathcal{S}$  and  $\mathcal{R}$ :

$$\mathcal{F}(u) = \mathcal{S}(u) + \mathcal{R}(u). \quad (3.2)$$

Please note that by means of the indicator function  $\iota_{\mathcal{K}}$ , defined as

$$\iota_{\mathcal{K}}(u) := \begin{cases} 0 & \text{for } u \in \mathcal{K}, \\ +\infty & \text{otherwise} \end{cases} \quad (3.3)$$

we can reformulate (3.1) as

$$\min_{u \in X} \mathcal{S}(u) + \tilde{\mathcal{R}}(u) \quad (3.4)$$

with  $\tilde{\mathcal{R}}(u) = \mathcal{R}(u) + \iota_{\mathcal{K}}(u)$ . Thus, in the following without loss of generality we focus on the case (3.4).

An important property of  $\mathcal{F}$  which eases theoretical treatment is convexity:

**Definition 3.1.** *The functional  $\mathcal{F}$  is convex, if*

$$\mathcal{F}(\lambda u + (1 - \lambda)v) \leq \lambda \mathcal{F}(u) + (1 - \lambda)\mathcal{F}(v), \quad \forall u, v \in X, \lambda \in (0, 1). \quad (3.5)$$

*The functional  $\mathcal{F}$  is strictly convex, if (3.5) is strict for every pair  $u \neq v$  with  $\mathcal{F}(u), \mathcal{F}(v) < \infty$ .*

Please note that  $\iota_{\mathcal{K}}(u)$  is convex if and only if  $\mathcal{K}$  is convex.

In the following sections, we recall basics on existence theory, convex optimization and optimality conditions from the literature. These basics will be used in the subsequent chapters.

### 3.3 Existence Theory

**Definition 3.2.**

- *The domain of  $\mathcal{F}$  is defined as the set*

$$\text{Dom}(\mathcal{F}) := \{u \in X \mid \mathcal{F}(u) \neq +\infty\}. \quad (3.6)$$

- *The functional  $\mathcal{F}$  is called proper, if  $\text{Dom}(\mathcal{F}) \neq \emptyset$ .*

**Definition 3.3** (Sublevel/superlevel sets). *Let  $\mathcal{F}(u) : X \rightarrow \overline{\mathbb{R}}$ . We define the sublevel set  $\text{level}_{\leq t}(\mathcal{F})$  of level  $t \in \mathbb{R}$  by*

$$\text{level}_{\leq t}(\mathcal{F}) := \{u \in X \mid \mathcal{F}(u) \leq t\}. \quad (3.7)$$

*and the superlevel set  $\text{level}_{\geq t}(\mathcal{F})$  of level  $t \in \mathbb{R}$  by*

$$\text{level}_{\geq t}(\mathcal{F}) := \{u \in X \mid \mathcal{F}(u) \geq t\}. \quad (3.8)$$

In the following, we require that the function space  $X$  is equipped with a weak topology, which guarantees that sublevel sets of  $\mathcal{F}$  are pre-compact. For the space  $W^{1,2}(\Omega)$  this is the standard weak topology for  $BV(\Omega)$  one has to choose the weak\* topology, cf. Chapter 2.

**Definition 3.4.** *[Coercivity] A functional  $\mathcal{F} : X \rightarrow \overline{\mathbb{R}}$  is called coercive, if every sublevel set  $\text{level}_{\leq t}$ ,  $t \in \mathbb{R}$ , is pre-compact with respect to the considered topology.*

**Definition 3.5.** [Lower semi-continuity] A functional  $\mathcal{F}$  is called weakly lower semi-continuous, if for every  $u \in X$  and every sequence  $u^k$  converging to  $u$  with respect to the weak topology considered,

$$\mathcal{F}(u) \leq \liminf_{u^k \rightarrow u} \mathcal{F}(u^k) \quad (3.9)$$

holds.

**Theorem 3.6** (Theorem 5.1 in [148]). Let  $\mathcal{F} : X \rightarrow \overline{\mathbb{R}}$  be proper, coercive and lower semi-continuous with respect to the considered topology. Then, there exists a minimizer of  $\mathcal{F}(u)$ .

Proof: see [148].

**Remark 3.7.** The major steps in proving existence are:

1. Choose a minimizing sequence. Without loss of generality the sequence lies in a sublevel set of  $\mathcal{F}(u)$ .
2. Due to coercivity, we can select a subsequence  $(u^k)_k$  converging to some  $\bar{u} \in X$  with respect to the considered topology.
3. Lower semi-continuity guarantees that

$$\mathcal{F}(\bar{u}) \leq \limsup_{k \rightarrow \infty} \mathcal{F}(u^k). \quad (3.10)$$

Thus, since  $(u^k)_k$  is a minimizing sequence,  $\bar{u}$  is a minimizer of  $\mathcal{F}(u)$ .

In the subsequent chapters, we will refer to this framework, when we discuss the required properties for individual functionals  $\mathcal{F}$ .

**Remark 3.8.**

- In the case of constraint minimization problems, the lower semi-continuity of  $\iota_{\mathcal{K}}$  is required, which induces that  $\mathcal{K}$  has to be closed with respect to the topology considered.
- A sufficient condition for uniqueness of the minimizer is, that  $\mathcal{F}(u)$  is strictly convex (requiring that  $\mathcal{K}$  is convex).

## 3.4 Fenchel Duality, Sub-Differential and Sub-Gradient

Please recall the definition for a proper function, see Definition 3.2.

**Definition 3.9** (Fenchel dual/conjugate). *The dual of a proper functional  $\mathcal{F} : X \rightarrow \overline{\mathbb{R}}$  is defined as*

$$\begin{aligned} \mathcal{F}^* : X^* &\rightarrow \overline{\mathbb{R}}, \\ u^* &\mapsto \mathcal{F}^*(u^*) := \sup_{u \in X} \langle u^*, u \rangle - \mathcal{F}(u), \end{aligned} \quad (3.11)$$

where  $\langle \cdot, \cdot \rangle$  denotes the dual pairing with respect to  $X$  and its dual space  $X^*$ .

**Definition 3.10** (Sub-differential). *Let  $\mathcal{F} : X \rightarrow \overline{\mathbb{R}}$  be convex. An element  $p \in X^*$  is called sub-gradient of  $\mathcal{F}$  at  $u$ , if*

$$\mathcal{F}(v) - \mathcal{F}(u) - \langle p, v - u \rangle \geq 0 \quad \forall v \in X. \quad (3.12)$$

We call the set  $\partial\mathcal{F}(u) \subseteq X^*$  of all sub-gradients of  $\mathcal{F}$  the sub-differential of  $\mathcal{F}$  at  $u$ . Similarly, we call an element  $u \in X$  sub-gradient of  $\mathcal{F}^*$  at  $p$ , if

$$\mathcal{F}^*(q) - \mathcal{F}^*(p) - \langle u, q - p \rangle \geq 0 \quad \forall q \in X^*, \quad (3.13)$$

and call the set  $\partial\mathcal{F}^*(p) \subseteq X$  of all sub-gradients of  $\mathcal{F}^*$  the sub-differential of  $\mathcal{F}^*$  at  $p$ . If the sub-differential of a convex function is single-valued, we identify this set with its single element.

**Remark 3.11** (Optimality condition). *The concept of sub-differentials enables us to state an optimality condition for a minimizer of a convex function: an element  $u^* \in X^*$  is minimizer of a convex functional if and only if*

$$0 \in \partial\mathcal{F}(u^*), \quad (3.14)$$

see e.g. [148][Lemma 10.15.].

**Definition 3.12.** *Let*

$$\mathcal{F}(u) = \mathcal{S}(u) + \mathcal{R}(u), \quad (3.15)$$

where  $\mathcal{S}, \mathcal{R} : X \rightarrow \overline{\mathbb{R}}$  are convex. The Fenchel transform of  $\mathcal{F}$  is defined as

$$\mathcal{F}^*(p) := \mathcal{S}^*(p) + \mathcal{R}^*(-p). \quad (3.16)$$

(Please note that the Fenchel transform depends strongly on the choice of the decomposition of  $\mathcal{F}$ .)

**Definition 3.13** (Dual gap). *Let  $\mathcal{S}, \mathcal{R} : X \rightarrow \overline{\mathbb{R}}$  be convex, lower semi-continuous and proper. Assume that*

$$\inf_{u \in X} (\mathcal{S}(u) + \mathcal{R}(u)) < +\infty, \quad \inf_{p \in X^*} (\mathcal{S}^*(p) + \mathcal{R}^*(-p)) > -\infty. \quad (3.17)$$

Then, we define the dual gap of  $\mathcal{F}$  as

$$\text{Gap}(\mathcal{F}) := \inf_{u \in X} (\mathcal{S}(u) + \mathcal{R}(u)) + \inf_{p \in X^*} (\mathcal{S}^*(p) + \mathcal{R}^*(-p)). \quad (3.18)$$

The following theorem shows that the dual gap is zero under certain conditions.

**Theorem 3.14.** *Let  $\mathcal{S}, \mathcal{R} : X \rightarrow \overline{\mathbb{R}}$  be convex, lower semi-continuous and proper. Assume that one of the functionals is continuous in one point and that there exist  $v \in X$  such that  $\mathcal{S}(v) + \mathcal{R}(v) < +\infty$ . Then,*

$$\inf_{u \in X} (\mathcal{S}(u) + \mathcal{R}(u)) = - \inf_{p \in X^*} (\mathcal{S}^*(p) + \mathcal{R}^*(-p)). \quad (3.19)$$

(cf. Theorem 10.22 in [148]).

If now  $\bar{u}$  is a minimizer of  $\mathcal{F}(u) = \mathcal{S}(u) + \mathcal{R}(u)$  and  $\bar{p}$  is a minimizer the Fenchel dual  $\mathcal{F}^*$ , then it obviously follows that

$$\mathcal{S}(\bar{u}) + \mathcal{R}(\bar{u}) + \mathcal{S}^*(\bar{p}) + \mathcal{R}^*(-\bar{p}) = 0. \quad (3.20)$$

It can be shown that (3.20) is a sufficient condition for minimizers of  $\mathcal{F}$  and  $\mathcal{F}^*$ :

**Theorem 3.15.** *Let  $\mathcal{S}, \mathcal{R} : X \rightarrow \overline{\mathbb{R}}$  be convex, lower semi-continuous and proper.*

(i) *If*

$$\mathcal{S}(u) + \mathcal{R}(u) + \mathcal{S}^*(p) + \mathcal{R}^*(-p) = 0 \quad (3.21)$$

*holds for some  $(u, p)$ , then  $u$  is a minimizer of  $\mathcal{F}(u) = \mathcal{S}(u) + \mathcal{R}(u)$ ,  $p$  is a minimizer of  $\mathcal{F}^*$  and (3.17) holds.*

(ii) *The condition (3.21) is equivalent to either of the Kuhn-Tucker conditions*

$$\bar{p} \in \partial \mathcal{S}(\bar{u}) \text{ and } -\bar{p} \in \partial \mathcal{R}(\bar{u}) \quad (3.22)$$

*or*

$$\bar{u} \in \partial \mathcal{S}^*(\bar{p}) \text{ and } \bar{u} \in \partial \mathcal{R}^*(-\bar{p}). \quad (3.23)$$

## 3.5 Constrained Convex Problems and Variational Inequalities

Let us now consider the constrained quadratic problem

$$\min_{u \in \mathcal{K}} \mathcal{S}(u) \quad (3.24)$$

where  $\mathcal{K} \subseteq X$  is a non-empty closed convex set and  $\mathcal{S}(u) : X \rightarrow \mathbb{R}$  is convex. The non-emptiness of  $\mathcal{K}$  asserts that  $\inf_{u \in \mathcal{K}} \mathcal{S}(u) < \infty$ .

In the following, we motivate a necessary and sufficient condition for  $u$  to be a minimizer of  $\mathcal{S}(u)$  in terms of a variational inequality. To this end, we reformulate (3.24) using the indicator function  $\iota_{\mathcal{K}}$ , see (3.3), and rewrite (3.24) as

$$\min_{u \in X} \mathcal{F}(u), \quad \mathcal{F}(u) := \mathcal{S}(u) + \iota_{\mathcal{K}}(u). \quad (3.25)$$

Applying Theorem 3.15 (ii) we find the following necessary and sufficient conditions for minimizers  $\bar{u}$  and  $\bar{p}$  of  $\mathcal{F}(u)$  and  $\mathcal{F}^*(p)$ ,

$$\bar{p} \in \partial\mathcal{S}(\bar{u}) \text{ and } -\bar{p} \in \partial\mathcal{R}(u). \quad (3.26)$$

In case of  $\mathcal{R}$  being the indicator function of a convex set,  $\partial\mathcal{R}(u)$  can be characterized as follows. By definition of the sub-differential, we see that  $q \in \partial\mathcal{R}(u)$ ,  $u \in \mathcal{K}$  is equivalent to

$$\mathcal{R}(v) - \mathcal{R}(u) - \langle q, v - u \rangle \geq 0 \quad \forall v \in X. \quad (3.27)$$

Recall that  $\mathcal{R}(v) = +\infty$  for  $v \notin \mathcal{K}$ . Moreover for any  $v \in \mathcal{K}$  we have  $\mathcal{R}(v) = 0$ . As a consequence, (3.27) is satisfied for any  $v \notin \mathcal{K}$ . Therefore, inequality (3.27) is equivalent to

$$\langle q, v - u \rangle \leq 0 \quad \forall v \in \mathcal{K}. \quad (3.28)$$

It follows that the sub-differential  $\partial\mathcal{R}(u)$  consists of all  $q$  satisfying (3.28).

Now assume that  $X$  is a Hilbert space. We then can make use of the isometry between  $X^*$  and  $X$ , so that the dual pairing  $\langle \cdot, \cdot \rangle$  can be replaced by the scalar product, also denoted by  $\langle \cdot, \cdot \rangle$ . In this case

$$\partial\mathcal{R}(u) = \{q \in X \mid \langle q, v - u \rangle \leq 0 \forall v \in \mathcal{K}\}, \quad (3.29)$$

where the right hand side of (3.29) is the normal cone  $N_{\mathcal{K}}(u)$  of the set  $\mathcal{K}$  at  $u$  (cf. [13, Example 16.12] or [136, Theorem 8.15] for the finite case). We immediately obtain the following optimality condition for the optimization problem (3.25):  $\bar{u}$  is a minimizer of (3.25) if and only if

$$0 \in \partial\mathcal{S}(\bar{u}) + N_{\mathcal{K}}(\bar{u}). \quad (3.30)$$

Thus, the above considerations provide an equivalent ansatz to the optimization problem (3.24) via the variational inequality

$$\begin{aligned} \text{Find } \bar{u} \in \mathcal{K} \text{ and } \bar{p} \in \partial\mathcal{S}(\bar{u}) \text{ such that} \\ \langle \bar{p}, v - \bar{u} \rangle \geq 0 \quad \forall v \in \mathcal{K}, \end{aligned} \quad (3.31)$$

which is a necessary and under mild conditions sufficient condition on  $\bar{u}$  to be a minimizer of  $\mathcal{F}(u)$  [44].

**Example 3.16.** *As an example, we consider the projection of  $f \in L^2(\Omega)$  onto a closed non-empty convex set  $\mathcal{K} \subseteq L^2(\Omega)$ . The corresponding optimization problem is given as*

$$\min_{u \in \mathcal{K}} \frac{1}{2} \|u - f\|_{L^2}^2. \quad (3.32)$$

*In Chapter 4 we will discuss theory, which provides that functional*

$$\mathcal{F}(u) := \frac{1}{2} \|u - f\|_{L^2}^2 + \iota_{\mathcal{K}}(u) \quad (3.33)$$

is proper, weakly lower semi-continuous with respect to the weak  $L^2$ -convergence and coercive with respect to the  $L^2$ -norm. Under these conditions Theorem 3.6 guarantees the existence of a minimizer  $\bar{u} \in \mathcal{K}$  of (3.33). From the strict convexity of  $\mathcal{S}(u)$  on  $\mathcal{K}$ , uniqueness of  $\bar{u}$  follows. For further theory on projections, we refer to [4, Chapter 2] and [13, Chapter 4].

In this particular example, using the fact that

$$\partial\mathcal{S}(\bar{u}) = \{u - f\}, \tag{3.34}$$

the variational inequality (3.31) simplifies to

$$\begin{aligned} &\text{Find } \bar{u} \in \mathcal{K} \text{ such that} \\ &\langle \bar{u} - f, v - \bar{u} \rangle \geq 0 \quad \forall v \in \mathcal{K}. \end{aligned} \tag{3.35}$$

◇

In Chapter 6 we will present a generalized approach, in which we consider quasi-variational inequalities instead of variational inequalities, where the constraint set  $\mathcal{K}$  is allowed to depend on the unknown  $u$ , i.e.  $\mathcal{K}$  in (3.31) is replaced by  $\mathcal{K}(u)$ .





# Chapter 4

## Convex Adaptive TV Regularization

### 4.1 Notation and Basic Definitions

#### 4.1.1 Multi-Indices for Partial Derivatives

In the following we consider functions  $u : \Omega \rightarrow \mathbb{R}$  which have partial derivatives in the weak sense of Sobolev functions or the classical sense. In order to refer to the partial derivatives of  $u$  in a compact way, we utilize multi-indices (cf.[148]) as follows: a multi-index  $\gamma \in \mathbb{N}_0^d$  is a  $d$ -tuple of nonnegative integers  $\gamma_1, \dots, \gamma_d$ . We then set

$$\partial^\gamma u := \frac{\partial^{|\gamma|} u}{\partial x_1^{\gamma_1} \dots \partial x_d^{\gamma_d}}. \quad (4.1)$$

For multi-indices, the following additional definitions will be required in this chapter: We define the length or order of  $\gamma$  as  $|\gamma| = \sum_{i=1}^d \gamma_i$ . Moreover, we define the number of multi-indices of length equal to  $l$  as

$$M(l) := \#\{\gamma \in \mathbb{N}_0^d \mid |\gamma| = l\} \quad (4.2)$$

and the number of multi-indices with length smaller or equal to  $l$  as

$$N(l) := \#\{\gamma \in \mathbb{N}_0^d \mid |\gamma| \leq l\}. \quad (4.3)$$

The set of multi-indices can be totally ordered as follows: For two multi-indices  $\gamma$  and  $\sigma$  we consider  $\gamma < \sigma$ , if either  $|\gamma| < |\sigma|$  or if there exists a  $1 \leq k \leq d$  such that  $\gamma_l = \sigma_l$  for  $l < k$  and  $\gamma_k < \sigma_k$ .

#### 4.1.2 Gradient and Divergence-Operator of Order $l$

We define the gradient of order  $l$  of  $u : \Omega \rightarrow \mathbb{R}$  as

$$\nabla^l u := (\partial^\gamma u)_{|\gamma|=l} : \Omega \rightarrow \mathbb{R}^{M(l)}, \quad (4.4)$$

and the divergence of order  $l$  as

$$\operatorname{div}^l u(x) := \sum_{|\gamma|=l} \partial^\gamma u_\gamma(x). \quad (4.5)$$

The gradient of order  $l$  consists of all partial derivatives of order  $l$ . As an alternative, we can consider all partial derivatives *up to* order  $l$ :

$$\bar{\nabla} u := (\nabla^1 u, \nabla^2 u, \dots, \nabla^l u) : \Omega \rightarrow \mathbb{R}^{N^{(l)}} \quad (4.6)$$

as well as the corresponding divergence operator

$$\bar{\operatorname{div}} u(x) := \sum_{|\gamma| \leq l} \partial^\gamma u_\gamma(x). \quad (4.7)$$

### 4.1.3 Higher Order Total Variation

For  $l \in \mathbb{N}$  and  $u \in W^{l-1,1}(\Omega)$  we define the total variation of order  $l$  as

$$\operatorname{TV}^l(u) := \sup \left\{ \int_{\Omega} u \operatorname{div}^l \varphi \, d\mathcal{L} \mid \varphi \in C_c^\infty(\Omega, \mathbb{R}^{N^{(l)}}), \|\varphi(x)\|_2 \leq 1 \right\}. \quad (4.8)$$

Moreover, let

$$\overline{\operatorname{TV}}^l(u) := \sum_{i=1}^l \operatorname{TV}^i(u) = \left( \sum_{i=1}^l \|\nabla^i u\|_{L^1} \right) + \operatorname{TV}^l(u). \quad (4.9)$$

We define the space of functions of bounded  $l$ -th order variation by

$$BV^l(\Omega) := \{u \in W^{l-1,1}(\Omega) \mid \operatorname{TV}^l(u) < \infty\}. \quad (4.10)$$

We equip  $BV^l(\Omega)$  with the norm  $\|\cdot\|_{BV^l} := \|\cdot\| + \overline{\operatorname{TV}}^l(u)$ .

**Definition 4.1** (Weak convergence). *We say that  $(u^k)_k \in BV^l(\Omega)$  weakly\* converges to some  $u \in BV^l(\Omega)$ , if*

$$\|u^k - u\|_{W^{l-1,1}} \rightarrow 0 \text{ and } \sup_{k \rightarrow \infty} \operatorname{TV}^l(u^k) < \infty, \quad (4.11)$$

(cf. [148]).

**Lemma 4.2.** *Let  $(u^k)_k \in BV^l(\Omega)$  with  $u^k \rightarrow u$  in  $L^1(\Omega)$ . Then, due to the convergence in  $L^1(\Omega)$ , we have*

$$\operatorname{TV}^l(u) \leq \liminf_{k \rightarrow \infty} \operatorname{TV}^l(u^k). \quad (4.12)$$

The proof is analogous to the proof for first-order TV, see e.g. [55, Sect. 5.2.1]. Since we require a generalization of this proof in the following sections, we provide the full proof together with a remark on this generalization:

*Proof.* Let  $\varphi \in C_c^\infty(\Omega, \mathbb{R}^{N(l)})$  with  $\|\varphi(x)\| \leq 1$  for every  $x \in \Omega$ , Then,

$$\int_{\Omega} u \operatorname{div}^l \varphi \, d\mathcal{L} = \lim_{k \rightarrow \infty} \int_{\Omega} u^k \operatorname{div}^l \varphi \, d\mathcal{L} := a. \quad (4.13)$$

Since we have  $\operatorname{TV}^l(u^k) \geq 0$ , there exists  $b := \liminf_{k \rightarrow \infty} \operatorname{TV}^l(u^k)$ . We consider a subsequence of  $(u^k)_k$ , also denoted by  $(u^k)_k$ , such that  $b = \lim_{k \rightarrow \infty} \operatorname{TV}^l(u^k)$ . Now let us assume that  $a > b$ . Since  $\int_{\Omega} u^k \operatorname{div}^l \varphi \, d\mathcal{L} \rightarrow a$  we can find a  $K$  such that for every  $k > K$

$$\int_{\Omega} u^k \operatorname{div}^l \varphi \, d\mathcal{L} > b. \quad (4.14)$$

Then, we have for all  $k > K$ , that

$$\operatorname{TV}^l(u^k) \geq \int_{\Omega} u^k \operatorname{div}^l \varphi \, d\mathcal{L} > b, \quad (4.15)$$

which is a contradiction. Thus,  $a \leq b$ . Using the definitions of  $a$  and  $b$  and (4.13), we find

$$\int_{\Omega} u \operatorname{div}^l \varphi \, d\mathcal{L} = a \leq b = \liminf_{k \rightarrow \infty} \operatorname{TV}^l(u^k). \quad (4.16)$$

Taking in (4.16) the supremum over all  $\varphi \in C_c^\infty(\Omega, \mathbb{R}^{N(l)})$  with  $\|\varphi(x)\|_2 \leq 1$  gives

$$\operatorname{TV}^l(u) \leq \liminf_{k \rightarrow \infty} \operatorname{TV}^l(u^k). \quad (4.17)$$

□

In particular, this shows that  $\operatorname{TV}^l(u)$  is weakly\* lower semi-continuous.

**Remark 4.3.** *Please note that the proof neither uses particular properties of  $\operatorname{div}^l$  nor the form of the constraint  $\|\varphi(x)\|_2 \leq 1$ . In the same way, we can proof*

$$\mathcal{R}(u) \leq \liminf_{k \rightarrow \infty} \mathcal{R}(u^k). \quad (4.18)$$

for

$$\mathcal{R}(u) := \sup \left\{ \int_{\Omega} u D\varphi \, d\mathcal{L} \mid \varphi \in C_c^\infty(\Omega, Y), \varphi(x) \in \mathcal{D} \right\} \quad (4.19)$$

with arbitrary differential operator  $D$  acting on functions  $\varphi : \Omega \rightarrow Y$  and arbitrary closed set  $\mathcal{D} \subset \mathbb{R}^d$ .

## 4.2 Motivation

In Chapter 2 we have considered standard TV regularization and variants of anisotropic and second-order TV. These regularization terms have in common that they are defined via a supremum over dual variables, which lie in convex subsets of spaces

of smooth functions and for which local constraints are imposed. For the standard TV, for example, these local constraints are given as  $\|\varphi(x)\|_2 \leq 1$ .

For each example considered in Chapter 2, the regularization strength can be incorporated into these local constraints. For the standard TV regularizer, this means that we consider

$$\mathcal{R}(u) = \sup \left\{ \int_{\Omega} u \operatorname{div} \varphi \, d\mathcal{L} \mid \varphi \in C_c^1(\Omega, \mathbb{R}^2), \|\varphi(x)\|_2 \leq \alpha \right\}. \quad (4.20)$$

We might also write these local constraints in terms of sets, e.g. for standard TV

$$\varphi(x) \in \mathcal{D}^{loc} := B_{\alpha}(0), \quad (4.21)$$

i.e.  $\mathcal{D}^{loc}$  are closed balls of radius  $\alpha$  centered at 0, or for the anisotropic TV semi-norm  $\operatorname{TV}_a$  (cf. (2.31))

$$\varphi(x) \in \mathcal{D}^{loc} := \{p \in \mathbb{R}^2 \mid \|p\|_{\infty} \leq \alpha\}, \quad (4.22)$$

where  $\|p\|_{\infty} = \max_{i \in \mathbb{R}^n} |p_i|$ , i.e.  $\mathcal{D}^{loc}$  are squares with side length  $\alpha$ . We refer to the sets  $\mathcal{D}^{loc}$  as *local constraint sets*. Analogously to the examples above the second-order TV approaches discussed in Chapter 2 can be defined via suitable local constraint sets.

We observe, that the local effect of the regularization, i.e. the regularization strength and an orientation dependent penalization is determined by the size and shape of the local constraint set. In turn, we can think of locally adapting the regularization term to improve the quality of the reconstructed image  $u$ . This, however, requires some additional information about expected structures in the unknown image  $u$ . We consider the following example:

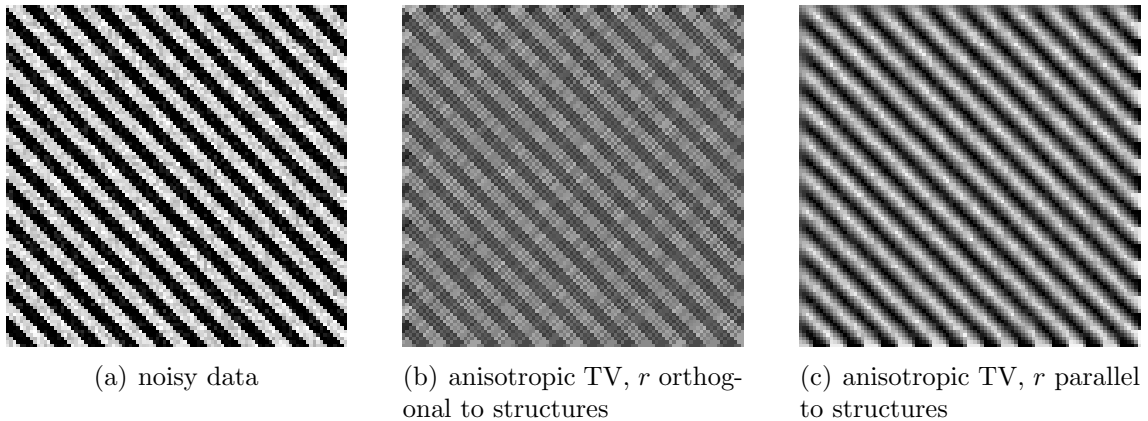


Figure 4.1: Illustration of applying anisotropic TV to data with oriented structures. (a) Input data with oriented parallel structures and Gaussian noise. (b) Smoothing orthogonal to image structures. (c) Smoothing parallel to image structures.

**Example 4.4.** We assume image data  $f$  with structures mainly in one direction  $n \in S^1$  and with additional Gaussian noise, see e.g. Fig. 4.1. In order to denoise these data, we can think of a TV variant which penalizes the total variation of the signal mainly in the direction perpendicular to  $n$ . To this end, we define

$$\mathcal{R}(u; r, \alpha, \beta) := \sup \left\{ \int_{\Omega} u \operatorname{div} \varphi \, d\mathcal{L} \mid \|r^{\top} \varphi\|_2 \leq \alpha, \|(r^{\perp})^{\top} \varphi\|_2 \leq \beta \right\} \quad (4.23)$$

for some  $0 < \alpha \ll \beta$  and arbitrary  $r \in S^1$ . Thus,  $\mathcal{R}(u; r, \alpha, \beta)$  penalizes jumps in direction of  $\pm r$  with factor  $\alpha$  and jumps in direction of  $\pm r^{\perp}$  with factor  $\beta$ . Therefore,  $\mathcal{R}(u; r, \alpha, \beta)$  is an anisotropic TV regularizer.

In order to denoise data with structures orientated in direction of  $n$ , one then should consider

$$\min_{u \in BV(\Omega)} \frac{1}{2} \|u - f\|_{L^2}^2 + \mathcal{R}(u; n, \alpha, \beta). \quad (4.24)$$

Fig. 4.1 shows the results of using  $\mathcal{R}(u; n^{\top}, \alpha, \beta)$  (b) and  $\mathcal{R}(u; n, \alpha, \beta)$  (c) with  $\alpha = 0.001$  and  $\beta = 0.5$  for denoising the test data (a). Clearly, using  $r = n$  produces preferable results. However, using this type of regularizer requires to know the orientation of present structures in advance.

Finally, we remark that  $\mathcal{R}(u; r, \alpha, \beta)$  can be re-written as

$$\mathcal{R}(u; r, \alpha, \beta) := \sup \left\{ \int_{\Omega} u \operatorname{div} \varphi \, d\mathcal{L} \mid \varphi(x) \in \mathcal{D}^{\text{loc}} \right\}, \quad (4.25)$$

where

$$\mathcal{D}^{\text{loc}} := \text{Rect}(r, \alpha, \beta) := \{ \varphi \in \mathbb{R}^2 \mid \|r^{\top} \varphi\|_2 \leq \alpha, \|(r^{\perp})^{\top} \varphi\|_2 \leq \beta \} \quad (4.26)$$

is a rectangle centered at 0 with one side parallel to  $r$  with length  $2\alpha$  and the other side perpendicular to  $r$  with length  $2\beta$ .

The point symmetry of  $\mathcal{D}^{\text{loc}}$  with respect to 0 asserts, that the gradient of  $u$  is penalized independent of its sign.  $\diamond$

In the case of higher order TV, constraint sets occur for the differential operators of different order independently. In view of unifying the notation, we remark that the differential operators and their respective constraint sets can be merged into one joint differential operator and one constraint set. We illustrate this procedure by means of the standard second-order TV model (2.36).

**Example 4.5.** We consider a regularization term combining the first- and second-order TV semi-norms:

$$\begin{aligned} \mathcal{R}(u) &:= \alpha \operatorname{TV}(u) + \beta \operatorname{TV}^2(u) \\ &= \sup \left\{ \int_{\Omega} u \operatorname{div} \varphi_1 \, d\mathcal{L} \mid \varphi_1 \in C_c^1(\Omega; \mathbb{R}^2), \|\varphi_1(x)\|_2 \leq \alpha \text{ a.e.} \right\} \\ &\quad + \sup \left\{ \int_{\Omega} u \operatorname{div}^2 \varphi_2 \, d\mathcal{L} \mid \varphi_2 \in C_c^2(\Omega; \mathbb{R}^4), \|\varphi_2(x)\|_2 \leq \beta \text{ a.e.} \right\}. \end{aligned} \quad (4.27)$$

Since  $C_c^\infty(\Omega; \mathbb{R}^n)$  is dense in  $C_c^q(\Omega; \mathbb{R}^n)$ ,  $q < \infty$ , we can restrict without loss of generality the sets of functions  $\varphi_i$ ,  $i = 1, 2$  in (4.27) to  $C_c^\infty(\Omega; \mathbb{R}^2)$  and  $C_c^\infty(\Omega; \mathbb{R}^4)$ . Moreover, we can combine the dual functions  $\varphi_i$  to  $\varphi := (\varphi_1, \varphi_2) \in C_c^\infty(\Omega; \mathbb{R}^6)$  and the two differential operators  $\text{div}$  and  $\text{div}^2$  to  $D\varphi := (\text{div}, \text{div}^2)^\top \varphi$ . With these definitions, we obtain

$$\mathcal{R}(u) = \sup \left\{ \int_{\Omega} u D\varphi \, d\mathcal{L} \mid \varphi \in C_c^\infty(\Omega; \mathbb{R}^6), \|\varphi_1(x)\|_2 \leq \alpha, \|\varphi_2(x)\|_2 \leq \beta \text{ a.e.} \right\}. \quad (4.28)$$

Finally, again by the denseness of  $C_c^\infty(\Omega; \mathbb{R}^6)$  we can switch back to  $\varphi \in C_c^2(\Omega; \mathbb{R}^6)$  in (4.28).  $\diamond$

In the next section, we will consider a general model, which includes arbitrary differential operators  $D$  and general local constraint sets  $\mathcal{D}^{loc}(x)$ , which depend on the image location  $x$ . The sets  $\mathcal{D}^{loc}(x)$  will be assumed to be closed and convex, but may have arbitrary shapes. Moreover, we will provide theory for existence and uniqueness of minimizers.

As already mentioned, when considering adaptive TV regularization, additional information about image structures is required. We discuss suitable approaches to gain such information in Section 4.6.

### 4.3 General Framework

In the following, we present a general framework of TV-based regularization. Compared to the previous section, we allow a wider class of data terms  $\mathcal{S}(u)$  as well as a wider class of TV-based regularization terms  $\mathcal{R}(u)$ .

Let us start with the data term  $\mathcal{S}(u)$ . In what follows, we assume that  $\mathcal{S}(u)$  is defined for functions  $u \in L^p(\Omega)$  for  $p \in \{1, 2\}$ , since these  $p$  cover the common cases used in most applications. The generalizations to larger  $p$ , however, is straightforward.

We assume that  $\mathcal{S}(u)$  is convex, lower semi-continuous with respect to the weak convergence in  $L^p(\Omega)$  and, in view of coercivity, is bounded from above by a term depending on  $\|u\|_{L^p}$ . The standard form we will use for our applications is

$$\mathcal{S}(u) := \frac{1}{p} \|A u - f\|_{L^p(\Omega_f)}^p, \quad (4.29)$$

where  $f \in L^p(\Omega_f)$  are the data,  $p \in \{1, 2\}$  and  $A : L^p(\Omega) \rightarrow L^p(\Omega_f)$  is a linear operator mapping to functions defined on  $\Omega_f \subseteq \Omega$ . Other forms could be the sum of two or more terms like (4.29), or, if a strict problem with unique solution is required, while  $A$  is not injective, to set

$$\mathcal{S}(u) := \frac{1}{2} \|A u - f\|_{L^p(\Omega_f)}^2 + \frac{\delta}{2} \|u - u_0\|_{L^2(\Omega)}^2 \quad (4.30)$$

with a small  $\delta > 0$ . We refer to Section 4.5 below for a discussion of data terms of such forms. In this section we will also provide the proofs of their required properties.

In addition to the general form of the data term, we allow certain convex constraints on the solution  $u$  of the considered minimization problem. To this end we define a weakly closed convex subset  $\mathcal{K} \subseteq L^2(\Omega)$  (which includes the case  $\mathcal{K} = L^2(\Omega)$ ). We incorporate this constraint in terms of the indicator function  $\iota_{\mathcal{K}}(u)$  as defined in (3.3).

Now, we turn to the regularization term. Our regularization framework on the one hand includes the examples discussed in the previous section, on the other hand provides a concept of adaptivity to locally steer the regularization. To this end, we consider an arbitrary differential operator  $D$  together with arbitrary closed convex and non-empty sets  $\mathcal{D}$ , again defined via local constraint sets  $\mathcal{D}^{loc}(x)$ . As a further generalization, we introduce a dependency of the local constraint sets on some arbitrary data  $v$ , i.e.  $\mathcal{D}^{loc} = \mathcal{D}^{loc}(x; v)$ . The general form we assume for regularization term  $\mathcal{R}$  is

$$\mathcal{R}(u; v) := \sup \left\{ \int_{\Omega} u D\varphi \, d\mathcal{L} \mid \varphi \in C_c^l(\Omega; Y), \varphi(x) \in \mathcal{D}^{loc}(x; v) \right\}, \quad (4.31)$$

where differential operator  $D$  acts on functions  $\varphi : \Omega \rightarrow Y$  (usually  $Y = \mathbb{R}^m$  for some  $m \in \mathbb{N}$ ). To assert well-posedness and convexity of  $\mathcal{R}(u; v)$ , we assume that every  $\mathcal{D}^{loc}(x; v)$  is closed, convex and non-empty. Throughout this chapter, we choose  $v = f$ , i.e.  $\mathcal{D}$  is depending on the data  $f$ . However, the theoretical results retrieved below are valid for arbitrary  $v$ . In particular in Chapter 6, we will consider such arbitrary functions  $v \in L^p(\Omega)$  and re-use the theoretical results from this chapter.

Let us now fix an appropriate function space and an appropriate topology. To have the data term well defined, we require  $u \in L^p(\Omega)$ , while for regularization we assume  $u \in BV^l(\Omega)$  for some  $l > 0$ . Thus we set

$$X := L^p(\Omega) \cap BV^l(\Omega), \quad (4.32)$$

which equipped with the norm

$$\|u\|_X := \|u\|_{L^p} + \sum_{i=1}^l \text{TV}^i(u) \quad (4.33)$$

is a Banach space. We introduce the following weak form of convergence:

**Definition 4.6.** *A sequence  $(u^k)_k \in X$  is said to weakly converge to  $u$  in  $X$ , denoted by  $u^k \xrightarrow{X} u$ , if*

1.  $(u^k)_k$  converges weakly\* to  $u$  in  $BV^l(\Omega)$  (cf. Definition 4.1), and
2.  $(u^k)_k$  weakly converges to  $u$  in  $L^p(\Omega)$

(in the case  $p = 1$  the second assumption is already covered by the first one).

**Remark 4.7** (Pre-compactness). *Any bounded set in  $X$  is pre-compact with respect to this weak convergence. This is due to the fact that we combine two types of weak pre-compactness of bounded sets, the weak pre-compactness in  $L^2(\Omega)$  and the weak\* pre-compactness in  $BV(\Omega)$ .*

**Remark 4.8** (Two-dimensional case). *In the two-dimensional case  $\Omega \subset \mathbb{R}^2$  (recall that  $\Omega$  is open, bounded and with Lipschitz-boundary) we note that the embedding from  $BV(\Omega)$  to  $L^2(\Omega)$  is continuous [148, Thm 9.78]. Thus  $\mathcal{S}(u)$ , originally defined on  $L^p(\Omega)$ ,  $p = 1, 2$ , is well-defined on  $BV(\Omega)$ . We therefore can choose  $X = BV(\Omega)$ , such that the considered weak convergence in  $X$  becomes the weak\* convergence in  $BV(\Omega)$ .*

With the generalizations mentioned above, our minimization problem becomes

$$\boxed{\min_{u \in \mathcal{K}} \mathcal{F}(u), \quad \mathcal{F}(u) := \mathcal{S}(u) + \mathcal{R}(u; f).} \quad (4.34)$$

It can be easily verified that this general approach covers all examples mentioned so far. We briefly list the corresponding settings for  $D$  and  $\mathcal{D}^{loc}(x; v)$ :

Example	cf.	$Y$	$D$	$\mathcal{D}^{loc}$
TV( $u$ )	(2.7)	$\mathbb{R}^2$	div	$B_\alpha(0)$
$\alpha$ TV $_a(u)$	(2.31)	$\mathbb{R}^2$	div	$Rect((1, 0)^\top, \alpha, \alpha)$
$\mathcal{R}(u; r, \alpha, \beta)$	Ex. 4.4	$\mathbb{R}^2$	div	$Rect(r, \alpha, \beta)$
TV $^2(u)$	(2.36)	$\mathbb{R}^6$	$\overline{\text{div}}^2$	$B_\alpha(0) \times B_\beta(0) \subset \mathbb{R}^2 \times \mathbb{R}^4$
TGV( $u$ )	(2.37)	$\text{Sym}^k(\mathbb{R}^d)$	$\widetilde{\text{div}}^k$	$\{p \mid \ \text{div}^l p\  \leq \alpha_l\}$
IC( $u$ )	(2.39)	$\mathbb{R}^2 \times \mathbb{R}^4$	div, div $^2$	$B_\alpha(0) \times B_\beta(0) \subset \mathbb{R}^2 \times \mathbb{R}^4$

**Remark 4.9** (Non-local Total Variation). *We remark that the above model of a regularization term can also be adapted to cover non-local TV approaches. For example, let us consider the non-local TV regularization proposed by Gilboa and Osher [65], which can be characterized as*

$$\mathcal{R}(u; v) := \sup \left\{ \int_{\Omega} \int_{\Omega} \varphi(x, y) (u(x) - u(y)) \, dy \, dx \mid \varphi(x, y) \in L^2(\Omega \times \Omega), \int_{\Omega} \frac{1}{w(x, y; v)} \varphi(x, y)^2 \, dy \leq 1 \text{ for a.e. } x \in \Omega \right\}, \quad (4.35)$$

where  $w : \Omega \times \Omega \rightarrow \mathbb{R}_+$  is a weighting function depending on the function  $v$ . The standard approach is to use  $v = f$ . A typical choice for a weighting function is

$$w(x, y; v) := \exp \left( - \int_{B_r(0)} \frac{1}{\sigma^2} |v(x+s) - v(y+s)|^2 \, ds \right) \quad (4.36)$$



(with suitable boundary conditions on  $f$ .) We can rewrite (4.35) as

$$\mathcal{R}(u; v) := \sup \left\{ \int_{\Omega} \int_{\Omega} u(x) D\varphi(x, y) \, dy \, dx \mid \varphi(x, y) \in L^2(\Omega \times \Omega), \right. \\ \left. \int_{\Omega} \frac{1}{w(x, y; v)} \varphi(x, y)^2 \, dy \leq 1 \text{ for a.e. } x \in \Omega \right\} \quad (4.37)$$

for  $(D\varphi)(x, y) := \varphi(x, y) - \varphi(y, x)$ . The main difference to our general form for regularizer  $\mathcal{R}$  in (4.40) is, that the test functions  $\varphi$  have to be augmented by a second argument. This is one reason why the theoretical approach to non-local TV differs from the one presented above. An revision of theory of non-local TV, however, is out of the cope of this thesis.

## 4.4 Theory

In the following, we focus on the case  $D = \overline{\text{div}}^l$  for  $l > 0$  and prove existence and uniqueness of a minimizer. For other differential operators, the theory is analogous, see Remark 4.17 at the end of this section.

In order to show existence of a minimizer, we make the following assumption:

**Assumption 4.10** (For Existence).

- (i)  $\mathcal{K}$  is non-empty, convex and closed with respect to the convergence from Definition 4.6.
- (ii)  $\mathcal{S} : L^p(\Omega) \rightarrow \mathbb{R}$ ,  $p = 1, 2$ , is convex and weakly lower semi-continuous in  $L^2(\Omega)$ .
- (iii) There exists non-negative constants  $q, C_1, C_2, C_3$  such that for all  $u \in \mathcal{K}$

$$\|u\|_{L^p} \leq C_1(\mathcal{S}(u))^q + C_2 \text{TV}(u) + C_3. \quad (4.38)$$

- (iv)  $\mathcal{D}^{\text{loc}}(x; v)$  for each  $x \in \Omega$  is closed, convex. Moreover, there exist constants  $0 < C_{\min} \leq 1 \leq C_{\max}$ , such that for every  $x \in \Omega$  and every  $v$

$$B_{C_{\min}}(0) \subseteq \mathcal{D}^{\text{loc}}(x; v) \subseteq B_{C_{\max}}(0). \quad (4.39)$$

(Recall that  $B_r(0)$  denotes the closed ball of radius  $r$  centered at 0.)

**Theorem 4.11** (Existence). *Let  $\mathcal{S}(u)$ ,  $\mathcal{K}$ ,  $A$  and  $\mathcal{D}^{\text{loc}}(x; v)$  satisfy Assumption 4.10. Then, the functional in (4.34) with*

$$\mathcal{R}(u; v) := \sup \left\{ \int_{\Omega} u \overline{\text{div}}^l \varphi \mid \varphi \in C_c^l(\Omega; \mathbb{R}^m), \varphi(x) \in \mathcal{D}^{\text{loc}}(x; v) \right\}, \quad (4.40)$$

has a minimizer in  $\mathcal{K}$ .

**Remark 4.12.**

1. We reformulate the constraint optimization (4.34) as

$$\min_{u \in X} \tilde{\mathcal{F}}(u), \quad \tilde{\mathcal{F}}(u) := \mathcal{S}(u) + \mathcal{R}(u; f) + \iota_{\mathcal{K}}(u), \quad (4.41)$$

where  $\iota_{\mathcal{K}}$  is the indicator function of  $\mathcal{K}$  (cf. (3.3)).

2. Note that the statement of Theorem 4.11 is different than that of Proposition 3.67 in [148], where  $u : \mathbb{R}^d \rightarrow \mathbb{R}$  is assumed to be constantly 0 outside  $\Omega$ , while here we consider functions  $u$  defined only on the open domain  $\Omega$ .

For the proof of Theorem 4.11 we require some properties of  $\mathcal{S}(u)$ ,  $\iota_{\mathcal{K}}$  and  $\mathcal{R}(u; v)$ , which are provided by the following lemma and proposition.

**Lemma 4.13.** *Let  $\mathcal{K}$  be convex and closed with respect to the weak convergence introduced in Definition 4.6. The indicator function  $\iota_{\mathcal{K}}$  is lower semi-continuous with respect to this weak convergence.*

*Proof.* Let  $(u^k)_{k=1}^K \in X$ ,  $u^k \xrightarrow{X} u \in X$ . First, assume  $u \notin \mathcal{K}$ . If there existed a subsequence of  $u^k$  in  $\mathcal{K}$ , from the closedness of  $\mathcal{K}$  it would follow that  $u \in \mathcal{K}$ , which is a contradiction. Therefore, any subsequence  $u^{k'}$  of  $u^k$  up to a finite set has to lie in  $X \setminus \mathcal{K}$ . Consequently,  $\lim_{k' \rightarrow \infty} \iota_{\mathcal{K}}(u^{k'}) = \infty$ . Since also  $\iota_{\mathcal{K}}(u) = \infty$  we have  $\iota_{\mathcal{K}}(u) = \lim_{k' \rightarrow \infty} \iota_{\mathcal{K}}(u^{k'})$  and thus  $\iota_{\mathcal{K}}(u) = \liminf_{u^k \xrightarrow{X} u} \iota_{\mathcal{K}}(u^k)$ . Second, assume  $u \in \mathcal{K}$ . Then  $\iota_{\mathcal{K}}(u) \leq \iota_{\mathcal{K}}(u^k)$  for every  $k$ , thus  $\iota_{\mathcal{K}}(u) \leq \liminf_{u^k \xrightarrow{X} u} \iota_{\mathcal{K}}(u^k)$ .  $\square$

**Proposition 4.14.** *Let Assumption 4.10 (iv) be satisfied. Then, the functional  $\mathcal{R}(u; v)$  defined in (4.31) with  $D = \overline{\text{div}}^l$ ,  $l > 0$ , and fixed  $v$  is convex, bounded from below and lower semi-continuous with respect to the weak\* convergence in  $BV^l(\Omega)$ . Moreover,*

$$\frac{C_{\min}}{\sqrt{l}} \overline{\text{TV}}^l(u) \leq \mathcal{R}(u; v) \leq C_{\max} \overline{\text{TV}}^l(u). \quad (4.42)$$

*Proof.* In the following, we consider test functions from the set

$$\mathcal{D} := \{\varphi \in C_c^l(\Omega; \mathbb{R}^{N(l)}) \mid \varphi(x) \in \mathcal{D}^{\text{loc}}(x; v)\} \quad (4.43)$$

with  $v$  fixed and denote the components of any such  $\varphi \in \mathcal{D}$  by  $\varphi_i : \Omega \rightarrow \mathbb{R}^{M(i)}$ ,  $i = 1, \dots, l$ , i.e.  $\varphi = (\varphi_1, \dots, \varphi_l)$ . We introduce the notation  $\mathcal{D}_0 := C_c^l(\Omega; \mathbb{R}^{N(l)})$ .

**Convexity of  $\mathcal{R}$ :** For fixed  $\varphi$  the term  $\int_{\Omega} u \text{div}^i \varphi_i \, dx$  for every  $i = 1, \dots, l$  is linear and thus convex. Since  $\int_{\Omega} u \overline{\text{div}}^l \varphi \, dx$  is the finite sum of terms  $\int_{\Omega} u \text{div}^i \varphi_i \, dx$ ,  $i = 1, \dots, l$ , it is also convex. By Lemma 10.5. in [148], the supremum over a family of convex functions is again convex.

**Lower bound for  $\mathcal{R}$ :** Due to Assumption 4.10 (iv) the constant function  $\varphi = 0$  lies in  $\mathcal{D}$ . With this particular  $\varphi$ , we find  $0 = \int_{\Omega} u \overline{\text{div}}^l \varphi \, dx \leq \mathcal{R}(u; v)$ .

**Weak\* lower semi-continuity of  $\mathcal{R}$ :** The proof of the weak\* lower semi-continuity is analogously to the proof of Lemma 4.2, see Remark 4.3.

**Lower  $\text{TV}^l$ -bound of  $\mathcal{R}$ :** To show the lower and upper  $\text{TV}^l$ -bound of  $\mathcal{R}$  we use the following relations:

$$\sup_{x \in A} f(x) \leq \sup_{x \in B} f(x) \text{ for } A \subseteq B \text{ and arbitrary } f, \quad (4.44)$$

$$\sup_{x \in B_r(0)} f(x) = r \sup_{x \in B_1(0)} f(x) \text{ for } r > 0 \text{ and arbitrary linear } f. \quad (4.45)$$

Let  $c := \frac{C_{min}}{\sqrt{l}}$ , where  $C_{min}$  is the constant from Assumption 4.10 (iv). Then,

$$\begin{aligned} c \overline{\text{TV}}^l(u) &= c \left( \sum_{k=1}^{l-1} |\nabla^k u| + \text{TV}^l(u) \right) \\ &= c \sup_{\varphi \in \mathcal{D}_0, \varphi_k(x) \in B_1(0), k=1, \dots, l} \sum_{k=1}^l \int_{\Omega} u \operatorname{div}^k \varphi_k \, dx \\ &= \sup_{\varphi \in \mathcal{D}_0, \varphi_k(x) \in B_c(0), k=1, \dots, l} \int_{\Omega} u \overline{\operatorname{div}}^l \varphi \, dx \\ &= \sup_{\varphi \in \mathcal{D}_0, \varphi(x) \in B_{C_{min}}(0)} \int_{\Omega} u \overline{\operatorname{div}}^l \varphi \, dx \\ &\leq \sup_{\varphi \in \mathcal{D}_0, \varphi(x) \in \mathcal{D}^{loc}(x)} \int_{\Omega} u \overline{\operatorname{div}}^l \varphi \, dx = \mathcal{R}(u; v), \end{aligned} \quad (4.46)$$

where we used (4.45) and the fact that

$$\underbrace{B_c(0) \times \dots \times B_c(0)}_{l \text{ times}} \subseteq B_{C_{min}}(0) \subseteq \mathcal{D}^{loc}(x) \subseteq \mathbb{R}^{N(l)} \quad (4.47)$$

together with (4.44).

**Upper  $\text{TV}^l$ -bound of  $\mathcal{R}$ :** Using again Assumption 4.10 (iv), we find

$$\begin{aligned} \mathcal{R}(u; v) &= \sup_{\varphi \in \mathcal{D}_0, \varphi(x) \in \mathcal{D}^{loc}(x; v)} \int_{\Omega} u \overline{\operatorname{div}}^l \varphi \, d\mathcal{L} \\ &\leq \sup_{\varphi \in \mathcal{D}_0, \varphi(x) \in B_{C_{max}}(0)} \int_{\Omega} u \overline{\operatorname{div}}^l \varphi \, d\mathcal{L} \\ &\leq \sup_{\varphi \in \mathcal{D}_0, \varphi_k(x) \in B_{C_{max}}(0)} \sum_{k=1}^l \int_{\Omega} u \operatorname{div}^k \varphi_k \, dx \\ &= C_{max} \sup_{\varphi \in \mathcal{D}_0, \varphi_k(x) \in B_1(0)} \sum_{k=1}^l \int_{\Omega} u \operatorname{div}^k \varphi_k \, dx \\ &= C_{max} \left( \sum_{k=1}^{l-1} |\nabla^k u| + \text{TV}^l(u) \right) \\ &= C_{max} \overline{\text{TV}}^l(u), \end{aligned} \quad (4.48)$$

where we used (4.44), (4.45) and that  $(v_1, \dots, v_l) \in B_r(0) \Rightarrow v_k \in B_r(0)$ ,  $k = 1, \dots, l$ .  $\square$

**Corollary 4.15.** *The functional  $\overline{\text{TV}}^l(u)$  is convex and weakly\* lower semi-continuous on  $BV^l(\Omega)$ .*

*Proof.* The corollary follows from Proposition 4.14 by using

$$\mathcal{D}^{loc}(x) := \underbrace{B_1(0) \times \dots \times B_1(0)}_{l \text{ times}}. \quad (4.49)$$

$\square$

*Proof of Theorem 4.11.* We show that the assumptions of Theorem 3.6 are satisfied.

**$\mathcal{F}(u)$  is proper:** Since  $\mathcal{K}$  is non-empty, there exists at least one  $u_0 \in \mathcal{K}$ . We have  $\mathcal{S}(u_0) < \infty$  and  $\iota_{\mathcal{K}}(u_0) < \infty$  by definition and  $\mathcal{R}(u_0; f) \leq C_{max} \overline{\text{TV}}^l(u_0) < \infty$ , since  $\mathcal{K} \subseteq X$ . Thus,  $\mathcal{F}(u_0) < \infty$ , i.e.  $\mathcal{F}$  is proper.

**$\mathcal{F}(u)$  is coercive:** We show that for any  $t \in \mathbb{R}$  the sublevel set  $\text{level}_{\leq t}(\mathcal{F})$  (cf. Definition 3.3) is pre-compact with respect to the weak convergence in  $X$ . To this end let  $u \in \text{level}_{\leq t}(\mathcal{F})$  for fixed  $t$ . Then,  $u \in \mathcal{K}$ . Moreover, using Assumption 4.10(iii), Proposition 4.14,  $\text{TV}(u) \leq \overline{\text{TV}}^l(u)$  we have

$$\begin{aligned} \|u\|_X = \|u\|_{L^p} + \overline{\text{TV}}^l(u) &\leq C_1 \mathcal{S}(u)^q + (1 + C_2) \overline{\text{TV}}^l(u) + C_3 \\ &\leq C_1 \mathcal{S}(u)^q + \frac{(1 + C_2) \sqrt{l}}{C_{min}} \mathcal{R}_v(u) + C_3 \\ &\leq C_1 t^q + \frac{(1 + C_2) \sqrt{lt}}{C_{min}} + C_3 \leq \infty. \end{aligned} \quad (4.50)$$

Thus, every sublevel set is bounded in  $\|\cdot\|_X$  and, by Remark 4.7, pre-compact with respect to the weak convergence in  $X$ .

**$\mathcal{F}(u)$  is weakly lower semi-continuous:**

From Lemma 4.19(i) we have that  $\mathcal{S}(u)$  is lower semi-continuous with respect to the weak convergence in  $X$ .  $\iota(u)$  and  $\mathcal{R}(u; v)$  by Lemma 4.13 and Proposition 4.14 are also lower semi-continuous with respect to this convergence. Since  $\mathcal{F}(u)$  is the sum of  $\mathcal{S}(u)$ ,  $\iota(u)$  and  $\mathcal{R}(u; v)$ , their weak lower semi-continuity carries over to  $\mathcal{F}(u)$ .  $\square$

**Theorem 4.16 (Uniqueness).** *Let Assumption 4.10 be satisfied. If  $\mathcal{S}$  is strictly convex, then functional (4.34) attains a unique minimizer.*

*Proof.* Theorem 4.11 guarantees existence of a minimizer of (4.34). Assume that there are two minimizers  $u_1, u_2 \in \mathcal{K}$ ,  $u_1 \neq u_2$ , with  $\mathcal{F}(u_i) = \inf_u \mathcal{F}(u)$ ,  $i = 1, 2$ .

Then for  $\tilde{u} := \frac{1}{2}(u_1 + u_2)$  we find that  $\tilde{u} \in \mathcal{K}$  since  $\mathcal{K}$  is convex and, since  $\mathcal{S}(u)$  is strictly convex and  $u \rightarrow \mathcal{R}(u, v)$  is convex, that

$$\begin{aligned}\mathcal{S}(\tilde{u}) &< \frac{1}{2}\mathcal{S}(u_1) + \frac{1}{2}\mathcal{S}(u_2), \\ \mathcal{R}(\tilde{u}, f) &< \frac{1}{2}\mathcal{R}(u_1, f) + \frac{1}{2}\mathcal{R}(u_2, f).\end{aligned}\tag{4.51}$$

Combining both inequalities gives

$$\mathcal{F}(\tilde{u}) < \frac{1}{2}(\mathcal{F}(u_1) + \mathcal{F}(u_2)) = \min_{u \in \mathcal{C}} \mathcal{F}(u),\tag{4.52}$$

which is a contradiction.  $\square$

**Remark 4.17.** *So far we have provided theory for the case  $D = \overline{\operatorname{div}}^l$ . The same theory holds in the arbitrary case, if we replace Assumption 4.10 (iv) by directly assuming*

$$\frac{C_{\min}}{\sqrt{l}} \overline{\operatorname{TV}}^l(u) \leq \mathcal{R}(u; v) \leq C_{\max} \overline{\operatorname{TV}}^l(u).\tag{4.53}$$

(cf. Proposition 4.14).

## 4.5 Examples of Data Terms

In view of the three applications of *image denoising*, *image deblurring* and *image inpainting*, we consider the following examples of linear operators.

For all examples we choose  $p = 2$ , thus  $X = L^2(\Omega) \cap BV(\Omega)$  and

$$\mathcal{S}(u) = \frac{1}{2} \|A u - f\|_{L^2(\Omega_f)}^2\tag{4.54}$$

where  $A : L^2(\Omega) \rightarrow L^2(\Omega_f)$  and  $\Omega_f \subseteq \Omega$ . In order to apply Theorem 4.11, which guarantees existence of a minimizer of  $\mathcal{F}(u)$ , we require  $\mathcal{S}(u)$  to be convex, lower semi-continuous with respect to the weak convergence in  $L^2(\Omega)$  and to fulfill Assumption 4.10 (iii) (providing coercivity of  $\mathcal{F}$ ). The convexity holds due to the fact that operator  $A$  is linear and  $\frac{1}{2} \|\cdot - f\|_{L^2(\Omega_f)}^2$  is convex. In order to show that  $\mathcal{S}(u)$  is weakly lower semi-continuous and satisfies Assumption 4.10(iii), we need the following Assumption on operator  $A$ . The subsequent lemma then provides both properties.

**Assumption 4.18** (Assumptions on  $A$ ).

(i) *There exists a constant  $C_a > 0$  such that for every constant  $u \in \mathcal{K}$*

$$\|u\|_{L^p(\Omega)} \leq C_a \|A u\|_{L^p(\Omega_f)}.\tag{4.55}$$

(ii)  $A$  is a bounded operator. There exists a constant  $C_b > 0$  such that for every  $u \in \mathcal{K}$

$$\|A u\|_{L^p(\Omega_f)} \leq C_b \|u\|_{L^p(\Omega)}. \quad (4.56)$$

**Lemma 4.19.** *Let operator  $A$  be linear and satisfy Assumption 4.18 (i) and (ii). Then, the following holds:*

(i)  $\mathcal{S}(u) := \frac{1}{2} \|A u - f\|_{L^p(\Omega_f)}^2$  is convex and lower semi-continuous with respect to the convergence from Definition 4.6.

(ii) There exists a constant  $C_A$  such that

$$\|u\|_{L^2(\Omega)} \leq C_A (\|A u\|_{L^2(\Omega_f)} + \text{TV}(u)); \quad (4.57)$$

Moreover, Assumption 4.10 (iii) holds.

*Proof.* Claim (i): Obviously,  $\mathcal{S}(u)$  is convex in  $u$ , since  $A u$  is linear and  $\|\cdot - f\|_{L^2(\Omega_f)}^2$  is convex. Next, we show that  $\mathcal{S}$  is lower semi-continuous with respect to the convergence from Definition 4.6. Let  $u^k \in X$  such that  $u^k \xrightarrow{X} u \in X$ . In the case  $p = 1$  the weak convergence induces  $\|u^k - u\|_{L^1(\Omega)} \rightarrow 0$ . The inverse triangle equation together with Assumption 4.18 (ii) gives

$$\|A u - f\|_{L^1(\Omega_f)} - \|A u^k - f\|_{L^1(\Omega_f)} \leq \|A(u^k - u)\|_{L^1(\Omega_f)} \leq C_b \|u^k - u\|_{L^1(\Omega)}. \quad (4.58)$$

Since the right hand side tends to zero for  $k \rightarrow \infty$ , and since we can restrict (4.58) to any subsequence for which  $\|A u^k - f\|_{L^1(\Omega_f)}$  converges, we find

$$\|A u - f\|_{L^1(\Omega_f)} \leq \liminf_{k \rightarrow \infty} \|A u^k - f\|_{L^1(\Omega_f)}. \quad (4.59)$$

In the case  $p = 2$  the weak convergence in  $X$  induces a weak convergence in  $L^2(\Omega)$ . Using Lemma 10.6 in [148] it suffices to show the (strong) lower semi-continuity of  $\mathcal{S}(u)$  in  $L^2(\Omega)$  to obtain the weak continuity. To this end, let  $u^k$  now converge to  $u$  in  $L^2(\Omega)$ . In particular  $\|u^k\|_{L^2(\Omega)}$  has to be bounded. Then,

$$\begin{aligned} \frac{1}{2} \|A u - f\|_{L^2(\Omega_f)}^2 - \frac{1}{2} \|A u^k - f\|_{L^2(\Omega_f)}^2 &= \frac{1}{2} \int_{\Omega_f} (A u - f)^2 - (A u^k - f)^2 d\mathcal{L} \\ &= \frac{1}{2} \int_{\Omega_f} (A u + A u^k - 2f)(A(u - u^k)) d\mathcal{L} \\ &\leq \frac{1}{2} C_b (C_b (\|u\|_{L^2(\Omega)} + \|u^k\|_{L^2(\Omega)}) + 2\|f\|_{L^2(\Omega)}) \|u - u^k\|_{L^2(\Omega)}. \end{aligned} \quad (4.60)$$

Since  $(\|u\|_{L^2(\Omega)} + \|u^k\|_{L^2(\Omega_f)} + 2\|f\|_{L^2(\Omega_f)})$  is bounded and  $\|u - u^k\|_{L^2(\Omega)} \rightarrow 0$ , we find

$$\frac{1}{2} \|A u - f\|_{L^2(\Omega_f)}^2 \leq \lim_{k \rightarrow \infty} \frac{1}{2} \|A u^k - f\|_{L^2(\Omega_f)}^2 = \liminf_{k \rightarrow \infty} \frac{1}{2} \|A u^k - f\|_{L^2(\Omega_f)}^2. \quad (4.61)$$

Thus, we have shown the strong lower semi-continuity of  $\mathcal{S}(u)$  in  $L^2(\Omega)$ .

Claim (ii): Let  $c := \int_{\Omega} u \, d\mathcal{L}$  and  $\tilde{u} := u - c$ . Then, using Assumption 4.18 (i) and (ii),

$$\begin{aligned}
 \|u\|_{L^p(\Omega)} &\leq \|\tilde{u}\|_{L^p(\Omega)} + \|c\|_{L^p(\Omega)} \leq \|\tilde{u}\|_{L^p(\Omega)} + C_a \|A c\|_{L^p(\Omega)} \\
 &\leq \|\tilde{u}\|_{L^p(\Omega)} + C_a \|A u\|_{L^p(\Omega_f)} + C_a \|A \tilde{u}\|_{L^p(\Omega_f)} \\
 &\leq \|\tilde{u}\|_{L^p(\Omega)} + C_a \|A u\|_{L^p(\Omega)} + C_a C_b \|\tilde{u}\|_{L^p(\Omega)} \\
 &= (1 + C_a C_b) \|\tilde{u}\|_{L^p(\Omega)} + C_a \|A u\|_{L^p(\Omega)}.
 \end{aligned} \tag{4.62}$$

We have from the Poincaré inequality [21] that there exists a constant  $C_P > 0$  such that

$$\|\tilde{u}\|_{L^2(\Omega)} \leq C_P \text{TV}(u). \tag{4.63}$$

Thus, (4.57) follows with  $C_A := \max((1 + C_a C_b)C_P, C_a)$ . With this estimate, we have

$$\begin{aligned}
 \|u\|_{L^2(\Omega)} &\leq C_A (\|A u\|_{L^2(\Omega_f)} + \text{TV}(u)) \\
 &\leq C_A (\|A u\|_{L^p(\Omega_f)} + \text{TV}(u)) \\
 &\leq C_A (\|A u - f\|_{L^p(\Omega_f)} + \text{TV}(u) + \|f\|_{L^p(\Omega_f)}) \\
 &\leq C_A (\sqrt{2\mathcal{S}(u)} + \text{TV}(u) + \|f\|_{L^p(\Omega_f)}).
 \end{aligned} \tag{4.64}$$

Thus 4.10 (iii) follows with  $q = \frac{1}{2}$ ,  $C_1 = C_A \sqrt{2}$ ,  $C_2 = C_A$  and  $C_3 = C_A \|f\|_{L^p(\Omega_f)}$ .  $\square$

**Example 4.20** (Denoising). *We consider the model of additive Gaussian noise introduced in Section 2.2. Let  $f \in L^2(\Omega)$  be the noisy data. To obtain the same data term as in the ROF functional (2.9), we set  $\Omega_f = \Omega$  and  $A = \text{Id}$ . Moreover, we do not require additional constraints on  $u$  representing the noise-free image, thus we set  $\mathcal{K} = X$ . We check that Assumption 4.10 is satisfied. Firstly, we observe that  $\mathcal{K}$  by definition is non-empty, closed and convex. Secondly, the weakly lower semi-continuity of norms (cf. [4]) carries over to squared norms. Finally, items (i) and (ii) hold with  $C_a = C_b = 1$ .  $\diamond$*

**Example 4.21** (Deblurring). *Let  $\Omega_f \subset \mathbb{R}^2$  be open, bounded and with Lipschitz boundary. We consider an operator  $A$  given by the convolution with a kernel  $K(x) : \mathbb{R}^2 \rightarrow \mathbb{R}$ , i.e.*

$$(A u)(x) := \int_{\mathbb{R}^2} u(y - x) K(x) \, d\mathcal{L}, \quad x \in \Omega_f, \quad u \in L^2(\mathbb{R}^2). \tag{4.65}$$

On  $K(x)$  we make the following assumptions:

1.  $K(x)$  is point-symmetric in 0 for all  $x \in \mathbb{R}^2$ ,
2.  $K(x)$  non-negative and bounded,
3. there exists  $r_1 > 0$  such that  $\text{supp } K(x) \subseteq \mathring{B}_{r_1}(0)$ , and
4. there exists  $r_2 > 0$  and  $C_K > 0$  such that  $K(x) \geq C_K$  on  $B_{r_2}(0)$ .

Now let  $u \in L^2(\Omega)$  with

$$\Omega := \{x \in \mathbb{R}^2 \mid \exists y \in \Omega_f \text{ with } |x - y| < r_1\}. \quad (4.66)$$

Please note that  $\Omega_f \subseteq \Omega$  and that  $\Omega$  is large enough to guarantee that  $A$  in (4.65) is well-defined. As a consequence of the above assumptions, we have  $K \in L^2(\mathbb{R}^2)$  and, due to the properties of convolutions (cf. [4]),  $A$  is a bounded linear operator from  $L^2(\Omega)$  to  $L^2(\Omega_f)$ . (In particular, 4.18 (ii) is fulfilled.) The motivation to distinguish between domains  $\Omega$  and  $\Omega_f$  is as follows. Let us assume that function  $u$  first is only defined on  $\Omega_f$ , i.e.  $u \in L^2(\Omega_f)$ . To still guarantee the well-posedness of  $A$  in (4.65), we have to prescribe values of  $u$  in  $\Omega \setminus \Omega_f$ . One alternative would be to extend any  $u \in BV(\Omega_f)$  outside  $\Omega_f$  with zero value. However, from the point of view of image processing, such an extension would lead to unnatural boundary effects. In order to constantly extend a function  $u \in BV(\Omega_f)$ , we could make use of a trace operator, see e.g. [55], to obtain values almost everywhere along the boundary  $\partial\Omega_f$  of  $\Omega_f$ . Then, extending  $u$  at points  $x \in \partial\Omega_f$  constantly in normal direction to the boundary does not fully define  $u$  on  $\Omega$  (problems arise at points  $x \in \partial\Omega_f$  where the boundary is curved or has corners). To conclude, extending  $u$  to the full domain  $\Omega$  is not straightforward and it is unclear how to assert  $u \in BV(\Omega)$ .

Here, we follow a different strategy by assuming that some data  $u_0 \in L^2(\Omega)$  are given. We set

$$f(x) := (A u_0)(x) = \int_{B_{r_1}(0)} u_0(y - x) K(x) d\mathcal{L}, \quad x \in \Omega_f. \quad (4.67)$$

(In addition, we might assume that  $f$  is distorted by additive Gaussian noise as considered in Example 4.20.)

Moreover, we set

$$\mathcal{K} := \{u \in X \mid u(x) \geq 0 \text{ a.e.}\}. \quad (4.68)$$

Obviously,  $\mathcal{K}$  is non-empty and convex. We show that  $\mathcal{K}$  is closed with respect to the weak convergence in  $X$ . To this end, let  $u^k \in \mathcal{K}$  weakly converge to some  $u \in X$  in the sense of Definition 4.6, which induces that  $u^k \rightarrow u$  in  $L^1(\Omega)$ . Lemma 1.22 in [4] provides that for a subsequence, also denoted by  $(u^k)_k$ ,  $u^k \rightarrow u$  a.e. in  $\Omega$ . Since  $u^k \geq 0$  a.e., it follows that  $u \geq 0$  a.e. and thus  $u \in \mathcal{K}$ . This shows that  $\mathcal{K}$  is closed with respect to the weak convergence in  $X$ .

Let us check that operator  $A$  fulfills the Assumption 4.18 (i). From the assumptions on  $K$  we find for  $u \equiv c$  being constant and non-negative that

$$\begin{aligned} (A u)(x) &= \int_{B_{r_1}(0)} u(y - x) K(x) d\mathcal{L} \\ &\geq \int_{B_{r_2}(0)} u(y - x) K(x) d\mathcal{L} \\ &\geq C_K c \pi r_2^2. \end{aligned} \quad (4.69)$$



Since  $\|u\|_{L^2(\Omega)}^2 = |\Omega|c^2$ , it follows from (4.69) that

$$\int_{\Omega_f} |A u|^2 d\mathcal{L} \geq |\Omega_f| C_K^2 \pi^2 r_2^4 c^2 \geq \frac{|\Omega_f| C_K^2 \pi^2 r_2^4}{|\Omega|} \|u\|_{L^2}^2. \quad (4.70)$$

Taking the square root on both sides of (4.70) gives

$$\|A u\|_{L^2(\Omega_f)} \geq \frac{\sqrt{|\Omega_f|} C_K \pi r_2^2}{\sqrt{|\Omega|}} \|u\|_{L^2(\Omega)}. \quad (4.71)$$

Thus, Assumption 4.18 (i) is fulfilled with  $C_a := \frac{\sqrt{|\Omega|}}{\sqrt{|\Omega_f|} C_K \pi r_2^2}$ .  $\diamond$

**Example 4.22** (Inpainting). *In this example, we consider the reconstruction of some data  $u_0$ , when only a part of these data is known. To this end, let  $\Omega_f \subseteq \Omega$  be open. For arbitrary  $u \in L^2(\Omega)$  we define*

$$A : L^2(\Omega) \rightarrow L^2(\Omega_f), \quad (A u)(x) := u(x), \quad \forall x \in \Omega_f. \quad (4.72)$$

With this definition, let  $f := A u_0$ . Then, the task of inpainting is to reconstruct the function  $u_0 \in L^2(\Omega)$  from data  $f \in L^2(\Omega_f)$  given only on  $\Omega_f$ .

To apply our approach, we set

$$\mathcal{K} := \{u \in L^2(\Omega) \mid u = f \text{ a.e. on } \Omega_f\}. \quad (4.73)$$

The set  $\mathcal{K}$  is non-empty, convex and, by an argumentation analogously to Example 4.21, closed with respect to the weak convergence in  $X$ .

We observe that with definitions (4.72) and (4.73)  $\mathcal{S}(u) = \frac{1}{2} \|A u - f\|_{L^2(\Omega_f)}^2$  becomes zero on  $\mathcal{K}$ . Thus we might alternatively set  $\mathcal{S}(u) \equiv 0$ . However, for the sake of a unified notation, we keep form (4.54) for  $\mathcal{S}$ .

We check that operator  $A$  fulfills the Assumption 4.18, (i) and (ii). Firstly, to show Assumption 4.18 (i), let  $u \equiv c$  be constant. We find

$$\int_{\Omega_f} u^2 d\mathcal{L} = |\Omega_f| c^2 = \frac{|\Omega_f|}{|\Omega|} \int_{\Omega} u^2 d\mathcal{L} \quad (4.74)$$

and thus

$$\|A u\|_{L^2(\Omega_f)} \geq \frac{\sqrt{|\Omega_f|}}{\sqrt{|\Omega|}} \|u\|_{L^2(\Omega)}. \quad (4.75)$$

Secondly, for arbitrary  $u \in L^2(\Omega)$ , we have

$$\|A u\|_{L^2(\Omega_f)}^2 = \int_{\Omega_f} |u|^2 d\mathcal{L} \leq \int_{\Omega} |u|^2 d\mathcal{L} = \|u\|_{L^2(\Omega)}^2, \quad (4.76)$$

thus 4.18 (ii) holds with  $C_b = 1$ .  $\diamond$

**Remark 4.23.** Please note that for the examples of deblurring and inpainting with the respective setting of operator  $A$ , the data term  $\mathcal{S}(u)$  is not strictly convex, so that uniqueness of a minimizer of  $\mathcal{F}(u)$  can not be guaranteed. If one requires uniqueness of a minimizer, the data term has to be slightly modified by adding a strictly convex term, for example  $\varepsilon\|u\|_{L^2}^2$ . In the case of inpainting, another alternative is to prescribe some value outside  $\Omega_f$ , e.g.

$$\mathcal{S}(u) := \int_{\Omega} w(u - \tilde{f})^2 d\mathcal{L} \quad (4.77)$$

with

$$w(x) = \begin{cases} 1 & \text{on } \Omega_f, \\ \varepsilon & \text{on } \Omega \setminus \Omega_f, \end{cases} \quad \tilde{f}(x) = \begin{cases} f(x) & \text{on } \Omega_f, \\ 0 & \text{on } \Omega \setminus \Omega_f. \end{cases} \quad (4.78)$$

for some small  $\varepsilon > 0$ , which is slightly different to using  $\varepsilon\|u\|_{L^2(\Omega)}^2$ .

## 4.6 Detecting Image Structures for Adaptivity

In order to set up an adaptive TV regularizer, some additional information is required to steer the adaptivity. For approaches in image processing and computer vision, typically information about image structures such as edges, corners and ridges can be used. Throughout this section, we revisit corresponding approaches from the literature [1, 22, 68, 154] to detect such structures.

In the following, we consider some grayscale image represented by a function  $u : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ , from which structural information is to be extracted. This information will be encoded by scalar or vector-valued functions defined on  $\Omega$ .

### 4.6.1 Edge Detection

The aim of this section is to derive information about edge location and edge orientation. This information will be encoded in terms of an edge indicator function  $\chi_e : \Omega \rightarrow [0, 1]$ , where values around 1 indicate the presence and values around 0 indicate the absence of an edge, and a vector field  $v_e : \Omega \rightarrow S^1$ . We remark that the information in  $v_e$  is assumed to be meaningful only in regions where  $\chi_e \approx 1$ .

We will use both  $\chi_e$  and  $v_e$  in Section 4.7 to define adaptive TV regularization approaches.

In order to detect edges, we utilize the structure tensor  $J(x, u)$  (cf. [57]), and its eigenvectors  $v_i(x, u)$ ,  $i = 1, 2$  and eigenvalues  $\lambda_i(x, u)$ ,  $i = 1, 2$ . We recall the basic definitions. Let

$$J_0(x; u) := \nabla u_{\sigma}(x) \nabla u_{\sigma}^{\top}(x), \quad (4.79)$$

where  $u_{\sigma} := u * K_{\sigma}$  is a smoothed version of  $u$ , obtained by convolution with a Gaussian kernel  $K_{\sigma}$  with standard deviation  $\sigma$ . The structure tensor  $J(x; u)$  is given as

$$J(x; u) = J_{\rho}(x; u) := (J_0(\cdot; u) * K_{\rho})(x), \quad (4.80)$$

with  $\rho > 0$ . For the simplicity of notation, we will leave out the subscript of  $J_\rho$  if it is clear from the context. Please note that in (4.80) the convolution is applied component-wise.

We assume without loss of generality that the eigenvalues of  $J(x; u)$  are ordered, i.e.  $\lambda_1(x; u) \geq \lambda_2(x; u) \geq 0$ , with the corresponding eigenvectors denoted by  $v_1(x; u)$  and  $v_2(x; u)$ .

The eigenvalues of the structure tensor possess the property, that in homogeneous regions  $\lambda_1 \approx \lambda_2 \approx 0$ , at edges  $\lambda_1 \gg \lambda_2 \approx 0$  and at corners  $\lambda_1 \approx \lambda_2 \gg 0$ . We therefore can use the difference of the eigenvalue to detect edges. To this end, we define

$$\text{coh}(x; u) := \lambda_1(x; u) - \lambda_2(x; u). \quad (4.81)$$

The notation refers to the term *coherence* used in the concept of anisotropic diffusion [175].

We set up a corresponding edge indicator function  $\chi_e(x; u) : \Omega \rightarrow [0, 1]$  by defining

$$\chi_e(x; u) := \begin{cases} \frac{\text{coh}(x)}{\lambda_1(x; u)} & \text{if } \lambda_1(x; u) > 0, \\ 0 & \text{if } \lambda_1(x; u) = 0, \end{cases} \quad (4.82)$$

or

$$\chi_e(x; u) := \min\{s \text{ coh}(x; u), 1\}, \quad (4.83)$$

While the first definition has the property of naturally normalizing the difference  $\lambda_1 - \lambda_2$  (recall that  $\lambda_1 \geq \lambda_2 \geq 0$ ), the second definition has the advantage of being Lipschitz-continuous with respect to  $\lambda_i$ ,  $i = 1, 2$ , a property which will be required for our regularization model in Chapter 6.

In order to retrieve the orientation of edges, we can make use of the eigenvalue  $v_1$  approximating the edge normals (at positions where  $\lambda_1 - \lambda_2$  is sufficiently large). Thus, we choose the vector field  $v_e$  to be

$$v_e(x; u) : \Omega \rightarrow S^1, \quad x \mapsto v_e(x; u) = v_1(x; u). \quad (4.84)$$

## 4.6.2 Corner Detection

The standard structure tensor  $J(x, u)$  can be used to detect the *location* of edge corners. However, it cannot detect the *directions* of the adjacent edges if they are not orthogonal to each other. This is due to the fact that the the eigenvectors of the structure tensor have to be orthogonal.

To resolve this problem, Mühlich and Aach [1, 117] have proposed elaborated structure tensors to detect *double* orientations. They consider two models, a transparent and an occlusion model, which we briefly recall below.

### Occlusion Model

Assume that  $u(x)$  can be locally decomposed into two functions  $u_i$ ,  $i = 1, 2$  as follows. There exists a surrounding  $U$  of  $x_0 \in \Omega$ , which can be split into two open

connected regions  $U_i$  with  $U_1 \cap U_2 = \emptyset$  and  $U = \overline{U_1 \cup U_2}$ , such that

$$u(x) = \begin{cases} u_1(x) & \text{if } x \in U_1 \\ u_2(x) & \text{if } x \in U_2. \end{cases} \quad (4.85)$$

Moreover, we assume that each  $u_i$  can be described by  $u_i(x) = \phi_i(\langle s_i, x \rangle)$  with functions  $\phi_i : \mathbb{R} \rightarrow \mathbb{R}$  and two orientations  $s_1 \perp s_2$ ,  $\|s_i\|_2 = 1$ ; for  $i = 1, 2$ , i.e.  $u_i$  is constant in direction of  $s_i$ . We consider the task of estimating  $r_i := s_i^\top$  from  $u$ . We denote the entries of  $r_i$  by  $r_{i1}$  and  $r_{i2}$ . Now let

$$\nu(u) := ((\partial_{x_1} u_\sigma)^2, \partial_{x_1} u_\sigma \partial_{x_2} u_\sigma, (\partial_{x_2} u_\sigma)^2)^\top, \quad (4.86)$$

where we applied the partial derivatives with respect to the two-dimensional coordinates  $x = (X_1, X_2)$  and where  $u_\sigma$  is obtained by convolution of  $u$  with a Gaussian kernel  $K_\sigma$ ,  $\sigma > 0$  as in the previous section. We then consider structure tensor

$$J_0(x; u) := \nu(x; u) \nu^\top(x; u), \quad J(x; u) := (J_0(\cdot; u) * K_\rho)(x), \quad (4.87)$$

where the convolution with kernel  $K_\rho$  is applied component-wise. It can be shown that under the above assumption on  $u$  the rank of  $J(x; u)$  is at most 2. The rank can be identified with the three cases of a homogeneous region (rank = 0, both  $\phi_i$  constant), an edge region (rank = 1, one  $\phi_i$  constant) and a corner region (rank = 2, both  $\phi_i$  non-constant). For each region, we have to determine the directions  $r_i$  in a different manner. If rank = 2, we even have to distinguish further between two special cases. In order to refer to different cases separately, we introduce a superscript  $r^{(j)}$ ,  $j = 1, \dots, 4$ .

Before we discuss the different kinds of regions, we introduce some necessary notations. For the sake of simplicity, we omit in the following the parameter  $x$  of all functions. We denote by  $\lambda_1(u) \geq \lambda_2(u) \geq \lambda_3(u) \geq 0$  the eigenvalues of  $J(u)$  and  $v_1(u)$ ,  $v_2(u)$  and  $v_3(u)$  the corresponding eigenvectors. Moreover, we require the following two values (the differences between the ordered eigenvalues):

$$\text{coh}_1(u) := \lambda_1(u) - \lambda_2(u), \quad \text{coh}_2(u) := \lambda_2(u) - \lambda_3(u). \quad (4.88)$$

The estimation of  $r_i^{(j)}$  then is as follows:

1. **Corners** ( $\text{coh}_2(u) > 0$ ): The rank of  $J(u)$  is 2. It can be shown that this is equivalent to the case that both  $\phi_i$  are non-constant, i.e. that the two  $\phi_i$  form a corner. Moreover, the vector

$$r = r_1 \otimes r_2 := (r_{11}r_{21}, r_{11}r_{22} + r_{12}r_{21}, r_{12}r_{22})^\top. \quad (4.89)$$

is element of the null space  $\mathcal{N}(J(x, u))$  of  $J$ . We can retrieve the pair  $(r_1, r_2)$  from the third eigenvector  $v_3 = (v_{3,1}, v_{3,2}, v_{3,3})^\top \in \mathcal{N}(J(x, u))$  as follows: One has to distinguish between the cases  $v_{3,1} \neq 0$  and  $v_{3,1} = 0$ . In the case  $v_{3,1} \neq 0$  we set

$$r_1^{(1)}(u) \parallel (v_{3,1}(u), y_1(u))^\top, \quad r_2^{(1)}(u) \parallel (v_{3,1}(u), y_2(u))^\top, \quad (4.90)$$

where  $y_1(u), y_2(u)$  are the solutions of the quadratic equation

$$y^2 + v_{3,2}(u)y + v_{3,1}(u)v_{3,3}(u) = 0. \quad (4.91)$$

In the case  $v_{3,1} = 0$ , the unit vectors

$$r_1^{(2)}(u) \parallel (v_{3,2}(u), v_{3,3}(u))^\top, \quad r_2^{(2)}(u) \parallel (-v_{3,3}(u), v_{3,2}(u))^\top \quad (4.92)$$

can be used.

2. **Edges** ( $\text{coh}_2(u) = 0, \text{coh}_1(u) > 0$ ): We determine  $r_1, r_2$  depending on the eigenvector  $v_1(u)$ . Along straight edges, the eigenvector  $v_1$  is parallel to  $(s_1^2, s_1 s_2, s_2^2)$  (cf. [154]). We can retrieve  $s_1 = r_1^\top$  up to its sign as  $s_1 = \text{sgn}(v_{1,2})\sqrt{v_{1,1}}, \sqrt{v_{1,3}}$  if  $v_{1,1} > 0$  and  $s_1 = \sqrt{v_{1,1}}, \text{sgn}(v_{1,2})\sqrt{v_{1,3}}$  otherwise. We then set

$$r_1^{(3)} := s_1^\top, \quad r_2^{(3)} := s_1. \quad (4.93)$$

3. **Homogeneous regions** ( $\text{coh}_1(u) = \text{coh}_2(u) = 0$ ): Since no detectable structures are present, the definition of  $r_i$  is ambiguous. We fall back to a default setting

$$r_1^{(4)}(u) := (1, 0)^\top, \quad r_2^{(4)}(u) := (0, 1)^\top. \quad (4.94)$$

Please note that superscripts (1) to (4) are used to distinguish between the different cases. In particular this becomes important, when the different vectors have to be interpolated between two distinct regions, as it will be required in Chapter 6.

### Transparent Model

Assume that  $u(x)$  can be locally decomposed into two functions  $u_i, i = 1, 2$  by

$$u(x) = u_1(x) + u_2(x), \quad (4.95)$$

where each  $u_i$  can be locally described by  $u_i(x) = \phi_i(\langle r_i, x \rangle)$  with functions  $\phi_i : \mathbb{R} \rightarrow \mathbb{R}$  and two orientations  $r_1 \nparallel r_2, \|r_i\| = 1$ ; for  $i = 1, 2$ . We denote the entries of  $r_i$  by  $r_{i1}$  and  $r_{i2}$ . The task is to estimate  $r_i$  from  $u$ .

We consider the structure tensor

$$J_0(x; u) := \nu(x; u)\nu^\top(x; u), \quad J(x; u) := (J_0(\cdot; u) * K_\rho)(x), \quad (4.96)$$

where  $\nu(x; u) \in \mathbb{R}^3$  is given as

$$\nu := (\partial_{x_1 x_1} u_\sigma, \partial_{x_1 x_2} u_\sigma, \partial_{x_2 x_2} u_\sigma)^\top, \quad (4.97)$$

i.e. different to the occlusion model with products of first-order derivatives, we now use second-order derivatives. It can be shown that the structure tensor has a rank of at most 2. As for the occlusion model, we again have to distinguish between the cases of a homogeneous region (rank = 0, two constant  $\phi_i$ ), an edge region (rank = 1, one  $\phi_i$  constant) or a vertex (rank = 2, two non-constant  $\phi_i$ ). The  $r_i^{(j)}$  then can be determined in the same way as in *occlusion model*.

### 4.6.3 Detection of Slope Discontinuities

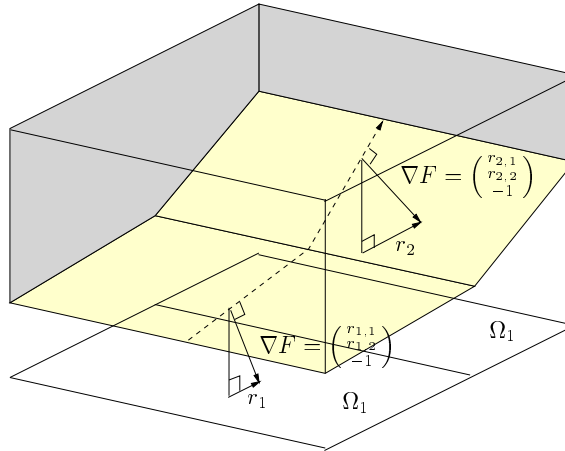


Figure 4.2: Graph  $\Gamma = (x, y, u(x, y))^T$  (yellow) of a continuous and piecewise affine function  $u$  with a discontinuity in the gradient (interface between  $\Omega_1$  and  $\Omega_2$ ). The epigraph of  $u$  is the volume above  $\Gamma$ , represented as the superlevel set of  $F(x, y, z) = u(x, y) - z$ . On the graph the gradient  $\nabla F$  of  $F$  coincides with the surface normal of  $\Gamma$ .

The structure tensors considered in Sections 4.6.1 and 4.6.2 are sufficient to identify discontinuities in  $u$  (edges and corners). We now focus on regions where  $u$  is continuous but has discontinuities in its first derivatives. In addition, we assume that  $u$  is piecewise affine. This assumption is in view of a first- and second-order TV regularization favoring such piecewise affine solutions. For the sake of simplicity, let us consider a prototypical function model with only one discontinuity, which locally on some  $U \subseteq \Omega$  represents a part of a larger image: we assume that  $U$  can be divided into two segments  $U_i, i = 1, 2$  such that  $u$  is affine in each segment, i.e.  $u$  can be represented as

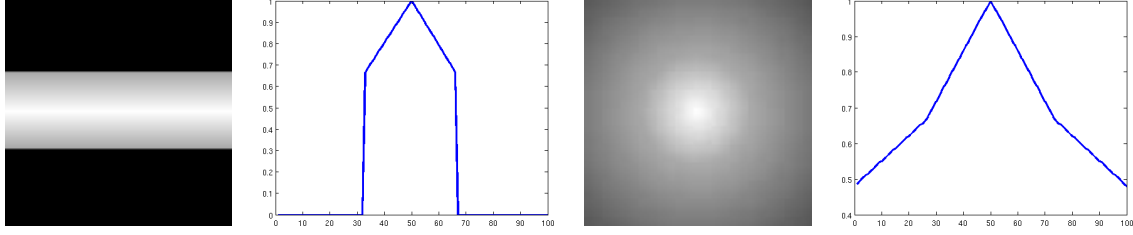
$$u(x, y) = \begin{cases} r_1^\top \begin{pmatrix} x \\ y \end{pmatrix} + b_1 & \text{if } (x, y) \in U_1, \\ r_2^\top \begin{pmatrix} x \\ y \end{pmatrix} + b_2 & \text{if } (x, y) \in U_2, \end{cases} \quad (4.98)$$

for  $U_i$  open, such that  $U_1 \cap U_2 = \emptyset$  and  $\overline{U_1 \cup U_2} = \Omega$ , and  $r_i \in \mathbb{R}^2, b_i \in \mathbb{R}$  for  $i = 1, 2$ . Please note that different from the previous sections, we represent two-dimensional coordinates in  $\Omega$  by  $(x, y)$  instead of  $x$  in the following. Fig. 4.2 illustrates a prototypical function  $u$  as considered in (4.98).

We now discuss a method to detect the case where  $r_1 \neq r_2$ . To this end, we consider the epigraph of  $u$  defined as the superlevel set

$$\text{level}_{\geq 0}(F) := \{(x, y, z) \mid F(x, y, z) \geq 0\} \quad (4.99)$$

of  $F(x, y, z) := u(x, y) - z$ . In order to detect (surface) edges of the graph (i.e., locations, where the slope changes), we now apply the three-dimensional structure


 Figure 4.3: Test images *roof* and *cone hat* for detecting slope discontinuities

tensor to  $F$ , i.e.  $J = ((\nabla F)(\nabla F)^\top)_\rho$ , where

$$\nabla F(x, y, z) = \frac{1}{\sqrt{(\partial_x u)^2 + (\partial_y u)^2 + 1}} (\partial_x u, \partial_y u, -1)^\top. \quad (4.100)$$

Note that  $\nabla F$  is constant in  $z$ . Since we are only interested in edges of the graph  $\Gamma = \{(x, y, z) \mid F(x, y, z) = 0\}$  (i.e., slope discontinuities), we define  $\nabla \tilde{F}_u(x, y) := \nabla F(x, y, u(x, y))$  and restrict structure tensor  $J$  to  $\Gamma$ :

$$J(x, y; u) := \left( (\nabla \tilde{F}_u(x, y)) (\nabla \tilde{F}_u(x, y))^\top \right)_\rho. \quad (4.101)$$

We observe that  $\nabla \tilde{F}_u(x, y)$  is the normal to the graph  $\Gamma$  at  $(x, y, u(x, y))$ .

**Remark 4.24.** *The following two scenarios are of particular interest:*

**a) Within an affine region:** *For an affine function  $u$ ,  $J(x, y; u)$  has exactly one non-zero eigenvalue. This is due to the fact that in this case  $\nabla \tilde{F}_u(x, y)$  is constant and the convolution of  $\nabla \tilde{F}_u \nabla \tilde{F}_u^\top$  does not change the rank.*

**b) Interface between two affine regions of different slope:** *For such  $u$ ,  $J(x, y; u)$  sums up two different directions  $(r_{1,1}, r_{1,2}, -1)$  and  $(r_{2,1}, r_{2,2}, -1)$ : Rewriting the convolution of the matrix entries as a weighted integral, we have*

$$\begin{aligned} J(x, y; u) &= (\nabla \tilde{F}_u \nabla \tilde{F}_u^\top)_\rho = \int_\Omega w(x) \nabla \tilde{F}_u \nabla \tilde{F}_u^\top dx \\ &= w_1 \begin{pmatrix} r_{1,1} \\ r_{1,2} \\ -1 \end{pmatrix} (r_{1,1}, r_{1,2}, -1) + w_2 \begin{pmatrix} r_{2,1} \\ r_{2,2} \\ -1 \end{pmatrix} (r_{2,1}, r_{2,2}, -1) \end{aligned} \quad (4.102)$$

with  $w_i := \int_{U_i} w(x) dx$ . We observe that in (4.102) two rank-1 matrices are added up. Each matrix has one non-zero eigenvalue  $w_i \cdot \|(r_{i,1}, r_{i,2}, -1)\|_2^2$  with corresponding eigenvector  $v_i = (r_{i,1}, r_{i,2}, -1)$ . Since the eigenvectors are linear dependent only if  $r_1 = r_2$ ,  $J(x, y; u)$  has rank 2 near the discontinuity, where  $r_1 \neq r_2$ .

In the following we denote by  $\lambda_i(x, y), i = 1, 2, 3$  the eigenvalues of  $J(x, y; u)$  in decreasing order. As an indicator for the existence of slope discontinuities we propose to use  $\lambda_2(x)$ . This is motivated by the fact that, similarly to the standard structure tensor in  $2D$ ,  $J(x, y; u)$  reveals two eigenvalues significantly larger than 0 at

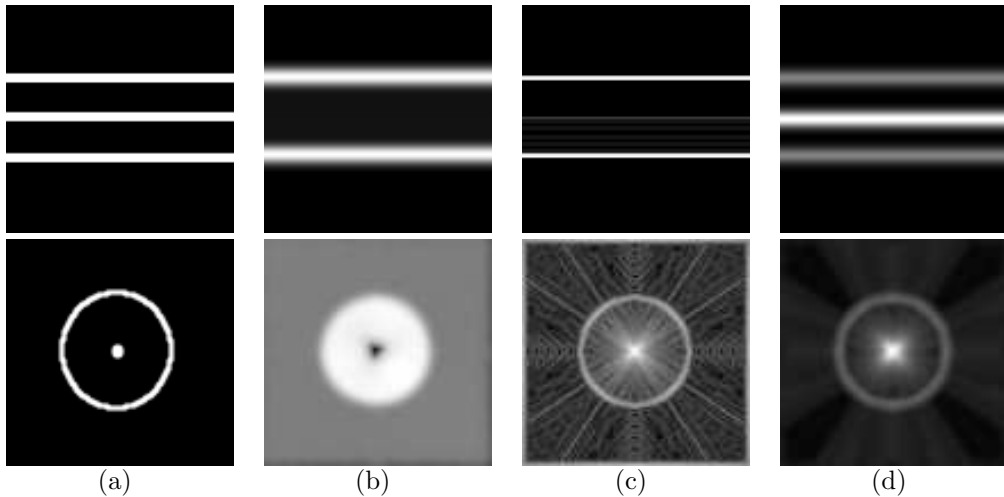


Figure 4.4: Detecting slope discontinuities using the standard structure tensor (b), a curvature based approach (c), and the proposed method (d) in the test images depicted in Fig. 4.3 (black=0, white=1). In both cases the standard structure tensor fails to detect the slope discontinuities as shown in the ideal result (a) (middle line in the first image, ring and center point in the second image). Only the proposed approach is able to detect the slope discontinuity in the first test image (top row). On the second test image (bottom row), the proposed approach provides a less noisy and more precise result than the curvature based approach.

edges of the graph, while in regions of constant slope the second eigenvalue becomes 0. Therefore the magnitude of the second eigenvalue can be used to distinguish between both cases. We propose

$$\begin{aligned} \chi_{sd} &: \Omega \rightarrow [0, 1], \\ \chi_{sd}(x) &:= \min(c\lambda_2(x), 1), \end{aligned} \tag{4.103}$$

with some constant  $c > 0$  as an indicator for regions of slope discontinuities. In order to be less sensitive to edges, which are already covered by the standard structure tensor, we use an upwind scheme to compute the gradient in (4.101). In practice, it is advisable to use the pre-smoothed  $u_\sigma$  (cf. Section 4.6.1) instead of  $u$  to be robust against noise.

To demonstrate the benefits of using  $\chi_{sd}$  to detect slope discontinuities, we compare our approach to one approach based on the standard structure tensor and one based on curvature, see Fig. 4.4. We observe that our approach detects slope discontinuities more reliably than the competitive methods.



## 4.7 Examples of Adaptive Regularization

### 4.7.1 First-Order TV Regularization

Let  $f \in L^\infty(\Omega_f)$  denote the input data. Throughout this section, we consider the differential operator in (4.31) to be  $D := \operatorname{div}$ . The regularizer then takes the form

$$\mathcal{R}(u; f) := \sup \left\{ \int_{\Omega} u \operatorname{div} \varphi \, d\mathcal{L} \mid \varphi \in C_c^1(\Omega; \mathbb{R}^2), \varphi(x) \in \mathcal{D}^{\operatorname{loc}}(x; f) \right\} \quad (4.104)$$

with local constraint sets  $\mathcal{D}^{\operatorname{loc}}(x; f) \subset \mathbb{R}^2$ .

**Example 4.25** (Adaptive regularization strength / weighted TV). *This TV regularization approach aims at reducing the regularization parameter at locations where edges are supposed to be present. Let  $\alpha_{\max} > 0$  be the regularization strength to be applied in homogeneous regions. We reduce this value at location where we expect edges to be. In order to identify such locations, we make use of the gradient of the input image. To this end we apply a pre-smoothing by convolving  $f$  with a Gaussian kernel  $K_\sigma$  with standard deviation  $\sigma > 0$ . We denote the result by  $f_\sigma$ . The pre-smoothing on the one hand guarantees that we can evaluate the gradient of  $f$  at every location  $x$  and, on the other hand, that our approach is robust against noise in the input data. Then, we set*

$$\alpha(x; f) = \max(\alpha_{\max}(1 - \kappa \|\nabla f_\sigma(x)\|_2), \varepsilon), \quad (4.105)$$

where  $\kappa > 0$  is a parameter for the edge sensitivity and  $\max(\cdot, \varepsilon)$  with a small  $\varepsilon > 0$  asserts the positive boundedness  $\alpha$  from below.

An alternative to (4.105), which originally was introduced for diffusion processes utilized for smoothing image data [132], is

$$\alpha(x; f) := \alpha_0 \frac{1}{1 + \kappa \|\nabla f_\sigma(x)\|_2^2}. \quad (4.106)$$

Such an adaptive regularization strength is e.g. used in [43, 80, 109, 170].

To define a isotropic TV regularization with regularization strength  $\alpha(x; f)$ , we choose  $\mathcal{D}^{\operatorname{loc}}(x; f)$  in (4.104) as

$$\mathcal{D}^{\operatorname{loc}}(x; f) = B_{\alpha(x; f)}(0), \quad (4.107)$$

i.e. we choose balls with varying radius  $\alpha(x; f)$  as local constraint sets.

For alternative approaches of spatially adaptive TV regularization proposed in the literature we refer to [100, 157, 158]. In the following, for the shortness of presentation, we refer to TV approaches with adaptive regularization strength as weighted TV.  $\diamond$

**Example 4.26** (Anisotropic TV with edge preservation). *This anisotropic approach requires information about the location and direction of edges in terms of an edge indicator function  $\chi_e : \Omega \rightarrow [0, 1]$ , where  $\chi_e(x; f) \approx 1$  indicates the presence and  $\chi_e(x; f) = 0$  the absence of an edge at position  $x$ , and a vector field  $v_e(\cdot; f) : \Omega \rightarrow S^1$  providing the edge normals.*

*One way to retrieve this information is to apply the structure tensor as described in Section 4.6.1 to the noisy data  $f$ , where a strong pre-smoothing of the input ( $\sigma \gg 0$ ) asserts robustness to the present noise. We choose  $\chi_e(x; f)$  to be one of the two alternatives (4.82) or (4.83) and  $v_e(x; f)$  as defined in (4.84). Please note that the extra argument  $f$  of  $\chi_e$  and  $v_e$  emphasizes that these values are obtained from the noisy data  $f$  and thus that this approach implements a data-driven adaptivity.*

*We will consider alternative ways to determine  $\chi_e(x; f)$  and  $v_e(x; f)$  in Chapter 6 and Chapter 8, Sections 8.3 and 8.4. In particular, Chapter 6 will consider a solution-driven instead of a data-driven adaptivity.*

*It remains to set the local constraint set  $\mathcal{D}^{\text{loc}}(x; f)$ . Our aim is to define a weak regularization by parameter  $\alpha_{\min}$  in normal direction to edges and a strong regularization  $\alpha_{\max}$  in tangential direction. Moreover, in homogeneous region (regions without edges), we want to have an isotropic regularization with parameter  $\alpha_{\max}$ . To achieve this, we define*

$$\alpha(x; f) = \alpha_{\max} + \chi_e(x; f)(\alpha_{\min} - \alpha_{\max}) \quad (4.108)$$

*and  $\mathcal{D}^{\text{loc}}$  as the ellipse with one half axis parallel to  $v_e$  and with length  $\alpha(x; f)$ , and the orthogonal half axis with length  $\beta(x; f) := \alpha_{\max}$ , i.e.*

$$\begin{aligned} \mathcal{D}^{\text{loc}}(x; f) &:= \text{Ell}(v_e(x; f), \alpha(x; f), \beta(x; f)) \\ &:= \left\{ p \in \mathbb{R}^2 \mid \frac{1}{\alpha(x; f)} \langle p, v_e(x; f) \rangle^2 + \frac{1}{\beta(x; f)} \langle p, v_e(x; f)^\top \rangle^2 \right\}. \end{aligned} \quad (4.109)$$

*Fig. 4.5 illustrates this ellipsoidal constraint set  $\mathcal{D}^{\text{loc}}(x; f)$ .*

*We originally proposed this example in [68]. For related models of anisotropic TV regularization we refer to [14, 53, 54, 91].*  $\diamond$

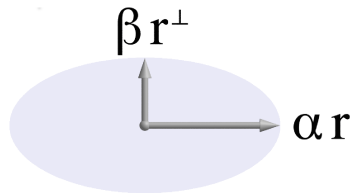


Figure 4.5: Illustration of an ellipsoidal local constraint set  $\mathcal{D}^{\text{loc}}(x; f)$ .

**Example 4.27** (Anisotropic TV with edge and corner preservation (double orientations)). *In [154] Steidl and Teuber proposed an anisotropic TV regularization*

approach, which is able to preserve both edges and corners. The approach uses the strategy proposed by [1, 117] to find double orientations  $r_i(x; f)$ ,  $i = 1, 2$ , see Section 4.6.2, which then are used to define the anisotropy. This approach fits well in the general concept (4.31) by choosing for  $\mathcal{D}^{loc}(x; f)$  a parallelogram with sides  $r_i(x; f)$ ,  $i = 1, 2$ :

$$\begin{aligned} \mathcal{D}^{loc}(x; f) &:= \text{Par}(r_1(x; f), r_2(x; f), \alpha) \\ &:= \{p \in \mathbb{R}^2 \mid |r_1(x; f)^\top p| \leq \alpha, |r_2(x; f)^\top p| \leq \alpha\}. \end{aligned} \quad (4.110)$$

Figure 4.6 illustrates  $\text{Par}(r_1, r_2, \alpha)$ . We refer to this example as adaptive TV regularization based on double orientations.  $\diamond$

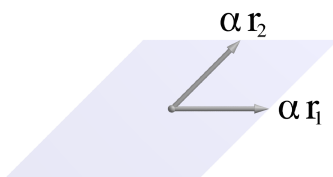


Figure 4.6: Shape of  $\mathcal{D}^{loc}(x; f)$  for the model of Steidl & Teuber [154]: a parallelogram with sides parallel to  $r_1$  and  $r_2$  with side length  $2\alpha$ .

## 4.7.2 Anisotropic TV of First- and Second-Order

Recall that  $f \in L^\infty(\Omega_f)$  denotes the input data. Throughout this section, we consider the differential operator in (4.31) to be  $D := \overline{\text{div}}^2$ . The regularizer then takes the form

$$\mathcal{R}(u; f) := \sup \left\{ \int_{\Omega} u \overline{\text{div}}^2 \varphi \, d\mathcal{L} \mid \varphi \in C_c^2(\Omega; \mathbb{R}^6), \varphi(x; f) \in \mathcal{D}^{loc}(x; f) \right\}. \quad (4.111)$$

For the adaptive regularization terms we discuss in this section, the local constraint sets  $\mathcal{D}^{loc}(x; f)$  are given as the cross product of two sets, i.e.

$$\mathcal{D}^{loc}(x; f) = \mathcal{D}_1^{loc}(x; f) \times \mathcal{D}_2^{loc}(x; f), \quad (4.112)$$

where the first set  $\mathcal{D}_1^{loc}$  defines adaptivity for the first-order TV regularization and the second set  $\mathcal{D}_2^{loc}$  defines adaptivity for the second-order TV regularization.

We study three examples:

**Example 4.28** (Anisotropic first-order TV plus weighted second-order TV for preserving edges). *In this example, we exploit the standard structure tensor, cf. Section 4.6.1 applied to input data  $f$ , to obtain information about the location and orientation of edges. To this end, we use a strong pre-smoothing of  $f$  ( $\sigma \gg 0$ ) to*

assert robustness to present noise. Based on the eigenvalues and eigenvectors of the structure tensor  $J(x; f_\sigma)$  in (4.80), we choose  $\chi_e(x; f)$  to be one of the two alternatives (4.82) or (4.83) and  $v_e(x; f)$  as defined in (4.84). As in the example before the extra argument  $f$  of  $\chi_e$  and  $v_e$  emphasizes that these values are obtained from the noisy data  $f$  and thus that this approach implements a data-driven adaptivity.

For first-order TV, we use an anisotropic approach as in Example 4.26. For the second-order TV regularization, we use an adaptive regularization strength, which is decreased at edges to avoid smoothing across the edges. To this end, let  $0 < \alpha_{\min} \leq \alpha_{\max}$  and  $0 < \gamma_{\min} \leq \gamma_{\max}$  be fixed. We set

$$\begin{aligned}\alpha(x; f) &= \alpha_{\max} - \chi_e(x; f)(\alpha_{\max} - \alpha_{\min}), \\ \beta(x; f) &= \alpha_{\max}, \\ \gamma(x; f) &= \gamma_{\max} - \chi_e(x; f)(\gamma_{\max} - \gamma_{\min}).\end{aligned}\tag{4.113}$$

For the sets  $\mathcal{D}_1^{\text{loc}}(x; f)$  (cf. (4.112)), which defines adaptivity for the first-order TV regularization, we choose a two-dimensional ellipse with one half axis of length  $\alpha(x; f)$  parallel to  $v_e(x; f)$  and the other half axis of length  $\beta(x; f)$  (cf. Example 4.26). The sets  $\mathcal{D}_2^{\text{loc}}(x; f)$  for the second-order TV regularization is set to be a four-dimensional ball of radius  $\gamma(x; f)$ . Combining the constraint sets, we have

$$\mathcal{D}^{\text{loc}}(x; f) := \text{Ell}(v_e(x; f), \alpha(x; f), \beta(x; f)) \times B_{\gamma(x; f)}(0).\tag{4.114}$$

Note that the above approach has the properties of

- a weak ( $\alpha_{\min}$ ) first-order regularization in normal direction at edges,
- a strong ( $\alpha_{\max}$ ) first-order regularization in tangential direction at edges,
- a weak ( $\gamma_{\min}$ ) second-order regularization at edges,
- a strong ( $\gamma_{\max}$ ) second-order regularization in planar regions, and
- a strong regularization ( $\alpha_{\max}, \gamma_{\max}$ ) in homogeneous (constant) regions.

◇

**Example 4.29** (Anisotropic first-order TV plus weighted second-order TV for preserving edges and slope discontinuities). For this adaptive regularization of first- and second-order TV, we require information about the location of edges and their orientation as well as the location of slope discontinuities. The edge information is assumed to be given by an edge indicator function  $\chi_e$  together with a normal vector field  $v_e$ . To obtain this information, e.g. the standard structure tensor (cf. Section 4.6) can be utilized. We assume the location of slope discontinuities to be indicated by a function  $\chi_{sd}(x; f) : \Omega \rightarrow [0, 1]$ , e.g. the indicator function proposed in Section 4.6.3.

For first-order TV, we use an anisotropic approach as in Example 4.26. For the second-order TV regularization, we use an adaptive regularization strength, which is

decreased at edges and slope discontinuities to avoid smoothing across these structures. To this end, let  $0 < \alpha_{\min} \leq \alpha_{\max}$  and  $0 < \gamma_{\min} \leq \gamma_{\max}$  be fixed. We set

$$\begin{aligned}\alpha(x; f) &= \alpha_{\max} - \chi_e(x; f)(\alpha_{\max} - \alpha_{\min}), \\ \beta(x; f) &= \alpha_{\max}, \\ \gamma(x; f) &= \gamma_{\max} - \min(\chi_e(x; f), \chi_{sd}(x; f))(\gamma_{\max} - \gamma_{\min}),\end{aligned}\tag{4.115}$$

and  $\mathcal{D}^{\text{loc}}$  as in the previous example as

$$\mathcal{D}^{\text{loc}}(x; f) := \text{Ell}(v_e(x; f), \alpha(x; f), \beta(x; f)) \times B_{\gamma(x; f)}(0).\tag{4.116}$$

Note that this approach shares the properties of the approach from Example 4.28, and in addition a weak ( $\gamma_{\min}$ ) second-order TV regularization at slope discontinuities.  $\diamond$

**Example 4.30** (Anisotropic first- and second-order TV). We recall the model we proposed in [98]. This adaptive regularization model, as the one before, takes into account edge information as well as information about the location of slope discontinuities. Different to Example 4.29, it is also anisotropic in the second-order regularization.

To motivate the approach, let us first consider the regularization of first- and second-order separately, i.e.  $\mathcal{R}(u; f)$  is split into

$$\mathcal{R}(u; f) = \mathcal{R}_1(u; f) + \mathcal{R}_2(u; f),\tag{4.117}$$

where

$$\mathcal{R}_i(u; f) := \sup \left\{ \int_{\Omega} u \operatorname{div}^i \varphi \, d\mathcal{L} \mid \varphi \in C_c^i(\Omega; \mathbb{R}^m), \varphi(x) \in \mathcal{D}_i^{\text{loc}}(x; f) \right\}\tag{4.118}$$

with  $m = 2$  for  $i = 1$  and  $m = 4$  for  $i = 2$ . In particular,  $\mathcal{R}_1(u; f)$  penalizes derivatives of  $u$  of first- and  $\mathcal{R}_2(u; f)$  derivatives of second-order.

For  $\mathcal{D}_1^{\text{loc}}(x; f)$  we choose the constraint sets from Example 4.26. The choice for the  $\mathcal{D}_2^{\text{loc}}(x; f)$  is motivated as follows. We first assume the smooth case  $u \in C^2(\Omega)$ . We are aiming at penalizing the second-order derivatives of  $u$  as  $\gamma_1 \|(Hu)\xi(f)\|_2 + \gamma_2 \|(Hu)\xi(f)^\perp\|_2$ , where  $\xi(f) = \xi(x; f)$  is some predefined vector field depending on  $f$  and

$$Hu(x) := \begin{pmatrix} \partial_{xx}u(x) & \partial_{xy}u(x) \\ \partial_{yx}u(x) & \partial_{yy}u(x) \end{pmatrix}\tag{4.119}$$

is the Hessian of  $u$ . Thus, we penalize the absolute values of the second-order derivative of  $u$  in direction of  $\xi(f)$  weighted by factor  $\gamma_1$  and in perpendicular direction weighted by  $\gamma_2$ .

Now, for arbitrary  $\varphi \in C_c^\infty(\Omega, \mathbb{R}^4)$ , let us consider

$$\int_{\Omega} (\operatorname{div}^2 \varphi)u \, d\mathcal{L} = \int_{\Omega} \langle \varphi, \nabla^2 u \rangle \, d\mathcal{L}.\tag{4.120}$$

Please recall from (4.4) that  $\nabla^2 u := (\partial_{x_1 x_1} u, \partial_{x_1 x_2} u, \partial_{x_2 x_1} u, \partial_{x_2 x_2} u)^\top$ . For a given normalized vector field  $\xi(x; f) = (\xi_1(x; f), \xi_2(x; f))^\top \in \mathbb{R}^2$ ,  $\|\xi(x; f)\|_2 = 1$ , we represent  $\varphi(f)$  as

$$\varphi(x) = t_1(x)w_1(x; f) + t_2(x)w_2(x; f) + s_1(x)w_3(x; f) + s_2(x)w_4(x; f), \quad (4.121)$$

where  $t, s \in \mathbb{R}^2$  and

$$\begin{aligned} w_1(x; f) &:= (\xi_1(x; f), \xi_2(x; f), 0, 0)^\top, \\ w_2(x; f) &:= (0, 0, \xi_1(x; f), \xi_2(x; f))^\top, \\ w_3(x; f) &:= (\xi_1^\perp(x; f), \xi_2^\perp(x; f), 0, 0)^\top, \\ w_4(x; f) &:= (0, 0, \xi_1^\perp(x; f), \xi_2^\perp(x; f))^\top. \end{aligned} \quad (4.122)$$

Note that  $\{w_i(x; f)\}_i$  form an orthonormal basis of  $\mathbb{R}^4$ . Then, standard calculus shows

$$\langle \varphi(x), \nabla^2 u(x) \rangle = \langle t(x), Hu(x)\xi(x; f) \rangle + \langle s(x), Hu(x)\xi^\perp(x; f) \rangle. \quad (4.123)$$

Now we calculate  $\gamma_1(x)\|(Hu(x))\xi(x; f)\|_2$  and  $\gamma_2(x)\|(Hu(x))\xi^\perp(x; f)\|_2$  for some weighting constants  $\gamma_1(x), \gamma_2(x) > 0$ . To this end, we take in (4.123) the point-wise supremum over all  $t(x) \in B_{\gamma_1}(0)$  and  $s(x) \in B_{\gamma_2}(0)$  and derive

$$\begin{aligned} \sup_{\substack{t(x) \in B_{\gamma_1}(x; f)(0), \\ s(x) \in B_{\gamma_2}(x; f)(0)}} \langle \varphi(x), \nabla^2 u(x) \rangle &= \gamma_1(x; f)\|(Hu(x))\xi(x; f)\|_2 + \gamma_2(x; f)\|(Hu(x))\xi^\perp(x; f)\|_2. \end{aligned} \quad (4.124)$$

The above considerations motivate the following definition for arbitrary  $u \in L^1(\Omega)$ :

$$\mathcal{R}_2(u; f) := \sup \left\{ \int_{\Omega} (\operatorname{div}^2 \varphi) u \, d\mathcal{L} \mid \varphi \in C_C^\infty(\Omega; \mathbb{R}^4), \varphi(x) \in \mathcal{D}^{loc}(x; f) \right\} \quad (4.125)$$

with

$$\begin{aligned} \mathcal{D}_2^{loc}(x; f) &:= \{q \in \mathbb{R}^4 : \langle q, w_1(x; f) \rangle^2 + \langle q, w_2(x; f) \rangle^2 \leq (\gamma_1(x; f))^2, \\ &\quad \langle q, w_3(x; f) \rangle^2 + \langle q, w_4(x; f) \rangle^2 \leq (\gamma_2(x; f))^2\}. \end{aligned} \quad (4.126)$$

Both constraint sets  $\mathcal{D}_1^{loc}(x; f) = \operatorname{Ell}(\xi(x; f), \alpha(x; f), \beta(x; f))$  and  $\mathcal{D}_2^{loc}(x; f)$  as in (4.126) fully determine the regularization term  $\mathcal{R}(u; f)$ .

Note that we used the same vector field  $\xi$  for first- and second-order regularization. One possible choice for  $\xi$  is  $\xi(x; f) = v_e(x; f)$  obtained from standard structure tensor. For the regularization strength, we use the same setting as in Example 4.29, see (4.115).  $\diamond$

Finally, we remark that the adaptive regularization models above include both first- and second-order terms. From a theoretical point of view the first-order term cannot be neglected, since otherwise coercivity of the functional would not hold. For the compactness of presentation, we will refer to these regularization approaches in the following chapters as *second-order* TV regularization, keeping in mind that they always include first-order regularization.

# Chapter 5

## Non-Convex Optimization

### 5.1 Motivation

In the previous sections we have considered convex minimization problems of the form

$$\min_{u \in \mathcal{K}} \mathcal{F}(u) \quad (5.1)$$

with a convex functional  $\mathcal{F} : X \rightarrow \mathbb{R}$  and a closed convex constraint set  $\mathcal{K}$ . In this case, extensive theory on existence and uniqueness of minimizers as been established in the literature.

In case that either the set  $\mathcal{K}$  or the functional  $\mathcal{F}(u)$  are not convex, existence of a minimizer in general is in question.

In the following, we focus on case that  $\mathcal{K}$  is convex but  $\mathcal{F}$  is not and discuss raising problems. In particular, we consider examples where the non-convexity is caused by a weighting of the data term with weights depending on  $\nabla u$ .

For compactness of presentation, let us assume that  $\mathcal{K} = X$ . Moreover, let us assume that  $\mathcal{F}(u)$  takes the form

$$\mathcal{F}(u) := \int_{\Omega} F(x, u(x), \nabla u(x)) \, dx \quad (5.2)$$

with  $F(x, \xi, \nu) : \Omega \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ . The form in (5.2) is a common assumption for variational problems, see e.g. [49].

The standard strategy to show existence of a minimizer of (5.2), cf. Remark 3.7 requires three main ingredients, which are (i) the *coercivity* of  $\mathcal{F}$  (cf. Definition 3.4), which guarantees boundedness of a minimizing sequence  $(u^k)_k$ , (ii) *pre-compactness* of sublevel sets in a suitable topology, which guarantees convergence of a subsequence of  $(u^k)_k$  to some function  $u^*$ , and (iii) *lower semi-continuity* of  $\mathcal{F}(u)$  (cf. Definition 3.5), which guarantees that  $u^*$  is a minimizer.

Let us recall a result from Dacorogna [49]:

**Theorem 5.1.** *Let  $\mathcal{F}(u) : W^{1,p}(\Omega) \rightarrow \mathbb{R}$  be of form (5.2). Then, if  $\mathcal{F}(u)$  is weakly lower semi-continuous, it follows that  $\nu \rightarrow F(x, \xi, \nu)$  is convex.*

(cf. Theorem 3.1 and the subsequent remark in [49]). A consequence of Theorem 5.1 is that functionals of the form (5.2), the property of  $\nu \rightarrow F(x, \xi, \nu)$  being convex is essential to prove existence of a minimizer, see e.g. Theorem 3.6.

Let us now assume that  $F$  does not have this property. Then, as a remedy, we can consider a relaxation (see e.g. [148]) of  $\mathcal{F}$ , which enlarges the set of minimizers in a way that their existence becomes provable. We can show that if minimizers of  $\mathcal{F}$  exist, that they are also minimizers of the relaxed functional. For this reason, we refer to the minimizers of the relaxed functional as *generalized* minimizers. We discuss this concept of relaxation in detail below.

## 5.2 Problem Relaxation

In the following, we focus on the functional

$$\begin{aligned} \mathcal{F} : W^{m,p}(\Omega) &\rightarrow \mathbb{R} \cup \{+\infty\}, \\ \mathcal{F}(u) &:= \int_{\Omega} F(x, u(x), \nabla u(x)) \, dx, \end{aligned} \tag{5.3}$$

where  $\Omega \subset \mathbb{R}^d$  and  $F(x, \xi, \nu)$  is non-convex with respect to  $\nu$ . Please note that we are assuming *scalar* functions  $u : \Omega \rightarrow \mathbb{R}$ . We will discuss vector-valued functions in Section 5.3. For the sake of compactness, we restrict the following presentation to the case  $m = 1$  and  $p = 1, 2$ . Sobolev spaces of higher order can be treated analogously. In case that a higher  $p$  is required for the well-posedness of a specific  $\mathcal{F}$ , the required coercivity and lower semi-continuity can be considered with respect to the weak topology of  $W^{m,2}(\Omega)$ .

As an example for an  $F(x, \xi, \nu)$ , which is non-convex with respect to  $\nu$ , we consider the following:

**Example 5.2.** For  $p \in \{1, 2\}$  and  $u^0 : \Omega \rightarrow \mathbb{R}$  let

$$F(x, \xi, \nu) := \frac{1}{2} \frac{(\xi - u^0)^2(x)}{\|\nu\|_2^p} + \alpha \|\nu\|_2^p. \tag{5.4}$$

The motivation for introducing (5.4) is as follows. Considering the functional

$$\mathcal{F}(u) = \int_{\Omega} F(x, u(x), \nabla u(x)) \, dx, \tag{5.5}$$

with  $F$  as in (5.4), we identify the first term  $\frac{1}{2} \frac{(\xi - u^0)^2}{\|\nu\|_2^p}$  as data term and the second term  $\alpha \|\nu\|_2^p$  as regularization term. While the regularization term is based on a penalization  $\|\nabla u\|_2^p$ , which is a standard approach, the data term differs from the standard form  $\frac{1}{2} \|u - f\|_{L^2}^2$  by the factor  $1/\|\nabla u(x)\|_2^p$ , which we interpret as an additional weighting. The weighting introduced is weak ( $\approx 0$ ) in regions where the gradient of  $u$  takes large (absolute) values and strong ( $\approx \infty$ ) in regions where  $u$  is



flat. In Chapter 8 we will consider an application, which requires such a weighting of the data term.

Please note that we allow  $\mathcal{F}(u)$  to take the value  $+\infty$ , such that with the particular  $F((x, u(x), \nabla u(x)))$  in (5.4) it is well-defined on  $W^{1,p}(\Omega)$ .  $\diamond$

Before we address the issue of weakly lower semi-continuity, let us first discuss appropriate function spaces. As mentioned before, for coercivity we require pre-compactness of sublevel sets of  $\mathcal{F}$ . Assuming that the sublevel sets are bounded in  $W^{1,p}(\Omega)$ , we require the pro-compactness of bounded sets. While in the case  $p = 2$  it suffices to consider  $W^{1,2}(\Omega)$  with the weak topology, in the case  $p = 1$  we have to extend  $W^{1,1}(\Omega)$  to  $BV(\Omega)$ , and use its weak\* topology, see Definition 2.2.

Recall that, since  $F(x, \xi, \nu)$  is non-convex with respect to  $\nu$ , functional  $\mathcal{F}(u)$  cannot be weakly lower semi-continuous on  $W^{1,p}(\Omega)$ . As a remedy, we consider the *relaxation* of  $\mathcal{F}$ . We briefly recall the main definitions and results below and refer to [148] for more details.

**Definition 5.3** (Weakly Lower Semi-Continuous Relaxation).

- **Relaxation in  $W^{1,p}(\Omega)$**  We define the relaxation of  $\mathcal{F} \in W^{1,p}(\Omega)$  in  $X \subseteq W^{1,p}(\Omega)$  as

$$\mathcal{F}_\diamond(u) := \inf \left\{ \liminf_k \mathcal{F}(u^k) \mid (u^k)_k \in X, u^k \rightharpoonup u \in W^{1,p}(\Omega) \right\}, \quad (5.6)$$

where we set  $\mathcal{F}_\diamond(u) := \infty$ , if no such sequence exists.

- **Relaxation in  $BV(\Omega)$**  We define the relaxation of  $\mathcal{F} \in W^{1,p}(\Omega)$  in  $X \subseteq BV(\Omega)$  as

$$\mathcal{F}_{\diamond BV}(u) := \inf \left\{ \liminf_k \mathcal{F}(u^k) \mid (u^k)_k \in W^{1,1}(\Omega), u^k \overset{*}{\rightharpoonup} u \in BV(\Omega) \right\}. \quad (5.7)$$

It can be shown that under certain growth conditions on  $F$  (yielding coercivity of  $\mathcal{F}$ ) that the relaxed functionals are weakly/ weakly\* lower semi-continuous (Lemma 5.4 and Theorem 5.8 in [148]) and that minimizers in the function spaces  $W^{1,2}(\Omega)$  ( $p = 2$ ) and  $BV(\Omega)$  ( $p = 1$ ) exist (Corollary 5.17 in [148] for  $p = 2$  and Lemma 5.20 for  $p = 1$ ). We refer to the minimizers of  $\mathcal{F}_\diamond(u)$  and  $\mathcal{F}_{\diamond BV}(u)$  as *generalized minimizers* of  $\mathcal{F}(u)$ , as the set of generalized minimizers contain minimizers of  $\mathcal{F}(u)$ , provided that the later exist. To show this, assume that a minimizer  $u^*$  of the functional  $\mathcal{F}(u)$  exists. In the case  $p = 2$ , we have  $u^* \in W^{1,2}(\Omega)$ . Considering the constant sequence  $u^k := u^*$ , we see that

$$\mathcal{F}(u^*) = \inf_{u \in W^{1,2}(\Omega)} \mathcal{F}(u) \leq \mathcal{F}_\diamond(u^*) \leq \liminf_{k \rightarrow \infty} \mathcal{F}(u^k) = \mathcal{F}(u^*), \quad (5.8)$$

and thus  $\mathcal{F}(u^*) = \mathcal{F}_\diamond(u^*)$ . Since  $\mathcal{F}_\diamond(u^*)$  by definition cannot attain values less than  $\inf_{u \in W^{1,2}(\Omega)} \mathcal{F}(u)$ ,  $u^*$  is a minimizer of  $\mathcal{F}_\diamond(u)$ , i.e. a generalized minimizer. This argumentation carries over to the case  $p = 1$  using the fact that

$$\mathcal{F}_{\diamond BV}(u) = \inf \left\{ \liminf_k \mathcal{F}(u^k) \mid (u^k)_k \in BV(\Omega), u \overset{*}{\rightharpoonup} u \in BV(\Omega) \right\}, \quad (5.9)$$

see [69, proof of Theorem 2].

It remains to provide analytic formulas for the relaxed functionals. For the relaxation in  $W^{1,2}(\Omega)$ , it can be shown that

$$\mathcal{F}_\diamond(u) = \int_{\Omega} F_c(x, u(x), \nabla u(x)) \, dx, \quad (5.10)$$

see e.g. [148, Theorem 5.5.], where  $F_c(x, \xi, \cdot)$  is the convex envelope of  $F(x, \xi, \cdot)$  given as

$$F_c(x, \xi, \nu) := \sup \{ F_c(x, \xi, \nu) \mid F_c \leq F \text{ and } F_c(x, \xi, \nu) \text{ is convex w.r.t. } \nu \}. \quad (5.11)$$

For our particular function  $F$  in Example (5.2),  $F_c$  takes the form

$$F_c(x, \xi, \nu) := \begin{cases} \frac{(\xi - u^{(0)}(x))^2}{2\|\nu\|_2^p} + \alpha\|\nu\|_2^p & \text{if } \sqrt{2\alpha}\|\nu\|_2^p > |\xi - u^{(0)}(x)|, \\ \sqrt{2\alpha}|\xi - u^{(0)}(x)| & \text{else,} \end{cases} \quad (5.12)$$

(cf. Lemma 5.12 in [148]).

For the relaxation in  $BV(\Omega)$  retrieving an analytic form for the general case is more involved and requires some additional assumptions on  $F$ . We refer to Definition 5.9, Theorem 5.10 and Corollary 5.11 in [148] for details. Applying this theory to our example (cf. Theorems 5.14 and 5.18 in [148]), one can show that

$$\mathcal{F}_{\diamond BV}(u) = \int_{\Omega} F_c(x, u(x), \nabla u(x)) \, dx + \alpha \, d|D^s(u)|(\Omega), \quad (5.13)$$

where  $|D^s(u)|(\Omega)$  is the singular part of the measure  $|Du|$  of  $u$  (see Section 2.7).

### 5.3 Casting Variational Problems into Gradient Flows

In the previous section we considered functionals  $\mathcal{F}(u)$  defined on scalar functions  $u : \Omega \rightarrow \mathbb{R}$ , where  $\Omega \subset \mathbb{R}^d$ . Let us now consider the vectorial case  $u : \Omega \rightarrow \mathbb{R}^m$ . On  $\mathcal{F}$  we assume the same form as before:

$$\mathcal{F}(u) := \int_{\Omega} F(x, u(x), \nabla u(x)) \, dx \quad (5.14)$$

with  $F(x, \xi, J) : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \rightarrow \mathbb{R}$ .

**Example 5.4.** *Our aim is to introduce an adaptive weighting of the data term*

$$\mathcal{S}(u) = \frac{1}{2} \|u - u^0\|_{L^2(\Omega)}^2 \quad (5.15)$$

similar to the scalar setting discussed above, but now with vector-valued functions  $u$  and  $u^0$ . The most general form of a weighting matching the form (5.2) would be

$$\frac{1}{2}(u(x) - u^0)^\top A(x, u, \nabla u)(u(x) - u^0(x)), \quad (5.16)$$

where  $\nabla u$  now is the Jacobian of  $u$  and  $A(x, \xi, J) \in \mathbb{R}^{m \times m}$  is positive semi-definite.

Combining such a data term with a regularization term on the Jacobian of  $u$ , we derive

$$F(x, u, \nabla u) := \frac{1}{2}(u(x) - u^0(x))^\top A(x, u, \nabla u)(u(x) - u^0(x)) + \frac{\alpha}{p} \|\nabla u\|_F^p, \quad (5.17)$$

where  $\|\cdot\|_F$  is the Frobenius norm. We are in particular interested in cases where  $A(x, \xi, J)$  is non-convex with respect to  $J$ . One particular example would be to set

$$A(x, \xi, J) = (J(x)^\top J(x))^\dagger, \quad (5.18)$$

where  $B^\dagger$  for arbitrary matrix  $B$  denotes the pseudo-inverse of  $B$ . In case of scalar functions  $u$  and  $u^0$ , this example simplifies to Example 5.2 with  $p = 2$ .  $\diamond$

In principle, we can follow the strategy of relaxing the functional as describe in the previous section. In order to derive an analytic form of the relaxed functional  $\mathcal{F}_\diamond(u)$  one then has to calculate the quasi-convex envelope of the function  $F(x, \xi, J)$  with respect to  $J$ .

However, for functions of the form (5.17) in the case  $m > 1$  and  $d > 1$ , an analytic form for the quasi-convex envelope is not known so far, and thus efficient numerical minimization based on this approach is not at hand.

To overcome this issue, we have proposed in [104, 105] to use semi-group theory to derive a related PDE approach. Below, we recall the general semi-group solution concept [30] as well as its special adaptation to handle the functional (5.14) with  $F$  as in (5.17)

The *convex semi-group solution concept*, cf. [30], is briefly summarized as follows. Let  $\mathcal{R} : H \rightarrow \mathbb{R} \cup \{\infty\}$  be a convex functional on a Hilbert space  $H$ , and let  $u_\alpha$  be a minimizer of the functional

$$\mathcal{G}_{u^{(0)}}(u) := \frac{1}{2} \|u - u^{(0)}\|_H^2 + \alpha \mathcal{R}(u). \quad (5.19)$$

Then, for a smooth initial function  $u^{(0)}$  and  $\alpha \rightarrow 0$ ,  $(u_\alpha - u^{(0)})/\alpha$  converges to an element of the subgradient  $\partial \mathcal{R}(u^{(0)})$  of  $\mathcal{R}$ . Performing an iterative minimization of  $\mathcal{G}_u$ , i.e.

$$u^{(k)} \in \arg \min_u \mathcal{G}_{u^{(k-1)}}(u), \quad (5.20)$$

yields an approximation of the solution of the flow

$$\partial_t u \in \partial \mathcal{R}(u) \quad (5.21)$$

at time  $t = k\alpha$ . In other words, the variational approach (5.19) approximates the gradient flow equation (5.21). For convex semi-groups the solutions of diffusion filtering (i.e. (5.21)) and iteratively applied variational methods (i.e. (5.20)) are comparable and look rather similar [149].

We expect a similar behavior for the non-convex functional

$$\mathcal{F}_{u^0}(u) := \int_{\Omega} \frac{1}{2} (u(x) - u^{(0)}(x))^{\top} A(x, u(x), \nabla u(x)) (u(x) - u^{(0)}(x)) + \frac{\alpha}{p} \|\nabla u(x)\|_F^p dx \quad (5.22)$$

( $u^{(0)} = u^0$ ) under the assumption that  $A(x, \xi, J)$  can be inverted.

To this end, let us assume that  $A(x, \xi, J)$  is indeed invertible. We comment on non-invertible  $A$  at the end of this section. We derive the according flow equation, which is the gradient flow associated with (5.22), as follows. Please note that this derivation is obtained in a purely formal manner. First, we introduce the abbreviations  $A(u) = A(x, u, \nabla u)$  and

$$S_{(k-1)}(u) := \frac{1}{2} \int_{\Omega} (u - u^{(k-1)})^{\top} A(u) (u - u^{(k-1)}) d\mathcal{L}. \quad (5.23)$$

The *directional derivative* of  $S_{(k-1)}$  at  $u$  in direction  $\phi$  (provided it exists) satisfies

$$\begin{aligned} \partial_{\tau} S_{(k-1)}(u + \tau\phi) &= \int_{\Omega} \phi^{\top} A(u) (u - u^{(k-1)}) d\mathcal{L} \\ &\quad + \frac{1}{2} \int_{\Omega} (u - u^{(k-1)})^{\top} \partial_{\phi} A(u) (u - u^{(k-1)}) d\mathcal{L}, \end{aligned} \quad (5.24)$$

where

$$\partial_{\phi} A(u) := \lim_{\tau \rightarrow 0} \frac{A(u + \tau\phi) - A(u)}{\tau}. \quad (5.25)$$

In a similar formal manner, the directional derivative of  $\mathcal{R}(u) = \frac{\alpha}{p} \int_{\Omega} \|\nabla u\|_F^p$  at  $u$  in direction  $\phi$  can be derived and reads

$$\partial_{\tau} \mathcal{R}(u + \tau\phi) = \alpha \int_{\Omega} \nabla \phi^{\top} \frac{\nabla u}{\|\nabla u\|_F^{2-p}} d\mathcal{L}. \quad (5.26)$$

Note, that by the right hand side of (5.26) in the case  $p = 1$  actually the sub-differential of the total variation semi-norm evaluated in direction  $\phi$  is meant.

Then, using (5.24) and (5.26), with  $v^{(k)} := u^{(k)} - u^{(k-1)}$  and dividing by  $\alpha > 0$ , formally gives the following optimality condition for the minimizer  $u^{(k)}$  of  $\mathcal{F}_{u^{(k-1)}}$ :

$$\int_{\Omega} \phi^{\top} A(u^{(k)}) \frac{v^{(k)}}{\alpha} + \frac{(v^{(k)})^{\top}}{2\alpha} \partial_{\phi} A(u^{(k)}) v^{(k)} d\mathcal{L} = \int_{\Omega} \nabla \phi^{\top} \frac{\nabla u^{(k)}}{\|\nabla u^{(k)}\|_F^{2-p}} d\mathcal{L} \quad (5.27)$$

for all  $\phi \in C_c^{\infty}(\Omega; \mathbb{R}^n)$ . Let  $t > 0$  be fixed and  $k = \lfloor t/\alpha \rfloor$  (where  $\lfloor x \rfloor$  denotes the largest integer values less than or equal to  $x$ ). Then, as in the convex case, we can

expect that  $v^{(k)}/\alpha$  converges to  $\partial_t u(t)$  for  $\alpha \rightarrow 0$ . From that, in turn, it follows that  $v^{(k)} \rightarrow 0$ . This together with (5.27) then shows that

$$\int_{\Omega} \phi^\top A(u(t)) \partial_t u(t) d\mathcal{L} = - \int_{\Omega} \nabla \phi^\top \frac{\nabla u(t)}{\|\nabla u(t)\|_F^{2-p}} d\mathcal{L}. \quad (5.28)$$

Applying in a formal way Green's formula and the fundamental lemma, from (5.28) the strong formulation is derived and reads

$$A(u(t)) \partial_t u(t) = \operatorname{div} \cdot \left( \frac{\nabla u(t)}{\|\nabla u(t)\|_F^{2-p}} \right), \quad (5.29)$$

where  $u(t)$  satisfies homogeneous Neumann boundary conditions. Multiplying both sides of (5.29) by  $(A(u(t)))^{-1}$ , we get

$$\partial_t u(t) = (A(u(t)))^{-1} \operatorname{div} \cdot \left( \frac{\nabla u(t)}{\|\nabla u(t)\|_F^{2-p}} \right). \quad (5.30)$$

Moreover, the initial condition associated with the flow is  $u(0) := u^0$ .

It remains one open issue, which is the required invertibility of matrix  $A(u(t))$ . In (5.18) we have considered a matrix-valued function  $A(u(t))$ , which does not provide invertible matrices. Let us revisit this example again to show how this issue can be resolved:

**Example 5.5.** *In the following, for compactness of notation, we omit the arguments  $t$  and  $x$  of  $u$ . Let*

$$A(u) := (\nabla u^\top \nabla u)^\dagger. \quad (5.31)$$

*To assert invertibility of  $A(u)$ , we regularize this matrix by adding  $\varepsilon \operatorname{Id}$  before applying the pseudo-inverse, which then can be replaced by the inverse:*

$$A_\varepsilon(u) := (\nabla u^\top \nabla u + \varepsilon \operatorname{Id})^{-1}. \quad (5.32)$$

*We can now apply the above concept with  $A_\varepsilon(u)$  instead of  $A(u)$  and derive*

$$\partial_t u = (\nabla u^\top \nabla u + \varepsilon \operatorname{Id}) \operatorname{div} \left( \frac{\nabla u}{\|\nabla u\|_F^{2-p}} \right). \quad (5.33)$$

*Now, we can formally let  $\varepsilon$  tend to zero and derive*

$$\partial_t u = (\nabla u^\top \nabla u) \operatorname{div} \left( \frac{\nabla u}{\|\nabla u\|_F^{2-p}} \right). \quad (5.34)$$

◇

**Remark 5.6.** For scalar data ( $m = 1$ ) and  $p = 1$  the equation (5.34) reads as

$$\partial_t u = \|\nabla u\|_2^2 \operatorname{div} \left( \frac{\nabla u}{\|\nabla u\|_2} \right). \quad (5.35)$$

One recognizes that (5.35) differs from the mean curvature flow (MCF) equation (see e.g. [112]), given as

$$\partial_t u = \|\nabla u\|_2 \operatorname{div} \left( \frac{\nabla u}{\|\nabla u\|_2} \right) \quad (5.36)$$

only by the power of the leading factor  $\|\nabla u\|_2$ .

As an alternative to  $A$  as in (5.31), we can consider

$$A(u) := ((\nabla u)^\top \nabla u)^{\frac{p}{2}\dagger}, \quad (5.37)$$

where the power  $\frac{p}{2}$  of the matrix  $(\nabla u)^\top \nabla u$  is defined via spectral decomposition. Of particular interest is the case  $p = 1$ , because then the functional (5.22) is invariant under affine rescaling of the image intensity. In this case the semi-group approach results in the gradient flow

$$\partial_t u = ((\nabla u)^\top \nabla u)^{\frac{1}{2}} \operatorname{div} \left( \frac{\nabla u}{\|\nabla u\|_2} \right), \quad (5.38)$$

which, in the scalar case, is exactly the MCF equation. For scalar radial-symmetric monotonic data an analytical comparison of the solutions of (5.38) and the MCF equation (5.36) has been provided in [50].

Despite the fact that it is not an *equivalent* reformulation of the variational problem (5.14), we propose to solve PDE (5.33) as an alternative to tackling the minimization of (5.14) with the considered non-convex  $F$ . We will discuss corresponding applications in Chapter 8.

# Chapter 6

## Quasi-Variational Inequalities for Regularization

### 6.1 Introduction

In this chapter, we generalize the variational approach from Chapter 4. Please recall, that in Chapter 4, we considered

$$\min_{u \in X} \mathcal{F}(u; v) := \min_{u \in X} \mathcal{S}(u) + \mathcal{R}(u; v), \quad (6.1)$$

to solve an inverse problem

$$A(u) = f \quad (6.2)$$

for some operator  $A$  and data  $f$ . The variable  $v$  in the regularization term  $\mathcal{R}(u; v)$  in (6.1) was introduced to allow adaptivity. In Chapter 4 we considered *data-driven* adaptive regularization, where the adaptivity was determined based on the (pre-smoothed) data  $v = f$ . The core idea of this data-driven adaptivity is that prominent image structures such as edges, corners and intensity gradients of the unknown function  $u_{true}$ , which has to be reconstructed from  $f$ , are estimated based on the given data  $f$ . These data, however, are only linked to  $u$  via the operator  $A$  (e.g. identity in the case of denoising, a blur operator in case of deblurring and a restriction to a subdomain of  $\Omega$  in case of inpainting) and in addition are effected by noise, so that the estimate of the sought image structures might be inaccurate.

Instead of using  $v = f$ , it would be more appropriate to determine the image structures directly on the reconstructed image  $u$  as an approximation of the unknown function  $u_{true}$ , i.e. to consider a solution-driven adaptivity instead of a data-driven one.

One alternative would be to directly replace the function  $f$  in  $\mathcal{R}(u; f)$  by  $u$ , such that the optimization problem becomes

$$\min_{u \in X} \mathcal{S}(u) + \mathcal{R}(u; u). \quad (6.3)$$

By this modification, however, the regularization problem in general becomes non-convex, so that existence and uniqueness of a minimizer as well as the numerical

treatment of the minimization problem become an issue. A remedy can be to consider a relaxation of this problem, as it is for example done in [68].

Here, we are heading for a different strategy: We first consider an arbitrary function  $u_0 \in X$  and the convex optimization problem

$$\bar{u} := \arg \min_{u \in X} \mathcal{S}(u) + \mathcal{R}(u; u_0). \quad (6.4)$$

Please note that in order to have  $\bar{u}$  well-defined, we required the strict-convexity of  $\mathcal{S}(u)$ , which guarantees the uniqueness of the minimizer.

We then search for a fixed point  $u^*$  of the mapping

$$u_0 \mapsto T(u_0) := \bar{u}. \quad (6.5)$$

It turns out that providing existence and uniqueness theory is easier for the fixed point problem than for the non-convex optimization problem (6.3). Moreover, the fixed point problem is the more natural approach: many existing methods in image processing and computer vision, such as iterated non-local means [32], consider an outer iteration for updating auxiliary or guiding variables and implicitly are expecting a fixed point.

We illustrate our approach by an example:

**Example 6.1.** *We revisit Example 4.25, where we considered an adaptive regularization for the task of denoising. The data term in this example is chosen as*

$$\mathcal{S}(u) := \frac{1}{2} \|u - f\|_{L^2}^2 \quad (6.6)$$

with data  $f \in L^2(\Omega)$ . As regularization term we use

$$\mathcal{R}(u; v) = \sup \left\{ \int_{\Omega} u \operatorname{div} \varphi \, d\mathcal{L} \mid \varphi \in C_c^1(\Omega; \mathbb{R}^2), \varphi(x) \in \mathcal{D}^{loc}(x; v) \right\}, \quad (6.7)$$

where each  $\mathcal{D}^{loc}(x; v)$  is a closed ball centered at 0 with radius

$$\alpha(x; v) = \max(\alpha_{max}(1 - \kappa \|\nabla v_{\sigma}\|), \varepsilon). \quad (6.8)$$

Recall that by  $v_{\sigma}$  we denote the convolution of  $v$  with a Gaussian kernel  $K_{\sigma}$  with standard deviation  $\sigma$ . Parameter  $\kappa > 0$  in (6.8) controls the edge sensitivity, while  $\max(\cdot, \varepsilon)$  with a small  $\varepsilon > 0$  asserts the positive boundedness  $\alpha$  from below, as required for existence theory (cf. Assumption 4.10 (iv)).

In Example 4.25, we have chosen an data-driven adaptivity by setting  $v = f$ . In order to cope with the noise in the data  $f$ , in general a strong pre-smoothing is required, which, however, besides eliminating noise blurs the input and removes weak edges. A pre-smoothing thus comes with a trade-off between being robust against noise and a preservation of edges.



A direct *solution-driven approach* would consist in first assuming  $u \in W^{1,1}(\Omega)$  and setting

$$\mathcal{R}(u) = \int_{\Omega} \alpha(x, u) \|\nabla u\|_2 \, d\mathcal{L}, \quad (6.9)$$

and then extending the functional to the space  $BV(\Omega)$ :

$$\mathcal{R}(u) := \sup \left\{ \int_{\Omega} u(x) \operatorname{div}(\alpha(x, u)\varphi(x)) \, d\mathcal{L} \mid \varphi \in C_c^l(\Omega; \mathbb{R}^2), \varphi(x) \in B_1(0) \right\}. \quad (6.10)$$

Please note that  $\mathcal{R}(u)$  in (6.9) and (6.10) is non-convex. Assuming a sufficiently high regularity of  $\alpha(x, u)$ , one can show weakly lower semi-continuity of  $\mathcal{R}(u)$  in (6.10). Together with coercivity, the existence of a minimizer can be shown. However, proving uniqueness of a minimizer of the corresponding  $\mathcal{F}(u)$  and providing an efficient numerical approach to minimize  $\mathcal{F}(u)$  are open issues.

At this point, one can think of applying a relaxation approach as discussed in Chapter 5. In [67] we have shown for one particular adaptive regularization approach that from a theoretical point of view this is an option. In particular, with this theory one can show existence of generalized minimizers. However, the optimization of the relaxed problems is an open issue.

We therefore propose a different approach.

In this approach we consider  $\alpha(x, u_0)$  for some given function  $u_0$  and  $\mathcal{D}^{loc}(x, u_0)$  and  $\mathcal{R}(u; u_0)$  as defined above. With  $u_0$  fixed, the functional  $\mathcal{F}(u; u_0)$  is strict-convex with respect to  $u$  (due to the strict-convexity of  $\mathcal{S}(u)$ ), and thus a unique minimizer  $\bar{u}$  exists. Then, we search for a fixed point of  $T : u_0 \mapsto \bar{u}$ .

We stress that one benefit of our solution-driven approach is, that the approach has not to be robust against noise, so that we can significantly reduce  $\sigma$ . For theoretical reasons, however, a small  $\sigma$ , is needed to guarantee that  $\nabla u(x)$  is well-defined. (In practice, after discretization, this is not an issue.)  $\diamond$

In the next section we provide a general framework for solution-driven adaptive TV regularization.

## 6.2 Approach

For solution-driven adaptive regularization we consider regularization terms of the form

$$\mathcal{R}(u; u_0) := \sup \left\{ \int_{\Omega} u D\varphi \, d\mathcal{L} \mid \varphi \in \mathcal{D}(u_0) \right\}, \quad (6.11)$$

with some differential operator  $D$ . The constraint set  $\mathcal{D}(u_0)$  is given as

$$\mathcal{D}(u_0) := \left\{ \varphi \in C_c^l(\Omega; \mathbb{R}^m), \varphi(x) \in \mathcal{D}^{loc}(x; u_0) \right\}, \quad (6.12)$$

where the local constraint sets  $\mathcal{D}^{loc}(x, u_0)$  depending on some arbitrary function  $u_0$ . Recall that we consider functionals of the form

$$\mathcal{F}(u; u_0) := \mathcal{S}(u) + \mathcal{R}(u; u_0), \quad (6.13)$$

where  $\mathcal{S} : L^2(\Omega) \rightarrow \mathbb{R}$  is the data term. One special case we focus on below is

$$\mathcal{S}(u) = \frac{1}{2} \|A u - f\|_{L^2(\Omega_f)}^2 \quad (6.14)$$

with some linear operator  $A : L^2(\Omega) \rightarrow L^2(\Omega_f)$ .

Now, we turn to the fixed point problem introduced above. To this end, let  $\bar{u}$  denote the minimizer of  $\mathcal{F}(u; u_0)$  (assuming that  $\mathcal{F}$  is strictly convex and thus  $\bar{u}$  is unique). We are interested in finding a fixed point  $u^*$  of

$$u_0 \mapsto T(u_0) := \bar{u} = \arg \min_{u \in X} \mathcal{F}(u; u_0). \quad (6.15)$$

We will provide existence theory for this fixed point problem in Section 6.3.

In the discretized setting and under sufficient conditions on the data term, it turns out that this fixed point problem is equivalent to a quasi-variational inequality problem (QVIP), for which comprehensive theory is available in the literature, see e.g. [12, 20, 37, 127]. In the case  $\mathcal{S}(u) = \frac{1}{2} \|A u - f\|_2^2$  we obtain this QVIP as follows. Moving to a discrete formulation of the problem, we now assume that  $u, f \in \mathbb{R}^n$  are the function values at the  $n$  nodes of an equidistant  $d$ -dimensional grid. Moreover, we replace the linear operator  $A$  by a  $n \times n$  matrix, also denoted by  $A$ . In what follows we assume that  $A$  is invertible, which is required for  $\mathcal{S}(u)$  to be strictly convex. We denote by  $L$  the discretization of the divergence operator  $D$  (with appropriate boundary conditions). The optimization problem we consider in the discrete setting is given as

$$\arg \min_{u \in \mathbb{R}^n} \mathcal{F}(u; u_0) = \frac{1}{2} \|A u - f\|_2^2 + \sup_{p \in \mathcal{D}(u_0)} (Lp)^\top u. \quad (6.16)$$

The constraint set  $\mathcal{D}(u_0)$  in (6.16) is given as

$$\mathcal{D}(u_0) = \{p \in \mathbb{R}^{nd}, p_i \in \mathcal{D}_i^{loc}(u_0), i = 1, \dots, n\}. \quad (6.17)$$

In (6.17) we use the notation  $p = (p_1, p_2, \dots, p_n)^\top$  with  $p_i \in \mathbb{R}^d$ ,  $i = 1, \dots, n$ . The local constraint sets  $\mathcal{D}_i^{loc}(u_0) \subset \mathbb{R}^d$  are provided now for every grid node  $i$ . Again, we assume them to be non-empty, closed, convex and depending on  $u_0$ .

To describe our ansatz, we consider first the dual problem of (6.16), which is derived as follows. The optimality condition for  $\bar{u}$  reads

$$A^\top (A \bar{u} - f) + Lp = 0. \quad (6.18)$$

We deduce from (6.18) that

$$\begin{aligned} \bar{u} &= (A^{-1} A^{-\top})(A^\top f - Lp) = A^{-1}(f - A^{-\top} Lp), \\ A \bar{u} &= f - A^{-\top} Lp, \end{aligned} \quad (6.19)$$

where  $A^{-\top} := (A^\top)^{-1}$ . Inserting (6.19) in (6.16) and using the abbreviation  $M := A^{-\top} L$ , we obtain

$$\begin{aligned}
 \mathcal{F}^*(p) &= \frac{1}{2} \|A\bar{u} - f\|_2^2 + (Lp)^\top \bar{u} \\
 &= \frac{1}{2} \|A^{-\top} Lp\|_2^2 + (A^{-1}f)^\top Lp - (A^{-1}A^{-\top}Lp)^\top Lp \\
 &= \frac{1}{2} \|A^{-\top} Lp\|_2^2 + f^\top A^{-\top} Lp - (A^{-1}A^{-\top}Lp)^\top Lp \\
 &= -\frac{1}{2} \|Mp\|_2^2 + f^\top Mp - (Mp)^\top Mp \\
 &= -\frac{1}{2} \|Mp - f\|_2^2 + \frac{1}{2} \|f\|_2^2.
 \end{aligned} \tag{6.20}$$

The dual formulation of (6.16) therefore reads as

$$\sup_{p \in \mathcal{D}(u_0)} \mathcal{F}^*(p) = \sup_{p \in \mathcal{D}(u_0)} -\frac{1}{2} \|Mp - f\|_2^2 + \frac{1}{2} \|f\|_2^2. \tag{6.21}$$

When maximizing  $\mathcal{F}^*(p)$  over  $\mathcal{D}(v)$ , the constant term  $\frac{1}{2} \|f\|_2^2$  can be omitted without changing the optimum. Moreover, by switching from the maximization of  $\mathcal{F}^*(p)$  to the minimization of  $-\mathcal{F}^*(p)$ , we can reformulate the dual problem (6.21) as

$$\arg \min_{p \in \mathcal{D}(u_0)} \mathcal{G}(p), \quad \mathcal{G}(p) := \frac{1}{2} \|Mp - f\|_2^2. \tag{6.22}$$

From a solution  $\bar{p}$  of the dual problem (6.22) we can retrieve the solution  $\bar{u}$  of the primal problem (6.16) by  $\bar{u} = A^{-1}(f - M\bar{p})$ . As necessary and sufficient optimality condition of (6.22) we obtain the following variational inequality:

$$\begin{aligned}
 &\text{find } \bar{p} \in \mathcal{D}(u_0) \text{ such that} \\
 &\quad \langle \nabla \mathcal{G}(\bar{p}), p - \bar{p} \rangle \geq 0 \quad \forall p \in \mathcal{D}(u_0) \\
 &\quad \Leftrightarrow \langle M^\top (M\bar{p} - f), p - \bar{p} \rangle \geq 0 \quad \forall p \in \mathcal{D}(u_0).
 \end{aligned} \tag{6.23}$$

Recall that we are interested in a fixed point  $u^*$  of

$$u_0 \mapsto T(u_0) := \arg \min_{u \in X} \mathcal{F}(u; u_0), \tag{6.24}$$

i.e.  $u^* = \bar{u} = T(u^*)$ . As we will see below, it makes sense to denote the corresponding dual variable, i.e. the solution  $\bar{p}$  of (6.23) with the choice  $u_0 = u^*$  by  $p^*$ . Inserting  $\bar{p} = p^*$  and  $u_0 = u^* = A^{-1}(f - Mp^*)$  into (6.23), we obtain the quasi-variational inequality problem (QVIP)

$$\begin{aligned}
 &\text{find } p^* \in \tilde{\mathcal{D}}(p^*) \text{ such that} \\
 &\quad \langle \nabla \mathcal{G}(p^*), p - p^* \rangle \geq 0 \quad \forall p \in \tilde{\mathcal{D}}(p^*) \\
 &\quad \Leftrightarrow \langle M^\top (Mp^* - f), p - p^* \rangle \geq 0 \quad \forall p \in \tilde{\mathcal{D}}(p^*),
 \end{aligned} \tag{6.25}$$

where  $\tilde{\mathcal{D}}(p) := \mathcal{D}(A^{-1}(f - Mp))$ . Please note that  $\tilde{\mathcal{D}}$  actually depends on  $Mp$ , a fact which we will make use of in the following sections.

Obviously,  $u^* = A^{-1}(f - Mp^*)$  is a fixed point of  $T$  if and only if  $p^* \in \tilde{\mathcal{D}}(p^*)$  solves the QVIP (6.25).

To illustrate our approach, we revisit Example 6.1 from above.

**Example 6.2.** Recall that in Example 6.1 we have considered the task of image denoising, where  $A = \text{Id}$ . We have chosen a regularization term  $\mathcal{R}(u; v)$  based on first-order total variation, i.e.  $D = \text{div}$ , where the regularization strength  $\alpha(x, v)$  was locally adapted based on the gradient of  $v$ .

Following the discretization procedure discussed above, we consider a regular grid with  $n$  nodes, on which we approximate the divergence operator by finite-differences. We denote the resulting discrete operator by  $L : \mathbb{R}^{nd} \rightarrow \mathbb{R}^n$ . Recall that

$$\mathcal{R}(u; v) := \sup_{p \in \mathcal{D}(v)} u^\top M p, \quad (6.26)$$

where  $\mathcal{D}(v)$  is the cross-product of the  $n$  local constraint sets  $\mathcal{D}_i^{\text{loc}}(v)$ ,  $i = 1, \dots, n$ , which in this example are chosen to be balls of radius  $\alpha_i(v)$ . Please note that we replaced the function  $\varphi \in C_c^1(\Omega; \mathbb{R}^d)$  by a discrete vector  $p \in \mathbb{R}^{nd}$ .

Similar to the continuous setting, by means of the adjoint we can shift the operator  $L$  in (6.26) to  $u$  and get

$$\mathcal{R}(u; v) = \sup_{p \in \mathcal{D}(v)} u^\top L p = \sum_{i=1}^n \alpha_i(v) \|(-L^\top u)_i\|_2, \quad (6.27)$$

where  $(x)_i$  denotes the  $i$ -th entry of  $x$ . We observe that  $-L^\top : \mathbb{R}^n \rightarrow \mathbb{R}^{nd}$  is a discretization of the gradient operator applied to all nodes  $i$  simultaneously. However, in order to avoid artifacts at the image borders, it is suitable to use homogeneous Neumann boundary conditions when discretizing the gradient. We therefore choose a discretization  $L$  of  $\text{div}$  such that  $-L^\top$  respects these boundary conditions, e.g. aiming at right-sided finite difference with homogeneous Neumann boundary conditions for the gradient we choose in left-sided finite-difference with homogeneous Dirichlet boundary conditions for  $\text{div}$ .

Recall that the adaptive regularization parameter  $\alpha(x, u)$  in the continuous setting was chosen as

$$\alpha(x; v) = \max(\alpha_{\max}(1 - \kappa \|\nabla v_\sigma(x)\|), \varepsilon). \quad (6.28)$$

When discretizing (6.28), we also make use of  $-L^\top$  as an approximation for the gradient. Therefore, (6.28) turns into

$$\begin{aligned} w &:= -L^\top(K_\sigma * v), \\ \alpha_i(v) &:= \max\{\alpha_0(1 - \kappa \|w_i\|_2), \varepsilon\}, \quad i = 1, \dots, n. \end{aligned} \quad (6.29)$$

where  $K_\sigma * v$  now denotes the discrete convolution with a Gaussian kernel  $K_\sigma$  and  $w = (w_1, \dots, w_n)^\top$  with  $w_i \in \mathbb{R}^d$ .

We now follow the proposed approach for a solution-driven adaptivity. In the case of denoising we have  $A = \text{Id}$  and  $M = L$ . We consider  $v = u_0$  for some fixed function  $u_0 \in BV(\Omega)$ . In order to switch to the dual formulation we have to express  $\mathcal{D}(u_0)$  in terms of  $p_0$ , where  $u_0 = f - L p_0$ . To this end, we set

$$\begin{aligned} w &:= -L^\top(K_\sigma * (f - L p_0)), \\ \tilde{\alpha}_i(p_0) &:= \max\{\alpha_0(1 - \kappa \|w_i\|_2), \varepsilon\}, \end{aligned} \quad (6.30)$$

and  $\tilde{\mathcal{D}}(p_0)$  to be the cross product of  $B_{\tilde{\alpha}_i(p_0)}(0)$ ,  $i = 1, \dots, n$ .

Finding a fixed point  $u^*$  of

$$u_0 \mapsto \arg \min_{u \in X} \mathcal{S}(u) + \mathcal{R}(u; u_0) \quad (6.31)$$

then is equivalent to solving the QVIP

$$\begin{aligned} & \text{find } p^* \in \tilde{\mathcal{D}}(p^*) \text{ such that} \\ & \Leftrightarrow \langle L^\top(Lp^* - f), p - p^* \rangle \geq 0 \quad \forall p \in \tilde{\mathcal{D}}(p^*), \end{aligned} \quad (6.32)$$

where we can retrieve  $u^*$  from  $p^*$  by  $u^* = f - Lp^*$ .  $\diamond$

## 6.3 Theory for the Continuous Case

Please recall that we assume  $\Omega \subset \mathbb{R}^d$  to be open, bounded and with Lipschitz-boundary. In this section, we focus on strict convex data terms  $\mathcal{S} : L^2(\Omega) \rightarrow \mathbb{R}$ . The appropriate function space for our variational approach is  $X = L^2(\Omega) \cap BV^l(\Omega)$ , together with the weak convergence from Definition 4.6, denoted by  $\overset{X}{\rightharpoonup}$ .

In the following, we show for a specific class of regularization terms that there exists a fixed point of the operator

$$\mathbb{T} : L^2(\Omega) \rightarrow X, \quad \mathbb{T}(v) := \arg \min_{u \in \mathcal{K}} \mathcal{F}(u; v), \quad (6.33)$$

where  $\mathcal{K} \subseteq X$  is some convex set. This restriction is twofold. Firstly, we focus on first-order regularization (i.e.  $l = 1$ ). The extension to higher order regularization terms is straightforward. Secondly, we rewrite the regularization term  $\mathcal{R}(u; v)$  as

$$\mathcal{R}(u; v) := \sup \left\{ \int u \operatorname{div}(Q(v)\varphi) \, d\mathcal{L} \mid \varphi \in C_c^1(\Omega; \mathbb{R}^2), \varphi(x) \in \mathcal{D}^{loc} \right\}, \quad (6.34)$$

for some function  $Q(v) = Q(x, v) : \Omega \times L^1(\Omega) \rightarrow \mathbb{R}^{d \times d}$  and with a  $d$ -dimensional constraint set  $\mathcal{D}^{loc}$  now being constant with respect to  $x$ . The standard examples for  $\mathcal{D}^{loc}$  are  $\mathcal{D}^{loc} = B_1(0)$  and  $\mathcal{D}^{loc} = \{x \in \mathbb{R}^d \mid \|x\|_\infty \leq 1\}$ . Obviously, with these restrictions on  $Q(v)$  and  $\mathcal{D}$ , we cannot represent every instance covered by (6.11) by the more specific form (6.34), in particular if  $\mathcal{D}^{loc}(x, v)$  has a complicated shape. However, (6.34) covers the first-order case examples mentioned in Chapter 4.

In the following, we will show that a fixed point of operator  $\mathbb{T}$  defined with regularizers of the form (6.34) exists. A generalization of this theory to arbitrary regularizers of the form (6.11) and proving uniqueness are subject to future work.

Before turning to the existence of a fixed point, let us start with the well-posedness operator  $\mathbb{T}$ .

### 6.3.1 Well-Posedness of Operator T

In order to make T well-defined, we require that a *unique* minimizer of  $\mathcal{F}(u; v)$  with respect to  $u$  exists. (If several minimizers of  $\mathcal{F}(u; v)$  existed, we could also define operator T as set-valued. However, it is unclear how adaptivity on a set of functions can be defined in a meaningful way.)

Since we changed the formulation of  $\mathcal{R}(u; v)$  in (6.34) compared to (6.11), the requirements for existence and uniqueness of a minimizer of  $\mathcal{F}(u; v)$  slightly change, compared to Assumption 4.10.

#### Assumption 6.3.

- (i)  $\mathcal{K} \subseteq X$  is non-empty, convex and closed with respect to the weak convergence in  $X$ .
- (ii)  $\mathcal{S} : L^2(\Omega) \rightarrow \mathbb{R}$  is strictly convex and lower semi-continuous with respect to the weak convergence in  $L^2(\Omega)$ .
- (iii) There exist constants  $c_1, c_2 > 0$  such that  $\|u\|_{L^2} \leq c_1 \mathcal{S}(u) + c_2$  for any  $u \in \mathcal{K}$ .
- (iv) There exist  $C_{min} > 0$  and  $C_{max} > 0$  independent from  $v \in L^1(\Omega)$  such that

$$C_{min} \text{TV}(u) \leq \mathcal{R}(u; v) \quad (6.35)$$

and

$$\|Q(x, v)\|_2 \leq C_{max}. \quad (6.36)$$

(In particular, this guarantees  $\mathcal{R}(u; v) \leq C \text{TV}(u)$  for some  $C > 0$ .)

**Remark 6.4.** We remark that compared to Assumption 4.10 two properties changed. Firstly, we added the strict convexity of  $\mathcal{S}$  in view of uniqueness of a minimizer. Secondly, we replaced Assumption 4.10 (iv) by Assumption 6.3 (iv), cf. also (4.42) in Proposition 4.14.

**Example 6.5.** The following two examples yield an anisotropic TV regularization:

1. Consider  $Q(x, v)$  for fixed  $x, v$  to be symmetric with eigenvalues  $\lambda_1, \lambda_2$  such that  $C_{min} \leq \lambda_i \leq C_{max}, i = 1, 2$  and  $\mathcal{D} = B_1(0) \subset \mathbb{R}^d$ , cf. Example 4.26.
2. Consider  $Q(x, v)$  to be a rotation matrix and  $\mathcal{D} = \{x \in \mathbb{R}^d \mid \|x\|_\infty \leq 1\}$ .

In both cases Assumption 6.3 (iv) is satisfied. ◇

**Proposition 6.6.** Let Assumption 6.3 be satisfied. Then, for given  $v \in L^2(\Omega)$ ,  $\mathcal{F}(\cdot; v)$  has a unique minimizer  $\bar{u} \in \mathcal{K} \subseteq BV(\Omega)$ . Moreover, there exists an  $R_{max} \geq 0$  independent from  $v$ , such that  $\|\bar{u}\|_{BV} \leq R_{max}$ . Thus  $\mathbb{T}(v) : L^2(\Omega) \rightarrow U$ , where

$$U := \{u \in \mathcal{K} \mid \|u\|_{BV} \leq R_{max}\}, \quad (6.37)$$

is well defined.

Analogously to Section 4.4, we require convexity, weakly \* lower semi-continuity of  $\mathcal{R}(u; v)$  and boundedness from below/above by the TV semi-norm:

**Lemma 6.7.** *Let Assumption 6.3 (iv) be satisfied. Then, the functional  $\mathcal{R}(u; v)$  defined in (6.34) with fixed  $v$  is convex and lower semi-continuous with respect to the weak\* convergence in  $BV(\Omega)$ . Moreover,*

$$\frac{C_{\min}}{\sqrt{l}}(\text{TV}(u)) \leq \mathcal{R}(u; v) \leq C_{\max} \text{TV}(u). \quad (6.38)$$

*Proof.* Convexity follows from the fact that  $\mathcal{R}(u; v)$  is the supremum of convex functions. The proof of the weak\* lower semi-continuity is analogous to the proof of Proposition 4.14. Finally, (6.38) follows directly from Assumption 6.3 (iv).  $\square$

**Remark 6.8.** *From the weak lower semi-continuity of  $\mathcal{S}$  in  $L^2(\Omega)$ , cf. Assumption 6.3 (ii), the weak\* lower semi-continuity of  $\mathcal{R}$  (Lemma 6.7) and of  $\iota_{\mathcal{K}}$  (cf. Remark 4.13) it follows the weak lower semi-continuity of  $\mathcal{F}$  with respect to the convergence in  $X$  (cf. Definition 4.6).*

*Proof of Proposition 6.6.* The proof of existence is analogous to the proof of Theorem 4.11: From strict convexity of  $\mathcal{F}(\cdot; v)$  uniqueness follows (cf. Proof of Theorem 4.16).  $\square$

### 6.3.2 Existence of a Fixed Point

Under the following assumption in addition to Assumption 6.3 we can prove existence of a fixed point of  $T$ :

**Assumption 6.9.**

- (i)  $\mathcal{D}$  takes the form  $\mathcal{D} = \{x \in \mathbb{R}^d \mid \|x\|_p \leq 1\}$  with  $p \in \{2, \infty\}$ .
- (ii)  $Q(x, v)$  is continuously differentiable with respect to  $x$  and Lipschitz continuous with respect to  $v$ , i.e.

$$\|(Q(x, v_1) - Q(x, v_2))w\|_q \leq C_l \|w\|_2 \|v_1 - v_2\|_{L^1} \quad (6.39)$$

for some  $C_l > 0$ , norm  $\|\cdot\|_q$  with  $\frac{1}{p} + \frac{1}{q} = 1$  on the left hand side and norm  $\|\cdot\|_2$  of  $w$  on the right hand side of (6.39).

**Remark 6.10.** *Assumption 6.9 (i) will allow us to express  $\mathcal{R}(u, v)$  for smooth functions  $u \in C^\infty(\Omega)$  as*

$$\mathcal{R}(u, v) = \int_{\Omega} \|Q(x, v)^\top \nabla u(x)\|_q \, dx. \quad (6.40)$$

*(We will prove this equality below.) Assumption 6.9 (ii) therefore will provide Lipschitz-continuity of  $v \mapsto \mathcal{R}(u, v)$ .*

**Theorem 6.11.** *Let Assumptions 6.3 and 6.9 be satisfied. Then, there exists a fixed point  $u^*$  of  $T(v)$  (cf. (6.33)), i.e.  $u^*$  minimizes  $\mathcal{F}(u; u^*)$  in (6.13).*

The proof of Theorem 6.11 requires the following definition, proposition and theorem.

**Definition 6.12.** *We consider the weak topology in  $L^2(\Omega)$  (cf. [4]). We call  $T : L^2(\Omega) \rightarrow L^2(\Omega)$  weakly upper semi-continuous if for any weakly closed subset  $B$  of  $\text{im}(T)$  the set  $T^{-1}(B)$  is weakly closed.*

**Proposition 6.13.** *Recall that  $U = \{u \in \mathcal{K} \mid \|u\|_{BV} \leq R_{max}\}$  (cf. (6.37)) with convex  $\mathcal{K} \subseteq L^2 \cap BV(\Omega)$ . The mapping  $v \mapsto T(v)$  from  $\mathcal{K}$  to  $U \subseteq \mathcal{K} \subseteq L^2(\Omega)$  is weakly upper semi-continuous with respect to the weak topology in  $L^2(\Omega)$ .*

The proof will be given below.

**Theorem 6.14** (Theorem 2.3 in [3]). *Let  $H$  be a Banach space and let  $\mathcal{K}$  be a closed convex subset of  $H$ . Then any weakly compact, weakly sequentially upper semi-continuous map  $T : U \rightarrow U$  has a fixed point.*

*Proof.* See proof of Theorem 2.3. in [3]. □

We will first prove Proposition 6.13, for which we require the following definition and lemmas, and then provide the proof of Theorem 6.11.

**Definition 6.15.** *Let  $\eta$  be the mollifier*

$$\eta(x) := \begin{cases} c \exp\left(-\frac{1}{1-|x|^2}\right) & \text{if } |x| < 1, \\ 0 & \text{else,} \end{cases} \quad (6.41)$$

*with the constant  $c > 0$  adjusted to guarantee  $\int_{\mathbb{R}^2} \eta(x) d\mathcal{L}$  (cf. [55, Sect. 4.2.1])). Moreover, we define for any  $s > 0$*

$$\eta_s(x) := \frac{1}{s^d} \eta\left(\frac{x}{s}\right). \quad (6.42)$$

**Lemma 6.16.** *Let  $Q(x) : \Omega \rightarrow \mathbb{R}^{d \times d}$  be Lipschitz-continuous with constant  $L$ . Then, for any  $\varphi \in L^\infty(\Omega; \mathbb{R}^2)$ ,*

$$\|\eta_\varepsilon * (Q\varphi) - Q(\eta_\varepsilon * \varphi)\|_{L^\infty} \leq CL\|\varphi\|_{L^\infty} \varepsilon \quad (6.43)$$

*holds for a constant  $C > 0$  independent from  $Q$  and  $\varphi$ .*



*Proof.*

$$\begin{aligned}
 & \|(\eta_\varepsilon * (Q\varphi))(x) - Q(\eta_\varepsilon * \varphi)(x)\|_2 \\
 &= \left\| \int_{\Omega} \eta_\varepsilon(x-y)Q(y)\varphi(y) - \eta_\varepsilon(x-y)Q(x)\varphi(y) dy \right\|_2 \\
 &\leq \int_{\Omega} \eta_\varepsilon(x-y)L\|x-y\|_2\|\varphi(y)\|_2 dy \\
 &\leq L\|\varphi\|_{L^\infty} \int_{\Omega} \eta_\varepsilon(x-y)\|x-y\|_2 dy \\
 &\leq L\|\varphi\|_{L^\infty} \int_{\mathbb{R}^d} \eta_\varepsilon(x-y)\|x-y\|_2 dy \\
 &= L\|\varphi\|_{L^\infty} \int_{\mathbb{R}^d} \eta_\varepsilon(y)\|y\|_2 dy \\
 &= L\|\varphi\|_{L^\infty} \int_{\mathbb{R}^d} \frac{1}{\varepsilon^d} \eta\left(\frac{y}{\varepsilon}\right) \|y\|_2 dy.
 \end{aligned} \tag{6.44}$$

We perform a change of variables  $\tilde{y} = \frac{y}{\varepsilon}$ ,  $d\tilde{y} = \frac{dy}{\varepsilon^d}$ :

$$L \int_{\mathbb{R}^d} \frac{1}{\varepsilon^d} \eta\left(\frac{y}{\varepsilon}\right) \|y\|_2 dy = L \int_{\mathbb{R}^d} \eta(\tilde{y})\|\tilde{y}\|_2 \varepsilon d\tilde{y} \leq L \int_0^\infty r^{d-1} V_d \tilde{\eta}(r) dr \varepsilon \tag{6.45}$$

with  $\tilde{\eta}(r) := \eta(r e_1)$  using the radial symmetry of  $\eta$  and  $V_d$  being the volume of the  $d$ -dimensional unit sphere  $S^d$ . Note that  $\int_0^\infty r^{d-1} V_d \tilde{\eta}(r) dr$  is bounded by some  $C_\eta > 0$ . Inserting (6.45) into (6.44), we obtain

$$\|\eta_\varepsilon * (Q\varphi) - Q(\eta_\varepsilon * \varphi)\|_{L^\infty} \leq L\|\varphi\|_{L^\infty} C_\eta \varepsilon. \tag{6.46}$$

□

**Lemma 6.17.** *For any  $u \in BV(\Omega)$  there exists  $(u^k)_k \in C^\infty(\Omega)$  such that  $u^k \xrightarrow{L^1} u$ ,  $\text{TV}(u^k) \rightarrow \text{TV}(u)$  and  $\mathcal{R}(u^k, v_i) \rightarrow \mathcal{R}(u, v_i)$  simultaneously for a finite set of  $\{v_i\}_{i=1}^n \subset L^1(\Omega)$ .*

*Proof.* For fixed  $v$  the proof follows the proof of [55, Thm. 2, Sect. 5.2], with some modifications. For the reader's convenience, instead of just listing the required modifications, we provide the full modified proof, also adapting the notation. We define  $\mathcal{P} := \{C_c^1(\Omega, \mathbb{R}^d), \|\varphi(x)\|_p \leq 1\}$ . From the equivalence of norms in  $\mathbb{R}^d$  we find that there exists a constant  $C_\varphi > 0$  such that for all  $\varphi \in \mathcal{P}$

$$\|\varphi(x)\|_2 \leq C_\varphi. \tag{6.47}$$

Now, let  $\varepsilon > 0$  be fixed. For a  $m \in \mathbb{N}, m > 0$  and  $k \in \mathbb{N}$  we define open sets

$$\Omega_k := \left\{ x \in \Omega \mid \text{dist}(x, \partial\Omega) > \frac{1}{m+k} \right\}, \tag{6.48}$$

where  $\text{dist}(x, A)$  is the distance of point  $x$  to set  $A$ , and choose  $m$  large enough to guarantee

$$\text{TV}(u)(\Omega \setminus \Omega_1) < \varepsilon, \tag{6.49}$$

where  $\text{TV}(u)(B)$  is the variation measure of  $u$  evaluated on the set  $B$ . Set  $\Omega_0 := \emptyset$  and define  $V_k := \Omega_{k+1} \setminus \overline{\Omega}_{k-1}$ . Please note that each  $x \in \Omega$  is contained in at most three sets  $V_k$ . Moreover, let  $\{\zeta_k\}_{k=1}^\infty$  be a sequence of smooth functions with

$$\zeta_k \in C_c^\infty(V_k), \quad 0 \leq \zeta_k \leq 1 \text{ and } \sum_{k=1}^\infty \zeta_k = 1 \text{ on } \Omega. \quad (6.50)$$

Let  $\eta_s$ ,  $s > 0$  be the mollifier from Definition 6.15. For each  $k$ , select  $\varepsilon_k > 0$  small enough such that for

$$\begin{aligned} \text{supp}(\eta_{\varepsilon_k} * (u\zeta_k)) &\subset V_k, \\ \int_{\Omega} |\eta_{\varepsilon_k} * (u\zeta_k) - u\zeta_k| \, d\mathcal{L} &< \frac{\varepsilon}{2^k}, \\ \int_{\Omega} \|\eta_{\varepsilon_k} * (u\nabla\zeta_k) - u\nabla\zeta_k\|_2 \, d\mathcal{L} &< \frac{\varepsilon}{2^k}. \end{aligned} \quad (6.51)$$

Define

$$u_\varepsilon := \sum_{k=1}^\infty \eta_{\varepsilon_k} * (u\zeta_k). \quad (6.52)$$

For each  $x \in \Omega$  there are only finitely many terms in (6.52), which are non-zero on a neighborhood of  $x$ . Thus,  $u_\varepsilon \in C^\infty(\Omega)$ . Since also  $u = \sum_{k=1}^\infty u\zeta_k$ , we find from (6.51) that

$$\|u_\varepsilon - u\|_{L^1(\Omega)} \leq \sum_{k=1}^\infty \int_{\Omega} |\eta_{\varepsilon_k} * (u\zeta_k) - u\zeta_k| \, d\mathcal{L} < \varepsilon. \quad (6.53)$$

Thus,  $u_\varepsilon \rightarrow u$  in  $L^1(\Omega)$ . Moreover, for  $\varphi \in \mathcal{P}$ , we have

$$\int_{\Omega} u_\varepsilon \operatorname{div} \varphi \, d\mathcal{L} = \int_{\Omega} u \left( \sum_k \zeta_k \operatorname{div}(\eta_{\varepsilon_k} * \varphi) \right) \, d\mathcal{L} \leq 3C_\varphi \text{TV}(u), \quad (6.54)$$

where we used that for every  $x \in \Omega$  there exist at most three  $\zeta_k(x) > 0$ . Taking the supremum over all such  $\varphi$  we find

$$\text{TV}(u_\varepsilon) \leq 3C_\varphi \text{TV}(u) < \infty. \quad (6.55)$$

Both convergence in  $L^1(\Omega)$  and boundedness of  $\text{TV}(u_\varepsilon)$  provide that  $u_\varepsilon \xrightarrow{*} u$ . Then, due to the weak\* lower semi-continuity of  $\mathcal{R}(u; v)$  (cf. Proposition 4.14) we have

$$\mathcal{R}(u; v) \leq \liminf_{\varepsilon \rightarrow 0} \mathcal{R}(u_\varepsilon; v). \quad (6.56)$$

For any  $\varphi \in \mathcal{P}$

$$\begin{aligned}
 \int_{\Omega} u_{\varepsilon} \operatorname{div}(Q(v)\varphi) \, d\mathcal{L} &= \sum_{k=1}^{\infty} \int_{\Omega} (\eta_{\varepsilon_k} * (u\zeta_k)) \operatorname{div}(Q(v)\varphi) \, d\mathcal{L} \\
 &= \sum_{k=1}^{\infty} \int_{\Omega} u\zeta_k \operatorname{div}(\eta_{\varepsilon_k} * (Q(v)\varphi)) \, d\mathcal{L} \\
 &= \sum_{k=1}^{\infty} \int_{\Omega} u \operatorname{div}(\zeta_k(\eta_{\varepsilon_k} * (Q(v)\varphi))) \, d\mathcal{L} - \sum_{k=1}^{\infty} \int_{\Omega} u \nabla \zeta_k(\eta_{\varepsilon_k} * (Q(v)\varphi)) \, d\mathcal{L} \\
 &= \underbrace{\sum_{k=1}^{\infty} \int_{\Omega} u \operatorname{div}(\zeta_k(\eta_{\varepsilon_k} * (Q(v)\varphi))) \, d\mathcal{L}}_{:=I_1^{\varepsilon}} - \underbrace{\sum_{k=1}^{\infty} \int_{\Omega} \varphi^{\top} Q(v)(\eta_{\varepsilon_k} * (u \nabla \zeta_k) - u \nabla \zeta_k) \, d\mathcal{L}}_{:=I_2^{\varepsilon}},
 \end{aligned} \tag{6.57}$$

where we used the fact  $\sum_{k=1}^{\infty} \nabla \zeta_k = 0$  on  $\Omega$ . Since  $\|Q(x, v)\varphi(x)\|_2 \leq C_{\max} C_{\varphi}$  (cf. Assumption 6.3 (iv) and (6.47)), we find that  $\|\zeta_k(x)(\eta_{\varepsilon_k} * (Q(v)\varphi))(x)\|_2 \leq C_{\max} C_{\varphi}$ . Thus, we can bound  $I_1^{\varepsilon}$  by

$$\begin{aligned}
 |I_1^{\varepsilon}| &\leq \left| \int_{\Omega} u \operatorname{div}(\zeta_1(\eta_{\varepsilon_1} * (Q(v)\varphi))) \, d\mathcal{L} \right| + \sum_{k=2}^{\infty} \left| \int_{\Omega} u \operatorname{div}(\zeta_k(\eta_{\varepsilon_k} * (Q(v)\varphi))) \, d\mathcal{L} \right| \\
 &\leq \left| \int_{\Omega} u \operatorname{div}(\zeta_1(\eta_{\varepsilon_1} * (Q(v)\varphi))) \, d\mathcal{L} \right| + \sum_{k=2}^{\infty} C_{\max} C_{\varphi} \operatorname{TV}(u)(V_k),
 \end{aligned} \tag{6.58}$$

Since each point in  $\Omega$  belongs to at most three sets  $V_k$ , together with  $V_k \subseteq \Omega \setminus \Omega_1$  and (6.49), (6.58) can be bounded by

$$\begin{aligned}
 |I_1^{\varepsilon}| &\leq \left| \int_{\Omega} u \operatorname{div}(\zeta_1(\eta_{\varepsilon_1} * (Q(v)\varphi))) \, d\mathcal{L} \right| + 3C_{\max} C_{\varphi} \operatorname{TV}(u)(V_k) \\
 &\leq \left| \int_{\Omega} u \operatorname{div}(\zeta_1(\eta_{\varepsilon_1} * (Q(v)\varphi))) \, d\mathcal{L} \right| + 3C_{\max} C_{\varphi} \operatorname{TV}(u)(\Omega \setminus \Omega_1) \\
 &\leq \left| \int_{\Omega} u \operatorname{div}(\zeta_1(\eta_{\varepsilon_1} * (Q(v)\varphi))) \, d\mathcal{L} \right| + 3C_{\max} C_{\varphi} \varepsilon.
 \end{aligned} \tag{6.59}$$

Next, we apply Lemma 6.16:

$$\begin{aligned}
 \left| \int_{\Omega} u \operatorname{div}(\zeta_1(\eta_{\varepsilon_1} * (Q(v)\varphi))) \, d\mathcal{L} \right| \\
 \leq \left| \int_{\Omega} u \operatorname{div}(\zeta_1(Q(v)(\eta_{\varepsilon_1} * \varphi))) \, d\mathcal{L} \right| + C_{\eta} C_{\varphi} L(v) \operatorname{TV}(u) \varepsilon,
 \end{aligned} \tag{6.60}$$

where  $L(v)$  is the Lipschitz-constant of  $x \mapsto Q(x, v)$ . Since  $\varphi \in \mathcal{P}$  and thus  $\eta_{\varepsilon_1} * \varphi \in \mathcal{P}$ , we have

$$\left| \int_{\Omega} u \operatorname{div}(\zeta_1(Q(v)(\eta_{\varepsilon_1} * \varphi))) \, d\mathcal{L} \right| \leq \mathcal{R}(u; v). \tag{6.61}$$

Combining (6.60) and (6.61), we obtain

$$\left| \int_{\Omega} u \operatorname{div}(\zeta_1(\eta_{\varepsilon_1} * (Q(v)\varphi))) d\mathcal{L} \right| \leq \mathcal{R}(u; v) + C_{\eta} C_{\varphi} L(v) \operatorname{TV}(u) \varepsilon. \quad (6.62)$$

Inserting (6.62) into (6.59), we end up with

$$|I_1^{\varepsilon}| \leq \mathcal{R}(u; v) + (C_{\eta} C_{\varphi} L(v) \operatorname{TV}(u) + 3C_{\max} C_{\varphi}) \varepsilon. \quad (6.63)$$

Now, we bound the term  $I_2^{\varepsilon}$ : Since  $\|Q(x, v)\varphi(x)\|_2 \leq C_{\max} C_{\varphi}$  we obtain from (6.51) that

$$|I_2^{\varepsilon}| = \left| \sum_{k=1}^{\infty} \int_{\Omega} \varphi^{\top} Q(v) (\eta_{\varepsilon_k} * (u \nabla \zeta_k) - u \nabla \zeta_k) d\mathcal{L} \right| \leq C_{\max} C_{\varphi} \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = C_{\max} C_{\varphi} \varepsilon. \quad (6.64)$$

Using bounds (6.63) and (6.64) in (6.57), we find

$$\int_{\Omega} u_{\varepsilon} \operatorname{div}(Q(v)\varphi) d\mathcal{L} \leq \mathcal{R}(u; v) + (4C_{\max} C_{\varphi} + C_{\eta} C_{\varphi} L(v) \operatorname{TV}(u)) \varepsilon. \quad (6.65)$$

Taking in (6.65) the supremum over all  $\varphi \in \mathcal{P}$ , we obtain

$$\mathcal{R}(u_{\varepsilon}; v) \leq \mathcal{R}(u; v) + (4C_{\max} C_{\varphi} + C_{\eta} C_{\varphi} L(v) \operatorname{TV}(u)) \varepsilon. \quad (6.66)$$

Both (6.56) and (6.66) together provide

$$|\mathcal{R}(u_{\varepsilon}; v) - \mathcal{R}(u; v)| \leq (4C_{\max} C_{\varphi} + C_{\eta} C_{\varphi} L(v) \operatorname{TV}(u)) \varepsilon \quad (6.67)$$

for  $\varepsilon$  small enough. Please note that the choice of  $\Omega_k, V_k, \varepsilon_k, \zeta_k$  and  $\eta_{\varepsilon_k}$  was independent from  $v$ . Thus, for the same sequence  $(u_{\varepsilon})_{\varepsilon}$  we get a uniform convergence for a finite set of  $\{v_i\}_{i=1}^n$ . Moreover, as we can express  $\operatorname{TV}(u)$  as  $\mathcal{R}(u; v)$  with  $Q(v) = \operatorname{Id}$  and  $p = 2$  (yielding  $C_{\varphi} = 1$  and  $L(v) = 1$ ), (6.67) provides also  $|\operatorname{TV}(u_{\varepsilon}) - \operatorname{TV}(u)| < (4C_{\max} + C_{\eta} \operatorname{TV}(u)) \varepsilon$ .  $\square$

**Lemma 6.18.** *Let  $\mathcal{R}(u; v)$  be as in (6.34) with  $\mathcal{D} = \{x \in \mathbb{R}^2 \mid \|x\|_p \leq 1\}$  for  $p \in \{2, \infty\}$ . Then, for every  $u \in C^{\infty}(\Omega)$  we have*

$$\mathcal{R}(u; v) = \int_{\Omega} \|Q(x, v)^{\top} \nabla u(x)\|_q dx \quad (6.68)$$

with  $\frac{1}{p} + \frac{1}{q} = 1$ .

*Proof.* We show the claim for  $p = 2$ . The case  $p = \infty$  is analogous. For the simplicity of notation, we drop the argument  $v$  of  $Q$  as well as  $x$  for any function, if it is clear from the context.

Recall that  $\mathcal{P} = \{C_c^1(\Omega, \mathbb{R}^d), \|\varphi(x)\|_2 \leq 1\}$  ( $p = 2$ ). For  $u \in C^{\infty}(\Omega)$  we have

$$\begin{aligned} \sup_{\varphi \in \mathcal{P}} \int_{\Omega} u \operatorname{div}(Q\varphi) d\mathcal{L} &= \sup_{\varphi \in \mathcal{P}} \int_{\Omega} \nabla u^{\top} (Q\varphi) d\mathcal{L} \\ &\leq \sup_{\varphi \in \mathcal{P}} \int_{\Omega} \|Q^{\top} \nabla u\|_2 \|\varphi\|_2 d\mathcal{L} \leq \int_{\Omega} \|Q^{\top} \nabla u\|_2 d\mathcal{L}. \end{aligned} \quad (6.69)$$

We show equality in (6.69). Define  $\varphi \in \mathcal{P}$  as

$$\varphi(x) := \begin{cases} \frac{Q(x)^\top \nabla u(x)}{\|Q(x)^\top \nabla u(x)\|_2} & \text{if } \nabla u(x) \neq 0, \\ 0 & \text{else,} \end{cases} \quad (6.70)$$

and

$$\zeta_k(x) = \begin{cases} 1 & \text{if } \text{dist}(x, \partial\Omega) \geq \frac{1}{k}, \\ 0 & \text{else.} \end{cases} \quad (6.71)$$

For fixed  $\varepsilon > 0$  choose  $k$  large enough such that

$$\left| \int_{\Omega} \zeta_k d\mathcal{L} - \int_{\Omega} 1 d\mathcal{L} \right| \leq \varepsilon. \quad (6.72)$$

Moreover, let  $\varphi_{j_k} := \eta_{\frac{1}{j_k}} * (\zeta_k \varphi)$  with  $j_k$  large enough such that  $\|\varphi_{j_k} - \zeta_k \varphi\|_{\infty} \leq \varepsilon$ . Then,  $\varphi_{j_k} \in \mathcal{P}$  and

$$\begin{aligned} & \left| \int_{\Omega} \|Q^\top \nabla u\|_2 d\mathcal{L} - \int_{\Omega} u \operatorname{div}(Q\varphi_{j_k}) d\mathcal{L} \right| \\ & \stackrel{(6.70)}{=} \left| \int_{\Omega} (Q^\top \nabla u)^\top \varphi d\mathcal{L} - \int_{\Omega} (Q^\top \nabla u) \varphi_{j_k} d\mathcal{L} \right| \\ & \leq C_{max} \|\nabla u\|_{\infty} \|\varphi - \varphi_{j_k}\|_{L^1}. \end{aligned} \quad (6.73)$$

We bound  $\|\varphi - \varphi_{j_k}\|_{L^1}$  by

$$\|\varphi - \varphi_{j_k}\|_{L^1} \leq \|\varphi - \zeta_k \varphi\|_{L^1} + \|\zeta_k \varphi - \varphi_{j_k}\|_{L^1} \leq 2\varepsilon. \quad (6.74)$$

Combining (6.73) and (6.74) gives

$$\left| \int_{\Omega} \|Q^\top \nabla u\|_q d\mathcal{L} - \int_{\Omega} u \operatorname{div}(Q\varphi_{j_k}) d\mathcal{L} \right| \leq 2C_{max} \|\nabla u\|_{\infty} \varepsilon. \quad (6.75)$$

Thus, we have found a sequence  $\varphi_{j_k} \in \mathcal{P}$  such that

$$\int_{\Omega} u \operatorname{div}(Q\varphi_{j_k}) d\mathcal{L} \rightarrow \int_{\Omega} \|Q^\top \nabla u\|_2 d\mathcal{L} \quad (6.76)$$

and equality in (6.69) holds.  $\square$

fformall

**Lemma 6.19.** *Let  $u \in U := \{u \in X \mid \|u\|_{BV} \leq R_{max}\}$  and let  $v^k \rightarrow v$  in  $L^1(\Omega)$ . Then, we have  $\mathcal{R}(u; v^k) \rightarrow \mathcal{R}(u; v)$  and  $\mathcal{F}(u; v^k) \rightarrow \mathcal{F}(u; v)$  uniformly for every  $u \in U$ .*

*Proof.* We show that for arbitrary  $u \in U$  and  $v_1, v_2 \in L^1(\Omega)$

$$|\mathcal{R}(u; v_1) - \mathcal{R}(u; v_2)| \leq C_l R_{max} \|v_1 - v_2\|_{L^1}, \quad (6.77)$$

from which the claim for  $\mathcal{R}(u; v)$  follows. Let us first consider a fixed  $u \in C^\infty(\Omega)$ . Then, by Assumption 6.9 and Lemma 6.18 we have

$$\begin{aligned} |\mathcal{R}(u; v_1) - \mathcal{R}(u; v_2)| &= \int_{\Omega} \| (Q(x, v_1) \nabla u \|_q - \| Q(x, v_2) \nabla u \|_q) d\mathcal{L} \\ &\leq C_l \|v_1 - v_2\|_{L^1} \text{TV}(u). \end{aligned} \quad (6.78)$$

Now, let  $u \in U$  be arbitrary. Applying Lemma 6.17, we can find  $\tilde{u} \in C^\infty(\Omega)$  such that for any  $\varepsilon > 0$

$$|\mathcal{R}(u; v_i) - \mathcal{R}(\tilde{u}; v_i)| \leq \varepsilon \text{ for } i = 1, 2, \quad |\text{TV}(u) - \text{TV}(\tilde{u})| \leq \varepsilon. \quad (6.79)$$

Then,

$$\begin{aligned} |\mathcal{R}(u; v_1) - \mathcal{R}(u; v_2)| &\stackrel{(6.79)}{\leq} |\mathcal{R}(\tilde{u}; v_1) - \mathcal{R}(\tilde{u}; v_2)| + 2\varepsilon \stackrel{(6.78)}{\leq} C_l \|v_1 - v_2\|_{L^1} \text{TV}(\tilde{u}) + 2\varepsilon \\ &\stackrel{(6.79)}{\leq} C_l \|v_1 - v_2\|_{L^1} (\text{TV}(u) + \varepsilon) + 2\varepsilon \stackrel{u \in U}{\leq} C_l \|v_1 - v_2\|_{L^1} (R_{max} + \varepsilon) + 2\varepsilon. \end{aligned}$$

Since we can find  $\tilde{u}$  such that  $\varepsilon$  becomes arbitrary small, (6.77) follows for fixed  $u$ . Since the right hand side of (6.77) is not depending on  $u$ , we achieve an uniform convergence of  $\mathcal{R}(\cdot; v^k) \rightarrow \mathcal{R}(\cdot; v)$  on  $U$  for  $v^k \rightarrow v$  in  $L^1(\Omega)$ . Since  $\mathcal{R}(u; v)$  and  $\mathcal{F}(u; v)$  differ only by the term  $\mathcal{S}(u)$  not depending on  $v$ , the uniform convergence  $\mathcal{F}(\cdot; v^k) \rightarrow \mathcal{F}(\cdot; v)$  also follows.  $\square$

**Lemma 6.20.** For  $v^k \rightarrow v^0$  in  $L^1(\Omega)$  and  $u^k := \text{T}(v^k) := \arg \min_u \mathcal{F}(u; v^k)$ :

$$\liminf_{k \rightarrow \infty} \mathcal{F}(u^k; v^0) \leq \limsup_{k \rightarrow \infty} \mathcal{F}(u^k; v^k). \quad (6.80)$$

*Proof.* For the sequence  $u^k := \text{T}(v^k)$ , let  $\liminf_{k \rightarrow \infty} \mathcal{F}(u^k; v^0) := c$ . For any  $\varepsilon > 0$  Lemma 6.19 guarantees the existence of a  $K > 0$  such that

$$|\mathcal{F}(u; v^k) - \mathcal{F}(u; v^0)| \leq \frac{\varepsilon}{2} \quad \forall u \in U, \forall k \geq K. \quad (6.81)$$

Moreover, we can find a  $k' \geq K$  such that

$$|\mathcal{F}(u^{k'}; v^0) - c| \leq \frac{\varepsilon}{2}. \quad (6.82)$$

Using (6.81) and (6.82), it follows that

$$|\mathcal{F}(u^{k'}; v^{k'}) - c| \leq |\mathcal{F}(u^{k'}; v^{k'}) - \mathcal{F}(u^{k'}; v^0)| + |\mathcal{F}(u^{k'}; v^0) - c| \leq \varepsilon. \quad (6.83)$$

In other words, there exists a sequence  $k' \rightarrow \infty$ , such that  $\mathcal{F}(u^{k'}; v^{k'}) \rightarrow c$  and thus

$$\liminf_{k \rightarrow \infty} \mathcal{F}(u^k; v^0) = c \leq \limsup_{k \rightarrow \infty} \mathcal{F}(u^k; v^k). \quad (6.84)$$

$\square$

*Proof of Proposition 6.13.* Let  $B$  be a weakly closed subset of  $U$  (recall that  $\text{im}(T) \subseteq U$ ). Let  $(v^k)_k$  be a sequence in  $T^{-1}(B) \subseteq V$  weakly converging in  $L^2(\Omega)$  to some  $v^0 \in V$ , i.e. there exists a sequence  $(u^k)_k \in U$  such that  $u^k = T(v^k)$ , and  $v^k \xrightarrow{L^2} v^0$ . Let  $u^0 := T(v^0)$ . Since  $U$  is weakly\* pre-compact in  $BV(\Omega)$  and weakly pre-compact in  $L^2(\Omega)$ , there exists a subsequence also denoted by  $(u^k)_k$ , such that  $u^k \xrightarrow{*} u$  for some  $u \in BV(\Omega)$  and  $u^k \xrightarrow{L^2} u$ . Since  $B$  is weakly closed in  $L^2(\Omega)$ , we find  $u \in B$ . Next, we show that  $u = u^0$ , i.e.  $u$  is the unique minimizer of  $\mathcal{F}(\cdot; v^0)$ . Since the embedding from  $L^2(\Omega)$  to  $L^1(\Omega)$  is compact (cf. [4, Theorem 8.9]), we can find a subsequence  $(v^{k'})_{k'}$ , of  $(v^k)_k$ , which converges to  $v^0$  strongly in  $L^1(\Omega)$ . Using the weakly lower semi-continuity of  $\mathcal{F}(\cdot; v^0)$  (cf. Remark 6.8) together with Lemma 6.20, we find

$$\begin{aligned} 0 \leq \mathcal{F}(u; v^0) - \mathcal{F}(u^0; v^0) &\leq \liminf_{k' \rightarrow \infty} \mathcal{F}(u^{k'}; v^0) - \mathcal{F}(u^0; v^0) \\ &\leq \limsup_{k' \rightarrow \infty} \mathcal{F}(u^{k'}; v^{k'}) - \mathcal{F}(u^0; v^0). \end{aligned} \quad (6.85)$$

Recall that  $u^{k'}$  is the minimizer of  $\mathcal{F}(\cdot; v^{k'})$  for all  $k' \geq 0$ , which induces

$$\mathcal{F}(u^{k'}; v^{k'}) \leq \mathcal{F}(u^0; v^{k'}). \quad (6.86)$$

Inserting (6.86) into (6.85), we obtain

$$0 \leq \mathcal{F}(u; v^0) - \mathcal{F}(u^0; v^0) \leq \limsup_{k' \rightarrow \infty} \mathcal{F}(u^0; v^{k'}) - \mathcal{F}(u^0; v^0). \quad (6.87)$$

Lemma 6.19 guarantees that the right hand side of (6.87) tends to zero, thus  $\mathcal{F}(u; v^0) = \mathcal{F}(u^0; v^0)$ . On the other hand,  $u^0$  by definition is the unique minimizer of  $\mathcal{F}(\cdot; v^0)$ , from which  $u = u^0$  follows. Since  $u \in B$  and  $u = u^0 = T(v^0)$ , we have shown that  $v^0 \in T^{-1}(B)$  and thus  $T^{-1}(B)$  is weakly closed.  $\square$

*Proof of Theorem 6.11.* The claim follows from Theorem 6.14, since  $\mathcal{K}$  by definition is convex and closed in  $L^2(\Omega)$  and  $v \rightarrow T(v)$  is weakly compact in  $L^2(\Omega)$  (it maps to the pre-compact set  $U$ , thus  $T(B)$  is pre-compact for any  $B \subseteq \mathcal{K}$ ) and weakly upper semi-continuous in  $L^2(\Omega)$  (Proposition 6.13).  $\square$

## 6.4 Theory for QVIs

### 6.4.1 Problem Statement

In Section 6.2 we have introduced the QVI problem

$$\begin{aligned} &\text{Find } p^* \in \tilde{\mathcal{D}}(p^*) \text{ such that} \\ &\quad \langle \nabla \mathcal{G}(p^*), p - p^* \rangle \geq 0 \quad \forall p \in \tilde{\mathcal{D}}(p^*), \\ &\Leftrightarrow \langle M^\top (Mp^* - f), p - p^* \rangle \geq 0 \quad \forall p \in \tilde{\mathcal{D}}(p^*), \end{aligned} \quad (6.88)$$

QVIs and their solutions are well studied in the literature (see e.g. [12, 20, 37, 127]) and various results on theory and existence are available. For the problem

(6.88), it turns out that the existence of a solution  $p^*$  follows from standard theory. Uniqueness results, however are only available for strongly monotone operators  $\nabla\mathcal{G}$ . In our case, where  $\nabla\mathcal{G}(p) = M^\top(Mp - f)$  and  $M = A^{-1}L$ , the non-trivial null space of  $L$  induces a non-trivial null space of  $M$ , so that  $M^\top M$  is not strictly monotone. As a remedy, a change of variables can be introduced. Let  $v := Mp$  and  $\mathcal{C}(v) := M\mathcal{D}(A^{-1}(f - v))$ , where the operation of  $M$  is applied element-wise on the set  $\mathcal{D}$  (recall that the sets  $\mathcal{D}$  and  $\tilde{\mathcal{D}}$  are linked via  $\tilde{\mathcal{D}}(p) = \mathcal{D}(A^{-1}(f - Mp))$ ). Then, we can rewrite (6.88) as the QVI problem

$$\begin{aligned} \text{Find } v^* \in \mathcal{C}(v^*) \text{ such that} \\ \langle v^* - f, v - v^* \rangle \geq 0 \quad \forall v \in \mathcal{C}(v^*). \end{aligned} \quad (6.89)$$

(If  $v \mapsto \mathcal{C}(v)$  was constant, this would be just the VI corresponding to the projection problem  $\min_{v \in \mathcal{C}} \|f - v\|_2^2$ .) Obviously, the mapping  $v \mapsto v - f$  is strictly monotone. Uniqueness of  $v$  follows e.g. from Theorem 3.1. in [122], provided that the mapping

$$v \mapsto \Pi_{\mathcal{C}(v)} w \quad (6.90)$$

is a contraction for every  $w$ .

There is one issue with this reformulation. Calculating the projection  $\Pi_{\mathcal{C}(v)} w$  onto the set  $\mathcal{C}(v)$  is equivalent to solving the original optimization problem with  $A = Id$  and thus in general can not be done analytically. Moreover, proving the contraction property of  $v \mapsto \Pi_{\mathcal{C}(v)} w$  is difficult. From a theoretical as well as numerical point of view it is easier to work with the pair of variables  $(p, v)$  with  $v = Mp$ , the set  $\hat{\mathcal{D}}(v) := \hat{\mathcal{D}}(A^{-1}(f - v))$  and the mapping  $v \mapsto \Pi_{\hat{\mathcal{D}}(v)} q$ . For this setup, we will show

1. Existence of a solution  $p^*$  of (6.88) (cf. Section 6.4.2),
2. Uniqueness of  $v^* := Mp^*$  (cf. Section 6.4.3).

We remark that uniqueness of  $p^*$  cannot be shown in general, since the part of  $p^*$  lying in the null space of  $M$  might be ambiguous, as can be seen in the following example:

**Example 6.21.** *We consider the discrete one-dimensional case, where  $u, p \in \mathbb{R}^n$ . As already discussed in Example 6.2, we like to discretize the gradient operator by right-sided finite differences with Neumann boundary conditions. In order to assert that the discrete divergence operator  $L$  is the negative adjoint to the discrete gradient, we have to set*

$$(Lp)_i = \begin{cases} p_1 & \text{if } i = 1, \\ p_i - p_{i-1} & \text{if } i \in \{2, \dots, n-1\}, \\ -p_{n-1} & \text{if } i = n. \end{cases} \quad (6.91)$$

(i.e. we use left-sided finite differences and a Dirichlet boundary condition). It can be easily checked that with this definition  $-L^\top$  is an approximation of the gradient with right-sided differences and homogeneous Neumann boundary conditions.



We observe that the null space  $\mathcal{N}(M) = \mathcal{N}(A^{-1}L)$  is spanned by the vector  $e_n := (0, \dots, 0, 1)^\top$ . Now, let  $v \in M\mathcal{D}(u_0)$  be fixed and  $p \in \mathcal{D}(u_0)$  such that  $v = Mp$ . Recall that  $\mathcal{D}(u_0) = \mathcal{D}_1^{loc} \times \dots \times \mathcal{D}_n^{loc}(u_0)$ . Assuming that the local constraint sets  $\mathcal{D}_n^{loc}$  is not a point set (cf. Assumption 4.10 (iv)), the entry  $p_n \in \mathcal{D}_n^{loc}$  is ambiguous, i.e.  $v = M(p + te_n)$  holds for fixed  $p_n$  and all  $t$  such that  $p_n + te_n \in \mathcal{D}_n^{loc}(u_0)$ .

In higher dimensions  $\mathcal{N}(M) \cap \mathcal{D}(p)$  is also not a point set, since it contains  $\mathcal{N}(M) \cap B_{C_{min}}(0)$ .  $\diamond$

Finally, we point out that the constraint sets  $\mathcal{D}(u)$ ,  $\tilde{\mathcal{D}}(p)$  and  $\hat{\mathcal{D}}(v)$  are identical as long as  $u, v$  and  $p$  are related to each other, i.e.  $v = Mp$  and  $u = A^{-1}(f - Mp)$ . In order to simplify the notation, we will in the following use the notation  $\mathcal{D}(u)$ ,  $\mathcal{D}(p)$  and  $\mathcal{D}(v)$ , respectively, whenever the parametrization of  $\mathcal{D}$  is clear from the context.

We slightly generalize the QVIP discussed above to allow a larger family of data terms. To this end, we assume an arbitrary quadratic convex function  $\mathcal{G}$ , which depends on  $Mp$  only, i.e.  $\mathcal{G}(p) = \tilde{\mathcal{G}}(Mp)$ , where  $\tilde{\mathcal{G}}(\cdot)$  is convex and quadratic. Please note that  $p \mapsto \tilde{\mathcal{G}}(Mp)$  is not strictly convex due to the null space of  $M$ . To simplify the notation, we use  $\mathcal{G}$  instead of  $\tilde{\mathcal{G}}$  from now on. The gradient of  $\mathcal{G}(Mp)$ , which appears in the QVIP, then takes the form  $\nabla\mathcal{G}(Mp) = M^\top g(Mp)$  for some affine  $g$ . In addition to the assumption on  $\mathcal{G}$ , we assume that the constraint set  $\mathcal{D}$  depends on  $Mp$ . With these generalization, the QVIP becomes

$$\begin{array}{l} \text{Find } p^* \in \mathcal{D}(Mp^*) \text{ such that} \\ \langle \nabla\mathcal{G}(Mp^*), p - p^* \rangle \geq 0 \quad \forall p \in \mathcal{D}(Mp^*) \\ \Leftrightarrow \langle M^\top g(Mp^*), p - p^* \rangle \geq 0 \quad \forall p \in \mathcal{D}(Mp^*). \end{array} \quad (6.92)$$

Concerning the primal solution  $u^*$  we assume that it can be retrieved from  $Mp^*$ . We retrieve the previous QVIP (6.88) with  $g(v) = v - f$ .

## 6.4.2 Existence

In this section we discuss existence results for the QVIP (6.92) with arbitrary continuous  $\nabla\mathcal{G}(p)$ . We utilize existence theory from [100].

First, we state necessary assumptions required for the existence of a solution. These assumptions will also be required for uniqueness results provided in Section 6.4.3.

**Assumption 6.22** (for existence). *Assume that*

$$\mathcal{D}: \bar{p} \quad \Rightarrow \quad \mathcal{D}(M\bar{p}) := \{p \in \mathbb{R}^{mn} : p_i \in \mathcal{D}_i^{loc}(M\bar{p}) \subseteq \mathbb{R}^{mn}, i = 1, \dots, n\}, \quad (6.93)$$

where we use the symbol  $\Rightarrow$  to denote a set-valued mapping. Each  $\mathcal{D}_i^{loc} : \mathbb{R}^{mn} \Rightarrow \mathbb{R}^{mn}, i = 1, \dots, n$  in (6.93) has the following properties:

- (i) For fixed  $p$  the set  $\mathcal{D}_i^{loc}(Mp)$  is a closed convex subset of  $\mathbb{R}^{mn}$ .

- (ii) There exists  $C_{max} > 0$ , such that for all  $i, p$ :  $\mathcal{D}_i^{loc}(\mathcal{M}p) \subseteq B_{C_{max}}(0)$  (closed ball with radius  $C_{max}$ ).
- (iii) There exists  $C_{min} > 0$ , such that for every  $p$  and every  $i$  we have  $B_{C_{min}}(0) \subseteq \mathcal{D}_i^{loc}(\mathcal{M}p)$ . In particular,  $\mathcal{D}_i^{loc}(\mathcal{M}p)$  is non-empty.
- (iv) The projection  $\Pi_{\mathcal{D}_i^{loc}(\mathcal{M}p)}(q)$  of  $q$  onto  $\mathcal{D}_i^{loc}(\mathcal{M}p)$  for a fixed  $q$  is continuous with respect to  $p$ .

(Please note that items (i) to (iii) are also required for existence for the TV-based data-driven adaptive regularization, cf. Assumption 4.10.)

**Remark 6.23.** Please note that, since the  $\mathcal{D}_i^{loc}$  are closed and convex, the projection  $\Pi_{\mathcal{D}_i^{loc}(v)}$  onto each  $\mathcal{D}_i^{loc}(v)$  is well defined. With the same argument  $\Pi_{\mathcal{D}(v)}$  is well-defined. Moreover, the projection  $\Pi_{\mathcal{D}(v)}$  can be expressed by the projections  $\Pi_{\mathcal{D}_i^{loc}(v)}$ . Thus Assumption 6.22(iv) together with the linearity of  $\mathcal{M}$  provides the continuity of  $p \mapsto \Pi_{\mathcal{D}(\mathcal{M}p)}$ .

**Proposition 6.24.** Let  $p \mapsto \mathcal{G}(\mathcal{M}p)$  continuously differentiable. Moreover, let  $\mathcal{D}(\mathcal{M}p)$  be defined as in (6.93), such that  $\mathcal{D}_i^{loc}(\mathcal{M}p), i = 1, \dots, n$  satisfy Assumption 6.22. Then the QVI problem (6.92) has a solution.

*Proof.* Since the mapping  $p \mapsto F(p) := \Pi_{\mathcal{D}(\mathcal{M}p)}(p - \nabla \mathcal{G}(\mathcal{M}p)) \in (B_{C_{max}}(0))^n$  is well-defined and continuous, Brouwer's fixed point theorem provides that a fixed point of  $F$  exists. Theorem. 5.1. in [37] shows that this fixed point solves the QVIP (6.92).  $\square$

We will see that for uniqueness results a higher regularity of  $p \mapsto \Pi_{\mathcal{D}(\mathcal{M}p)}(q)$ , namely Lipschitz-continuity is required.

**Remark 6.25** (A-priori bounds). From Assumption 6.22 (ii) we derive an a-priori bound for  $\mathcal{D}(\mathcal{M}p)$  independent from  $p$ :

$$\mathcal{D}(\mathcal{M}p) \subseteq (B_{C_{max}}(0))^n \subseteq B_{\sqrt{n}C_{max}}(0) \quad \forall p \in \mathbb{R}^{mn}. \quad (6.94)$$

We define  $R := \sqrt{n}C_{max}$ . In particular, (6.94) provides a bound for a solution  $p^*$  of (6.92):

$$p^* \in \mathcal{D}(\mathcal{M}p^*) \subseteq B_R(0), \text{ i.e. } \|p^*\|_2 \leq R = \sqrt{n}C_{max}. \quad (6.95)$$

### 6.4.3 Uniqueness

Now we provide uniqueness results for  $v^* := \mathcal{M}p^*$ , where  $p^*$  is a solution to the QVIP (6.92). Recall that for these results we required that  $\nabla \mathcal{G}(p)$  has the specific form

$$\nabla \mathcal{G}(\mathcal{M}p) = \mathcal{M}^\top g(\mathcal{M}p) \quad (6.96)$$

with some function  $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ .

We prove uniqueness of  $v^*$  under the following assumption:

**Assumption 6.26** (for uniqueness).

(i)  $g(v)$  is Lipschitz-continuous with Lipschitz-constant  $\mu$ .

(ii)  $g(v)$  is strongly monotone with constant  $\nu$ : For all  $v_1, v_2 \in \mathbb{R}^n$  we have

$$\langle g(v_1) - g(v_2), v_1 - v_2 \rangle \geq \nu \|v_1 - v_2\|_2^2. \quad (6.97)$$

(iii) The set  $\mathcal{D}(p)$  depends only on  $v = Mp$ . We consider the variation rate  $\eta$  of  $v \mapsto \mathcal{D}(v)$ , i.e. the smallest number, such that for all  $v, v' \in \mathbb{R}^n$  and  $q \in (B_{C_{max}}(0))^n$

$$\|\Pi_{\mathcal{D}(v)}q - \Pi_{\mathcal{D}(v')}q\|_2 \leq \eta \|v - v'\|_2, \quad \forall q \in B_R(0). \quad (6.98)$$

We assume that  $\eta < \infty$ .

(iv) The variation rate  $\eta$  is less than  $\frac{1}{\|M\|_2} \frac{\nu}{\mu}$ .

**Remark 6.27.**

1. When considering the QVIP (6.88), we have  $g(v) = v - f$  and thus  $\mu = \nu = 1$ .
2.  $\|M\|_2$  is the operator norm of  $M$  and thus  $\|M\|_2 = \sqrt{\lambda_{max}}$ , where  $\lambda_{max}$  is the largest eigenvalue of  $M^T M$ .

Before showing uniqueness, let us first define the set-valued operator  $\hat{T}(v)$  mapping from  $\mathbb{R}^n$  to  $B_R(0)$  as follows: Let  $\bar{p} \in \hat{T}(v)$  if and only if  $\bar{p} \in \mathcal{D}(v)$  and it solves the VI

$$\langle \nabla \mathcal{G}(M\bar{p}), p - \bar{p} \rangle \geq 0 \quad \forall p \in \mathcal{D}(v). \quad (6.99)$$

We remark that due to our assumptions of  $\mathcal{G}$  being quadratic and convex and  $g$  being strongly monotone,  $\mathcal{G}(v)$  is actually strictly convex. It follows that operator  $M \circ \hat{T}(v) = M \arg \min_{\tilde{v} \in M\mathcal{D}(v)} \mathcal{G}(\tilde{v})$  (where  $M\mathcal{D}(v) := \{q \mid \exists p \in \mathcal{D}(v) \text{ with } q = Mp\}$ ) is single-valued.

We find that for any solution  $p^*$  to QVIP (6.92)  $v^* = Mp^*$  is a fixed point of  $M \circ \hat{T}$ .

**Theorem 6.28** (Uniqueness).

1. Under Assumptions 6.22 and 6.26 (i)-(iii), the mapping  $M \circ \hat{T}$  is Lipschitz-continuous with constant  $\gamma := \eta \|M\|_2 \frac{\mu}{\nu}$ .
2. Let  $p^*$  be a solution of the QVI (6.92) (cf. Proposition 6.24). If in addition to the assumptions above, Assumption 6.26 (iv) holds, then  $v^* := Mp^*$  is unique.

Please note that Theorem 6.28 is a generalization of our theory in [103].

*Proof.* The proof follows the proof of Theorem 6 in the paper by Nesterov & Scriali [122], but uses the specific form of  $\nabla\mathcal{G}(p) = M^\top g(Mp)$ . In particular, we do not require  $\nabla\mathcal{G}$  to be a strongly monotone operator as in [122].

*Claim 1:* We fix two different points  $v_1, v_2 \in \text{im}(M)$ . Let  $\mathcal{D}_i := \mathcal{D}(v_i)$ ,  $p_i \in \hat{\mathcal{T}}(v_i)$  and  $G_i = \nabla\mathcal{G}(Mp_i)$ . If  $M(p_1 - p_2) = 0$ , we immediately find

$$\|M\hat{\mathcal{T}}(v_1) - M\hat{\mathcal{T}}(v_2)\|_2 = \|Mp_1 - Mp_2\|_2 = 0 < c\|v_1 - v_2\|_2 \quad (6.100)$$

for any  $c \geq 0$ .

Let us now assume  $M(p_1 - p_2) \neq 0$ . Since  $p_i$ ,  $i = 1, 2$  solve

$$\arg \min_{p \in \mathcal{D}(v_i)} \mathcal{G}(Mp), \quad (6.101)$$

the VI

$$\langle \nabla\mathcal{G}(Mp_i), q - p_i \rangle = \langle G_i, q - p_i \rangle \geq 0 \quad \forall q \in \mathcal{D}_i \quad (6.102)$$

holds, and, for arbitrary large  $\tau \geq 0$ ,

$$p_i = \Pi_{\mathcal{D}_i}(p_i - \tau G_i). \quad (6.103)$$

For  $q := \Pi_{\mathcal{D}_2}(p_1 - \tau G_1)$ , we find from (6.103) and the condition on the variation rate (6.98) that

$$\|p_1 - q\|_2 = \|\Pi_{\mathcal{D}_1}(p_1 - \tau G_1) - \Pi_{\mathcal{D}_2}(p_1 - \tau G_1)\|_2 \leq \eta\|v_1 - v_2\|_2. \quad (6.104)$$

On the other hand,  $q$  minimizes the distance to  $p_1 - \tau G_1$  within  $\mathcal{D}_2$ :

$$q = \arg \min_{p \in \mathcal{D}_2} \frac{1}{2} \|p - p_1 + \tau G_1\|^2. \quad (6.105)$$

Since  $p_2 \in \mathcal{D}_2$ , it follows from the VI corresponding to (6.105) that

$$\langle q - (p_1 - \tau G_1), p_2 - q \rangle \geq 0. \quad (6.106)$$

From (6.102) and (6.106) it follows that

$$\begin{aligned} \langle q - p_1, p_2 - q \rangle &\geq \tau \langle G_1, q - p_2 \rangle \\ &= \tau \langle G_1, q - p_1 \rangle + \tau \langle G_2, p_1 - q \rangle + \underbrace{\tau \langle G_2, q - p_2 \rangle}_{\geq 0} + \tau \langle G_1 - G_2, p_1 - p_2 \rangle \\ &\geq \tau \langle G_1 - G_2, q - p_1 \rangle + \tau \langle G_1 - G_2, p_1 - p_2 \rangle. \end{aligned} \quad (6.107)$$

Now we use the specific form assumed on  $\nabla\mathcal{G}(p_i)$ , cf. (6.96):

$$G_i = M^\top g(Mp_i) =: M^\top g_i, \quad (6.108)$$

and, using the strong monotonicity of  $g$ , further bound

$$\begin{aligned}
 & \tau \langle G_1 - G_2, q - p_1 \rangle + \tau \langle G_1 - G_2, p_1 - p_2 \rangle \\
 &= \tau \langle M^\top (g_1 - g_2), q - p_1 \rangle + \tau \langle M^\top (g_1 - g_2), p_1 - p_2 \rangle \\
 &= \tau \langle M^\top (g_1 - g_2), q - p_1 \rangle + \tau \langle g(M p_1) - g(M p_2), M p_1 - M p_2 \rangle \\
 &\geq \tau \langle M^\top (g_1 - g_2), q - p_1 \rangle + \tau \nu \|M p_1 - M p_2\|_2^2.
 \end{aligned} \tag{6.109}$$

Combining (6.107) and (6.109), rearranging the terms and using the Lipschitz-continuity of  $g$  we find

$$\begin{aligned}
 \tau \nu \|M p_1 - M p_2\|_2^2 &\leq \langle q - p_1, p_2 - q \rangle + \tau \langle M^\top (g_1 - g_2), p_1 - q \rangle \\
 &= \langle q - p_1, p_2 - q \rangle + \tau \langle g(M p_1) - g(M p_2), M p_1 - M q \rangle \\
 &\leq \|q - p_1\|_2 \cdot \|q - p_2\|_2 + \tau \mu \|M p_1 - M p_2\|_2 \cdot \|M p_1 - M q\|_2.
 \end{aligned} \tag{6.110}$$

Next, we divide both sides of (6.110) by  $\tau \nu \|M p_1 - M p_2\|_2 > 0$  and derive

$$\|M p_1 - M p_2\|_2 \leq \frac{\|q - p_1\|_2 \cdot \|q - p_2\|_2}{\tau \nu \|M p_1 - M p_2\|_2} + \frac{\mu}{\nu} \|M p_1 - M q\|_2. \tag{6.111}$$

Since  $\tau$  can be chosen arbitrarily large, we find

$$\|M p_1 - M p_2\|_2 \leq \frac{\mu}{\nu} \|M p_1 - M q\|_2. \tag{6.112}$$

Please observe that by (6.104) we have

$$\|M p_1 - M q\|_2 \leq \|M\|_2 \|p_1 - q\|_2 \leq \|M\|_2 \eta \|v_1 - v_2\|_2. \tag{6.113}$$

Combining (6.112) and (6.113), we find

$$\|M \hat{T}(v_1) - M \hat{T}(v_2)\|_2 = \|M p_1 - M p_2\|_2 \leq \|M\|_2 \eta \frac{\mu}{\nu} \|v_1 - v_2\|_2. \tag{6.114}$$

This proves claim 1.

*Claim 2:* Let  $v_1 \neq v_2$  be two fixed points of  $\hat{T}$ . Since with Assumption 6.26 (iv)  $\|M\|_2 \eta \frac{\mu}{\nu} < 1$ , we obtain from (6.114)

$$\|v_1 - v_2\|_2 = \|M \hat{T}(v_1) - M \hat{T}(v_2)\|_2 < \|v_1 - v_2\|_2, \tag{6.115}$$

which is a contradiction.  $\square$

We remark that alternatively to the Lipschitz-constant  $\gamma$  of  $M \hat{T}$  one often considers the *contraction gap* defined as

$$\delta := 1 - \gamma = 1 - \frac{\mu}{\nu} \|M\|_2 \eta, \tag{6.116}$$

which has to be positive to guarantee uniqueness.

**Remark 6.29.** We revisit QVIP (6.88), where  $\mathcal{G}(p) = \frac{1}{2}\|Mp - f\|^2$  and thus  $g(v) = v - f$  and  $\mu = \nu = 1$ . On this special case we remark:

- An alternative condition to Assumption 6.26 (iv) to guarantee uniqueness is that the Lipschitz-constant of  $v \mapsto M \circ \Pi_{\mathcal{D}(v)}(q)$  is less than 1 for all  $q \in B_R(0)$ . Note that this mapping is different to  $v \mapsto \Pi_{M\mathcal{D}(v)}(q)$  (cf. Section 6.4.1).
- We recall that we are actually interested in the primal variable  $u := A^{-1}(f - Mp)$  (which in our applications is the restored image). It follows from Theorem 6.28 that this variable is unique under Assumptions 6.22 and 6.26.

To guarantee uniqueness of the fixed point problem for specific examples of adaptive TV regularization, it remains to provide a sufficiently small variation rate. The variation rate, on the other hand, typically is related to the regularization strength, as we illustrate by revisiting Example 6.2:

**Example 6.30.** For the weighted TV regularization in Example 6.2 we considered the constraint set  $\mathcal{D}(p)$  determined by the local constraint sets  $\mathcal{D}_i^{loc}$  given as

$$\mathcal{D}_i^{loc} = B_{\alpha_i}(0), \quad \alpha_i := \max\{\alpha_0(1 - \kappa(|(L^\top u)_i|), \varepsilon\}. \quad (6.117)$$

Due to the relation  $u = A^{-1}(f - v)$  (recall that  $v = Mp$ ) each  $\alpha_i$  depends on  $v$  by

$$\alpha_i(v) = \max\{\alpha_0(1 - \kappa(|(L^\top A^{-1}(f - v))_i|), \varepsilon\} \quad (6.118)$$

$$= \max\{\alpha_0(1 - \kappa(|(M^\top(f - v))_i|), \varepsilon\}, \quad (6.119)$$

where we used  $M = A^{-\top}L$ . We calculate the variation rate  $\eta$  of  $\mathcal{D}(v)$ . Let  $v, \tilde{v} \in \mathbb{R}^n, q \in \mathbb{R}^{mn}$  be arbitrary. Since the projection of  $q$  onto  $\mathcal{D}(v)$  is a scaling of the  $n$  components  $q_i \in \mathbb{R}^m$  to at most length  $\alpha_i(v)$ , we find

$$\|\Pi_{\mathcal{D}(v)}q - \Pi_{\mathcal{D}(\tilde{v})}q\|_2^2 \leq \sum_{i=1}^n |\alpha_i(v) - \alpha_i(\tilde{v})|^2 \quad (6.120)$$

$$\leq \alpha_0^2 \kappa^2 \sum_{i=1}^n \left| |(M^\top(f - v))_i| - |(M^\top(f - \tilde{v}))_i| \right|^2 \quad (6.121)$$

$$\leq \alpha_0^2 \kappa^2 \sum_{i=1}^n |(M^\top(f - v) - M^\top(f - \tilde{v}))_i|^2 \quad (6.122)$$

$$= \alpha_0^2 \kappa^2 \|M^\top(v - \tilde{v})\|_2^2 \quad (6.123)$$

$$\leq \alpha_0^2 \kappa^2 \|M\|_2^2 \|v - \tilde{v}\|_2^2. \quad (6.124)$$

Thus  $\eta = \alpha_0 \kappa \|M\|_2$ .

Considering a data term of the form  $\mathcal{S}(u) = \frac{1}{2}\|Au - f\|_2^2$  with invertible  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , we retrieve a dual formulation with  $\mathcal{G}(p) = \frac{1}{2}\|f - Mp\|_2^2$  and  $M = A^{-1}L$ . Since then  $\mu = \nu = 1$ , Theorem 6.28 guarantees a unique solution if

$$\alpha_0 \kappa \|M\|_2^2 < 1. \quad (6.125)$$

For the task of denoising in 2D ( $\Omega \subset \mathbb{R}^2$ ), where  $A = \text{Id}$ ,  $M = L$  and  $\|M\|_2^2 = \|L\|_2^2 = 8$  (assuming an equi-distant grid with grid size 1), condition (6.125) becomes  $\alpha_0 \kappa < \frac{1}{8}$ . Given a fixed maximal regularization strength  $\alpha_0$  we thus can determine feasible values for  $\kappa$  to guarantee uniqueness of the solution.

For the task of deblurring, where  $M = A^{-\top} L$ , we expect that in practical applications  $\|M\|_2 = \|A^{-\top} L\|_2^2 \gg 1$  due to small eigenvalues of  $A$  and thus that uniqueness can be guaranteed only for very small  $\alpha_0$  (weak smoothing) or small  $\kappa$  (weak adaptivity).  $\diamond$

#### 6.4.4 An Analytic Example

We revisit Example 6.30 and focus on the task of denoising in the one-dimensional setting. Recall that in this example the objective function is given as

$$\mathcal{F}(u; v) := \sum_{i=1}^n \left( \frac{1}{2} (u_i - f_i)^2 + \alpha_i(v) \| (L^\top u)_i \|_2 \right), \quad (6.126)$$

with  $\alpha(v) \in \mathbb{R}^n$  for fixed  $v \in \mathbb{R}^n$  defined as

$$\begin{aligned} \alpha_i(v) &:= \max(\alpha_0(1 - \kappa|v_{i+1} - v_i|), \varepsilon) & i < n, \\ \alpha_n(v) &:= 0. \end{aligned} \quad (6.127)$$

Recall that we are searching for a fixed point of

$$u^0 \mapsto \arg \min_u \mathcal{F}(u, u^0). \quad (6.128)$$

We consider data  $f$  to be given as follows: We assume  $n = 3N$  for some  $N > 0$ , such that the grid nodes can be divided into three disjoint sets  $I_1 := \{1, \dots, N\}$ ,  $I_2 := \{N + 1, \dots, 2N\}$  and  $I_3 := \{2N + 1, \dots, 3N\}$ . Now let

$$f_i = \begin{cases} 0 & \text{if } i \in I_1 \cup I_3, \\ 1 & \text{if } i \in I_2. \end{cases} \quad (6.129)$$

It can be shown that any solution of the inner problem asserts  $u_i \in [0, 1]$  and takes the form  $u(a, b) := (u_1(a, b), \dots, u_n(a, b))^\top$  with

$$u_i(a, b) := \begin{cases} a & \text{if } i \in I_1 \cup I_3, \\ b & \text{if } i \in I_2, \end{cases} \quad (6.130)$$

for some  $0 \leq a \leq b \leq 1$ . Below we find a fixed point  $u^*$  of (6.128) of this form, i.e.  $u^* = u(a^*, b^*)$  for some  $0 \leq a^* \leq b^* \leq 1$ . Assuming form (6.130) on  $u$  and analogously on  $u^0$ , the objective function in (6.126) simplifies to

$$\mathcal{F}(u(a, b); u^0(a^0, b^0)) = Na^2 + \frac{N}{2}(b-1)^2 + 2\tilde{\alpha}(a^0, b^0)(b-a), \quad (6.131)$$

where

$$\tilde{\alpha}(a^0, b^0) := \max\{\alpha_0(1 - \kappa(b^0 - a^0)), \varepsilon\}. \quad (6.132)$$

**Lemma 6.31.** *Let  $\kappa \leq 1 - \frac{\varepsilon}{\alpha_0}$ . Then, a fixed point  $u^*$  of  $u^0 \mapsto \arg \min_u \mathcal{F}(u; u_0)$  of the form  $u^* = u^*(a^*, b^*)$ , (cf. (6.130)) is given by:*

$$a^* := \begin{cases} \frac{1}{3} & \text{if } \alpha_0 \geq \frac{N}{3}, \\ \frac{\tilde{\alpha}}{N} & \text{else,} \end{cases} \quad b^* := \begin{cases} \frac{1}{3} & \text{if } \alpha_0 \geq \frac{N}{3}, \\ 1 - \frac{2\tilde{\alpha}}{N} & \text{else,} \end{cases} \quad (6.133)$$

where  $\tilde{\alpha} = \tilde{\alpha}(a^*, b^*) = \alpha_0 N \frac{1-\kappa}{N-3\alpha_0\kappa}$ .

*Proof.* Let  $X = \{(a, b) \mid 0 \leq a \leq b \leq 1\}$ . We calculate a fixed point of  $(a^0, b^0) \mapsto \arg \min_{(a,b) \in X} \mathcal{F}(a, b; a^0, b^0)$ , where

$$\mathcal{F}(a, b; a^0, b^0) := Na^2 + \frac{N}{2}(b-1)^2 + 2\tilde{\alpha}(a^0, b^0)(b-a) \quad (6.134)$$

and

$$\tilde{\alpha}(a^0, b^0) = \max\{\alpha_0(1 - \kappa(b^0 - a^0)), \varepsilon\}. \quad (6.135)$$

With the assumption that  $\kappa \leq 1 - \frac{\varepsilon}{\alpha_0}$ , since  $0 \leq b^0 - a^0 \leq 1$ , (6.135) simplifies to

$$\tilde{\alpha}(a^0, b^0) = \alpha_0(1 - \kappa(b^0 - a^0)). \quad (6.136)$$

For a fixed  $\tilde{\alpha}(a^0, b^0)$ , we find

$$\frac{\partial \mathcal{F}(a, b)}{\partial a} = 2Na - 2\tilde{\alpha}(a^0, b^0), \quad (6.137)$$

$$\frac{\partial \mathcal{F}(a, b)}{\partial b} = N(b-1) + 2\tilde{\alpha}(a^0, b^0). \quad (6.138)$$

From  $\nabla \mathcal{F}(\bar{a}, \bar{b}) = 0$  for optimal  $\bar{a}, \bar{b}$  it follows

$$\bar{a} = \frac{\tilde{\alpha}(a^0, b^0)}{N}, \quad \bar{b} = 1 - \frac{2\tilde{\alpha}(a^0, b^0)}{N}. \quad (6.139)$$

These values  $\bar{a}, \bar{b}$  define an admissible  $\bar{u}(\bar{a}, \bar{b})$  only if  $\bar{b} \geq \bar{a}$ . This condition is satisfied if  $\tilde{\alpha}(a^0, b^0) \leq \frac{N}{3}$ . If  $\tilde{\alpha}(a^0, b^0) \geq \frac{N}{3}$  (including equality), we obtain a 'constant' minimizer of (6.126), i.e.  $\bar{u}_i = (\frac{1}{3}, \dots, \frac{1}{3})^\top$ , from which  $\tilde{\alpha}(\bar{u}) = \alpha_0$  follows. Consequently, we have that  $u^* = (\frac{1}{3}, \dots, \frac{1}{3})^\top$  is a fixed point if  $\alpha_0 \geq \frac{N}{3}$ .

Let us now assume that  $\alpha_0 < \frac{N}{3}$ . For a fixed point, we require  $(\bar{a}, \bar{b}) = (a^0, b^0)$ . Using this inequality and inserting (6.139) into (6.136), we derive

$$\tilde{\alpha}(\bar{a}, \bar{b}) = \alpha_0 \left(1 - \kappa \left(1 - \frac{3\tilde{\alpha}(a^0, b^0)}{N}\right)\right). \quad (6.140)$$

Thus, with  $(a^*, b^*) := (\bar{a}, \bar{b}) = (a^0, b^0)$  we have

$$\tilde{\alpha}(a^*, b^*) = \alpha_0 \left(1 - \kappa \left(1 - \frac{3\tilde{\alpha}(a^*, b^*)}{N}\right)\right) \quad (6.141)$$

$$\Leftrightarrow \tilde{\alpha}(a^*, b^*) = \alpha_0 N \frac{1-\kappa}{N-3\alpha_0\kappa}. \quad (6.142)$$

Please note that (6.142) is well defined with  $\tilde{\alpha} > 0$ , since we assumed  $\kappa \leq 1 - \frac{\varepsilon}{\alpha_0} < 1$  and  $\alpha_0 \leq \frac{N}{3}$ .  $\square$



We remark that in the case  $\alpha_0 \geq \frac{N}{3}$ , depending on the choice of  $\kappa$ , there exist other fixed points than the one provided above. In the case that

$$\alpha_0 < \frac{N}{3} \text{ and } \kappa \leq 1 - \frac{\varepsilon}{\alpha_0}, \quad (6.143)$$

it follows from the proof that the fixed point is unique. Then

$$\alpha_0 \kappa \leq \alpha_0 - \varepsilon \leq \frac{N}{3} - \varepsilon. \quad (6.144)$$

Let us now check uniqueness by means of the theory presented in Section 6.4. Theorem 6.28 provides uniqueness under the assumption that

$$\kappa \alpha_0 \|M\|_2^2 = \kappa \alpha_0 \|L\|_2^2 < 1 \quad (6.145)$$

(cf. Example 6.30). Since in the one-dimensional case four is a tight upper bound for  $\|L\|_2^2$ , we find in our particular example that

$$\kappa \alpha_0 \leq \frac{1}{4} \quad (6.146)$$

guarantees uniqueness. As one would expect, condition (6.146) is stronger than the condition (6.143), as it covers a larger class of fixed point problems.

## 6.5 Data Terms for Image Restoration

When we proposed our ansatz for solution-driven adaptivity in Section 6.1 (in the continuous setting) we assumed the strict convexity of the inner problem. This assumption is required to assert well-definedness of the dependence of  $\mathcal{D}$  on  $u$ .

In the derivation of the QVIP (6.88) we made the assumption that  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is invertible, which besides allowing the equivalent formulation as QVIP guarantees the strict convexity of the inner problem.

Let us discuss this issue in view of the applications of image denoising, image deblurring and image inpainting:

**Denoising:** Since we use  $A \equiv \text{Id}$ , the inner problem is strictly convex.

**Deblurring:** In the continuous setting, the operator  $A$  given by the convolution by some kernel might not be injective, so that the inner problem is not strictly-convex. In the discrete setting, depending on the discretization, the resulting matrix might be invertible.

**Inpainting:** The operator  $A : L^2(\Omega) \rightarrow L^2(\Omega_f)$ ,  $(Au)(x) = u(x)$  for  $x \in \Omega_f$  is not injective, thus the inner problem is not strictly convex.

For the tasks of deblurring and inpainting, we suggest three alternative modifications to circumvent the problem of non-strict convexity:

1. We augment operator  $A$  with  $\varepsilon \text{Id}$  for some small  $\varepsilon > 0$ .
2. We consider the sum of two convex data terms  $\mathcal{S}_1(u)$  and  $\mathcal{S}_2(u)$ , one of which is strictly convex.
3. For the task of inpainting, we can think of a third alternative, namely to use the operator

$$\mathcal{S}(u) = \frac{1}{2} \|w\|_{L^2(\Omega)}^2 \text{ with } w(x) = \begin{cases} u(x) - f(x) & \text{if } x \in \Omega_f, \\ \varepsilon(u(x) - u_0(x)) & \text{else,} \end{cases} \quad (6.147)$$

for a small  $\varepsilon > 0$  and some arbitrary function  $u_0 \in L^2(\Omega)$ .

On the second alternative we remark the following. We consider the discrete setting and the QVIP resulting from the proposed data term of the form

$$\mathcal{S}(u) = \mathcal{S}_1(u) + \mathcal{S}_2(u) \quad (6.148)$$

with  $\mathcal{S}_i(u) = \frac{a_i}{2} \|A_i u - f_i\|_2^2$  for some  $a_i > 0$  and arbitrary  $f_i$ ,  $i = 1, 2$ . Then, the optimality condition for  $u$  minimizing  $\mathcal{S}_1(u) + \mathcal{S}_2(u) + \langle Lp, u \rangle$  for fixed  $p \in \mathcal{D}$  reads

$$\begin{aligned} a_1 A_1^\top (A_1 u - f_1) + a_2 A_2^\top (A_2 u - f_2) + Lp &= 0 \\ \Leftrightarrow (a_1 A_1^\top A_1 + a_2 A_2^\top A_2)u &= a_1 A_1 f_1 + a_2 A_2 f_2 - Lp. \end{aligned} \quad (6.149)$$

Assuming that one operator is positive semi-definite and the other is positive definite,  $(a_1 A_1^\top A_1 + a_2 A_2^\top A_2)$  is invertible and we have

$$u = (a_1 A_1^\top A_1 + a_2 A_2^\top A_2)^{-1} (A_1 f_1 + A_2 f_2 - Lp). \quad (6.150)$$

Thus, following the derivation in Section 6.2 we obtain a QVIP, where  $\nabla G$  depends on  $Mp := (a_1 A_1^\top A_1 + a_2 A_2^\top A_2)^{-1} Lp$ . Moreover, the primal variable  $u$  depends on  $Mp$ . We therefore can apply our theory from Section 6.4.

# Chapter 7

## Numerics

The aim of this section is to discuss numerical algorithms for the solution of convex, non-convex and quasi-variational inequality problems (QVIPs). We will start with algorithms for convex minimization and recall standard methods from the literature, in particular primal-dual algorithms [36] and FISTA [15]. For the solution of QVIPs we will revisit an approach we have proposed in [103]. This approach consists of an outer iteration, in which a sequence of convex minimization problems is solved. For these convex problems, we will utilize the standard algorithms mentioned above. Different to standard approaches for QVIPs, our approach is able to cope with non-strictly monotone operators in the QVIPs.

Finally, we will describe a Newton-like algorithm for solving non-convex optimization problems on manifolds, which we have proposed in [16, 17].

### 7.1 Algorithms for Convex Problems

In this section we recall numerical algorithms from the literature to solve convex optimization problems of the form

$$\begin{aligned} \min_{u \in \mathbb{R}^n} \mathcal{F}_1(u) + \mathcal{F}_2(u), \text{ where} \\ \mathcal{F}_1(u) := \frac{1}{2} \|A u - f\|_2^2, \\ \mathcal{F}_2(u) := \sup_{p \in \mathcal{D}} \langle L p, u \rangle, \end{aligned} \tag{7.1}$$

with some linear operators  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$  and  $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$  and a compact convex set  $\mathcal{D} \subset \mathbb{R}^m$ . Please note that in (7.1) we have set  $X = \mathbb{R}^n$  for the space to minimize over, which means that we assume that the problem under consideration is already discretized.

In the following, we recall two state-of-the-art algorithms for solving (7.1), for which a-posteriori error bounds on the numerical solution are available.

### 7.1.1 Primal-Dual

An extensive study of algorithms for solving (7.1) can be found in the paper by Chambolle and Pock [36]. Starting with a basic algorithm, they present several accelerated variants exploiting higher regularity of  $\mathcal{F}_i, i = 1, 2$ .

Here, we focus on Algorithm 2 from [36], which is suited best under the assumption that  $A$  is invertible. We adopt this algorithm to the special form of  $\mathcal{F}(u)$  in (7.1):

---

**Algorithm 2:** A primal-dual algorithm for minimizing functionals of the form (7.1).

---

**Input:**  $N \geq 1$ , data  $f$ , initial guess  $u^{(0)} \in \mathbb{R}^n, p^{(0)} \in \mathbb{R}^m$ ,

**Output:**  $u = u^{(N)}$

Let  $\tilde{u}^{(0)} = u^{(0)}$  and  $\tau^{(0)}, \sigma^{(0)} > 0$  such that  $\tau^{(0)}\sigma^{(0)}\|L\|_2^2 \leq 1$ .

**begin**

**for**  $k = 1, \dots, N$  **do**

$p^{(k)} := \Pi_{\mathcal{D}}(p^{(k-1)} + \sigma(k-1)L^\top \tilde{u}^{(k-1)})$       // update dual variable

$u^{(k)} := (\text{Id} + \tau^{(k-1)}\partial\mathcal{F}_1)^{-1}(u^{(k-1)} - \tau^{(k-1)}Lp^{(k)})$       // update primal

**variable**

$\theta^{(k)} := \frac{1}{\sqrt{1+2\tau^{(k-1)}}}$

$\tau^{(k)} := \tau^{(k-1)}\theta^{(k)}$

$\sigma^{(k)} := \frac{\sigma^{(k-1)}}{\theta^{(k)}}$

$\tilde{u}^{(k)} := u^{(k)} + \theta^{(k)}(u^{(k)} - u^{(k-1)})$       // extrapolation step

---

Please note that  $(\text{Id} + \tau\partial\mathcal{F})^{-1}$  in Algorithm 2 is the *resolvent* or *proximity* operator defined as

$$(\text{Id} + \tau\partial\mathcal{F})^{-1}(u_0) := \arg \min_u \frac{1}{2\tau}\|u - u_0\|_2^2 + \mathcal{F}(u). \quad (7.2)$$

In the case  $A = \text{Id}$  we have  $(\text{Id} + \tau\partial\mathcal{F}_1)^{-1}(u_0) = \frac{1}{1+\tau}(u_0 + \tau f)$ . In the case of  $A \neq \text{Id}$  we refer to [36, Section 6] for a discussion on efficient implementations of the resolvent operator. Convex constraints on  $u$  can be incorporated into  $\mathcal{F}_1$  using the indicator function  $\iota_{\mathcal{K}}$ .

The following error estimate holds for Algorithm 2 (cf. [36, Theorem 2]):

**Theorem 7.1.** *Let  $(u^{(k)}, p^{(k)})_k$  be the sequence generated by Algorithm 2 with the choice  $\tau_0 := \tau^{(0)} > 0$  and  $\sigma^{(0)} = \frac{1}{\tau^{(0)}\|L\|_2^2}$ . Then, for any  $\varepsilon > 0$  there exists  $N_0$  depending on  $\tau_0$  and  $\varepsilon$ , such that for any  $N \geq N_0$*

$$\|u^{(N)} - \bar{u}\|_2^2 \leq \frac{1 + \varepsilon}{N^2} \left( \frac{\|u^{(0)} - \bar{u}\|_2}{\tau_0^2} + \|L\|_2\|p^{(0)} - \bar{p}\|_2^2 \right), \quad (7.3)$$

where  $\bar{u}$  is the exact solution to (7.1) and  $\bar{p} := \arg \max_{p \in \mathcal{D}} \langle Lp, \bar{u} \rangle$ .

*Proof:* see [36].

## 7.1.2 FISTA

The *Fast Iterative Shrinking and Thresholding Algorithm* (FISTA) was proposed by Beck and Teboulle in [15]. FISTA can be applied in the situation where at least one of the functionals  $\mathcal{F}_i(u)$ ,  $i = 1, 2$  is differentiable with Lipschitz-continuous gradient.

Here, we consider applying FISTA to the dual problem of (7.1), which under the assumption that  $A$  is invertible can be reformulated as

$$\min_{p \in \mathcal{D}} \mathcal{G}(p) := \min_{p \in \mathcal{D}} \frac{1}{2} \|Mp - f\|_2^2 \quad (7.4)$$

with  $M = A^{-\top} L$ .

Given a solution  $\bar{p}$  of the dual problem (7.4), the solution of the primal problem then is retrieved via  $\bar{u} := A^{-1}(f - M\bar{p})$ .

By means of the indicator function  $\iota_{\mathcal{D}}$  (cf. (3.3)) of set  $\mathcal{D}$  problem (7.4) can be re-written as

$$\min_{p \in \mathbb{R}^m} \mathcal{G}_1(p) + \mathcal{G}_2(p), \quad (7.5)$$

where

$$\begin{aligned} \mathcal{G}_1(p) &:= \frac{1}{2} \|Mp - f\|_2^2, \\ \mathcal{G}_2(p) &:= \iota_{\mathcal{D}}(p). \end{aligned} \quad (7.6)$$

We observe that  $\nabla \mathcal{G}_1(p)$  is Lipschitz-continuous with constant  $\lambda := \|M\|_2^2$ .

The minimization procedure of FISTA makes use of an auxiliary strict convex problem

$$\min_{u \in \mathbb{R}^n} \frac{\Lambda}{2} \|u - (v - \frac{1}{\Lambda} \nabla \mathcal{G}_1(v))\|_2^2 + \mathcal{G}_2(u) \quad (7.7)$$

for fixed  $v \in \mathbb{R}^n$  and  $\Lambda > 0$ . We denote the unique minimizer of (7.7) by  $\rho_{\Lambda}(v)$ .

The FISTA algorithm (with constant step size) is given as

---

### Algorithm 3: FISTA.

---

**Input:**  $N \geq 1$ , initial guess  $p^{(0)} \in \mathcal{D} \subset \mathbb{R}^m$ , Lipschitz constant  $\lambda$  of  $\nabla \mathcal{G}_1$

**Output:**  $p = p^{(N)}$ .

Let  $\tau^{(0)} = 1$ ,  $\tilde{p}^{(1)} = p^{(0)}$

**begin**

<p><b>for</b> <math>k = 1, \dots, N</math> <b>do</b></p>	<p><math>p^{(k)} := \rho_{\lambda}(\tilde{p}^{(k-1)})</math></p>
<p style="padding-left: 20px;"><math>\tau^{(k)} := \frac{1}{2}(1 + \sqrt{1 + 4(\tau^{(k-1)})^2})</math></p>	
<p style="padding-left: 20px;"><math>\tilde{p}^{(k)} := p^{(k)} + \frac{\tau^{(k-1)} - 1}{\tau^{(k)}}(p^{(k)} - p^{(k-1)})</math></p>	

---

We briefly recall the convergence results for Algorithm 3, cf. [15, Theorem 4.4]:

**Theorem 7.2.** *Let  $\bar{p}$  be the solution of (7.4) and let  $(p^{(k)})_k$  be the sequences generated by Algorithm 3. Then for any  $k \geq 1$*

$$\mathcal{G}(p^{(k)}) - \mathcal{G}(\bar{p}) \leq \frac{2\lambda}{(k+1)^2} \|p^{(0)} - \bar{p}\|_2^2. \quad (7.8)$$

From Theorem 7.2, we can also obtain an error bound on the numerical solution of the primal problem  $\mathcal{F}(u)$  as follows:

**Proposition 7.3** (Error bound). *Let  $p^{(N)}$  be the result of Algorithm 3 applied to the dual problem (7.5). For  $Mp^{(N)}$  we have the following error estimate:*

$$\|Mp^{(N)} - M\bar{p}\|_2 \leq \frac{2\|M\|_2}{N+1} \|p^{(0)} - \bar{p}\|_2, \quad (7.9)$$

Using the boundedness of  $\mathcal{D}$  and defining  $R := \max\{\|p\|_2 \mid p \in \mathcal{D}\}$ , it follows that

$$\|Mp^{(N)} - M\bar{p}\|_2 \leq \frac{4\|M\|_2 R}{N+1}. \quad (7.10)$$

**Remark 7.4.** *Let us consider the primal variable  $u$  given as  $u = A^{-1}(f - Mp^{(N)})$ . Due to the estimate*

$$\|u^{(N)} - \bar{u}\|_2 \leq \|A^{-1}\|_2 \|Mp^{(N)} - M\bar{p}\|_2 \quad (7.11)$$

we obtain from (7.9) and (7.10) corresponding error bounds for  $u$  with an additional factor  $\|A^{-1}\|_2$ .

The proof of Proposition 7.3 requires the following additional lemma.

**Lemma 7.5.** *Let  $\mathcal{D}$  be a non-empty, closed and convex set and  $\mathcal{G}(p) := \frac{1}{2}\|Mp - f\|_2^2 + \iota_{\mathcal{D}}(p)$ . The minimizer  $\bar{p}$  of the problem  $\min_{p \in \mathbb{R}^m} \mathcal{G}(p)$  satisfies*

$$\frac{1}{2}\|Mp - M\bar{p}\|_2^2 \leq \mathcal{G}(p) - \mathcal{G}(\bar{p}), \text{ for every } p \in \mathcal{D}. \quad (7.12)$$

*Proof.* The proof follows [56, Eqn. (20)-(25)]. We consider the decomposition  $\mathcal{G}(p) = \mathcal{G}_1(p) + \mathcal{G}_2(p)$  as in (7.5) and re-write  $\mathcal{G}_1(p) = H(Mp)$ , where  $H(v) := \frac{1}{2}\|v - f\|_2^2$ . Moreover, we define

$$I(p) := H(Mp) - H(M\bar{p}) - \langle M^\top \left( \frac{\partial}{\partial v} H(M\bar{p}) \right), p - \bar{p} \rangle, \quad (7.13)$$

$$J(p) := \mathcal{G}_2(p) - \mathcal{G}_2(\bar{p}) + \langle M^\top \left( \frac{\partial}{\partial v} H(M\bar{p}) \right), p - \bar{p} \rangle, \quad (7.14)$$

where  $\frac{\partial}{\partial v} H$  is the gradient of  $H$ . Then, by definition,

$$I(p) + J(p) = \mathcal{G}(p) - \mathcal{G}(\bar{p}). \quad (7.15)$$

Since  $H(v)$  is strongly convex with parameter 1, i.e.  $H(v) - H(v') - \langle \frac{\partial}{\partial v} H(v'), v - v' \rangle \geq \frac{1}{2} \|v - v'\|_2^2$ , we find

$$H(Mp) - H(M\bar{p}) - \langle \frac{\partial}{\partial v} H(M\bar{p}), Mp - M\bar{p} \rangle \geq \frac{1}{2} \|Mp - M\bar{p}\|_2^2 \quad (7.16)$$

$$\Leftrightarrow H(Mp) - H(M\bar{p}) - \langle M^\top (\frac{\partial}{\partial v} H(M\bar{p})), p - \bar{p} \rangle \geq \frac{1}{2} \|Mp - M\bar{p}\|_2^2 \quad (7.17)$$

$$\Leftrightarrow I(p) \geq \frac{1}{2} \|Mp - M\bar{p}\|_2^2. \quad (7.18)$$

Now, we show that  $J(p) \geq 0$ . Since  $\bar{p}$  is the minimizer of  $\mathcal{G}$ , we have

$$0 \in \partial(H \circ M)(\bar{p}) + \partial\mathcal{G}_2(\bar{p}) \quad (7.19)$$

$$\Leftrightarrow -M^\top (\frac{\partial}{\partial v} H(M\bar{p})) \in \partial\mathcal{G}_2(\bar{p}), \quad (7.20)$$

where  $\partial H = \{M^\top (\frac{\partial}{\partial v} H(M\bar{p}))\}$  and  $\partial\mathcal{G}_2$  are the sub-differentials of  $H(v)$  and  $\mathcal{G}_2(p)$ , respectively. By definition of the sub-differential,

$$-M^\top (\frac{d}{dp} H(M\bar{p})) \in \partial\mathcal{G}_2(\bar{p}) \quad (7.21)$$

$$\Leftrightarrow \mathcal{G}_2(p) \geq \mathcal{G}_2(\bar{p}) + \langle -M^\top (\frac{\partial}{\partial v} H(M\bar{p})), p - \bar{p} \rangle, \quad \forall p \in \mathbb{R}^m \quad (7.22)$$

$$\Leftrightarrow J(p) \geq 0, \quad \forall p \in \mathbb{R}^m. \quad (7.23)$$

Inserting (7.18) and (7.23) into (7.15) shows the claim.  $\square$

*Proof of Proposition 7.3.* By applying Theorem 7.2, we obtain

$$\mathcal{G}(p^{(N)}) - \mathcal{G}(\bar{p}) \leq \frac{2\|M\|_2^2}{(N+1)^2} \|p^{(0)} - \bar{p}\|_2^2. \quad (7.24)$$

Combining (7.12) from Lemma 7.5 and (7.24) we find

$$\|Mp^{(N)} - M\bar{p}\|_2^2 \leq \frac{4\|M\|_2^2}{(N+1)^2} \|p^{(0)} - \bar{p}\|_2^2. \quad (7.25)$$

Taking the square root on both sides of (7.25), we derive the first claim

$$\|Mp^{(N)} - M\bar{p}\|_2 \leq \frac{2\|M\|_2}{(N+1)} \|p^{(0)} - \bar{p}\|_2. \quad (7.26)$$

The inequality (7.10) follows from (7.26) using

$$\|p^{(0)} - \bar{p}\|_2 \leq \|p^{(0)}\|_2 + \|\bar{p}\|_2 \leq 2R. \quad (7.27)$$

$\square$

## 7.2 An Algorithm for Solving QVIP

In the following, we discuss an algorithm for solving QVIPs of the form

$$\begin{aligned}
 &\text{Find } p^* \in \mathcal{D}(Mp^*) \text{ such that} \\
 &\quad \langle \nabla \mathcal{G}(Mp^*), p - p^* \rangle \geq 0 \quad \forall p \in \mathcal{D}(Mp^*) \\
 &\Leftrightarrow \langle M^\top g(Mp^*), p - p^* \rangle \geq 0 \quad \forall p \in \mathcal{D}(Mp^*).
 \end{aligned} \tag{7.28}$$

we have introduced in Section 6.4. We proposed this algorithm for the special case  $g(v) = v - f$  originally in [103].

Before we describe the algorithm in detail, we require to recall some definitions from Section 6.4. In this section we have introduced the solution operator  $\hat{T}(v)$  to the variational equation

$$\begin{aligned}
 &\text{Find } p^* \in \mathcal{D}(v) \text{ such that} \\
 &\quad \langle M^\top g(Mp^*), p - p^* \rangle \geq 0 \quad \forall p \in \mathcal{D}(v),
 \end{aligned} \tag{7.29}$$

i.e. any  $p^* \in \hat{T}(v)$  solves (7.29).

Throughout this section, we assume that Assumptions 6.22 and 6.26 from Section 6.4 are satisfied. In particular, these assumptions imply that the operator  $M \circ \hat{T}$  is single-valued and a contraction. We denote the Lipschitz constant of  $M \circ \hat{T}$  by  $\gamma$ . Consequently,  $\gamma < 1$ .

### 7.2.1 Proposed Algorithm

In the following, we present an algorithm to solve the QVIP (7.28),

This algorithm consists of two nested iterations. In the outer loop we update the value  $v$ , which defines the constraint set  $\mathcal{D}(v)$ . The inner step consists in solving the *inner problem*, which thanks to its convex nature can be stated in three equivalent forms (with fixed constraint set  $\mathcal{D}(v)$ ):

**1. Saddle-point problem:**

$$\min_{u \in \mathbb{R}^n} \max_{p \in \mathcal{D}(v)} \mathcal{S}(u) + u^\top Lp, \tag{7.30}$$

**2. Dual Problem:**

$$\min_{p \in \mathcal{D}(v)} \mathcal{G}(p), \tag{7.31}$$

where  $\mathcal{G}(p)$  is the Fenchel dual to  $\mathcal{F}(u) = \max_{p \in \mathcal{D}(v)} \mathcal{S}(u) + u^\top Lp$ .

**3. Variational Inequality (VI):**

$$\begin{aligned}
 &\text{find } \bar{p} \in \mathcal{D}(v) \text{ such that} \\
 &\quad \langle \nabla \mathcal{G}(\bar{p}), p - \bar{p} \rangle \geq 0, \quad \forall p \in \mathcal{D}(v).
 \end{aligned} \tag{7.32}$$



In the outer iteration, see Algorithm 4, we update variable  $v^{[k]}$ , which defines the constraint set  $\mathcal{D}(v^{[k]})$  and apply a numerical solver to the convex inner problem (7.30) - (7.31). The outer iteration is fairly independent from the formulation of the inner problem and the solver used, except that in order to guarantee convergence we have assume some a-priori bounds on the error of the numerical solution of the inner problem. We outline this error bound below. For the moment, we consider some arbitrary numerical solver providing some  $\bar{p}_{approx} \approx \bar{p}$ , where  $\bar{p}$  is the exact solution of the inner problem. We denote this solver by  $sol(\mathcal{D}, p_0, N)$ , where  $\mathcal{D}$  is the fixed constraint set,  $p_0$  is an initial value and  $N$  is the number of inner iteration steps. (In order to distinguish outer iterates from inner iterates, we use squared brackets when indexing the former, and parentheses when indexing the latter, i.e.  $p^{[k]}$  versus  $p^{(k)}$  for variable dual  $p$ .)

---

**Algorithm 4:** Outer Iteration

---

**Output:** Sequence  $(p^{[k]})_k$  converging to a solution  $p^*$  of (6.92).  
 Choose arbitrary  $p^{[0]} \in B_R(0) \subset \mathbb{R}^{mn}$ . // initialization  
**begin**  
     **for**  $k = 0, \dots, K - 1$  **do**  
          $v^{[k]} = M p^{[k]}$   
          $p^{[k+1]} = sol(\mathcal{D}(v^{[k]}), p^{[k]}, N)$  //  $\rightarrow$  solving inner problem  
     **end for**

---

In Section 7.2.2 below we show that under sufficient conditions the sequence  $(p^{[k]})_k$  produced by Algorithm 4 converges to a fixed point  $p^*$  of (6.92). These conditions include that Assumptions 6.22 and 6.26 are satisfied (which in particular guarantees a unique  $p^*$ ) and the following a-priori error bound on the solution of each inner problem. Recall that the operator which maps  $v$  to an *exact* solution  $\bar{p}$  of (7.30) - (7.31) is denoted by  $\hat{T}(v)$ . We assume that for any  $\varepsilon > 0$  we can find  $N$  large enough and independent of  $p^0 \in B_R(0)$  and  $\mathcal{D}$ , such that the inner problem in terms of  $v = Mp$  can be solved up to an error less than  $\varepsilon$ :

$$\|M(sol(\mathcal{D}(v_0), p_0, N) - M\hat{T}(v))\|_2 \leq \varepsilon, \quad \forall p_0 \in B_R(0), v_0 \in \mathbb{R}^{mn}, \quad (7.33)$$

where  $R$  is the a-priori bound on  $p$  (cf. Remark 6.25).

Numerical methods for solving (7.30)-(7.31) with the required error bound have been discussed in Section 7.1.

### 7.2.2 Convergence

In the following, we show convergence of Algorithm 4 and provide convergence *rates* for a special case.

**Proposition 7.6** (Convergence). *Let  $M \circ \hat{T}$  be a contraction (with Lipschitz constant  $\gamma < 1$ ). For fix  $\varepsilon > 0$ , we assume that the number of iterations of the solver for the inner problem (7.30)–(7.31) is chosen large enough to guarantee*

$$\|M \text{sol}(\mathcal{D}(v), p, N) - M \hat{T}(v)\|_2 \leq \varepsilon \quad (7.34)$$

(independent of  $p \in B_R(0)$  and  $v$ ).

Then, the following holds:

(i) for the sequence  $v^{[k]}$  resulting from Algorithm 4 we have

$$\|v^{[k]} - v^*\|_2 \leq \varepsilon \frac{1}{1 - \gamma} + \gamma^k \|v^0 - v^*\|_2, \quad (7.35)$$

where  $v^*$  is the unique fixed point of  $M \circ \hat{T}$ .

(ii)  $\{v^{[k]}\}_k$  converges to the fixed point  $v^*$  for  $\varepsilon \rightarrow 0$  and  $k \rightarrow \infty$ .

*Proof.* We have

$$\begin{aligned} \|v^{[k]} - v^*\|_2 &= \|M p^{[k]} - M p^*\|_2 = \|M p^{[k]} - M \hat{T}(v^*)\|_2 \\ &\leq \|M p^{[k]} - M \hat{T}(v^{[k-1]})\|_2 + \|M \hat{T}(v^{[k-1]}) - M \hat{T}(v^*)\|_2 \\ &\leq \|M \text{sol}(\mathcal{D}(v^{[k-1]}), p^{[k-1]}, N) - M \hat{T}(v^{[k-1]})\|_2 + \gamma \|v^{[k-1]} - v^*\|_2 \\ &\leq \varepsilon + \gamma \|v^{[k-1]} - v^*\|_2. \end{aligned} \quad (7.36)$$

Recursion gives

$$\begin{aligned} \|v^{[k]} - v^*\|_2 &\leq \varepsilon + \gamma \|v^{[k-1]} - v^*\|_2 \\ &\leq \varepsilon + (\gamma \varepsilon + \gamma^2 \|v^{[k-2]} - v^*\|_2) \leq \dots \\ &\leq \varepsilon (1 + \gamma + \gamma^2 + \dots + \gamma^{k-1}) + \gamma^k \|v^{[0]} - v^*\|_2. \end{aligned} \quad (7.37)$$

Using the limit of the geometric series, we deduce claim (i). Claim (ii) follows from (i) under Assumption 6.26 (iv), since then  $\gamma < 1$  and thus  $\gamma^k \rightarrow 0$  for  $k \rightarrow \infty$ .  $\square$

We remark that literature on numerical methods for solving QVIs provides convergence rates (see e.g. [122]) in the case of strongly monotone operators  $\nabla \mathcal{G}$ . In our case, where  $\nabla \mathcal{G}$  is not strongly monotone, we can provide convergence rates only if we have control over the component of  $p^{[k]}$  which lies in the null space  $\mathcal{N}(M)$ .

We therefore consider the following special case: Assume that  $\hat{T}(v^{[k]}) \in \mathcal{N}^\perp(M)$  and that the result of  $\text{sol}(\mathcal{D}(v^{[k]}, p^{[k]}, N)$ , when starting with a value  $p^{[k]} \in \mathcal{N}^\perp(M)$ , stays in this subspace.

Exemplarily, we focus on the FISTA algorithm (Section 7.1.2) to solve the dual problem (7.31).

Let  $\lambda_{min}$  be the smallest positive (recall that  $\mathcal{N}(M) \neq \{0\}$ ) and  $\lambda_{max}$  be the largest eigenvalue of  $M^\top M$ . Using the fact that  $\|x\|_2 \leq \frac{1}{\sqrt{\lambda_{min}}} \|Mx\|_2$  for  $x \in \mathcal{N}^\perp(M)$ , it follows from (7.9) (cf. Proposition 7.3), that

$$\|M \text{sol}(\mathcal{D}(v^{[k]}, p^{[k]}, N) - M \hat{T}(v^{[k]}))\|_2 \leq \frac{2\sqrt{2}\varrho}{N+1} \|Mp^{[k]} - M \hat{T}(v^{[k]})\|_2, \quad (7.38)$$

where  $\varrho := \frac{\lambda_{max}}{\lambda_{min}}$  is the condition number of  $M^\top M$  restricted to  $\mathcal{N}^\perp(M)$ . (Recall that  $p^{(0)} = p^{[k]}$  is the initial value for the inner problem.) Please note that standard error bounds consider  $\|p^{[k]} - \hat{T}(v^{[k]})\|_2$  as initial error on the right hand side of (7.38), while we consider  $\|Mp^{[k]} - M \hat{T}(v^{[k]})\|_2$  and require in (7.38) that  $p^{[k]}, \hat{T}(v^{[k]}) \in \mathcal{N}^\perp(M)$ .

**Proposition 7.7** (Convergence rates). *Let the Lipschitz constant  $\gamma$  of  $M \circ \hat{T}$  be less than 1. Moreover, assume that the inner problem (7.31) can be solved with an error bound*

$$\|Mp^{(N)} - M \hat{T}(v^{[k]})\|_2 \leq \frac{\delta}{4} \|Mp^{(0)} - M \hat{T}(v^{[k]})\|_2, \quad (7.39)$$

where  $\delta := 1 - \gamma$  is the contraction gap (cf. (6.116)) of the problem (7.28). Recall that  $\hat{T}(v^{[k]})$  in (7.39) is the exact solution of the inner problem. Consider a solution  $p^*$  of (6.92) and  $v^* := Mp^*$ . Then, Algorithm 4 converges according to

$$\|v^{[K]} - v^*\|_2 \leq \frac{1}{\delta} \exp\left(-\frac{\delta}{2}K\right) \|v^{[0]} - M \hat{T}(v^{[0]})\|_2, \quad (7.40)$$

where  $K$  is the number of outer iterations.

*Proof.* Recall that in the  $k$ -th outer iteration we solve the inner problem (7.31) with initial value  $p^{(0)} = p^{[k]}$ . The numerical solution  $p^{(N)}$  of this inner problem is denoted as  $p^{[k+1]}$  in the outer loop. Thus, the required error bound (7.39) can be rewritten as

$$\|Mp^{[k+1]} - M \hat{T}(v^{[k]})\|_2 \leq \frac{\delta}{4} \|Mp^{[k]} - M \hat{T}(v^{[k]})\|_2, \quad (7.41)$$

or equivalently, with  $v^{[k]} := Mp^{[k]}$ ,

$$\|v^{[k+1]} - M \hat{T}(v^{[k]})\|_2 \leq \frac{\delta}{4} \|v^{[k]} - M \hat{T}(v^{[k]})\|_2 = \frac{\delta}{4} r_k, \quad (7.42)$$

where  $r_k := \|v^{[k]} - M \hat{T}(v^{[k]})\|_2 = \|Mp^{[k]} - M \hat{T}(v^{[k]})\|_2$ . The proof of Proposition 7.7 follows the lines of the proof of Theorem 4.4 by Nesterov & Scramali. Let  $v^*$  be the unique fixed point of  $M \hat{T}(v)$  provided by Theorem 6.28 (2). From the Lipschitz-continuity of  $M \hat{T}(v)$  we have

$$\begin{aligned} r_k &\geq \|v^{[k]} - v^*\|_2 - \|M \hat{T}(v^*) - M \hat{T}(v^{[k]})\|_2 \\ &\geq \|v^{[k]} - v^*\|_2 - \gamma \|v^{[k]} - v^*\|_2 = \delta \|v^{[k]} - v^*\|_2. \end{aligned} \quad (7.43)$$

Now we show  $r_k \leq \exp(-\frac{\delta}{2}k)r_0$ :

$$\begin{aligned}
 r_{k+1} &= \|v^{[k+1]} - M \hat{T}(v^{[k+1]})\|_2 \\
 &\leq \|v^{[k+1]} - M \hat{T}(v^{[k]})\|_2 + \|M \hat{T}(v^{[k+1]}) - M \hat{T}(v^{[k]})\|_2 \\
 &\stackrel{(7.42)}{\leq} \frac{\delta}{4} \|v^{[k]} - M \hat{T}(v^{[k]})\|_2 + \gamma \|v^{[k+1]} - v^{[k]}\|_2 \\
 &\leq \frac{\delta}{4} \|v^{[k]} - M \hat{T}(v^{[k]})\|_2 + \gamma \|v^{[k+1]} - M \hat{T}(v^{[k]})\|_2 \\
 &\quad + \gamma \|v^{[k]} - M \hat{T}(v^{[k]})\|_2 \\
 &\stackrel{(7.42)}{\leq} (\frac{\delta}{4} + \frac{\delta}{4}\gamma + \gamma)r_k.
 \end{aligned} \tag{7.44}$$

Inserting  $\gamma = 1 - \delta$  we derive

$$\begin{aligned}
 r_{k+1} &\leq (\frac{\delta}{4} + \frac{\delta}{4}(1 - \delta) + 1 - \delta)r_k \\
 &= (1 - \frac{\delta}{2} - \frac{\delta^2}{4})r_k.
 \end{aligned} \tag{7.45}$$

Applying (7.45) recursively, using  $(1 - s)^n \leq \exp(-sn)$  and the monotonicity of the exponential function, we find

$$r_k \leq \exp(-(\frac{\delta}{2} + \frac{\delta^2}{4})k)r_0 \leq \exp(-\frac{\delta}{2}k)r_0. \tag{7.46}$$

Combining (7.43) and (7.46), we finally obtain

$$\delta \|v^{[k]} - v^*\|_2 \leq \exp(-\frac{\delta}{2}k)r_0 = \exp(-\frac{\delta}{2}k)\|v^{[0]} - v^*\|_2. \tag{7.47}$$

Please note that by assumption we have  $\gamma < 1$  and consequently  $\delta = 1 - \gamma > 0$ . Dividing (7.47) by  $\delta$  gives the claimed error estimate.  $\square$

**Remark 7.8.**

**1D case:** *In the one-dimensional case the null space  $\mathcal{N}(M)$  of  $M$  is spanned by the vector  $(0, \dots, 0, 1)^\top$ . When considering Example 6.30 and the FISTA Algorithm (Algorithm 3 above), we can guarantee that the sequence  $(p^{(k)})_k$  generated by FISTA stays in  $\mathcal{N}(M)^\perp$ , provided that the initial value  $p^{(0)} = p_0$  is chosen in  $\mathcal{N}(M)^\perp$ . This is due to the fact that the projection onto  $\mathcal{D}(v^{[k]})$  decouples into independent projections to 1D intervals for each coordinate. In particular, if  $p_n^{(0)} = 0$ , the constraint  $p_n^{(0)} \in [-\alpha_n, \alpha_n]$  is fulfilled and thus  $p_n^{(0)}$  is not changed by the projection  $\Pi_{\mathcal{D}(v^{[k]})}(p)$ , i.e.  $p_n^{(0)} = p_n^{(1)} = \dots = p_n^{(N)} = 0$ . The sequence  $p^{(k)}$  then converges to the solution  $\bar{p} \in \hat{T}(v^{[k]}) \cap \mathcal{N}(M)^\perp$ , which is unique in this subspace. As a consequence, starting with  $p^{[0]} \in \mathcal{N}(M)^\perp$  for the outer iteration, we can guarantee also  $p^{[k]} \in \mathcal{N}(M)^\perp$ , such that the initial value of every subsequent inner problem is again in  $\mathcal{N}^\perp(M)$ . The error estimate (7.38) then allows to apply Proposition 7.7.*

**2D case:** *In the two-dimensional case there exist counter examples indicating that (7.39) does not hold in general. The reason is that the convergence depends on the component  $\|p^{[k]} - \bar{p}\|_{\mathcal{N}(M)}$  of the initial error  $\|p^{[k]} - \bar{p}\|_2$  ( $\bar{p} \in \hat{T}(v^{[k]})$  fixed).*

### 7.2.3 Numerical Experiments

In order to verify the convergence of our algorithm, we revisit the example discussed in Section 6.4.4, for which an analytic formula for a fixed point is at hand. Recall that in this example the functional takes the form

$$\mathcal{F}(u, v) := \frac{1}{2} \sum_{i=1}^n |u_i - f_i|^2 + \sum_{i=1}^{n-1} \alpha_i(v) |u_{i+1} - u_i|. \quad (7.48)$$

for  $u, f \in \mathbb{R}^n$ , where the data  $f$  are given as

$$f_i = \begin{cases} 0 & \text{if } i \in I_1 \cup I_3, \\ 1 & \text{if } i \in I_2. \end{cases} \quad (7.49)$$

For this particular case, the analytic form of a fixed point of  $u^0 \rightarrow \arg \min_u \mathcal{F}(u; u_0)$  is provided by Lemma 6.31.

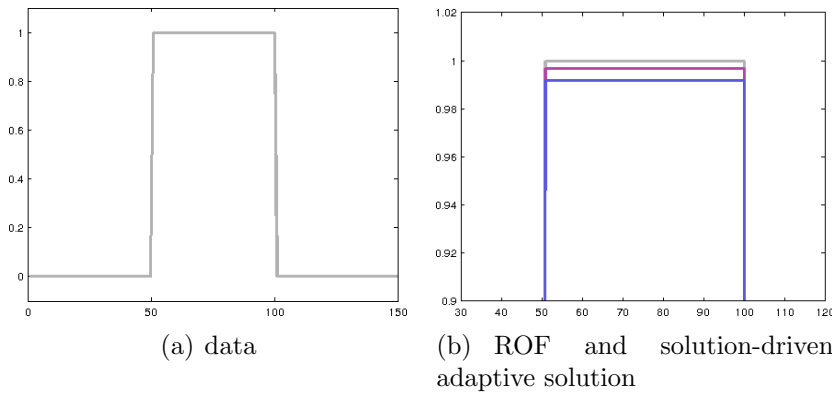


Figure 7.1: *1D example with analytic solution (cf. Section 6.4.4).* Left: data  $f$ . Right: solution  $u^*$  of (7.48) (zoom, red line) for 150 grid nodes and parameters  $\alpha_0 = 0.2$ ,  $\kappa = 0.6$  compared to the solution of the standard ROF model (blue line) with  $\alpha = 0.2$ . Smoothing with the proposed approach provides a result with higher contrast compared to the standard ROF model.

By means of this example, we experimentally verify the convergence rate provided by Proposition 7.7. To this end, we solve the corresponding QVI numerically with the proposed algorithm and investigate the experimental error. Fig. 7.2 shows the theoretical and experimental convergence rates (logarithmic error over time steps) for this example and different contraction gaps  $\delta = 1 - \lambda_2 = 1 - \alpha_0 \kappa \|L\|_2^2$ . (Please note that the positive contraction gaps assert uniqueness). The experimental errors  $\|u^{[K]} - u^*\|_2 = \|Lp^{[K]} - Lp^*\|_2$  stay significantly below the theoretical bound and also show an exponential decay.

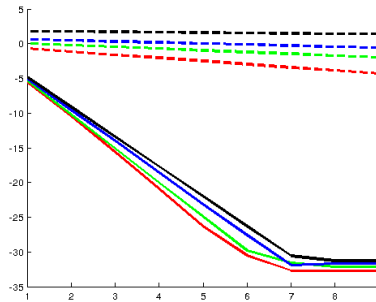


Figure 7.2: *Convergence rates.* Plot of logarithmic numerical error (solid lines) to analytic solution over outer iteration steps for different contraction gaps  $\delta = 1 - \kappa\alpha_0\mu_2 = 0.1, 0.3, 0.5, 0.9$  given by  $\alpha_0 = 2$ ,  $\kappa = \frac{t}{4\alpha_0}$  and  $t \in \{0.9, 0.7, 0.5, 0.1\}$  (black, blue, green, red). The plot shows an exponential error decay, which stays well below the theoretical bound (dashed lines). The bendings between step 6 and 8 is caused by the fact that the point-wise errors reach machine accuracy.

## 7.3 A Newton-like Algorithm for Non-convex Optimization Problems

### 7.3.1 Introduction

In this section, we consider the optimization problem

$$\min_{u \in G} \mathcal{F}(u), \quad (7.50)$$

where  $G$  is a Lie group. We assume that  $G$  is embedded in an Euclidean space  $X$ . As an example, we consider  $G = SE(3) \subset \mathbb{R}^3 \times \mathbb{R}^{3 \times 3}$ . We assume that  $\mathcal{F}(u)$  is differentiable but non-convex. The above setting is chosen in view of an application presented in Section 8.6, namely the joint estimation of scene and egomotion from monocular image sequences, where such an optimization problem arises.

In the following, we will discuss a Newton-like method to solve the problem (7.50). Newton-like methods in the more general case of Riemannian manifolds, but with the restriction on convex minimization problems, have already been studied in the literature. An early reference to this topic is [60]. Convergence of such methods has been intensively studied in [2, 77, 168, 181] and recently in [135]. The approach we present below, however, considers the more special case of Lie groups and make use of the group structure. Moreover, we consider also *non-convex functionals*  $\mathcal{F}$ .

### 7.3.2 Preliminaries

In the following, we recall some facts about manifolds and Lie groups needed for the presentation of our approach. To keep the presentation compact, we refer to [2] for

the basic definitions of differential geometry and focus on some specific definitions required in the context of Lie groups.

To begin with, let  $u \cdot v$  denote the group operation on two elements  $u, v \in G$  and  $e$  be the neutral element of this operation. We consider  $G$  acting on the manifold  $\mathcal{M} = G$  by defining

$$l_u(v) := u \cdot v. \quad (7.51)$$

As already mentioned, we assume that  $\mathcal{M}$  is embedded in an Euclidean space. We denote the inner product of this space by  $\langle \cdot, \cdot \rangle$ .

The tangent space at point  $u$  is denoted by  $T_u G$ . From the operator  $l_u$ , we derive its differential  $L_u$ , which maps bijectively from  $T_e G$  to  $T_u G$ . Thus, we can identify each tangent space  $T_u G$  with  $T_e G$ . Moreover, the tangent space  $T_e G$  can be identified with a Lie algebra  $\mathfrak{g}$ . The collection of all pairs  $(u, T_u G)$ ,  $u \in G$  is referred to as the tangent bundle  $T\mathcal{M}$ .

We equip the manifold  $G$  with a Riemannian metric by defining

$$\langle v, w \rangle_u := \langle (L_u)^{-1}v, (L_u)^{-1}w \rangle_{\mathfrak{g}} \quad (7.52)$$

for some metric  $\langle \cdot, \cdot \rangle_{\mathfrak{g}}$  on  $T_e G = \mathfrak{g}$ . Identifying  $T_u G$  with  $T_e G$ , i.e.  $v = L_u \tilde{v}$  and  $w = L_u \tilde{w}$  for some  $\tilde{v}, \tilde{w}$ , we see that the metric does not depend on the point  $u$ . For the remainder of this section, we thus denote the Riemannian metric by  $\langle \cdot, \cdot \rangle_G$ . Furthermore, we require the exponential map from  $T_e G$  to  $G$ , which we denote by  $\text{Exp}$ .

Next, we define the manifold gradient. We start with the classical gradient  $\nabla \mathcal{F}(u)$  in the embedding  $X$ . Based on this gradient we define the manifold gradient  $\nabla_G \mathcal{F}$  implicitly by the equation

$$\langle \nabla_G \mathcal{F}(u), v \rangle_G = \langle \nabla \mathcal{F}(u), v \rangle, \quad \forall v \in T_u G, \quad (7.53)$$

(cf. [2]) or, equivalently,

$$\langle \nabla_G \mathcal{F}(u), L_u v \rangle_G = \langle \nabla \mathcal{F}(u), L_u v \rangle, \quad \forall v \in T_e G. \quad (7.54)$$

Using the adjoint  $L_u^*$  of  $L_u$ , (7.54) can be rewritten as

$$\langle L_u^* \nabla_G \mathcal{F}(u), v \rangle_G = \langle L_u^* \nabla \mathcal{F}(u), v \rangle, \quad \forall v \in T_e G. \quad (7.55)$$

Thus, we can retrieve  $\nabla_G \mathcal{F}(u)$  by projecting  $L_u^* \nabla \mathcal{F}(u)$  onto  $T_e G$ :

$$\nabla_G \mathcal{F}(u) := L_u^* \Pi_{T_e G}(L_u^* \nabla \mathcal{F}(u)), \quad (7.56)$$

where  $\Pi_{T_e G}(v)$  denotes the projection of  $v$  onto  $T_e G$ .

Finally, we require the Levi-Civita connection  $\bar{\nabla}$  of the manifold  $\mathcal{M}$ , which is defined as the unique torsion-free and metric preserving affine connection ([2])  $\bar{\nabla} : C^\infty(\mathcal{M}, T\mathcal{M}) \times C^\infty(\mathcal{M}, T\mathcal{M}) \rightarrow C^\infty(\mathcal{M}, T\mathcal{M})$ , given as

$$\bar{\nabla}_W V := \lim_{t \rightarrow 0} \frac{1}{t} \left( L_{u(0)} L_{u(t)}^{-1} V(u(t)) - V(u(0)) \right) \quad (7.57)$$

for any trajectory  $u(t) : \mathbb{R} \rightarrow \mathcal{M}$  with  $\frac{d}{dt}u(t) = W(u(t))$ .

**Example 7.9.** *We provide two basic examples.*

**$\mathbb{R}^n$ :** *We consider  $\mathcal{M} = \mathbb{R}^n$  and identify it with the Lie group  $G = \mathbb{R}^n$ . We represent an element  $x \in \mathbb{R}^n$  by a standard column vector  $x = (x_1, \dots, x_n)^\top$ . Then the group operation is given as the element-wise addition,  $x \cdot y = (x_1 + y_1, \dots, x_n + y_n)^\top$  and the neutral element is  $(0, 0, \dots, 0)^\top$ . The tangent space  $T_e G$  can be identified with the space  $\mathbb{R}^n$  itself. On  $T_e G$ , we consider the canonical basis  $\{e_k\}_{k=1, \dots, n}$  of  $\mathbb{R}^n$  and the standard scalar product  $\langle x, y \rangle = x^\top y$ . The mappings  $l_x, L_x, \text{Exp}$  and  $\Pi_{T_e G}$  are just the identity operators. The Levi-Civita connection  $\bar{\nabla}_v w$  for  $v = (v_1, \dots, v_n)^\top$  simplifies to*

$$\bar{\nabla}_v w = \sum_{i=1}^n \langle w, \partial v_i \rangle e_i. \quad (7.58)$$

**SE(3):** *The Lie group  $SE(3)$  is embedded in  $\mathbb{R}^{4 \times 4}$ . Since it is common to refer to matrices with capital letters, we change our notation accordingly. The group operation is the matrix product  $A \cdot B$ . The neutral element is the identity matrix. The operator  $l_Q$  is given as  $l_Q W = QW$ . The Lie algebra  $\mathfrak{g} = \mathfrak{se}(3)$  is given as*

$$\mathfrak{g} = \left\{ W = \begin{pmatrix} [w]_\times & v \\ 0^\top & 0 \end{pmatrix} \mid w, v \in \mathbb{R}^3 \right\}, \quad (7.59)$$

where

$$[w]_\times := \begin{pmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{pmatrix} \quad (7.60)$$

represents the cross product of vectors through  $[w]_\times v = w \times v$ . The mapping  $L_Q$  from  $T_{\text{Id}} G$  to  $T_W G$  is given as  $L_Q W = QW$  with the standard matrix product and thus takes the same form as  $l_q$ . Closed forms are available for the projection onto the tangent space, and for  $\text{Exp} : T_{\text{Id}} G \rightarrow G$ .

The Riemannian metric we consider is given as

$$\langle W^1, W^2 \rangle_G := \langle [w^1]_\times, [w^2]_\times \rangle + \langle v^1, v^2 \rangle. \quad (7.61)$$

The Levi-Civita connection  $\bar{\nabla}_W$  can be described, after choosing a basis for  $\mathfrak{g}$ , by means of the Christoffel symbols. We refer to [17] for details.

◇

### 7.3.3 Optimization

We consider subsequent update steps to derive a sequence  $(u^k)$  such that  $\mathcal{F}(u^k)$  decreases with  $k$ . Given  $u^k$ , we determine  $u^{k+1}$  by

$$u^{k+1} = \phi(t^k, u^k, d^k) \quad (7.62)$$



with search direction  $d^k$ , step size  $t^k \in \mathbb{R}_+$  and an update function  $\phi$  given as

$$\phi(t, u, d) := l_u \text{Exp}(t \cdot d). \quad (7.63)$$

Let us now consider an operator  $B : T_e G \rightarrow T_e G$  such that

$$\langle v, Bv \rangle_G > 0 \quad \forall v \in T_e G, v \neq 0. \quad (7.64)$$

Then,

$$d^k := -B \nabla_G \mathcal{F}(e) \in T_e G \quad (7.65)$$

is a descent direction [17, Proposition 3], i.e.

$$\frac{d}{dt} \mathcal{F}(\phi(t, u^k, d^k))|_{t=0} < 0, \quad (7.66)$$

as long as  $\nabla_G \mathcal{F}(e) \neq 0$ . Thus, since due to our assumption  $\mathcal{F}(u)$  is smooth, it is possible to find a  $t^k > 0$  such that  $\mathcal{F}(u^{k+1}) < \mathcal{F}(u^k)$ .

Now we are going to use information from second-order derivatives to find a suitable  $B$ . Our approach follows the classical Newton method for vector spaces  $\mathbb{R}^n$ , where one would choose  $B$  as the inverse of the Hessian of  $\mathcal{F}(u^k)$ . On a Riemannian manifold, the Levi-Civita connection  $\bar{\nabla}$  takes the role of the Hessian. Applying the Levi-Civita connection to the manifold gradient, i.e.  $\bar{\nabla}_w \nabla_G \mathcal{F}(u^k)$ , describes the change of  $\nabla_G \mathcal{F}(u^k)$  in the tangent space when moving from the point  $u^k$  to the point  $\phi(1, u^k, d)$ . We are looking for a  $d^k$  such that we can reach a critical point  $\nabla_G \mathcal{F}(\phi(1, u^k, d^k)) = 0$ . Using

$$0 \stackrel{!}{=} \nabla_G \mathcal{F}(\phi(1, u^k, d^k)) \approx \nabla_G \mathcal{F}(u^k) + \bar{\nabla}_{d^k} \nabla_G \mathcal{F}(u^k), \quad (7.67)$$

we see that  $d^k$  should be chosen such that the right hand side of (7.67) also vanishes, i.e. that

$$\bar{\nabla}_{d^k} \nabla_G \mathcal{F}(u^k) = -\nabla_G \mathcal{F}(u^k). \quad (7.68)$$

Note that the Levi-Civita connection is linear in  $d^k$ , thus we can find a linear operator  $A : T_e G \rightarrow T_e G$ , such that

$$A d^k = \bar{\nabla}_{d^k} \nabla_G \mathcal{F}(u^k). \quad (7.69)$$

Using (7.69), we reformulate (7.68) as

$$A d^k = -\nabla_G \mathcal{F}(u^k). \quad (7.70)$$

We first focus on the case that  $A$  is positive definite. Then,  $A$  is invertible and (7.70) can be equivalently written as

$$d^k = -(A)^{-1} \nabla_G \mathcal{F}(u^k). \quad (7.71)$$

Thus, choosing  $B := A^{-1}$  in (7.65), which is also positive definite, provides a descent direction  $d^k$ . We remark that in the case  $G = \mathbb{R}^n$  the matrix  $A$  just becomes the

Hessian  $H_{\mathcal{F}}$  of  $\mathcal{F}$  at  $u^k$ , see Example 7.10 below. Thus the proposed approach in the case  $G = \mathbb{R}^n$  equals the classical Newton methods.

It remains to provide a suitable  $B$  when  $A$  is not positive-definite. In this case, we consider an eigenvalue decomposition of  $A$  and propose to modify  $A$  by adding a positive definite  $M$ , where the eigenvalues are chosen large enough to guarantee that  $A + M$  is positive definite. Denoting the smallest eigenvalue of  $A$  by  $\sigma_{min} < 0$ , one choice would be to choose  $M = c\text{Id}$  with  $c = \sigma_{min} + \varepsilon$  for some small  $\varepsilon > 0$ . In general, the choice of  $M$  is depending on the specific application at hand. For a particular example, where a specific choice of  $M$  is motivated, we refer to [17].

We remark that modifying matrix  $A$  in this way can be motivated by applying a proximal point algorithm to the problem

$$\min_{d \in T_{u^k}G} h(d), \quad h(d) := \mathcal{F}(u^k) + \langle \nabla_G \mathcal{F}(u^k), d \rangle_G + \frac{1}{2} \langle d, Md \rangle_G, \quad (7.72)$$

such that  $h(d)$  is augmented by an additive quadratic term  $\frac{1}{2} \langle d, Md \rangle_G$ .

**Example 7.10.** *We revisit the specific case  $G = \mathbb{R}^n$  in Example 7.9. From (7.58) we obtain with  $v := \nabla_G \mathcal{F}(u)$*

$$\overline{\nabla}_w \nabla \mathcal{F}(u) = H_{\mathcal{F}}(u)w, \quad (7.73)$$

where  $H_{\mathcal{F}}(u)$  is the Hessian of  $\mathcal{F}$  at  $u$ . In particular, the linear operator  $A$  with  $Aw = \overline{\nabla}_w \nabla \mathcal{F}(u) = H_{\mathcal{F}}(u)w$  is just the Hessian  $H_{\mathcal{F}}(u^k)$  of  $\mathcal{F}(u)$ , and thus, the above choice of  $B = (A)^{-1}$  leads to the standard Newton method.  $\diamond$

In the case that  $A$  is not invertible and is augmented by an additional term to guarantee invertibility, it is not clear that the time step  $t^k = 1$  in (7.62) is optimal for an efficient descent step. We therefore propose to use a standard line search to determine the approximate optimal  $t^k$  in each time step. In particular, we choose Wolfe's Rule, [27] as criterion for the line search. The line search is applied to approximately solve

$$t^k \approx \arg \min_{t > 0} \mathcal{F}(\phi(t, u^k, d^k)). \quad (7.74)$$

We refer to Section 8.6 for an application, which utilizes the above optimization strategy.

# Chapter 8

## Applications

### 8.1 Image Restoration

In the following, we evaluate the adaptive regularization methods proposed in Chapters 4 and 6 on three image restoration problems, namely image denoising, deblurring and inpainting.

Our evaluation is based on the five test images depicted in Fig. 8.1, referred to as *cameraman*, *Lena*, *peppers*, *boat* and *mandrill*. All undistorted images (referred to as the *original* images in the following) are scaled to the range  $[0, 1]$  before a distortion is applied.

For each restoration task we consider variational approaches of the form

$$\bar{u} = \arg \min_u \mathcal{S}(u) + \mathcal{R}(u; v), \quad (8.1)$$

where  $\bar{u}$  represents the restored image,  $\mathcal{S}(u)$  is an application specific data term as detailed in the individual subsections below and  $\mathcal{R}(u; v)$  is the regularization term depending on some parameter  $v$ . Here,  $v$  is chosen to be the input data  $f$  for the data-driven approaches or – for the solution-driven approaches – found by solving a fixed point problem (i.e.  $\bar{u} = v$ ), cf. Chapter 6. For the data-driven methods, the numerical minimization of the variational problem (8.1) after discretization is performed by means of the primal-dual method as described in Section 7.1.1. To solve the fixed point problem of the solution-driven variants, i.e.  $\bar{u} = v$ , we utilize the algorithm proposed in Section 7.2.

The regularization methods we consider in our evaluation are:

1. weighted TV, see Example 4.25 with regularization strength  $\alpha(x, v)$ ,
2. anisotropic TV, see Example 4.26 with edge indicator  $\chi_e(x; v)$  as in (4.83) and edge normals  $v_e(x; v)$ ,
3. TV with double orientations, see Example 4.27 with two non-orthogonal orientations  $r_1(x; v)$  and  $r_2(x; v)$ ,



Figure 8.1: Test images used to evaluate adaptive TV regularization for the applications of denoising (2nd row), deblurring (3rd row) and inpainting (4th row). The first row shows the original undistorted images. For the application of deblurring, due to the support of kernel function, the data are given only on  $\Omega_f \subseteq \Omega$ . In the test images for inpainting the regions colored in white indicate the regions to be inpainted.

4. anisotropic second-order TV, see Example 4.30.

For the first three methods we consider both the data-driven (d.d.) and solution-driven (s.d.) variants. On anisotropic second-order TV we remark that the solution-driven variant does not show significant differences to the data-driven variant in all three tasks. For this reason we only consider the latter in our evaluation.

We compare to the following state-of-the-art regularization methods:

1. standard TV [142] see Section 2.3,
2. standard second-order TV [147] see (2.36) in Section 2.8.3, and
3. Total Generalized Variation (TGV) [28] see (2.37) in Section 2.8.3.

In addition, for each restoration task we choose one application-specific state-of-the-art method to compare with, as detailed in the following sections.

We compare all methods qualitatively as well as quantitatively. For the latter evaluation, we make use of the mean SSIM index proposed by Wang et al. [171], which for each method under consideration is applied to the pair of original and reconstructed image. The optimal value for similarity is 1. In addition, for the sake of completeness, we also provide the peak-signal-to-noise ratio (PSNR) for each reconstruction.

Both quantitative and qualitative comparison are fair only if for each method suitable parameters are chosen. In order to determine (approximately) optimal parameters, we utilize an automated search as discussed in the appendix at the end of this section. Such an automated search is possible because of the availability of the undistorted image.

### 8.1.1 Denoising

In the following, we evaluate the data-driven and solution-driven adaptive regularization methods, introduced in Chapters 4 and 6, respectively, for denoising grayscale images.

Our evaluation is based on the five test images depicted in Fig. 8.1, top row. We consider the model of additive Gaussian noise with zero mean and standard deviation  $\sigma = 0.1$ , cf. Section 2.2. The noisy versions of the five test images are depicted in Fig. 8.1, second row.

In the variational approaches, we choose the data term

$$\mathcal{S}(u) := \frac{1}{2} \|u - f\|_{L^2(\Omega)}^2, \quad (8.2)$$

as the data term, where  $f \in L^2(\Omega)$  represents the noisy input data.

In addition to the methods listed above, we choose BM3D [47] as application-specific state-of-the-art methods.

For the ROF model, we use the primal-dual algorithm discussed in Section 7.1.1. For the TGV approach we make use of code provided by the authors of [28].

Method	cameraman	Lena	peppers	boat	mandrill
std. TV	0.8078	0.7934	0.7941	0.7181	0.6467
TV <sup>2</sup>	0.8160	0.8143	0.8188	0.7334	0.6561
TGV	0.8143	0.7982	0.8101	0.7231	0.6490
BM3D	<b>0.8556</b>	<b>0.8450</b>	<b>0.8346</b>	<b>0.7620</b>	<b>0.6895</b>
weighted TV(d.d.)	0.8215	0.7934	0.7941	0.7181	0.6466
weighted TV(s.d.)	0.8243 (4)	0.7934 (2)	0.7967 (2)	0.7181 (4)	0.6466 (2)
anisotr. TV(d.d.)	0.8320	0.8283	0.7945	0.7181	0.6570
anisotr. TV(s.d.)	0.8324 (4)	0.8254 (4)	0.8073 (4)	0.7272 (4)	0.6595 (4)
double orient.(d.d.)	0.8254	0.8083	0.8056	0.7283	0.6459
double orient.(s.d.)	0.8254 (4)	0.8082 (4)	0.8055 (4)	0.7283 (4)	0.6471 (2)
anisotr. TV <sup>2</sup> (d.d.)	0.8403	0.8361	0.8282	0.7501	0.6646

Table 8.1: *Denoising*. Quantitative comparison based on the similarity between the denoised and original images. The ideal value is 1. For the solution-driven approaches we choose the number  $K \in \{2, 3, 4\}$  of outer iterations, for which we obtain the highest SSIM index after parameter optimization ( $K$  given in brackets). We observe that BM3D [47] gives the best results. The best method among the regularization approaches is the anisotropic second-order TV. For the weighted TV approach in some cases the automated parameter optimization chooses  $\kappa = 0$  as optimal value, so that the method becomes similar to standard TV.

Method	cameraman	Lena	peppers	boat	mandrill
std. TV(d.d.)	26.483	26.905	27.424	26.062	24.159
TV <sup>2</sup> (d.d.)	26.568	27.294	27.921	26.370	24.311
TGV(d.d.)	26.491	26.936	27.730	26.135	23.969
BM3D(d.d.)	<b>28.550</b>	<b>28.446</b>	<b>28.527</b>	<b>27.022</b>	<b>24.845</b>
weighted TV(d.d.)	27.081	26.901	27.421	26.061	24.194
weighted TV(s.d.)	27.077 (3)	26.899 (2)	27.512 (3)	26.063 (2)	24.195 (2)
anisotr. TV(d.d.)	27.100	27.957	27.407	26.063	24.338
anisotr. TV(s.d.)	27.489 (4)	27.729 (4)	27.940 (4)	26.338 (3)	24.355 (4)
double orient.(d.d.)	26.988	27.558	27.872	26.312	24.115
double orient.(s.d.)	26.995 (4)	27.555 (4)	27.870 (4)	26.313 (4)	24.101 (2)
anisotr. TV <sup>2</sup> (d.d.)	27.336	27.923	28.220	26.690	24.479

Table 8.2: *Denoising*. Quantitative comparison based on the peak-signal-to-noise-ratio (PSNR) of the denoised images. Also w.r.t. PSNR, BM3D [47] gives the best results. When switching from the SSIM index to PSNR the ranking in most cases stays the same, except e.g. for Lena, where data-driven anisotropic TV becomes the second best, replacing the anisotropic second-order TV.

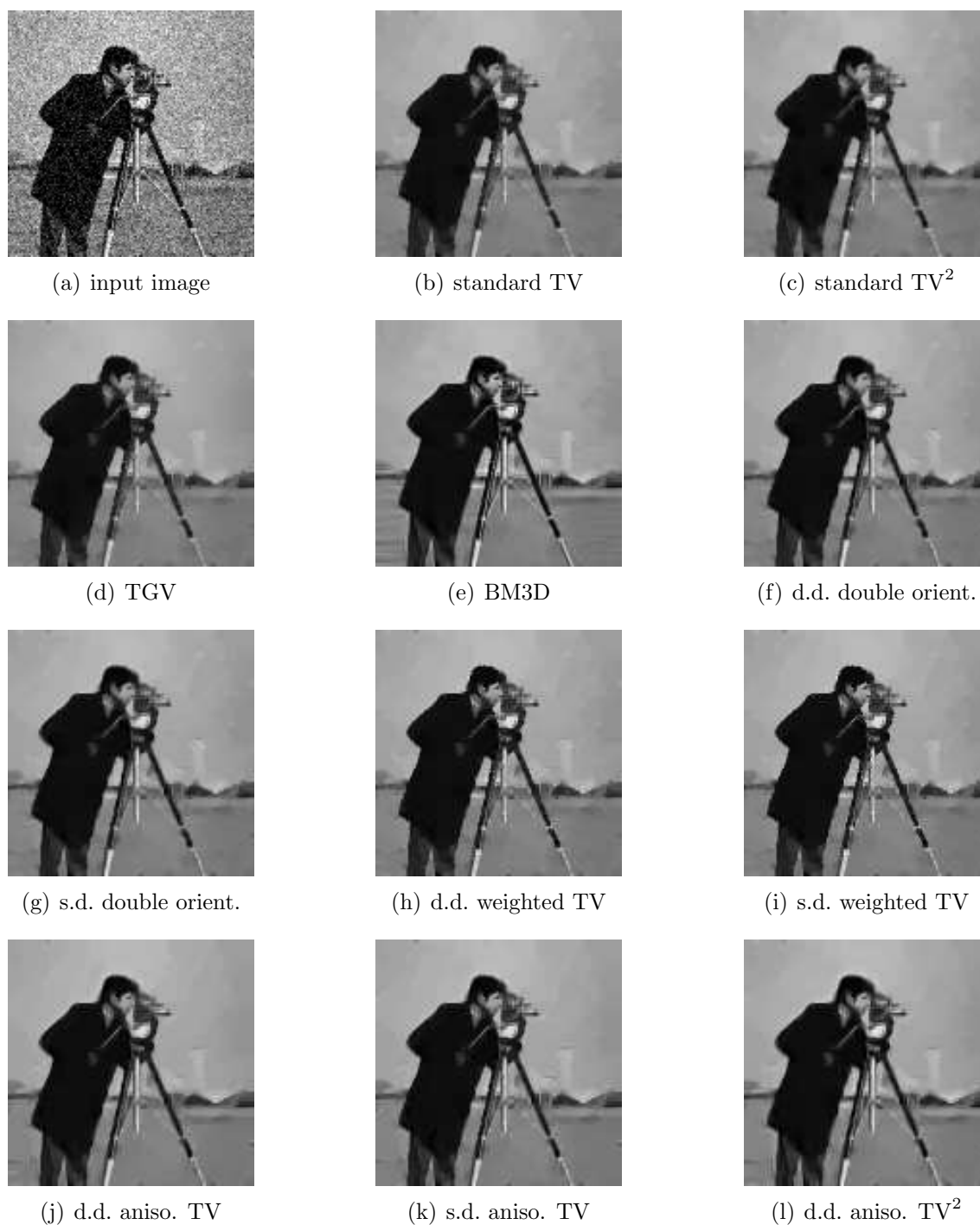


Figure 8.2: *Denoising of the cameraman image.* Qualitative comparison of the considered denoising methods on the representative results. In order to highlight differences, we provide a close-up of the head region in Fig. 8.3.



Figure 8.3: *Denoising of the cameraman image.* Qualitative comparison of the considered denoising methods on a close-up of the head region of the cameraman image. (For the sake of completeness, we provide the SSIM indices, now determined on the specific image region. The PSNR gives the same ranking.) BM3D [47] also in terms of visual appearance provides the best result. Weighted TV produces an almost as sharp result as BM3D. The other methods produce smoother results.





(a) data-driven aniso. TV



(b) solution-driven aniso. TV

(c) data-driven aniso. TV<sup>2</sup>

(d) BM3D

Figure 8.4: *Denoising of the Lena image* We evaluate the performance of the anisotropic TV models to the best competing method, namely BM3D. In this example we observe a slightly worse reconstruction from the solution-driven anisotropic TV compared to the data-driven one. This is in agreement with the SSIM indices, see Table 8.1. For the readers convenience, we also provide the results from anisotropic second-order TV and BM3D. The latter gives the best results.



(a) data-driven aniso. TV



(b) solution-driven aniso. TV

(c) data-driven aniso. TV<sup>2</sup>

(d) BM3D

Figure 8.5: *Denoising of the peppers image.* In this example we observe an improvement by solution-driven adaptivity, which provides a more regular result with less remaining noise than the data-driven variant. For the readers convenience, we also provide the results from anisotropic second-order TV and BM3D. The latter gives the best result.



(a) data-driven aniso. TV



(b) solution-driven aniso. TV

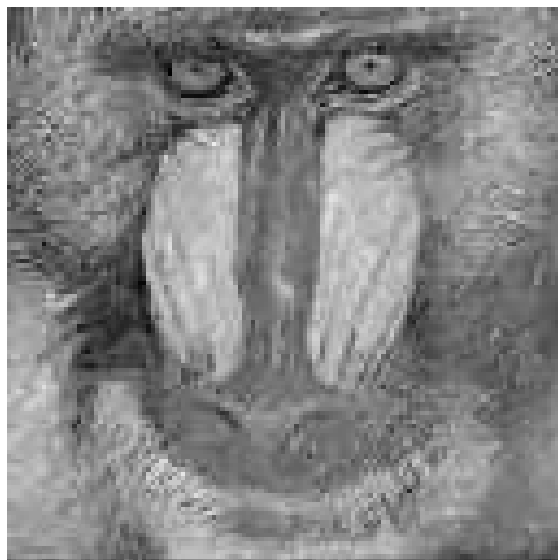
(c) data-driven aniso. TV<sup>2</sup>

(d) BM3D

Figure 8.6: *Denoising of the boat image.* We again observe an improvement by solution-driven adaptivity, which compared to the data-driven variant provides a more regular result with less remaining noise. For the readers convenience, we also provide the results from anisotropic second-order TV and BM3D. The latter two give the best results.



(a) data-driven aniso. TV



(b) solution-driven aniso. TV

(c) data-driven aniso. TV<sup>2</sup>

(d) BM3D

Figure 8.7: *Denoising of the mandrill image.* We compare the four best performing methods on the mandrill image. A visual comparison shows that as in the quantitative evaluation BM3D is the best performing method. In particular, it shows a better reconstruction of the image textures than the other methods. The results of the latter are of similar visual quality.

Concerning the BM3D method we use the original MATLAB source, version 1.9, from [http://www.cs.tut.fi/~foi/GCF-BM3D/index.html#ref\\_software](http://www.cs.tut.fi/~foi/GCF-BM3D/index.html#ref_software).

We compare all methods qualitatively as well as quantitatively. As mentioned above, to provide a fair comparison, we perform an automated parameter selection, in which for each method we optimize the results with respect to the mean SSIM index proposed by Wang et al. [171], which for each method under consideration is applied to the pair of original and reconstructed image.

We then compare the methods under consideration with respect to the SSIM index and, for the sake of completeness, also with the peak-signal-to-noise-ratio (PSNR). The results are listed in Tables 8.1 (SSIM) and 8.2 (PSNR), respectively.

Let us first consider the SSIM indices of the results, see Table 8.1. The ideal value would be 1. Clearly, the BM3D method outperforms the other methods. The second best method and in particular the best regularization method is the anisotropic second-order TV regularization.

Concerning the first-order methods we observe the following.

Firstly, in most cases we see an improvement of the data-driven adaptive methods compared to standard TV model. Here, the *weighted* TV method is an exception, since for four of five images the SSIM index is the same as for standard TV. The reason is, that the automated parameter selection process chooses the value  $\kappa = 0$ , which means that there is actually no adaptivity applied. In contrast, data-driven *anisotropic* TV and TV regularization based on *double directions* compared to standard TV improve the results in four out of five cases.

Secondly, we focus on the improvement between data-driven and solution-driven methods. We observe that solution-driven TV with *double orientations* in four of five cases does not provide an improvement (with the mandrill image being an exception). The reason might be that we do only choose the double directions  $r_i$  in an adaptive fashion, while parameter  $\alpha$  is kept fix. It seems that these orientations are quite well estimated already on the noisy input data. Solution-driven *weighted* TV gives better results than data-driven weighted TV in two of five cases. In the other three cases, as for the data-driven variants, the non-adaptive mode ( $\kappa = 0$ ) was chosen during parameter optimization. Solution-driven *anisotropic* TV improves the results of the data-driven variant in four out of five cases. Compared to standard TV we see that solution-driven anisotropic TV always gives better results.

Finally, we discuss the results of *anisotropic second-order* TV. As mentioned above, we consider only the data-driven variant, since we do not observe significant changes with the solution-driven counterpart. For four of the test images we observe that anisotropic second-order TV provides the second best results after BM3D. The exception is the mandrill image, which contains more textures than the other test images. Consequently, none of the TV-based methods is able to provide a good result. The fact that anisotropic second-order TV performs well in the other cases demonstrates the importance of second-order regularization for the task of image denoising.

Considering the PSNR instead of the SSIM index, we observe that the ranking generally does not change, except for the first-order anisotropic TV, which now in

two of five cases becomes the second best method after BM3D.

We also compare the results qualitatively. For the shortness of presentation, we show the results of *all* considered methods only for the cameraman image, while for the other test images we only compare a subset of the considered methods. The depicted results all match those listed in Tables 8.1 and 8.2.

The results for the cameraman image are shown in Fig. 8.2. Since details are only visible after zooming-in, we provide a close-up of the head region in Fig. 8.3. On these results we observe that BM3D, as in the quantitative evaluation, gives the best results with a lot of details. The adaptive methods in general provide better results than standard TV, in particular with respect to contrast. Moreover, the solution-driven variants slightly improve over the data-driven methods, e.g. in preserving edges. The results of TGV as well as the anisotropic second-order TV are slightly smoother, but not as blurry as the result from standard second-order TV. In particular, we observe that introducing second-order TV prevents stair-casing effects.

For the remaining test images, we concentrate on the data- and solution-driven first-order anisotropic TV approach, data-driven second-order anisotropic TV and BM3D as competing method, as those methods performed best in the quantitative comparison. The corresponding results are depicted in Figs. 8.4 to 8.7. We observe that the solution-driven approach in general is able to produce more regular results compared to the data-driven approach. In the case of the Lena image, this comes with a slightly worse reconstruction of details. For the mandrill image we do not observe a significant improvement of the result. This might be due to the fact that the image contains large textured areas. In the other three cases (cameraman, peppers, boat) solution-driven adaptive regularization can significantly improve the results with respect to contrast and image details. We currently are not aware of an criterion to distinguish these cases. Further investigating this issue is planned for future work. Comparing visually with the results from anisotropic second-order TV and BM3D, we see that solution-driven anisotropic TV cannot cope with the former ones, as we have already observed with the SSIM index.

### 8.1.2 Non-blind Deblurring

In this section we consider the application of non-blind deblurring. Here, the term *non-blind* refers to the assumption that the blur kernel is known beforehand and can be used in the restoration approach.

We generate artificial test data as follows. Let  $K(x)$  be a truncated Gaussian kernel,  $K(x) \in [0, 1]$ , such that the support  $\text{supp}(K)$  of  $K$  is bounded, i.e. there exist some open ball  $\mathring{B}_r(0)$  of radius  $r$  centered at 0 such that  $\text{supp}(K) \subseteq \mathring{B}_r(0)$ . Given a (undistorted) image represented by  $u_0 : \Omega \rightarrow \mathbb{R}$ , we obtain data  $f$  on the domain

$$\Omega_f := \{x \in \mathbb{R}^2 \mid \exists y \in \Omega \text{ such that } \|x - y\| \leq r\} \quad (8.3)$$

Method	cameraman	Lena	peppers	boat	mandrill
std. TV(d.d.)	0.8348	0.8349	0.8493	0.7664	0.6533
$TV^2$ (d.d.)	0.8015	0.8021	0.8173	0.6961	0.6253
TGV(d.d.)	0.8337	0.8367	0.8519	0.7629	0.6083
Schmidt(d.d.)	0.6123	0.6971	0.6823	0.7054	0.6140
weighted TV(d.d.)	0.8324	0.8376	0.8491	0.7667	0.6537
weighted TV(s.d.)	<b>0.8754</b> (4)	0.8602 (2)	0.8704 (4)	0.7152 (4)	0.6803 (4)
anisotr. TV(d.d.)	0.8336	0.8000	0.8406	0.7662	0.6528
anisotr. TV(s.d.)	0.8708 (4)	<b>0.8794</b> (4)	<b>0.8894</b> (3)	0.8203 (4)	<b>0.6788</b> (4)
double orient.(d.d.)	0.8324	0.8249	0.8382	0.7621	0.6506
double orient.(s.d.)	0.8718 (4)	0.8752 (4)	0.8827 (4)	<b>0.8260</b> (4)	0.6743 (4)
anisotr. $TV^2$ (d.d.)	0.7967	0.7956	0.8424	0.7638	0.6199

Table 8.3: *Deblurring*. Quantitative comparison based on the similarity between the deblurred and original images. For the solution-driven approaches we choose the number  $K$  (in brackets) of outer iterations out of  $\{2, 3, 4\}$ , for which we obtain the the highest SSIM index after parameter optimization. The solution-driven approaches in most cases outperform the data-driven as well as the state-of-the-art methods from the literature (with the only exception being weighted TV on the boat image, where the data-driven variant outperforms the solution-driven variant).

Method	cameraman	Lena	peppers	boat	mandrill
std. TV	26.013	27.506	28.354	26.463	24.6890
$TV^2$	24.793	26.286	27.211	25.283	23.1949
TGV	25.970	27.529	28.399	26.427	24.0133
Schmidt	24.045	24.972	26.095	25.688	24.0586
weighted TV(d.d.)	26.157	27.580	28.363	26.656	24.6958
weighted TV(s.d.)	<b>28.235</b> (4)	28.525 (2)	29.360 (4)	26.124 (4)	25.1403 (4)
anisotr. TV(d.d.)	25.937	26.277	28.103	26.456	24.6060
anisotr. TV(s.d.)	27.877 (4)	<b>29.348</b> (4)	<b>30.074</b> (4)	27.775 (4)	<b>25.1894</b> (4)
double orient.(d.d.)	25.940	27.312	28.146	26.422	24.6964
double orient.(s.d.)	27.903 (4)	29.236 (4)	30.043 (4)	<b>28.045</b> (4)	25.1147 (4)
anisotr. $TV^2$ (d.d.)	24.662	26.223	28.116	26.393	24.5884

Table 8.4: *Deblurring*. Quantitative comparison based on the PSNR between the deblurred and original images. For the solution-driven approaches we choose the number  $K$  (in brackets) of outer iterations out of  $\{2, 3, 4\}$ , for which we obtain the the highest SSIM index after parameter optimization. The first-order solution-driven approaches clearly outperform the data-driven as well as the state-of-the-art methods from the literature.

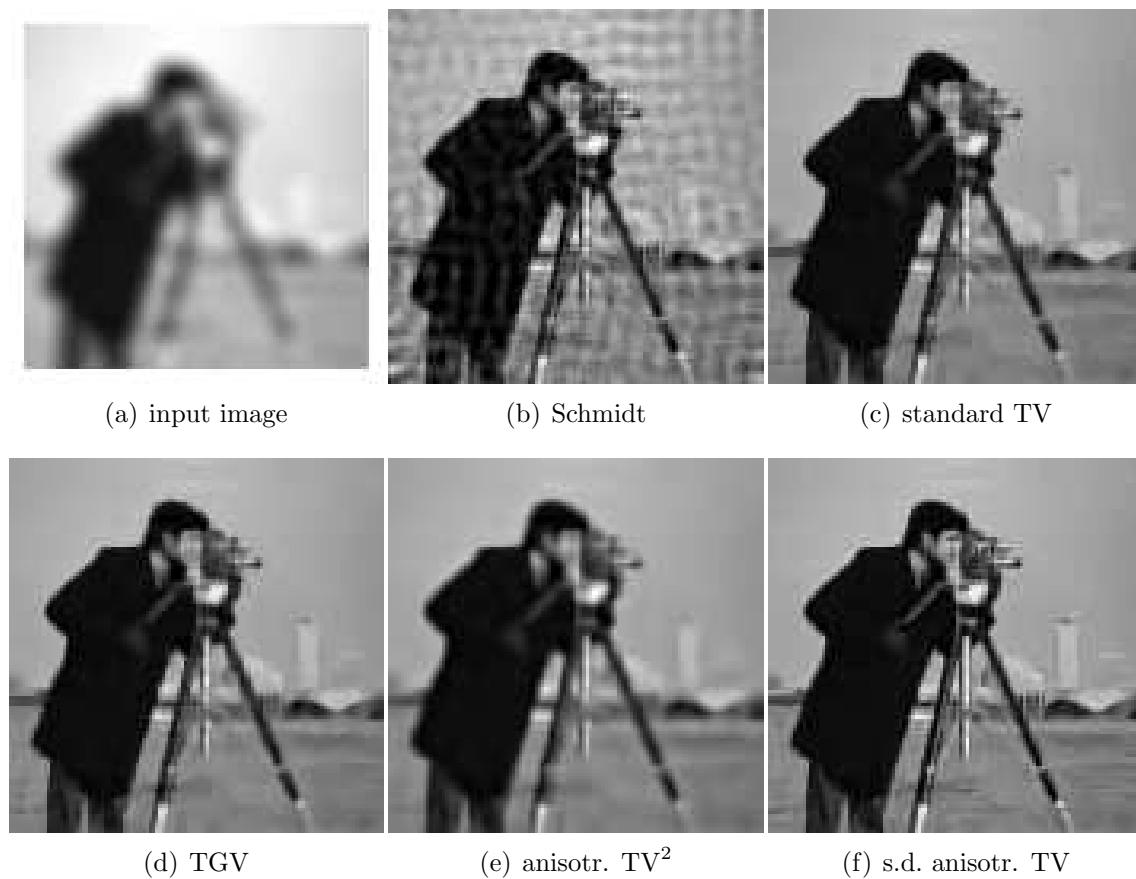


Figure 8.8: *Deblurring of the cameraman image.* Qualitative comparison of solution-driven anisotropic first-order TV (f) as the best performing method w.r.t. the SSIM index with the method by Schmidt et al., standard TV, TGV and anisotropic second-order TV (b)–(e). Visual inspection confirms the finding from the quantitative comparison that anisotropic first-order TV performs best.



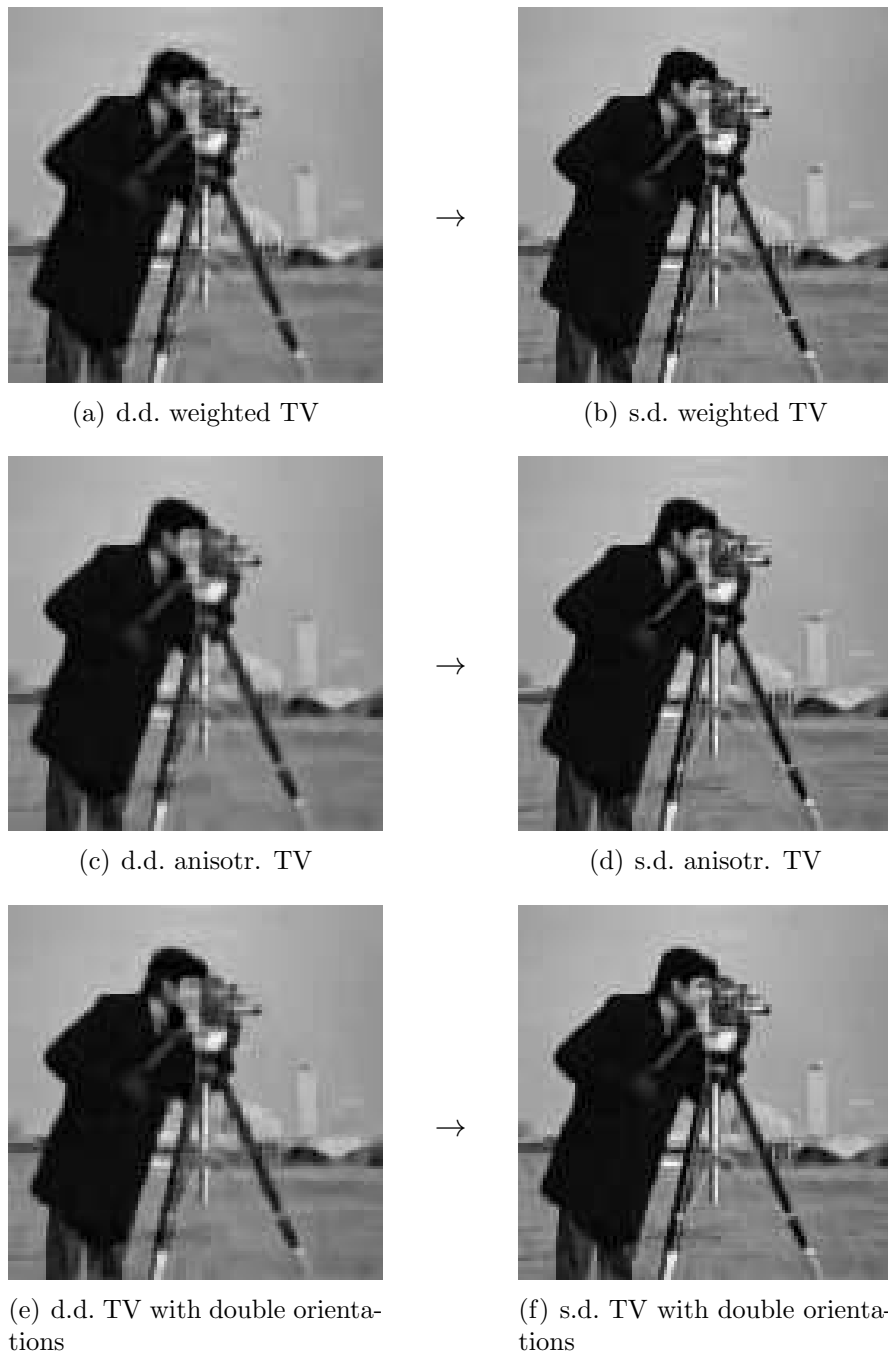


Figure 8.9: *Deblurring of the cameraman image.* Comparison of data-driven (left column) and solution-driven (right column) adaptive regularization methods. All three solution-driven approaches improve the results compared to the data-driven variants.

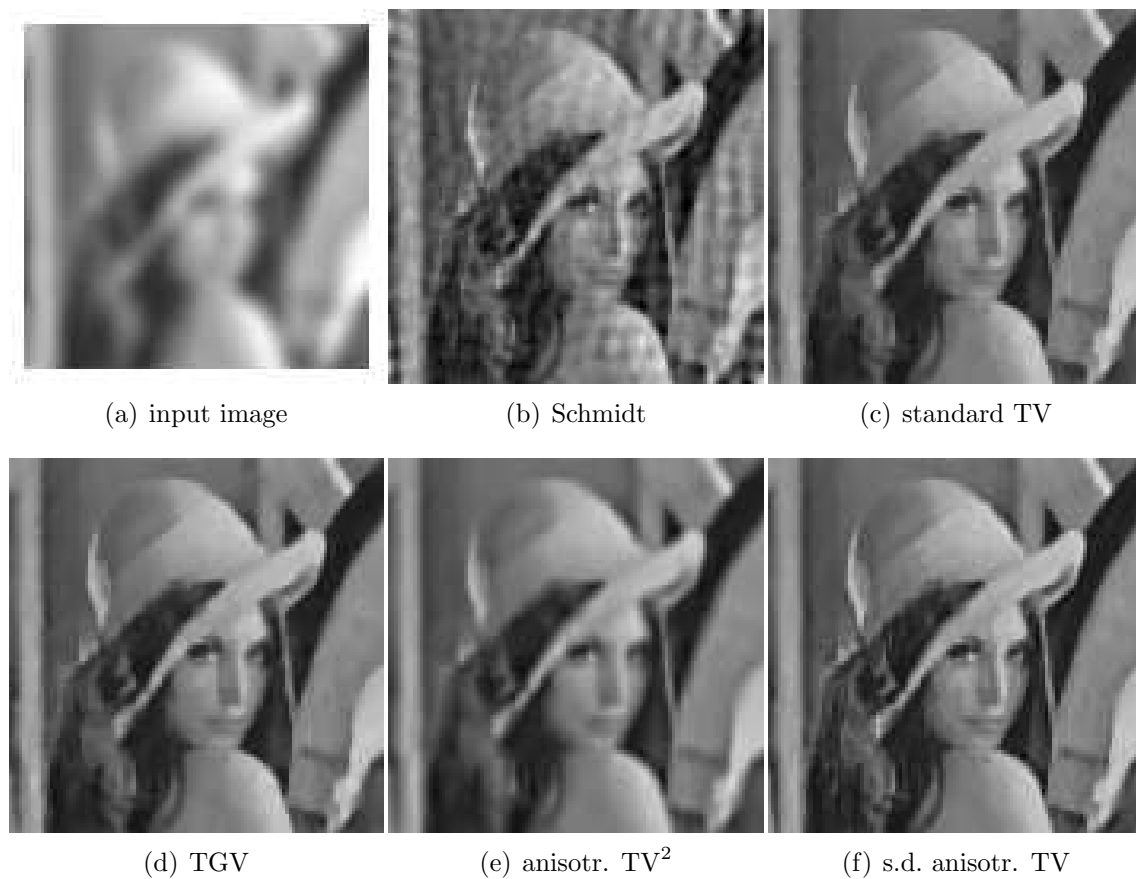


Figure 8.10: *Deblurring of the Lena image* Qualitative comparison of solution-driven anisotropic first-order TV (f) as the best performing method w.r.t. the SSIM index with the method by Schmidt et al., standard TV, TGV and anisotropic second-order TV (b)–(e). Visual inspection confirms the finding from the quantitative comparison that anisotropic first-order TV performs best.

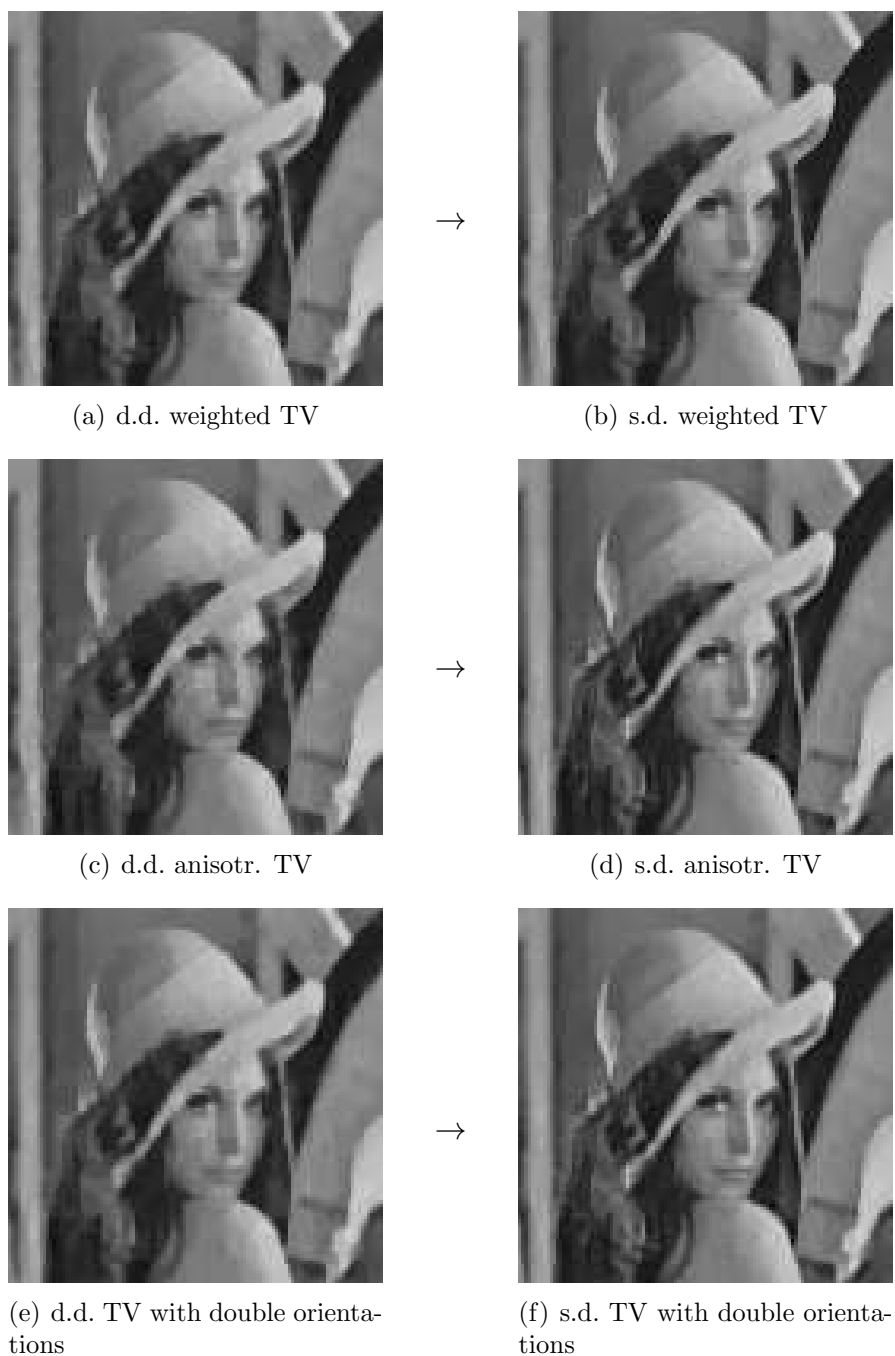


Figure 8.11: *Deblurring of the Lena image.* Comparison of data-driven (left column) and solution-driven (right column) adaptive regularization methods. All three solution-driven approaches improve the results compared to the data-driven variants.

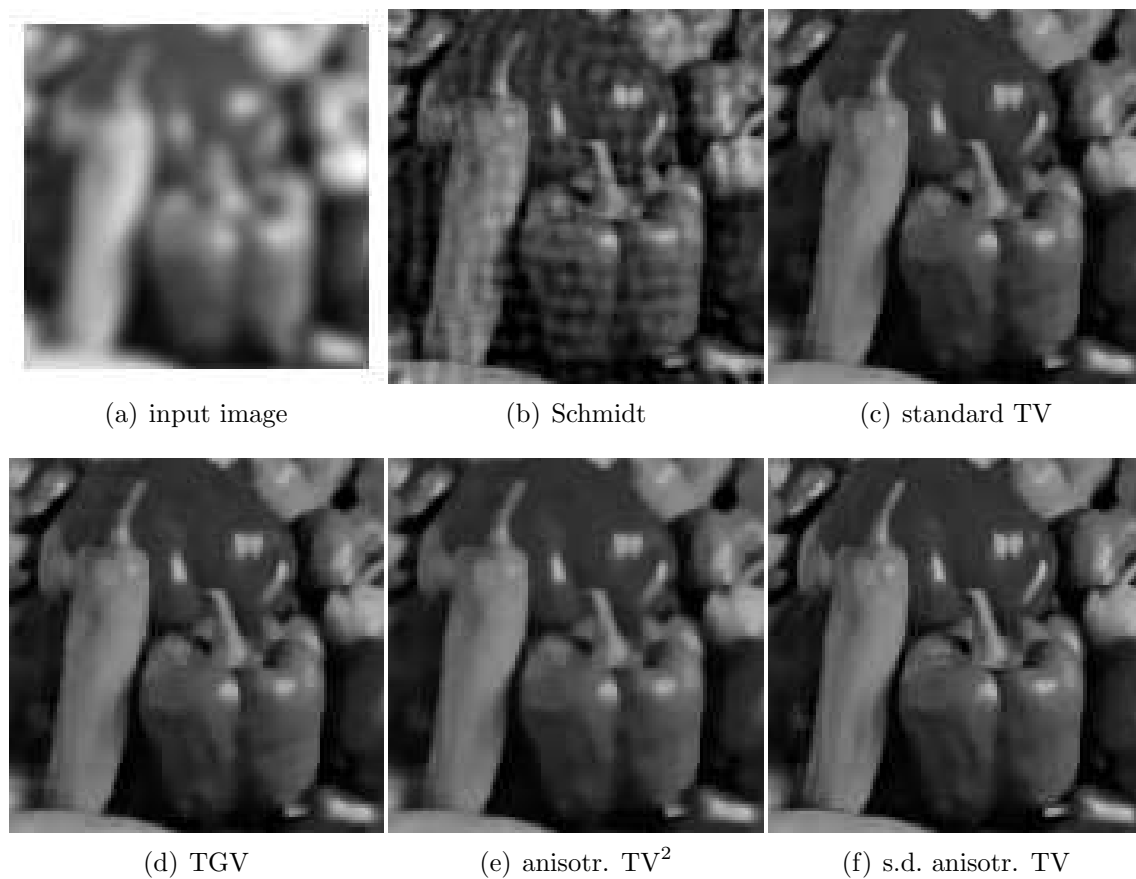


Figure 8.12: *Deblurring of the peppers image.* Qualitative comparison of solution-driven anisotropic first-order TV (f) as the best performing method w.r.t. the SSIM index with the method by Schmidt et al., standard TV, TGV and anisotropic second-order TV (b)–(e). Visual inspection confirms the finding from the quantitative comparison that anisotropic first-order TV performs best.

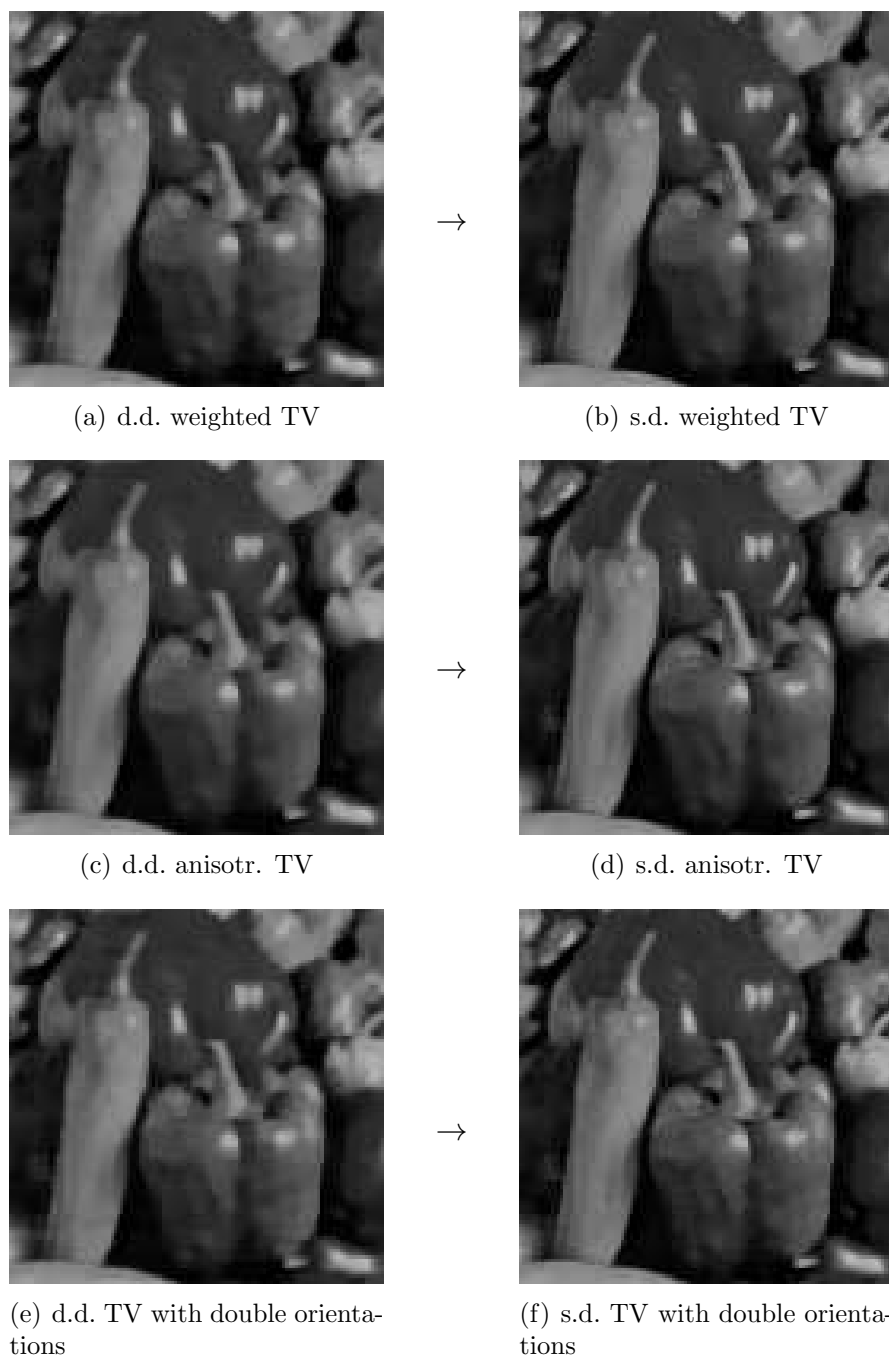


Figure 8.13: *Deblurring of the peppers image.* Comparison of data-driven (left column) and solution-driven (right column) adaptive regularization methods. All three solution-driven approaches improve the results compared to the data-driven variants.

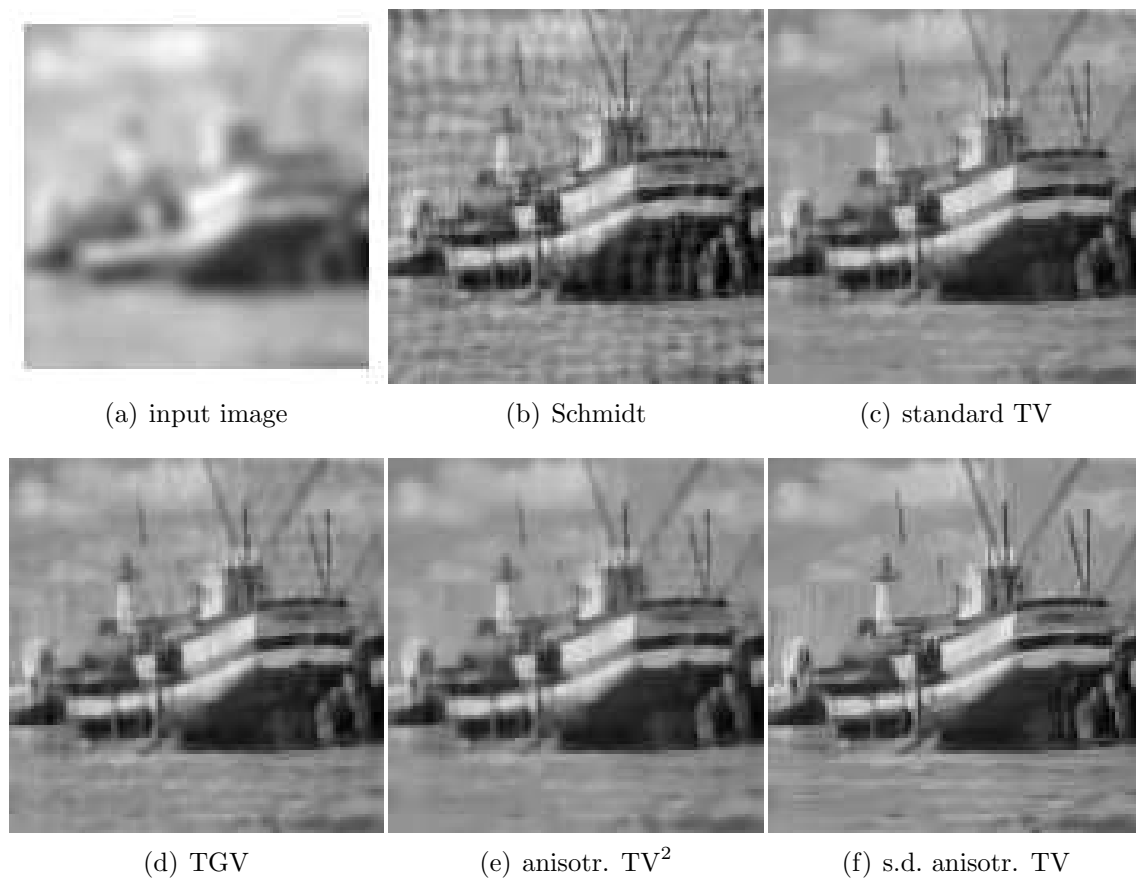


Figure 8.14: *Deblurring of the boat image.* Qualitative comparison of solution-driven anisotropic first-order TV (f) as the best performing method w.r.t. the SSIM index with the method by Schmidt et al., standard TV, TGV and anisotropic second-order TV (b)–(e). Visual inspection confirms the finding from the quantitative comparison that anisotropic first-order TV performs best.

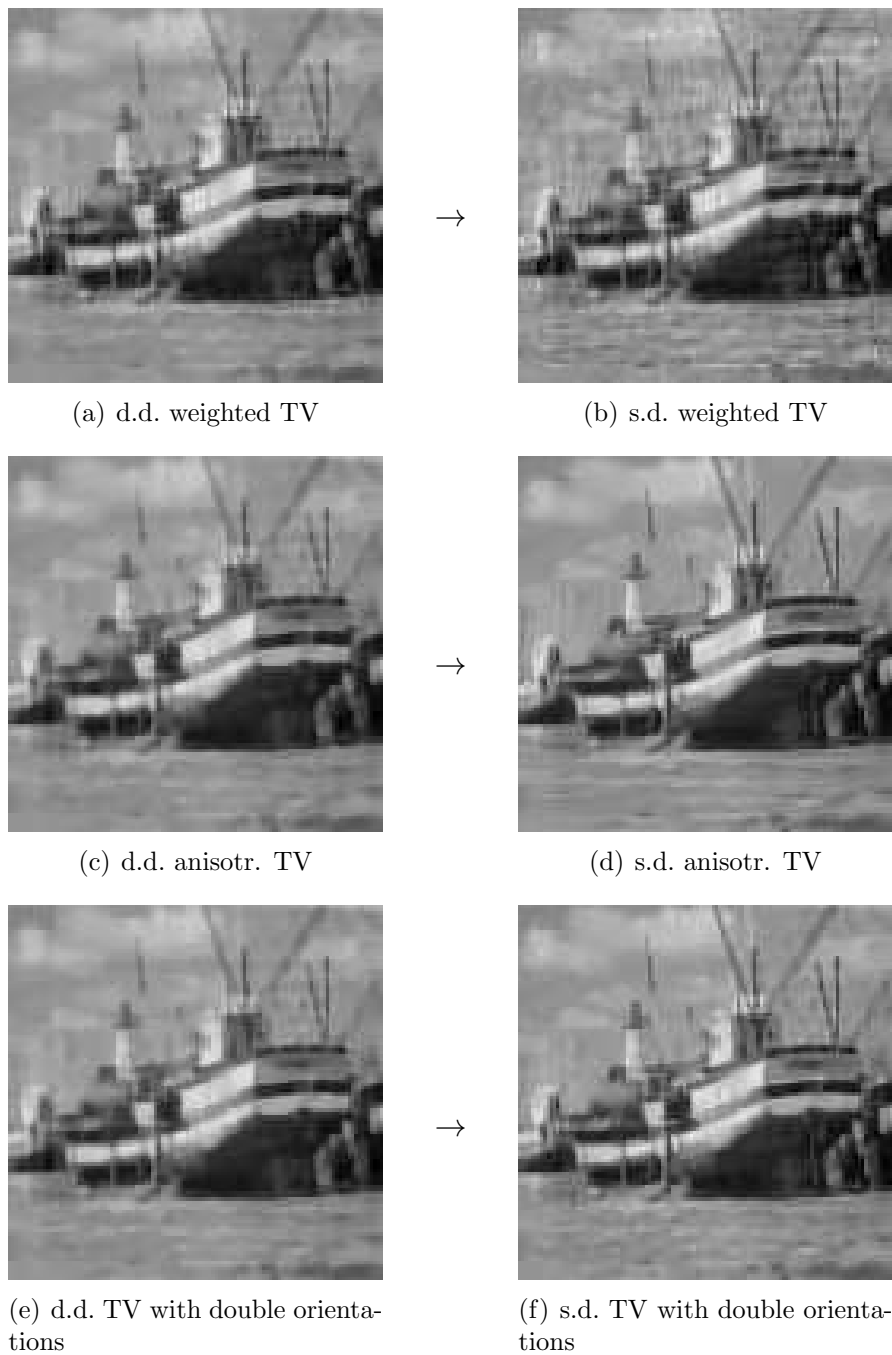


Figure 8.15: *Deblurring of the boat image.* Comparison of data-driven (left column) and solution-driven (right column) adaptive regularization methods. All three solution-driven approaches improve the results compared to the data-driven variants.

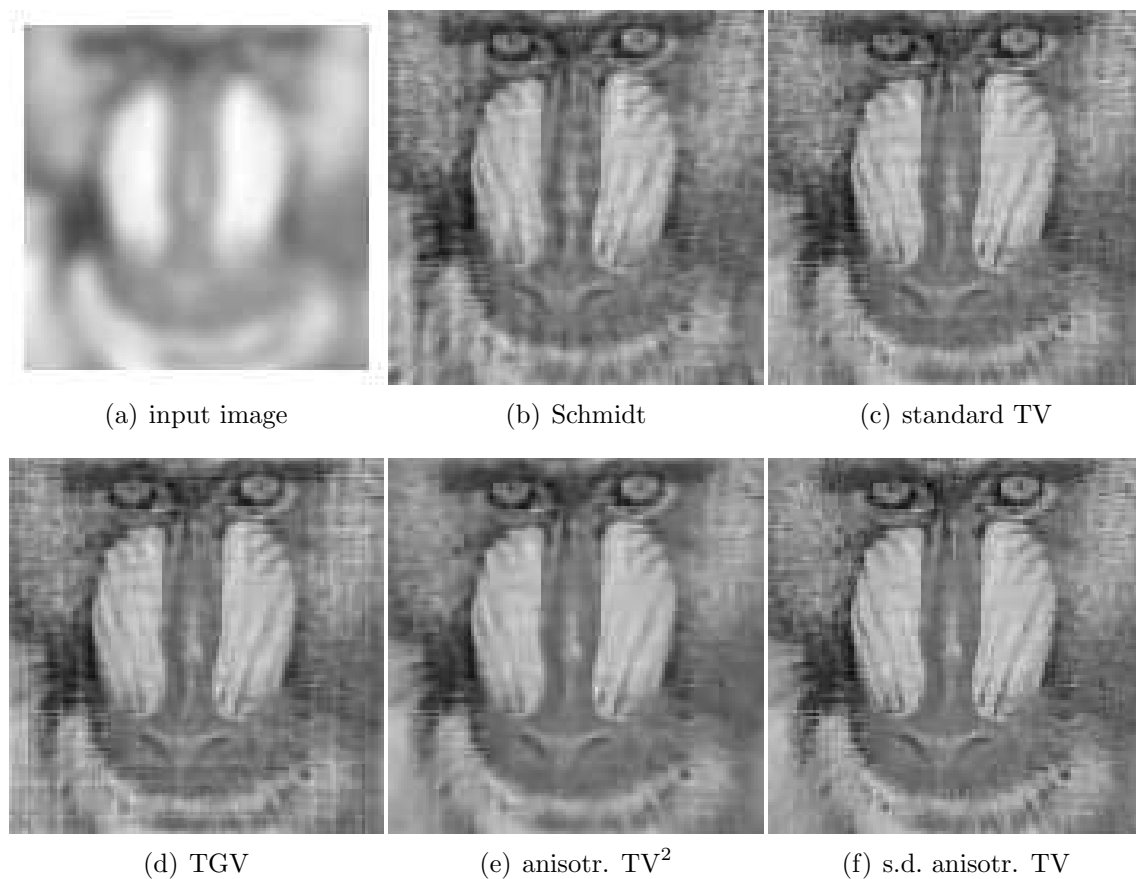


Figure 8.16: *Deblurring of the mandrill image*. Qualitative comparison of solution-driven anisotropic first-order TV (f) as the best performing method w.r.t. the SSIM index with the method by Schmidt et al., standard TV, TGV and anisotropic second-order TV (b)–(e). Visual inspection confirms the finding from the quantitative comparison that anisotropic first-order TV performs best, since it provides a much sharper results with less artifacts.



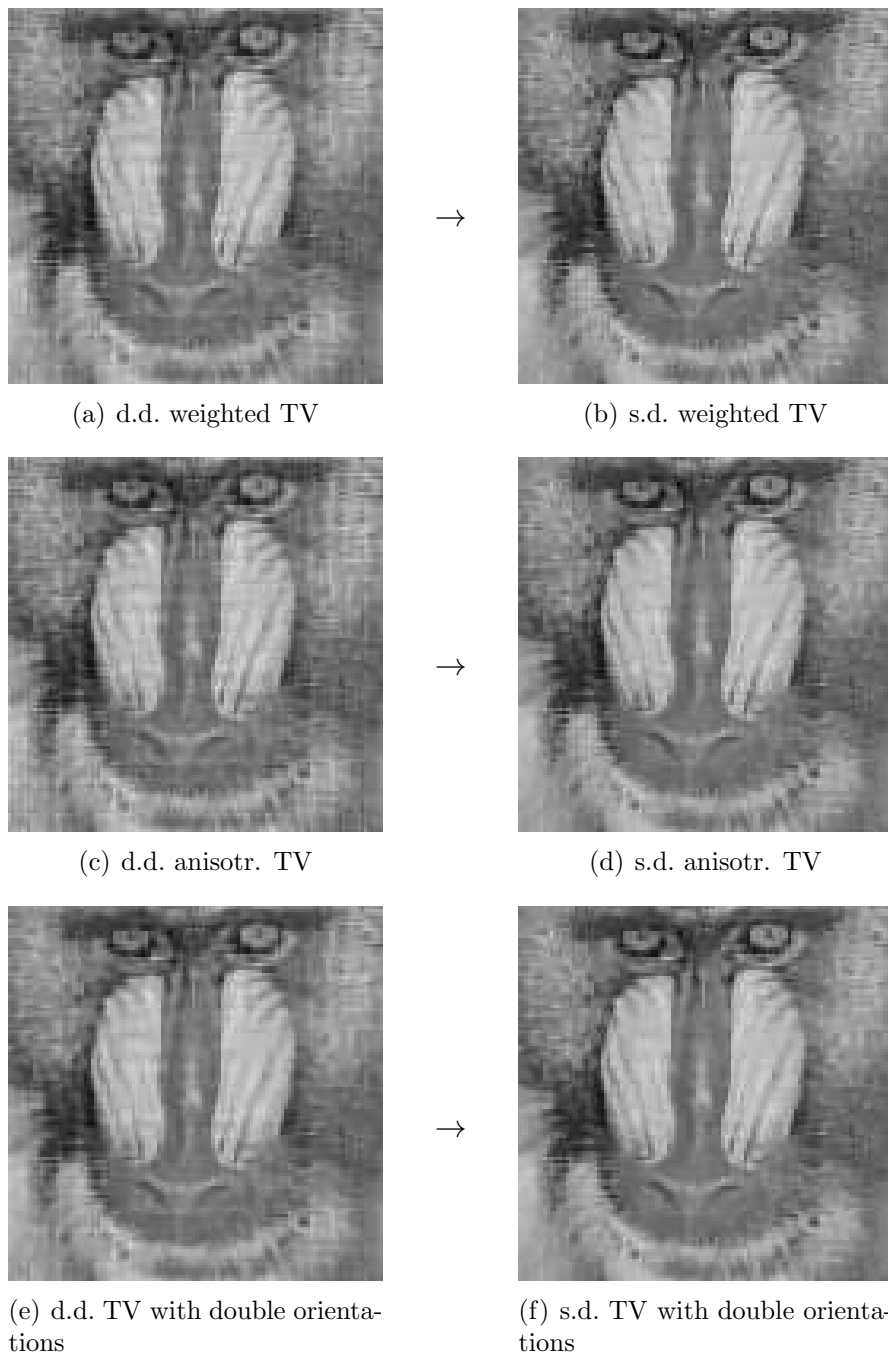


Figure 8.17: *Deblurring of the mandrill image.* Comparison of data-driven (left column) and solution-driven (right column) adaptive regularization methods. All three solution-driven approaches improve the results compared to the data-driven variants. In particular, much less artifacts from the deblurring occur.

given by

$$f(x) := \int_{B_r(0)} u_0(x-y)K(y) dy. \quad (8.4)$$

In addition, to show robustness of our method with respect to noise, we assume that  $f$  is affected by additive Gaussian noise with zero mean and standard deviation  $\sigma = 0.01$  (cf. the noise model in Section 2.2). Test data generated in this way are depicted in Fig. 8.1, third row.

The data term we consider is  $\frac{1}{2}\|Au - f\|_{L^2}^2$ , see Example 4.21. We observe that after discretization operator  $A$  is positive definite. For this reason, we omit an additional term  $\varepsilon\|u\|_2^2$ , which would be required in case of non-strict convex data terms (cf. Chapter 6). We obtain deblurring approaches by combining this data term with each of the regularizers mentioned above.

As additional state-of-the method, we evaluate the method proposed by Schmidt et. al. [151]. We use the original MATLAB source code for 'Discriminative Non-blind Deblurring' from the website <http://www.gris.tu-darmstadt.de/research/visinf/software/index.en.htm>. In this method, we use the pre-learned MRF model, which is provided in the software package.

For the variational approaches, as before, we perform an automated parameter selection, cf. the appendix below. For the method by Schmidt, there are no additional parameters to be optimized over.

Again, we compare the results of the method under consideration quantitatively, see Tables 8.3 (SSIM index) and 8.4 (PSNR) as well as qualitatively by visual inspection, see Figs. 8.10 to 8.17. We observe that both quantitatively and qualitatively the solution-driven approaches clearly outperform the state-of-the-art as well as the data-driven adaptive approaches. In case of the test images *Lena*, *peppers* and *mandrill* solution-driven *anisotropic* TV gives the best results, while for the *cameraman* image *weighted* TV and for the *boat* image TV with *double orientations* perform best.

### 8.1.3 Inpainting

In order to define an inpainting task, we provide a mask  $m : \Omega \rightarrow \{0, 1\}$ , which indicates regions of missing information by 0. In turn, the domain  $\Omega_f$  of valid information is given as  $\Omega_f := \{x \in \Omega \mid m(x) = 1\}$ . For our test images we consider regions of missing data, which regularly arranged squares of  $7 \times 7$  pixels. These regions are indicated by white color in Fig. 8.1, bottom row.

The data term we consider is given as

$$\mathcal{S}(u) = \varepsilon\|u\|_{L^2(\Omega \setminus \Omega_f)}^2 + \iota_{\mathcal{K}}(u), \quad (8.5)$$

where  $\varepsilon > 0$  is a small constant, so that the first term guarantees strict convexity of the optimization problem. The constraint set  $\mathcal{K}$  is chosen to be

$$\mathcal{K} := \{u \in BV(\Omega) \mid u = f \text{ a.e. on } \Omega_f\}, \quad (8.6)$$

Method	cameraman	Lena	peppers	boat	mandrill
std. TV(dd)	0.9693	0.9573	0.9642	0.9502	0.9317
$TV^2$ (dd)	0.9709	0.9646	0.9719	0.9625	0.9379
TGV(dd)	0.9689	0.9639	<b>0.9722</b>	0.9606	0.9349
Garcia(dd)	0.9660	0.9656	0.9693	0.9613	0.9346
weighted TV(dd)	0.9693	0.9573	0.9642	0.9502	0.9317
weighted TV(sd)	0.9694 (4)	0.9574 (2)	0.9643 (2)	0.9504 (2)	0.9336 (4)
aniso. TV(dd)	0.9680	0.9654	0.9695	0.9504	0.9317
aniso. TV(sd)	<b>0.9771</b> (3)	<b>0.9754</b> (2)	0.9712 (3)	<b>0.9648</b> (4)	<b>0.9445</b> (4)
double orient.(dd)	0.9651	0.9534	0.9615	0.9500	0.9219
double orient.(sd)	0.9730 (4)	0.9623 (4)	0.9672 (3)	0.9528 (3)	0.9363,(4)
aniso. $TV^2$ (dd)	0.9708	0.9660	0.9716	0.9635	0.9378

Table 8.5: *Inpainting*. Quantitative comparison based on the SSIM index. The ideal value is 1. We observe that in four of five cases, solution-driven anisotropic regularization outperforms the other approaches. On the *peppers* image, Total Generalized Variation (TGV) provides the best result. Moreover, we observe that second-order regularization often provides the second best result, indicating that higher-order is advantageous for the task of inpainting.

Method	cameraman	Lena	peppers	boat	mandrill
std. TV(dd)	29.528	30.079	32.952	29.660	29.7239
$TV^2$ (dd)	30.084	31.240	34.210	31.364	30.3503
TGV(dd)	30.033	31.326	<b>34.540</b>	31.302	29.8381
Garcia(dd)	29.212	31.457	33.719	31.473	29.8643
weighted TV(dd)	29.528	30.079	32.952	29.660	29.9084
weighted TV(sd)	29.741 (4)	30.084 (3)	32.953 (4)	29.663 (2)	30.0613 (4)
aniso. TV(dd)	29.712	31.223	34.144	29.457	29.9103
aniso. TV(sd)	<b>31.910</b> (3)	<b>33.024</b> (2)	34.470 (3)	<b>32.106</b> (4)	<b>31.0784</b> (4)
double orient.(dd)	28.743	29.460	32.911	29.846	28.8730
double orient.(sd)	31.279 (4)	31.132 (4)	33.694 (4)	30.334 (4)	30.3306 (4)
aniso. $TV^2$ (dd)	29.9890	31.405	34.092	31.499	30.3565

Table 8.6: *Inpainting*. Quantitative comparison based on the PSNR of the inpainted images. The ranking of the methods is similar to those base on the SSIM index. In four of five cases, solution-driven anisotropic regularization outperforms the other approaches.



(a) d.d. weighted TV

→



(b) s.d. weighted TV



(c) d.d. anisotr. TV

→



(d) s.d. anisotr. TV



(e) d.d. TV with double orientations

→



(f) s.d. TV with double orientations

Figure 8.18: *Inpainting of the cameraman image.* Comparison of data-driven (left column) and solution-driven (right column) approaches. Only the solution-driven approaches are able to transport geometric information into the regions to be inpainted. Solution-driven anisotropic TV provides the best result.

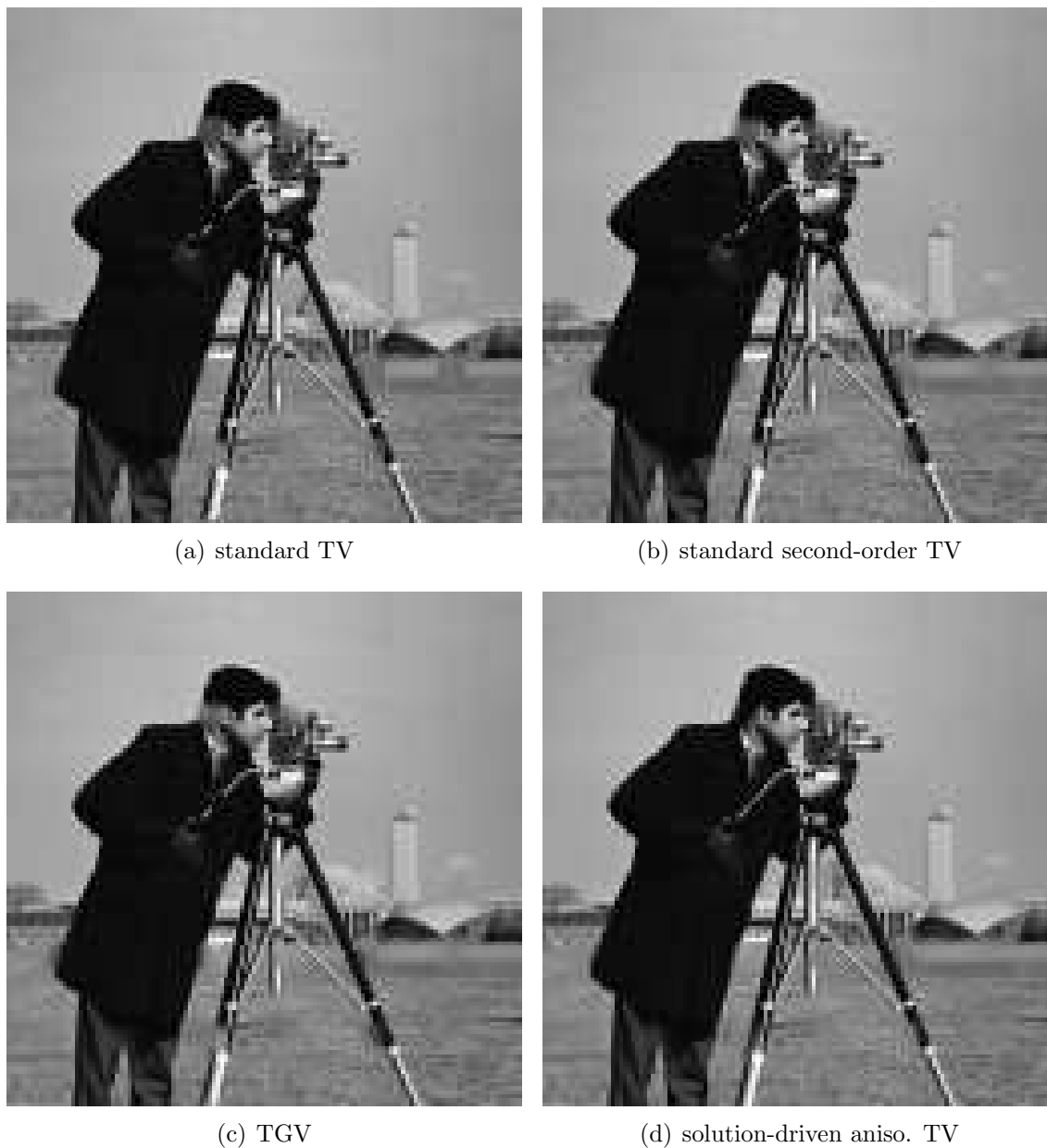


Figure 8.19: *Inpainting of the cameraman image.* Comparison of the solution-driven anisotropic TV (d) to the three best performing non-adaptive approaches, which are standard TV (a), second-order TV (b) and TGV (c). The solution-driven approach is much better able to transport geometric information into the areas of missing information (see e.g. neck or camera). Textures (see. e.g. the grass) are not well reconstructed by any method.



Figure 8.20: *Inpainting of the Lena image.* Comparison of solution-driven anisotropic TV (d) to standard second-order TV (a), TGV (b) and Garcia's method (c). The solution-driven approach is much better able to transport geometric information into the areas of missing information (see e.g. Lena's hair streaks). Textures (e.g. the stola) are not well reconstructed by any method.



(a) standard second-order TV



(b) TGV



(c) Gracia



(d) solution-driven aniso. TV

Figure 8.21: *Inpainting of the peppers image.* Comparison of solution-driven anisotropic TV (d) with standard second-order TV (a), TGV (b) and Garcia's method (c). For these results differences between the methods are not that obvious as for the other examples considered. Critical areas are e.g. along the silhouettes of the long small pepper and the stem of the pepper in the image center.



(a) standard second-order TV



(b) TGV



(c) Garcia



(d) solution-driven aniso. TV

Figure 8.22: *Inpainting of the boat image.* Comparison of solution-driven anisotropic TV (d) with standard second-order TV (a), TGV (b) and Gracia's method (c). The solution-driven approach is much better able to transport geometric information into the areas of missing information (see e.g. the masts).



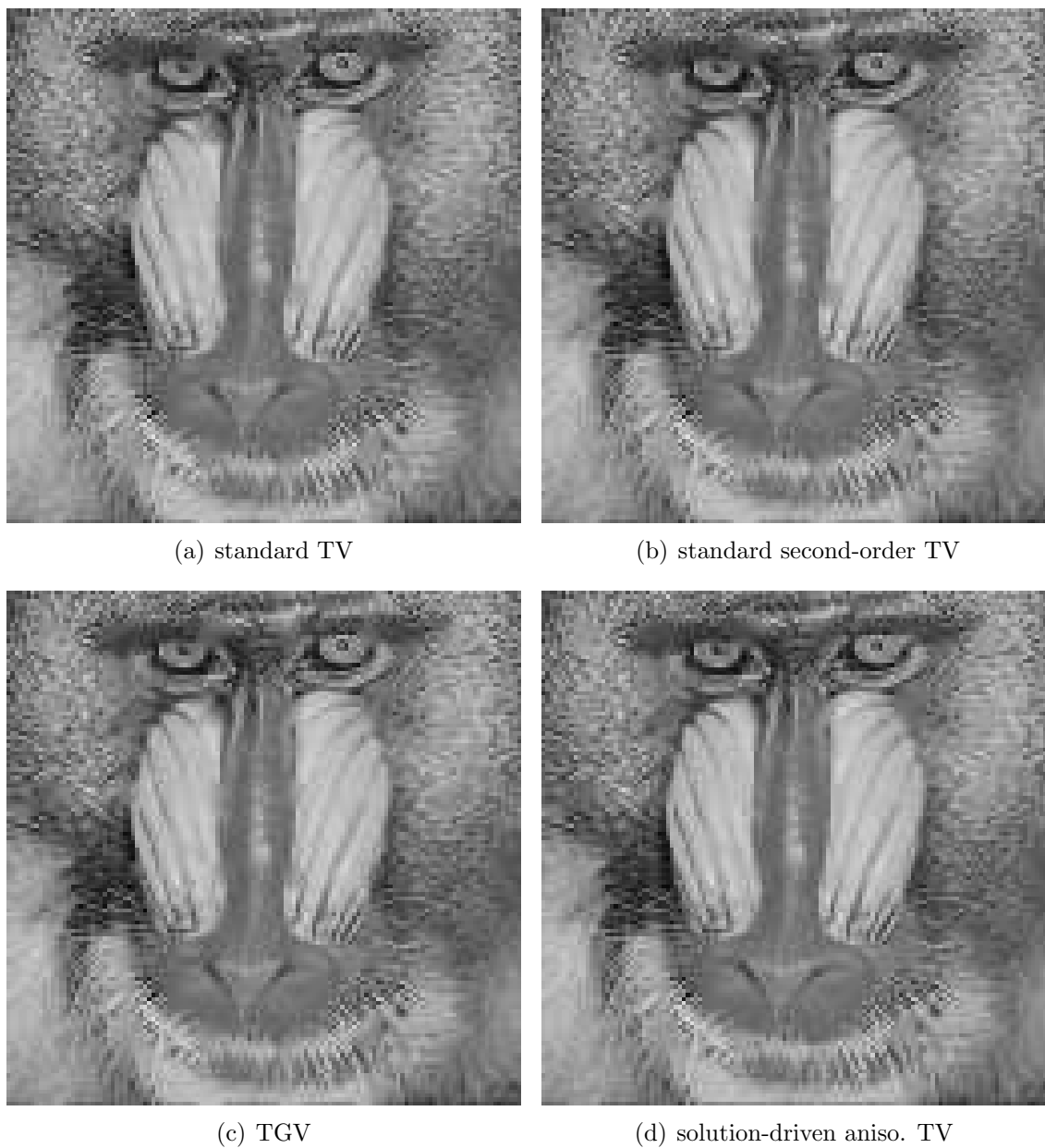


Figure 8.23: *Inpainting of the mandrill image.* Comparison of solution-driven anisotropic TV (d) to standard TV (a), standard second-order TV (b), and TGV (c). The solution-driven approach is much better able to transport geometric information into the areas of missing information (see e.g. the borders of the nose bridge). Textures (e.g. hairs) are not well reconstructed by any method.

where  $f : \Omega_f \rightarrow \mathbb{R}$  are the known image data. We combine this data term with each of the regularization terms listed above.

In the following we compare the adaptive regularization approaches to standard TV regularization, standard second-order TV and Total Generalized Variation (TGV) [28] and, in addition, to the inpainting method proposed by Garcia [61]. For the latter, we use the MATLAB source code available at the website <http://www.mathworks.com/matlabcentral/fileexchange/27994-inpaint-over-missing-data-in-1-d--2-d--3-d--n-d-arrays/content/inpaintn.m>.

As in the applications before, we evaluate the methods under consideration both quantitatively and qualitatively. From the qualitative comparison, see Tables 8.5 and 8.6, we observe that the *data-driven* approaches are not able to cope with the state-of-the-art methods. Concerning the *solution-driven* approaches, we observe that *weighted* TV does not provide a significant improvement compared to its data-driven variant. In contrast, *anisotropic* TV and TV with *double directions* show a clear improvement. In particular, solution-driven anisotropic TV clearly outperforms the other methods in four of five cases. The only exception is the peppers image, on which the second-order methods headed by TGV perform better.

Next, we focus on the qualitative evaluation. Let us first compare the data-driven with their solution-driven counterparts. For this purpose, we focus on the *cameraman* image. Fig. 8.18 shows the results of the adaptive methods. We observe that the solution-driven approaches different to the data-driven adaptive methods are able to transport geometric information into the regions to be inpainted, which explains their favorable performance. Solution-driven *anisotropic* TV provides the best result.

We now turn to a comparison of adaptive TV regularization to state-of-the-art methods. For the compactness of presentation, we focus on solution-driven *anisotropic* TV compared to the *three best performing* non-adaptive state-of-the-art methods. The corresponding results are depicted in Figs. 8.19 to 8.23. A visual inspections shows that solution-driven anisotropic TV outperforms the other approaches, which supports the findings from the quantitative evaluation. Finally, we remark that, as regularization approaches with local differential operators in general restore geometric (cartoon-like) data much better than textures, non of the considered regularization methods is able to sufficiently restore textured regions.

#### 8.1.4 Image Evolution by Solution-driven Adaptivity

In the previous section we compared non-adaptive, data-driven and solution-driven TV regularization approaches. For this comparison, we optimized the parameter setting for each approach separately. As a consequence, when directly comparing a data-driven approach to its solution-driven counterpart, the parameters sets in general differ not only by the number of outer iterations ( $K = 1$  for data-driven and  $K > 1$  for solution-driven), but potentially in all parameters. (Once again we stress that such an parameter optimization is required for a fair comparison.)

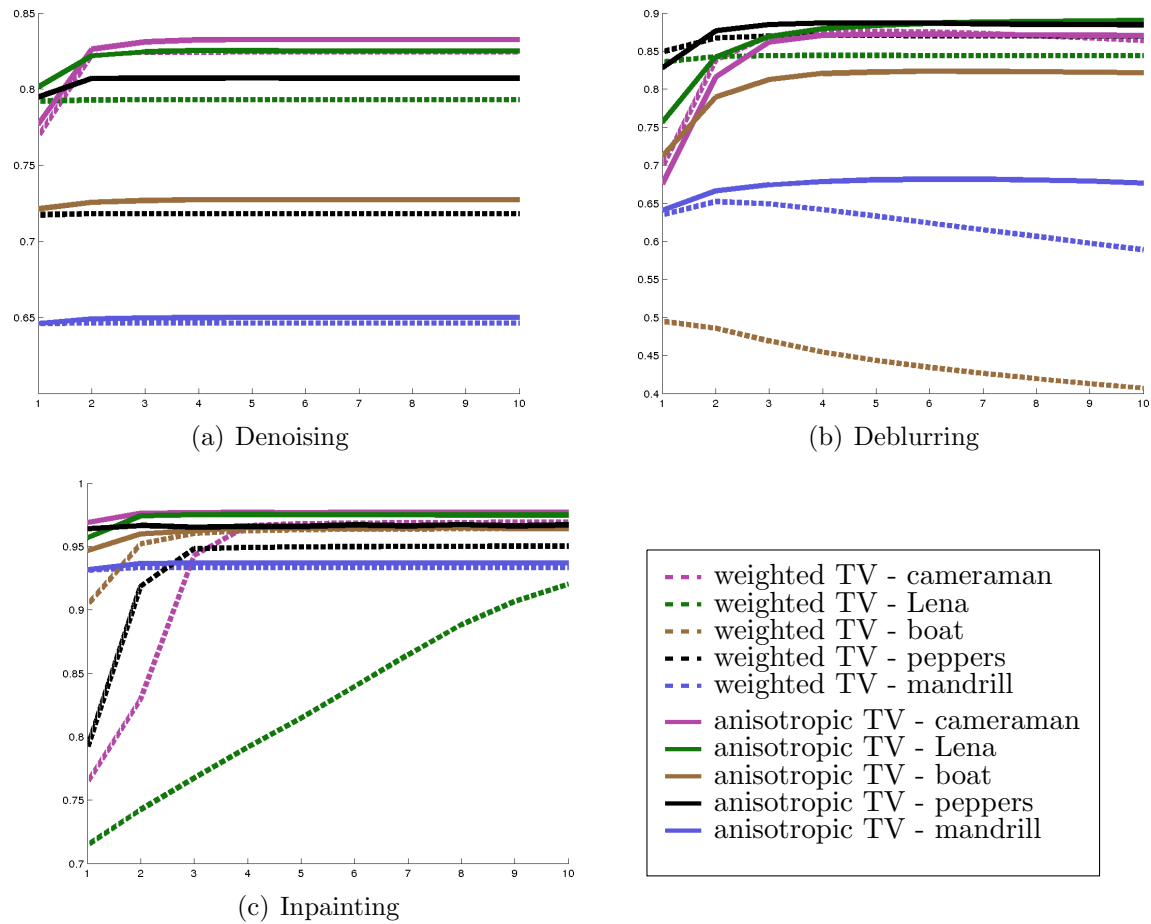


Figure 8.24: [Best viewed in color.] Evolution of the SSIM index during the outer iteration of solution-driven first-order *weighted* TV (dashed lines) and *anisotropic* TV (solid lines). In most cases we see an increase during the first three to four outer iterations. After five iterations usually the SSIM index becomes constant.



Figure 8.25: Inpainting. We depict the result of solution-driven anisotropic TV after  $k = 1, 2, 5$  and  $10$  outer iteration steps. We observe that geometric information is transported into the areas to be inpainted. Due to the nature of the regularization term, textures are less well reconstructed than cartoon-like image regions. The results typically stabilize during the first 10 iterations.

On the other hand, we can also investigate how the results evolve during the outer iteration, while the rest of the parameters is kept fixed. We refer to the sequence of results generated by the outer loop as the *evolution* of the solution, having in mind that we expect convergence to a fixed point.

Figure 8.24 shows the evolution of the SSIM index during the outer iteration of solution-driven adaptive regularization for the three different image restoration tasks. As parameters we have chosen the ones found by parameter optimization for the case of  $K = 4$  outer iterations. However, for the weighted TV approaches, in case that this parameter optimization led to  $\kappa = 0$  (i.e. the non-adaptive case), in order to demonstrate the effect of adaptivity, we consider  $\kappa = 0.1$  instead.

We observe a relatively smooth behavior of the SSIM index, which in most cases increases during the two to three outer iterations and then stays fairly constant. We remark that in general there is no guarantee of an improvement of the SSIM index with increasing  $k$ . We observed cases, where the optimal SSIM value already was reached after the first outer iteration, followed by a slight decrease afterwards. We currently are not aware of an criterion on the parameter selection to guarantee an improvement of the SSIM measure during the outer iteration. Answering this question is part of future work.

Finally, we want to discuss the visual changes during the outer iterations. For the applications of *denoising* and *deblurring*, these changes mainly occur between the first and second iteration. The change between the second and third iteration are less significant and the subsequent changes are even too weak to observe them visually. Typically after 5 iterations there is basically no change, which supports experimentally that the sequence converges to a fixed point. The overall change during the outer iterations is of the same quality as the difference of the results depicted in Figs. 8.2 and 8.3, (f) versus (g), (h) versus (i) and (j) versus (k), Figs. 8.4 to 8.7, (a) versus (b) and Figs. 8.9, 8.11, 8.13, 8.15 and 8.17, left versus right column. (Recall that those results different parameter settings where used for data-driven and solution-driven variants.) For this reason we omit a visual inspection of the evolution in the cases of denoising and deblurring.

For the task of *inpainting*, using anisotropic TV, we remark that significant changes take place not only between the first and second iteration, but also in the subsequent ones. We depict the corresponding results after  $K = 1, 2, 5$  and 10 outer iterations for exemplarily selected image regions in Fig. 8.25. It can be observed how geometric information is transported into the regions with missing data. The sequences typically stabilize during the first 10 outer iterations.

### 8.1.5 Relation to Non-Convex Regularization

Introducing adaptivity in TV regularization locally changes the way how the gradient (or higher derivatives) of the final solution is penalized. To gain insight into this effect, with a given solution  $u^*$ , one can study the empirical distribution of  $|\nabla u^*(x_i)|$  versus  $\langle \nabla u^*(x_i), p(x_i) \rangle$  (borrowing the  $\nabla$ -notation from the continuous setting), where  $p$  is the dual variable determined during the optimization process.

We do this exemplarily for the case of denoising the cameraman and Lena images and for weighted TV regularization, where  $\langle \nabla u^*(x_i), p(x_i) \rangle = \alpha_i \|\nabla u^*(x_i)\|$ .

Studying the distribution of the pairs  $(\|\nabla u^*(x_i)\|, \alpha_i \|\nabla u^*(x_i)\|)_i$ , which describes how the individual discrete gradients are penalized in the regularization term, see Fig. 8.26, one recognizes that the adaptive approaches mimic a non-convex regularizer. For the other test images, we observe similar distributions.

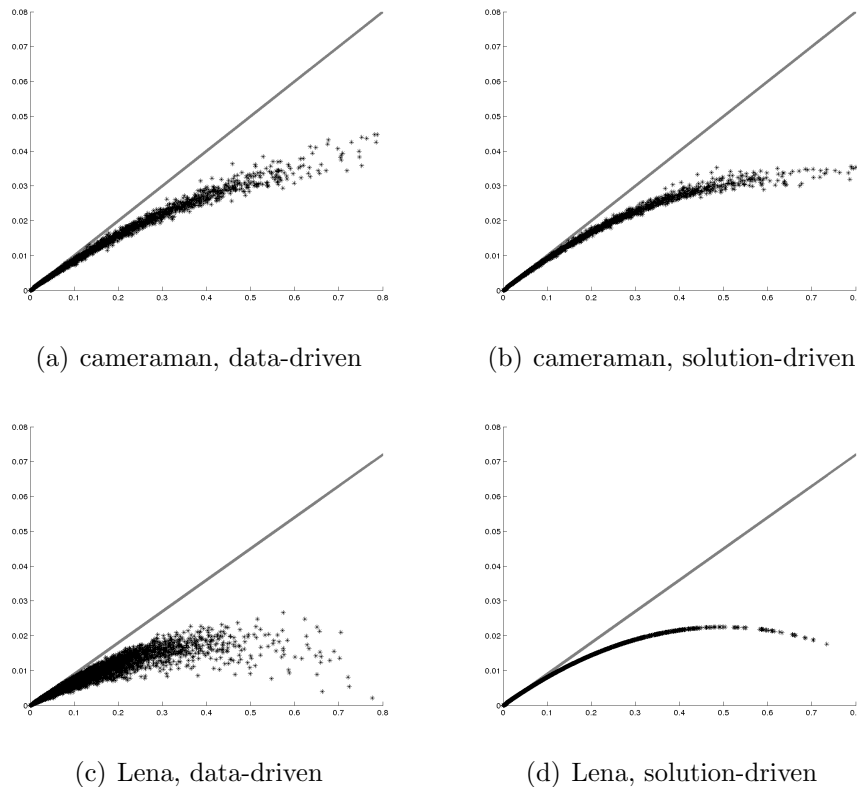


Figure 8.26: A-posteriori distribution of  $(\|\nabla u^*(x_i)\|, \alpha_i \|\nabla u^*(x_i)\|)_i$  (black dots) from the data-driven and solution-driven weighted TV regularization with  $\alpha_0 = 0.1$  and  $\kappa = 1$  when denoising the cameraman image and the Lena image, respectively. We compare these distributions to  $(\|\nabla u^*(x_i)\|, \alpha_0 \|\nabla u^*(x_i)\|)$  (gray line), where  $\alpha_0$  is the maximum value of the regularization strength. We observe that the weighted TV approach mimics non-convex regularizers. Increasing  $\kappa$  leads to more concave shapes.

### 8.1.6 Appendix: Parameter Selection

When aiming at a quantitative comparison of image restoration methods, it raises the question of how the parameters for each considered method should be chosen. In the following we discuss the approach we apply prior to such quantitative comparison in order to guarantee an objective parameter selection.



Figure 8.27: Influence of quality measure on automated parameter selection: We consider the applications of denoising, deblurring and inpainting on the cameraman image and focus on standard TV regularization. We optimize the parameter  $\alpha$  of the TV regularizer by a hierarchical grid search using the PSNR (left column) and the SSIM index (right column). Obviously the SSIM index is the more suitable quality measure for the task of finding optimal parameters.

Method	main	secondary	fixed
std. TV	$\alpha$		$N$
double orient.	$\alpha$	$\sigma, \rho$	$N$
weighted TV	$\alpha, \kappa$	$\sigma$	$N$
anisotr. TV	$\alpha_{min}, \alpha_{max}$	$s, \sigma, \rho$	$N$
TV <sup>2</sup>	$\alpha, \beta$		$N$
TGV	$\alpha_0, \alpha_1$		$N$
anisotr. TV <sup>2</sup>	$\alpha_{min}, \alpha_{max}, \gamma_{max}$	$s_1, s_2, \sigma, \rho_1, \rho_2$	$N, \gamma_{min} = \alpha_{min},$
BM3D	$\sigma$ ( noise level)		
Schmidt [151]	-	-	learned pw. MRF model
Garcia	-	-	$N$

Table 8.7: Parameter selection: for each considered method we perform a hierarchical grid search on the main parameter, followed by a coordinate descent method on all parameters except for the fixed ones. For each candidate set of parameters, we evaluate the SSIM index [171] between reconstructed and original image.

Only if two methods are closely related to each other, one may consider to use the same parameter set on each method to demonstrate differences in their behavior. The methods we consider, however, are too different in the nature of their parameters to follow this strategy. Instead and in view of practical applications, where the objective is to find the best image reconstruction possible, we consider an individual tuning of the parameters for each method.

Our strategy is based on the requirement that the original undistorted images ('ground truth') are available, so that we can utilize an arbitrary distance or similarity measure between reconstructed and original image to determine optimal parameters. Here, we focus on two commonly used quality measures, the *peak-signal-to-noise-ratio* (PSNR) and the mean SSIM index proposed by Wang et al. [171].

Since a manual tuning of the parameters is too tedious, an automated approach is preferable. One approach is a hierarchical grid search. In this approach a region of interest in the parameter space is overlay with a regular grid and the optimal parameter choice over all grid nodes is determined by an exhaustive search. After finding the optimal grid node, we overlay the neighborhood of this node by a finer grid and re-start the search for the optimal parameters on this finer grid. This process is iterated several times. We observe that the effort of this search grows exponentially with the dimension of the search space, i.e. the number of parameters to optimize over. Therefore a hierarchical grid search is tractable in reasonable time for up to three parameters.

An alternative to a hierarchical grid search is a coordinate descent approach. In this approach one performs a loop of all parameters and for each parameter  $i$  tries to optimize the output by changing the parameter by a small offset  $\pm h_i$ . If this try was successful, one continues with the same parameter. Otherwise one moves on to



the next parameter. In the situation that no further optimization is possible with the current vector  $h = (h_1, \dots, h_n)^\top$ , this vector is reduced in length and the loop is continued until  $h$  falls below a given threshold.

Our experiments show, that it is advantageous to select (up to) three main parameters, on which a hierarchical grid search is performed. and afterwards proceed with a coordinate descent approach on all parameters. (We refer to parameters where only coordinate search is applied to as *secondary* parameters.) As main parameters we choose those determining the regularization strength and for weighted TV regularization also the variable for the edge sensitivity. The number of iteration steps is excluded from the optimization, and instead set to a sufficiently large value (= 5000). The reason for this fixed large number is that we want to guaranteed convergence in any case. We refer to Table (8.7) for a detailed list of the main, secondary and fixed parameters.

On our parameter selection we remark the following.

1. Due to the fact that the quality measure on the results as a function of the parameters may have local minima, and that our parameter selection approach only tests a finite number of parameters, there is no guarantee that the found parameters are optimal. However, as a visual inspections shows, we can consider them as a good approximation. For the shortness of presentation we refer to those as the *optimal* parameters, referring to the optimality among the checked set of parameters.
2. Concerning the non-adaptive regularization approaches and the state-of-the-art methods from the literature, the main parameters actually comprise *all* parameters. Consequently, the parameter optimization by hierarchical grid search addresses all parameters.

It remains to choose the measure to compare reconstructed and original image, where as candidates we consider the PSNR (which in view of optimization is equivalent to using the mean-square error) and the mean SSIM index. To illustrate the difference, we consider standard TV regularization for the applications denoising, deblurring and inpainting on the cameraman image. To this end we optimize the parameter  $\alpha$  by means of a hierarchical grid search both with respect to PSNR and SSIM. The results obtain with the found parameters are depicted in Fig. 8.27. We observe by visual inspection that the parameters being optimal with respect to the PSNR lead to results with lower quality (noise artifacts, blurriness, inefficient inpainting) compared to parameters optimized with respect to the SSIM index. We therefore select the SSIM index for determining optimal parameters in our experiments.

## 8.2 Displacement Regularization

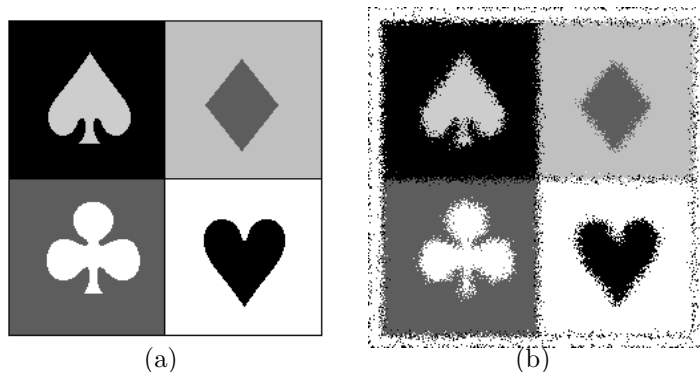


Figure 8.28: Test image *cards* used for the evaluation of displacement regularization. The original image (a) is distorted by artificial displacements errors (b).

A standard approach for retrieving discrete data  $(u_i)_i$  from a continuous signal  $u$  is the process of *sampling*, in which the signal is evaluated at pre-defined sampling points  $(x_i)_i \in \Omega$  for some domain  $\Omega \subset \mathbb{R}^d$ :

$$u_i := u(x_i). \quad (8.7)$$

A commonly used noise model for this process is to assume noise on the sampled data  $u_i$ . In the following we consider a different noise model, where the noise is affecting the position of the sampling points  $x_i$  instead of the data  $u_i$ .

To this end, let  $u : \mathbb{R}^d \rightarrow \mathbb{R}^m$  be a function representing *continuous* scalar- or vector-valued data on the open bounded domain  $\Omega \subset \mathbb{R}^d$ .

We study the following problem: Assume that function  $u(x)$  has to be sampled at positions  $(x_i)_{i=1}^N \subset \Omega$ , and that during the sampling a certain error in the location of the sampling point is induced, i.e.  $u$  is sampled at position  $\Phi(x_i)$  instead of  $x_i$ , where  $\Phi : \Omega \rightarrow \mathbb{R}^d$ . Please note that, despite the fact, that a sampling process is typically understood as a process of collecting data at *discrete* positions, we prefer to stay with a continuous formulation of the problem. In the following we denote the sampled function by

$$\begin{aligned} u^0 : \Omega &\rightarrow \mathbb{R}^m, \\ x &\mapsto u^0(x) := u(\Phi(x)). \end{aligned} \quad (8.8)$$

We refer to  $\phi(x) - x$  as a *sampling error* or *displacement error*. Fig. 8.28 shows an example of an image with sampling errors.

Given data  $u^0$  corrupted with displacement errors, we consider the task of reconstructing  $u$  from  $u^0$ . To this end we assume that  $u$  is a smooth function, so that

we can make a first order Taylor series expansion on  $u^0(x)$ ,

$$\begin{aligned} u^{(0)}(x) &= (u \circ \Phi)(x) = u(x + (\Phi(x) - x)) \\ &\approx u(x) + \nabla u^\top(x) (\Phi(x) - x), \end{aligned} \quad (8.9)$$

where

$$\nabla u = \begin{pmatrix} \partial_1 u_1 & \partial_1 u_2 & \cdots & \partial_1 u_m \\ \partial_2 u_1 & \partial_2 u_2 & \cdots & \partial_2 u_m \\ \vdots & \vdots & \ddots & \vdots \\ \partial_d u_1 & \partial_d u_2 & \cdots & \partial_d u_m \end{pmatrix} \quad (8.10)$$

is the Jacobian of  $u$  and  $\approx$  in (8.9) symbolizes that the left hand side approximates the right hand side for small displacements  $\Phi - \text{Id}$ . In the following, we neglect the approximation error in (8.9), so that equality holds.

We remark that in the case of scalar data. i.e. for  $m = 1$ , (8.9) approximates the standard *optical flow* equation:

$$\frac{\partial u}{\partial t}(x, t) + (\nabla u)^\top(x, t)(\Phi(x, t) - x) = 0. \quad (8.11)$$

The relation is immediate if one considers  $\Phi - \text{Id}$  as the optical flow and  $u - u^{(0)}$  as a time discretization of a time dependent function  $u(t)$ . The optical flow problem, however, consists in the task of finding  $\Phi$  for a given pair  $u, u^0$ , while we consider restoring an unknown  $u$  from a given  $u^0$ . In the scalar case the system is *underdetermined*. However, it is solvable at positions, where  $\nabla u$  does not degenerate. For vector-valued functions instead, the system becomes *overdetermined* in the case  $m > d$  and a sufficiently large column rank of  $\nabla u$ . Then, to determine  $\Phi$ , the least squares problem

$$\|\nabla u^\top(x)(\Phi(x) - x) - u^{(0)}(x) + u(x)\|_2^2 \quad (8.12)$$

has to be considered. A minimizer of (8.12) is given by

$$\Phi(x) - x = (\nabla u^\top(x))^\dagger (u^{(0)}(x) - u(x)), \quad (8.13)$$

where  $(\nabla u^\top(x))^\dagger$  denotes the *Moore–Penrose* or *pseudo-inverse* (see [121]) of  $\nabla u^\top(x)$ .

The above ansatz to use the pseudo-inverse to estimate a displacement  $\Phi$  from a given pair of images, provides already a simplistic solution to the optical flow problem. However, this approach assumes small displacements and is only capable to capture components of the displacement orthogonal to the null space of  $\nabla u$ . Another issue is, that it is very sensitive to noise.

Let us now consider the problem of finding the pairs  $(u(x), \Phi(x))$  of minimal energy satisfying the constraint  $u(\Phi(x)) = u^0(x)$  for given  $u^0$ .

For solving this problem approximately, we propose a variational method consisting in minimization of the functional

$$\frac{1}{2} \int_{\Omega} \|\Phi(x) - x\|_2^2 d\mathcal{L} + \mathcal{R}(u) \quad (8.14)$$

over the set of functions satisfying  $u^{(0)} = u \circ \Phi$ . The regularization term  $\mathcal{R}(u)$  is defined as

$$\mathcal{R}(u) := \alpha \int_{\Omega} \|\nabla u(x)\|_F^p dx, \quad (8.15)$$

where  $\alpha > 0$ ,  $\|\cdot\|_F$  is the Frobenius norm and  $p$  is some positive integer.

In the following, for notational convenience, we omit parameter  $x$  indicating the spatial dependency of functions and write e.g.  $u$  instead of  $u(x)$ . Inserting (8.13) into (8.14), we obtain

$$\begin{aligned} \mathcal{F}_{u^{(0)}}(u) &:= \frac{1}{2} \int_{\Omega} (u - u^{(0)})^\top (\nabla u^\top \nabla u)^\dagger (u - u^{(0)}) d\mathcal{L} + \mathcal{R}(u) \\ &= \frac{1}{2} \int_{\Omega} (u - u^{(0)})^\top A^\top(u) A(u) (u - u^{(0)}) d\mathcal{L} + \mathcal{R}(u), \end{aligned} \quad (8.16)$$

where we introduced the abbreviation  $A(u) := (\nabla u^\top)^\dagger$ .

**Remark 8.1.** *We remark that the weighting of the data term by a factor  $1/\|\nabla u(x)\|^p$ , which introduces non-convexity of the functional, can also be modeled by means of the strategies discussed in Chapters 4 and 6, i.e. as a data-driven weighting of the data term with weight  $1/\|\nabla u^0(x)\|^p$ , where (assuming  $f \approx u$ ) or by a solution-driven fixed point approach. We will discuss and compare the different strategies in the experimental part below.*

### Scalar Data ( $m = 1$ )

In the scalar case, we can utilize the fact that an analytical form of the pseudo-inverse  $A(u)$  of  $(\nabla u)^\top$  can be provided as

$$A(u) = \frac{1}{\|\nabla u\|_2^2} \begin{pmatrix} \partial_x u \\ \partial_y u \end{pmatrix}. \quad (8.17)$$

Using (8.17), functional (8.16) becomes

$$\mathcal{F}_{u^{(0)}}(u) := \frac{1}{2} \int_{\Omega} \frac{(u - u^{(0)})^2}{\|\nabla u\|_2^2} + \alpha \|\nabla u\|_2^p d\mathcal{L}. \quad (8.18)$$

Note that for scalar data  $u$ , the Frobenius norm of  $\nabla u$  equals the Euclidean norm. In [97] we have provided a motivation for generalizing this functional to

$$\mathcal{F}_{u^{(0)}}(u) := \frac{1}{2} \int_{\Omega} \frac{(u - u^{(0)})^2}{\|\nabla u\|_2^p} + \alpha \|\nabla u\|_2^p d\mathcal{L} \quad (8.19)$$

for arbitrary integer  $p > 0$ , defined on the function space  $L^2(\Omega) \cap W^{1,p}(\Omega)$ . Please note that the functional (8.19) is allowed to take the value  $+\infty$  and moreover, that the functional (8.19) takes the form

$$\mathcal{F}_{u^{(0)}}(u) := \int_{\Omega} F(x, u(x), \nabla u(x)) dx, \quad (8.20)$$

where  $F(x, \xi, \nu)$  is non-convex with respect to  $\nu$ . Recalling the discussion of functionals of this form in Section 5.1, we observe that  $\mathcal{F}_{u^{(0)}}(u)$  is not weakly lower semi-continuous. We therefore can only guarantee existence of generalized minimizers, i.e. minimizers of the relaxed functional  $\mathcal{F}_\diamond$  ( $p = 2$ ) or  $\mathcal{F}_{\diamond BV}(u)$  ( $p = 1$ ) (cf. Chapter 5), respectively. We refer to  $\mathcal{F}_\diamond$  as the non-convex quadratic (NCQ) functional and to  $\mathcal{F}_{\diamond BV}(u)$  as the non-convex bounded-variation (NCBV) functional. We propose to use these relaxed functionals to retrieve variational methods to remove displacement errors in scalar data. For the numerical treatment of both we refer to [97].

We remark, that the non-convex weight introduced in (8.19) could also be treated in a fixed point approach as considered in Chapter 6: find a fixed point of

$$u_0 \rightarrow \arg \min_u \mathcal{F}(u; u_0) := \arg \min_u \int_{\Omega} \frac{\|u(x) - f(x)\|^2}{\|\nabla(u_0)_\sigma(x)\|_2^p + \varepsilon^p} dx + \alpha \mathcal{R}(u), \quad (8.21)$$

for  $p \in \{1, 2\}$  and

$$\mathcal{R}(u) = \begin{cases} TV(u) & \text{for } p = 1, \\ \int_{\Omega} \|\nabla u\|^2 d\mathcal{L} & \text{for } p = 2. \end{cases} \quad (8.22)$$

Please note that in (8.21) we introduced some regularization in the denominator to avoid division by zero. In addition, if  $p = 1$ , we require a pre-smoothing of  $u_0$  denoted by  $(u_0)_\sigma$  to have  $\nabla u_0(x)$  well-defined.

However, our experiments show that such an approach for the proposed weighting of the data term does not lead to satisfactory results, since the reconstruction quality decreases during the fixed point iteration. Consequently, stopping after the first iteration is advisable, i.e. it becomes a data-driven approach and relates to the minimization problem

$$\min_u \int_{\Omega} \frac{\|u(x) - f(x)\|^2}{\|\nabla f(x)\|_2^p + \varepsilon^p} dx + \alpha \mathcal{R}(u). \quad (8.23)$$

We refer to this approach as weighted  $L^2 - L^p$  and take it into account for comparison with NCQ and NCBV.

Let us now experimentally compare the NCQ and NCBV approaches to approaches with standard data  $L^2$ -term combined with TV (ROF) or TGV regularization and the weighted  $L^2 - L^p$  approach, respectively.

To this end we consider six test images, the *cards* image introduced in Fig. 8.28 and the five images depicted in Fig. 8.29. These images are distorted by artificial displacement errors. Analogously to our evaluation in Section 8.1, we apply an automated parameter selection to determine approximately optimal parameters for each method with respect to the SSIM measure [171]. Table 8.8 shows the resulting SSIM values. We observe that the NCBV approach outperforms the other methods in four of five cases. Only for the *cards* image the situation is different, which might be due to the cartoon nature of the image. Here, the weighted  $L^2 - L^2$  approach provides the best result. However, we will see that the visual impression of the corresponding result is not that satisfactory.



Figure 8.29: Further test images with artificially added displacement errors used for the evaluation of displacement regularization.

For the sake of compactness, we depict the corresponding results only for the *cards* and *cameraman* image, see Figs. 8.30 and 8.31, respectively. We observe that the visual quality of the results, except for the *cards* image, is in agreement to the retrieved SSIM values, i.e. the proposed approaches outperform the approaches, in which the standard  $L^2$  data term is combined with TV/ TGV regularization as well as the weighted  $L^2 - L^p$  approaches.

**Remark 8.2.** *In the above considerations we have assumed small displacement errors. Studying the example in Fig. 8.28 in detail, shows that sampling errors are dominant in the edge regions of images. We distinguish between small displacements, which lead to noisy contours in the image, and large displacements, by which gray values are shifted into different regions of the image. In the latter case, the displacement error leads to isolated noise peaks. We observe that even in this case, despite the violation of our assumption our approach provides satisfactory results, cf. Fig. 8.30.*

### Vector-Valued Data ( $m > 1$ )

In the case of vector-valued data, we face the problem that no analytical solution of the pseudo-inverse is available. In order to avoid the numerical computation of the pseudo-inverse, we additionally regularize the possibly singular matrix  $(\nabla u)^\top \nabla u$  by the regular, symmetric, and strictly positive definite matrix  $((\nabla u)^\top \nabla u + \varepsilon I)$  with some  $\varepsilon > 0$ . Therefore, in the sequel, we consider minimizing of the regularized

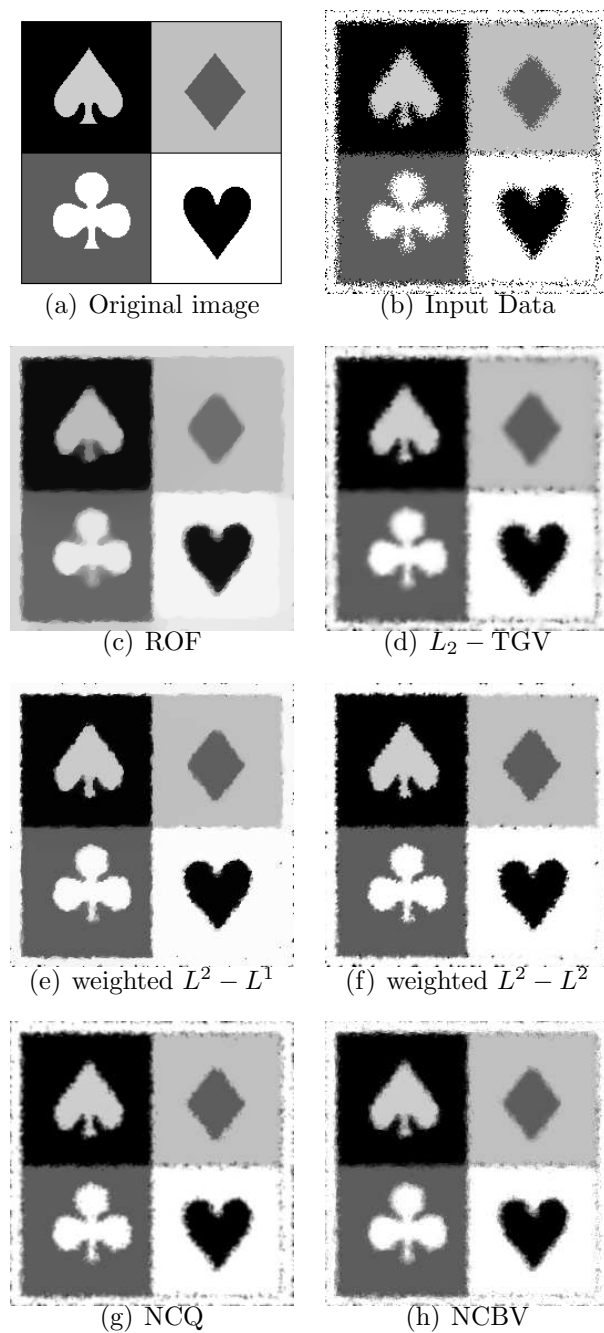


Figure 8.30: Results for smoothing the *cards* with displacement errors. The weighted  $L^2 - L^1$  approach provides significantly sharper results than the competing approaches.

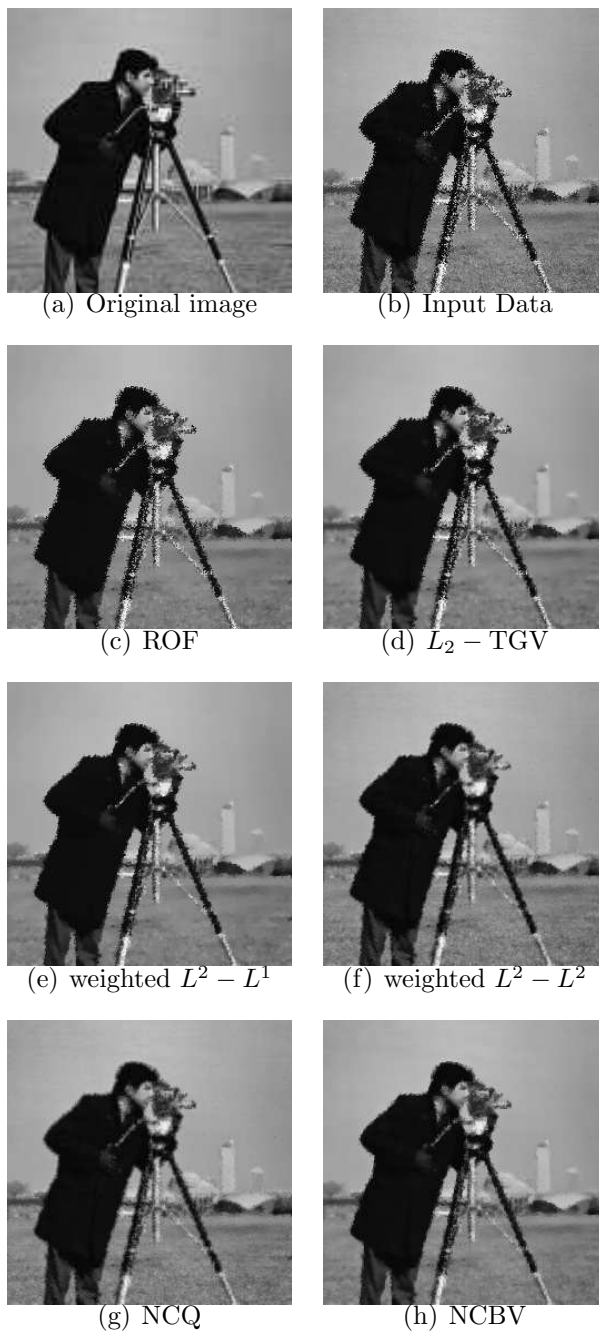


Figure 8.31: Results for smoothing the *cameraman* with displacement errors. In contrast to the standard approaches with  $L^2$  data term and TV/TGV regularization, NCQ and NCBV are able to remove the displacement errors. The NCBV approach provides the result with the visually best quality.



Method	cards	cameraman	Lena	peppers	boat	mountain
$L^2$ -TV	0.6184	0.7316	0.7788	0.7869	0.6628	0.7424
$L^2$ -TGV	0.6756	0.7498	0.8103	0.8253	0.6934	0.7686
weighted $L^2 - L^1$	0.7263	0.7576	0.8040	0.6889	0.6912	0.7594
weighted $L^2 - L^2$	<b>0.7535</b>	0.7522	0.7927	0.6253	0.5931	0.7417
NCBV	0.6916	<b>0.7646</b>	<b>0.8132</b>	<b>0.8290</b>	<b>0.6972</b>	<b>0.7687</b>
NCQ	0.7037	0.7401	0.7862	0.7976	0.6749	0.7473

Table 8.8: *Displacement regularization*. Quantitative comparison of  $L^2$  - TV (ROF),  $L^2$ -TGV, NCQ and NCBV for removing sampling point errors. We compare the SSIM values of the four different methods after automatically retrieving approximately optimal parameters as described in Section 8.1. The proposed approaches clearly outperform the methods based on standard  $L^2$  data term in combination with TV and TGV regularization.

functional

$$\mathcal{F}_{u^{(0)}}^\varepsilon(u) := \frac{1}{2} \int_{\Omega} (u - u^{(0)})^\top ((\nabla u)^\top \nabla u + \varepsilon \text{Id})^{-1} (u - u^{(0)}) + \|\nabla u\|_F^p d\mathcal{L}. \quad (8.24)$$

For this functional, the existence theory within the classical framework of the *calculus of variations* [48, 49] due to the missing weak lower semi-continuity is not applicable. Moreover, we face the problem that to retrieve a relaxed functional of (8.24) an analytic formula for the quasi-convex envelope of

$$J \rightarrow (u - u^{(0)})^\top (J^\top J + \varepsilon \text{Id})^{-1} (u - u^{(0)}) + \|J\|_F^p \quad (8.25)$$

is not at hand.

As a remedy, we propose to alter the problem by applying the concept of semi-group theory as discussed in Section 5.3 to derive a gradient flow of the form (5.30). Inserting

$$A(u) = (\nabla u^\top \nabla u + \varepsilon \text{Id})^{-1} \quad (8.26)$$

into (5.30), we derive the PDE

$$\partial_t u = ((\nabla u)^\top \nabla u + \varepsilon \text{Id}) \nabla \cdot \left( \frac{\nabla u}{\|\nabla u\|_F} \right) \quad (8.27)$$

with initial value  $u(0) = f$  and homogeneous Neumann boundary conditions. Please note that in (8.27) for compactness of notation we omitted the dependence of  $u$  with respect to time  $t$ .

We propose to use the PDE (8.27) to remove displacement errors in vector-valued data such as color images. In the following we consider two examples.

The first one is an artificial test image with a color gradient, cf. Fig. 8.32 (a), distorted by sampling errors as shown in Fig. 8.32 (b). Please note that in the

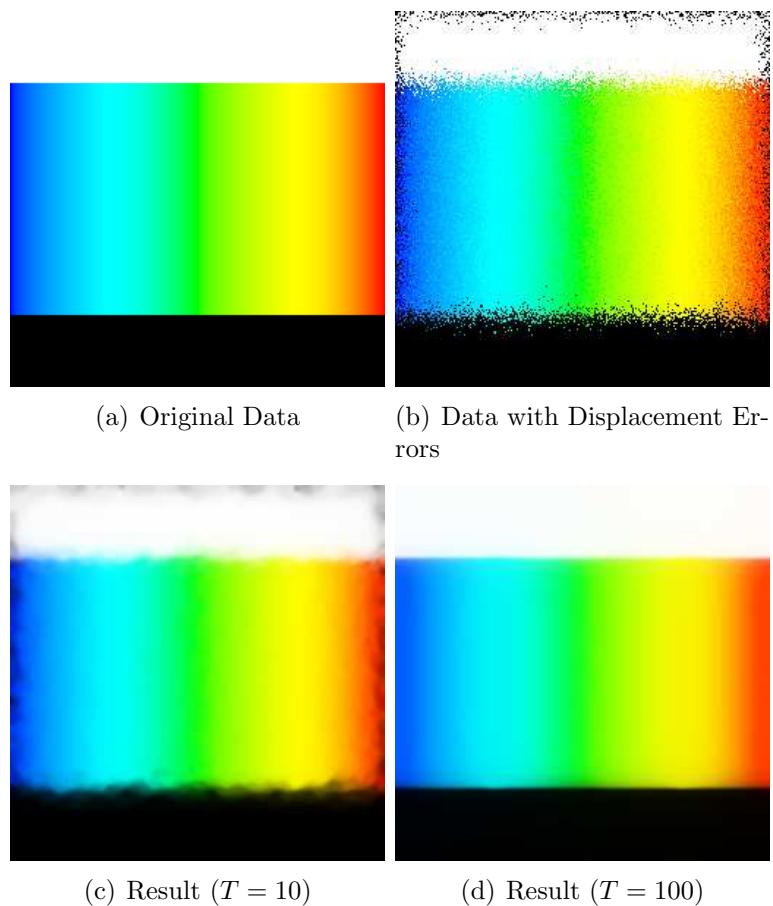


Figure 8.32: [Best viewed in color.] Removing displacement errors in color images using PDE (8.27). We consider the artificial test image (a) distorted by displacement errors (b). The numerical solution to the PDE (8.27) provides a sequence of smoother and smoother images. Here we depict the solution at times  $T = 10$  (c) and  $T = 100$  (d). At the second instance of time, the displacement errors are completely removed without any artifacts arising in the color gradient.



Figure 8.33: [Best viewed in color.] Removing JPEG artifacts in color images using PDE (8.27): (a) Image with JPEG artifacts. (b) Result of the proposed method based on PDE (8.27). Our approach is able to properly remove these artifacts.

generation of our synthetic data, to be able to evaluate pixels which by the random offset are shifted out of the given image domain, we extend the image by black color outside  $\Omega$ . By solving PDE (8.27) numerically, the proposed method produces a sequence of images which become smoother over time. The solution at two instances of time are depicted in Fig. 8.32 (c) and (d). At the second instance of time, the displacement errors are completely removed. It can be observed that the proposed methods does not produce any color artifacts in the color gradient.

In our second example, we consider a test image with JPEG compression artifacts. The motivation behind this is, that such artifacts appear as blocks, in which constant colors are used, where slopes occur in the original image. This actually means that the original color at a certain pixel location is replaced by a color from a near by position. Thus, we can interpret these artifacts as sampling errors. Fig. 8.33 (a) shows a close-up of a real digital photograph with such JPEG artifacts produced by the standard compression of the camera software. It turns out that the proposed method is well suited for removing such artifacts, as we observe from the results depicted in Fig. 8.33 (b).

## 8.3 Denoising Time-of-Flight Data

### 8.3.1 Introduction

Time-of-Flight (ToF) cameras have been designed to capture the depth of a scene with respect to the camera position. The principle of a time-of-Flight camera is briefly summarized as follows:

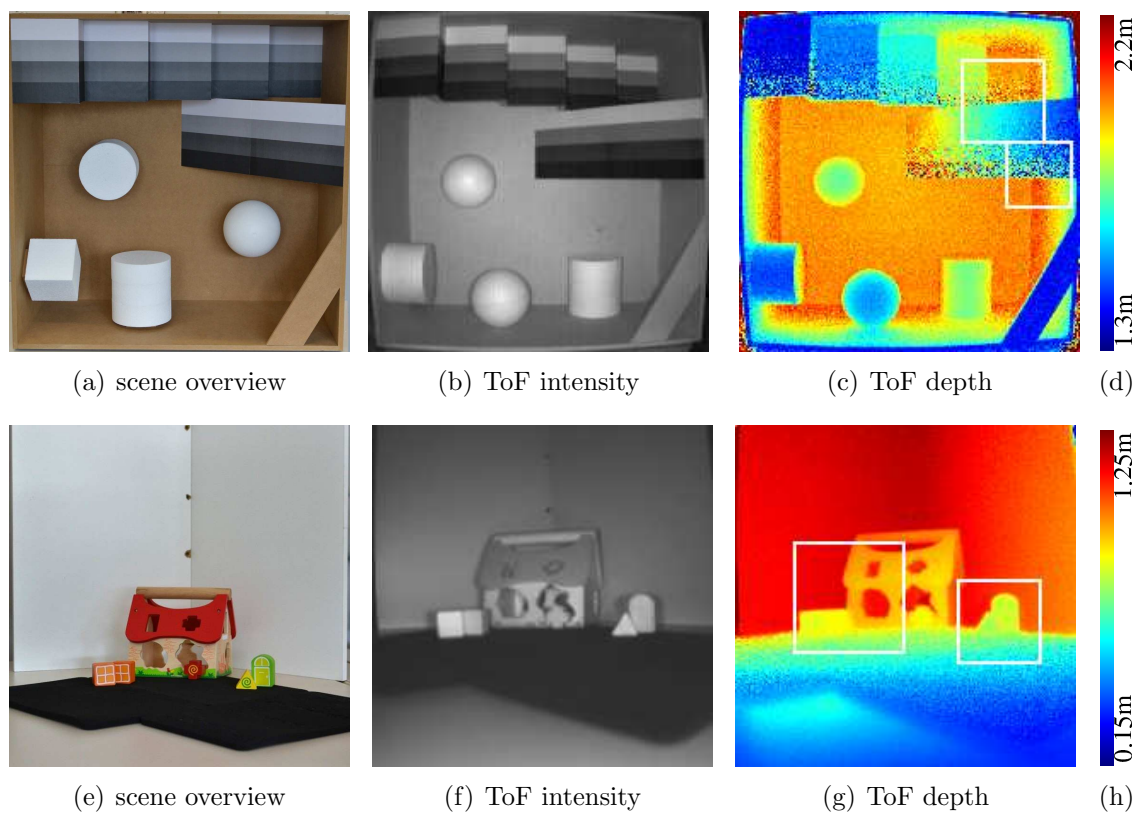


Figure 8.34: Test data sets *HCI box* (top) and *shapes* (bottom). White rectangles indicate regions further investigated in the experimental section.

- The scene is illuminated by an infra-red light source. The emitted IR signal is amplitude modulated.
- Ideally, the light of source is scattered once at the surface of an object and than captured by the camera sensor.
- The camera sensor is able to determine intensity, amplitude and phase of the incoming signal. This is typically done based on of four sequential recordings, denoted by  $I_i : \Omega \rightarrow \mathbb{R}$ ,  $i = 1, \dots, 4$ . We refer to these recordings as the *raw data*.
- By determining the correlation to the emitted signal, a phase shift can be estimated. Formulated in terms of  $(I_i)_i$ , this phase shift is given as

$$\phi(x) = \arctan \left( \frac{I_4(x) - I_2(x)}{I_3(x) - I_1(x)} \right). \quad (8.28)$$

The phase shift  $\phi$  is proportional to travel time of the light from the emitting source to the sensor, and thus, due to the constant speed of light, (up to some systematic and statistical errors) proportional to the radial scene depth. i.e. the distance of each object to the camera position.

- Besides the phase, also the intensity

$$I(x) = \frac{1}{4} \sum_{j=1}^4 I_j(x) \quad (8.29)$$

and the amplitude

$$A(x) = \frac{1}{2} \sqrt{(I_1(x) - I_3(x))^2 + (I_4(x) - I_2(x))^2} \quad (8.30)$$

of the IR signal are determined. These data can also be exploited when further processing the depth data, as we will see below.

By repeating the above steps, ToF cameras provide sequences of depth maps with a constant frame rate. For details on the technology of time-of-flight we refer to the Ph.D. thesis by Mirko Schmidt [150] and the survey chapter in [92].

ToF depth imaging is prone to several error sources, including systematic errors such as multi-scattering, phase-dependent error (wiggling), internal scattering as well as the statistical error of photon counting by the sensor (noise) [146, 150].

The noise can be approximately described by the following model: Assuming that the raw data are affected by independent and identically distributed Gaussian noise with variance  $\sigma$ , it can be shown that the distribution of the estimated depth is approximately Gaussian with a variance varying locally, depending on the signals amplitude  $A(x)$ :

$$\phi(x) \sim \mathcal{N} \left( 0, \frac{\sigma^2}{2A^2(x)} \right). \quad (8.31)$$

In contrast, the noise in the intensity and amplitude data is still i.i.d., cf. [58]:

$$I(x) \sim \mathcal{N}(0, \frac{\sigma^2}{2}), \quad A(x) \sim \mathcal{N}(0, \frac{\sigma^2}{4}). \quad (8.32)$$

Due to the strong noise (cf. Fig. 8.34, right column) it is inevitable to apply a denoising technique before further processing the depth data. In [102] we have surveyed state-of-the-art techniques for denoising ToF data and investigated different strategies from smoothing the raw data to smoothing the final output. It turned out that a favorable strategy is to apply a variational approach with adaptive second-order total variation regularization at the very end of the depth acquisition pipeline. For the specific task of denoising ToF data, TV regularization has several advantages:

1. The noise model can easily be incorporated into the data term by introducing a weighting factor.
2. Second-order TV favors piecewise smooth solutions with jumps and thus is suitable for the reconstruction of depth maps.
3. There is only little texture contained in the depth data, thus the known lack of TV to handle textures is less relevant.
4. Reducing the regularization strength at edges, as it is done in adaptive regularization approaches, prevents a loss of contrast, which other methods including standard TV show.
5. When determining the adaptivity for the TV regularizer, we can make use of the fact that not only depth information, but also the intensity at a far lower noise level is available.

We will show in the experimental section, that the approach presented here produces better results than applying state-of-the-art methods (off the shelf) on depth data only.

### 8.3.2 Approach

#### Edge Detection

In order to apply second-order anisotropic total variation to denoise ToF data, we first have to determine edge locations and edge positions. Here, we use both the intensity and the depth data, to which we apply an edge detection algorithm as follows.

The core of our edge detection is based on the structure tensor approach proposed by Köthe [87], combined with Canny's edge detector and hysteresis thresholding [33].

1. As suggested by Köthe, the sampling frequency for the image gradient is doubled by evaluating finite differences at every full and half pixel position.

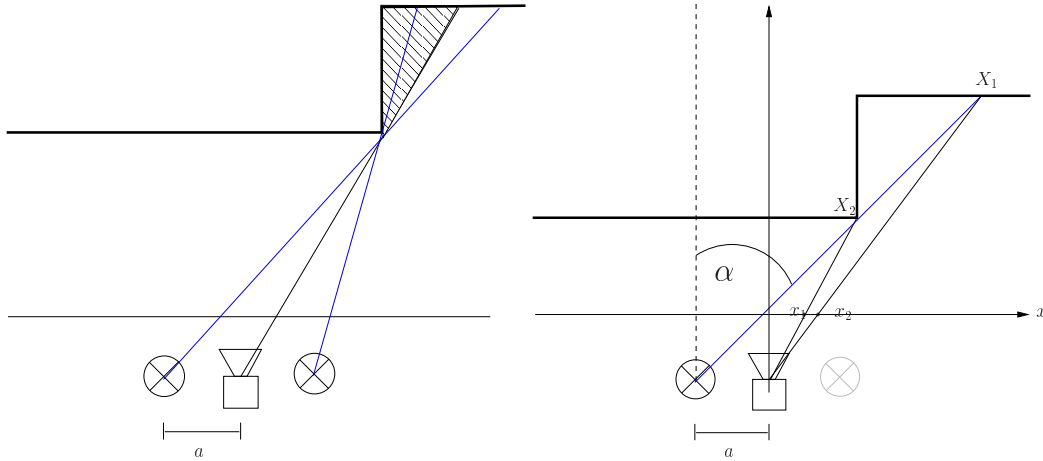


Figure 8.35: Geometry of camera and light paths.

2. The structure tensor (cf. Section 4.6) is calculated from the image gradient.
3. The structure tensor is locally averaged using an hour glass shaped Gaussian. The eigenvector to the larger eigenvalue provides approximate edge normals.
4. The eigenvalues and eigenvectors serve as input for Canny's algorithm.
5. A hysteresis thresholding with parameters  $\tau_{low}$  and  $\tau_{high}$  is performed.
6. In order to separate adjacent edges, we look for discontinuities of the edge normals. If the scalar product of two adjacent edge normal exceeds a given threshold  $\tau_s$ , the corresponding edge is split up into two parts. We will see that for post-processing of the edge data, such a splitting is eligible.
7. Edges, which are shorter than a given threshold, are removed.

When applying the above edge detection approach to the depth data only, due to the strong noise it is inevitable to have false positive results.

As a remedy, we consider the following strategy. The above procedure is applied independently to both the intensity and depth data as follows.

1. **High-Confident Edges:** We apply the detection procedure with large thresholds  $\tau_{high}$  and  $\tau_{low}$  to the depth data to find strong edges. Due to the high thresholds we expect very rare false positives.
2. **Intensity Edges:** We apply the above detection procedure to the intensity image. Since the noise of the intensity image is not as strong as in the depth data, we expect less false positives due to noise. There is, however, the issues that an intensity edge might be caused by a varying reflectance of the objects surface, and not by the overlapping of two different objects. We refer to

such edges as *texture edges*. Another important source for intensity edges are shadows. To eliminate both kinds of intensity edges, we apply appropriate post-processing steps as detailed in the next paragraphs.

3. **Low-Confident Edges:** The detection procedure is again applied to the depth data, but with smaller thresholds  $\tau_{low}$  and  $\tau_{high}$  as in the first step. As a result, we can expect to find more edges as before with a higher risk of false positive detections. We use this edge set in the subsequent step to detect and remove shadow edges in the intensity image.

Since ToF cameras rely on active illumination of the scene and the light sources are fixed parts of the camera, the geometry of lightening of a ToF camera is known in advance. However, it depends on the camera model. Typical light sources are single LEDs, LED clusters or LED rings. In this section, we make use of the known geometry to remove shadow edges detected in the intensity data. We restrict our approach to the geometry of the PMD Cam Cube 3, where the scene is lighted by two LED arrays of size  $42\text{mm} \times 43\text{mm}$  mounted at the left and right side of the camera sensor. With the scene typically being in a distance of a few meters, we may treat the two LED arrays as two point sources. We remark that our approach can be easily adapted for other camera models.

The strategy to detect shadow edges is the following: Since the light sources are located to the left and right of the sensor, we restrict our approach to non-horizontal edges. Considering an edge detected in the intensity image as a candidate for a shadow edge, we have to identify a depth edge causing this shadow. We do this separately for each light source. The search can be narrowed down due to the following considerations:

- the depth edge has to lie in close vicinity of the shadow,
- the depth edge should be parallel to the intensity edge, and
- after fixing the light source (left/right) producing the shadow, the depth edge has to lie either to the right or to the left side of the shadow, otherwise it is occluded in the camera view.

We search for such depth edge in the third data set containing a higher number of edges than the first set. For a candidate pair of depth and shadow edge, we check if the corresponding 3D points share the same optical ray from the light source. For simplicity, let us assume for the moment that we search for a vertical shadow along the line  $y = 0$ . We consider a shadow casting edge at position  $(x_1, 0)$  in the image plane and the position of the shadow at  $(x_2, 0)$ . The corresponding 3D points are  $X_i := (d_i x_i, 0, d_i)$ ,  $i = 1, 2$  (cf. Fig. 8.35), where  $d_i$  is the Cartesian depth of the two points. Note that we can retrieve the Cartesian depth from the ToF data providing the radial depth of the scene. From the 3D points, the angle  $\alpha_i$  of each optical ray to the light source is given by

$$\tan(\alpha_i) = \frac{d_i}{a + d_i x_i}, \quad i = 1, 2, \quad (8.33)$$



where  $a$  is the horizontal distance between camera sensor and light source. For non-vertical edges, due to the fact that relative to typical scene depths, the distance between light source and camera sensor is small, we assume (8.33) to hold approximately. If both angles  $\alpha_i$  are approximately equal, we regard the found pair of edges as shadow casting edge and corresponding shadow. In this case, we remove the found shadow edge from the set of intensity edges. Moreover, the presence of this shadow asserts that the casting edge (from edge set 3) is indeed a depth edge. We therefore add this edge to the first set of depth edges.

As mentioned before, the set of intensity edges may contain edges which are caused by changing surface properties of the objects in the scene (*texture edges*). We can detect such texture edges by comparing the average depth on both sides of the edge. To this end, we calculate for each intensity edge

$$d_{\text{diff}} = \frac{1}{|E|N} \sum_{x \in E} \sum_{j=1}^N |d(x + j\nu(x)) - d(x - j\nu(x))|, \quad (8.34)$$

where  $\nu$  is the known edge normal, and discard intensity edges if  $d_{\text{diff}}$  lies below a given threshold.

This ansatz might fail at depth ridges coinciding with intensity edges, where the depth varies on both sides of the edge, leading to a higher value of  $d_{\text{diff}}$ .

If such ridges should be discarded as well, one can in addition calculate the average absolute value of the depth gradient and threshold its difference

$$\text{grad}_{\text{diff}} = \frac{1}{|E|N} \sum_{x \in E} \sum_{j=1}^N ||\nabla d(x + j\nu(x))| - |\nabla d(x - j\nu(x))||. \quad (8.35)$$

We remark that for denoising with adaptive TV it might be advantageous to adapt the regularization also to depth ridges in order to avoid a flattening of these ridges. We will study such an approach in future work.

The final step of the edge detection process consists in merging the set of high-confident edges with the intensity edges. Note that the set of high-confident edges includes also low-confident depth edges of the third set, which were found as shadow casting edges during the shadow removal process. To sum up, the final detection result contains edges if they are either

- high-confident depth edges,
- image edges excluding texture and shadow edges,
- shadow casting edges,

where for each edge its location  $\chi_e(x)$  and orientation  $v_e(x)$  (edge normal) are stored.

### Denoising with Anisotropic Second-Order TV

We use a variational approach with anisotropic second-order TV regularization, where we incorporate the edge information found in the detection phase.

We consider a weighted data term of the form

$$\mathcal{S}(u) := \int_{\Omega} w(x)(u - f)^2 dx, \quad (8.36)$$

where  $f$  are the data to be denoised and the weighting function  $w(x) : \Omega \rightarrow \mathbb{R}^+$  is chosen to take into account the locally varying variance of the noise. Assuming that  $f$  is affected by Gaussian noise zero mean and standard deviation  $\sigma(x)$ , the weighting term should be chosen as  $w(x) \propto \frac{1}{\sigma(x)^2}$ . In the case of ToF depth data, where we assume the variance to depend on the amplitude  $A(x)$ , cf. (8.32), this leads to  $w(x) \propto A^2(x)$ . It turns out, that in practice it is suitable to bound the weighting term from above. From our experiments we found that the choice

$$w(x) := \min \left\{ \tilde{A}^2(x), 1 \right\}, \quad \tilde{A} = C \frac{A(x) - \min_x A(x)}{\max_x A(x) - \min_x A(x)} \quad (8.37)$$

with some  $C > 0$  performs well for our approach.

**Remark 8.3.** *Introducing the above weighting of the data term, we assume that besides the depth also the amplitude signal is provided by the camera. For the PMD cam cube, which we used to record our test data, the amplitude can be calculated from the accessible raw data  $I_i$ . In case that a ToF camera does not provide access to either amplitude or raw data, the intensity image might serve as an approximation for the amplitude data.*

The regularization term  $\mathcal{R}(u)$  for our approach is chosen to be the one from Example 4.28. Recall that this TV regularization is anisotropic in the first- and adaptive in the second-order, as it reduces the first-order regularization strength in normal direction  $v_e$  and the second-order regularization strength isotropically at edges.

Please note that it is easily checked that, if  $w(x)$  in (8.36) is bounded from below by a positive constant, Assumption 4.10 is satisfied. Consequently, Theorem 4.11 ensures existence of a minimizer of

$$\mathcal{F}(u) := \mathcal{S}(u) + \mathcal{R}(u) \quad (8.38)$$

for  $\mathcal{S}$  and  $\mathcal{R}$  defined as above.

Finally, we remark that adapting the ideas of slope detection (Section 4.6) for denoising ToF data is planned as future work.

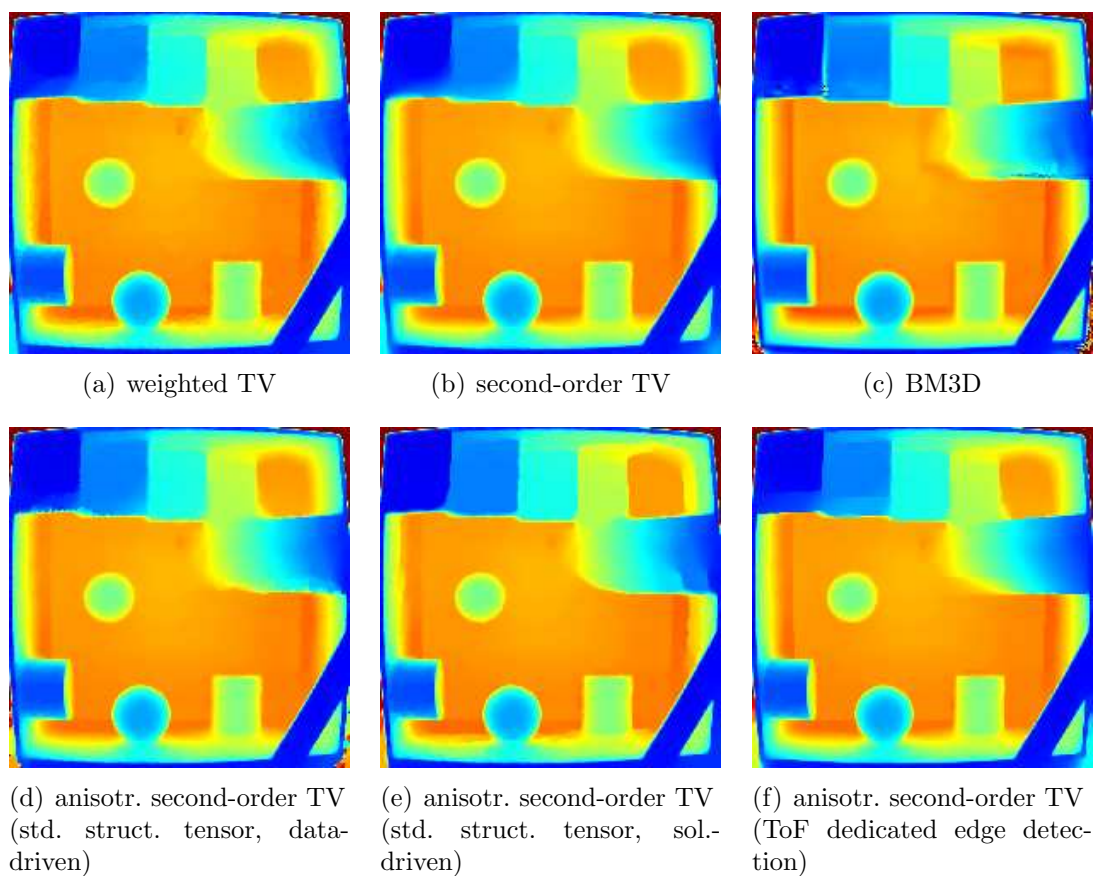


Figure 8.36: *Denoising ToF data.* Comparison of different denoising approaches on the *HCI box* data set. The proposed anisotropic second-order TV regularization (f) based on the specialized edge detection process performs best concerning both issue of edge sharpness and surface regularity.

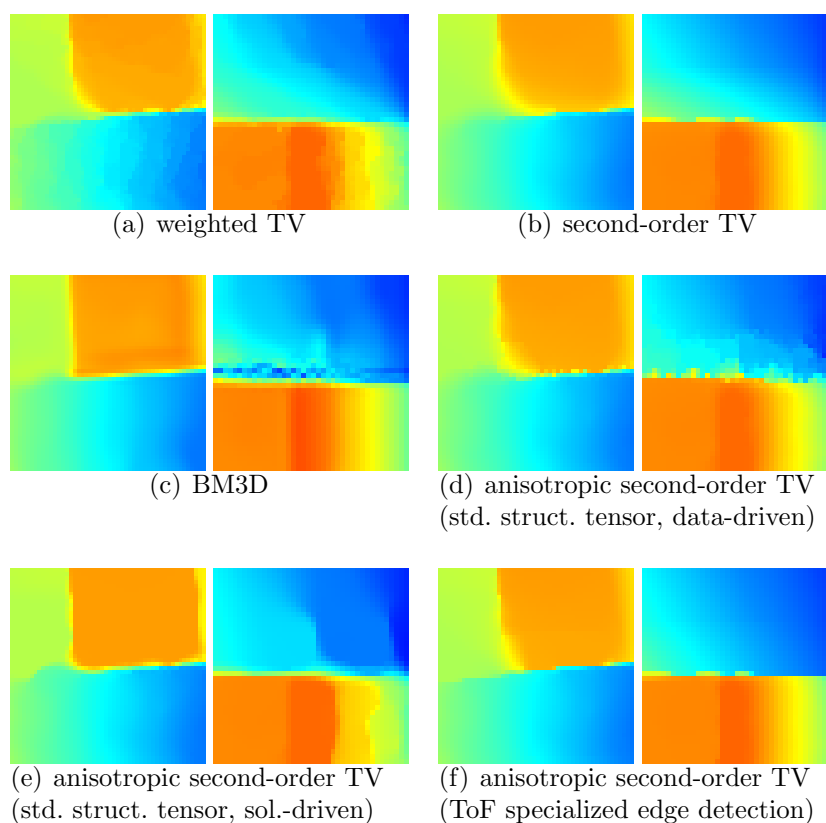


Figure 8.37: *Denosing ToF data*. Close-ups of results in Fig. 8.36, showing piecewise affine regions with sharp edges. The TV approach based on the ToF specific edge detection shows for most regions except for the edge of the lowest stair step (top part of first close-up) the best performance concerning regularity of edges and slopes.

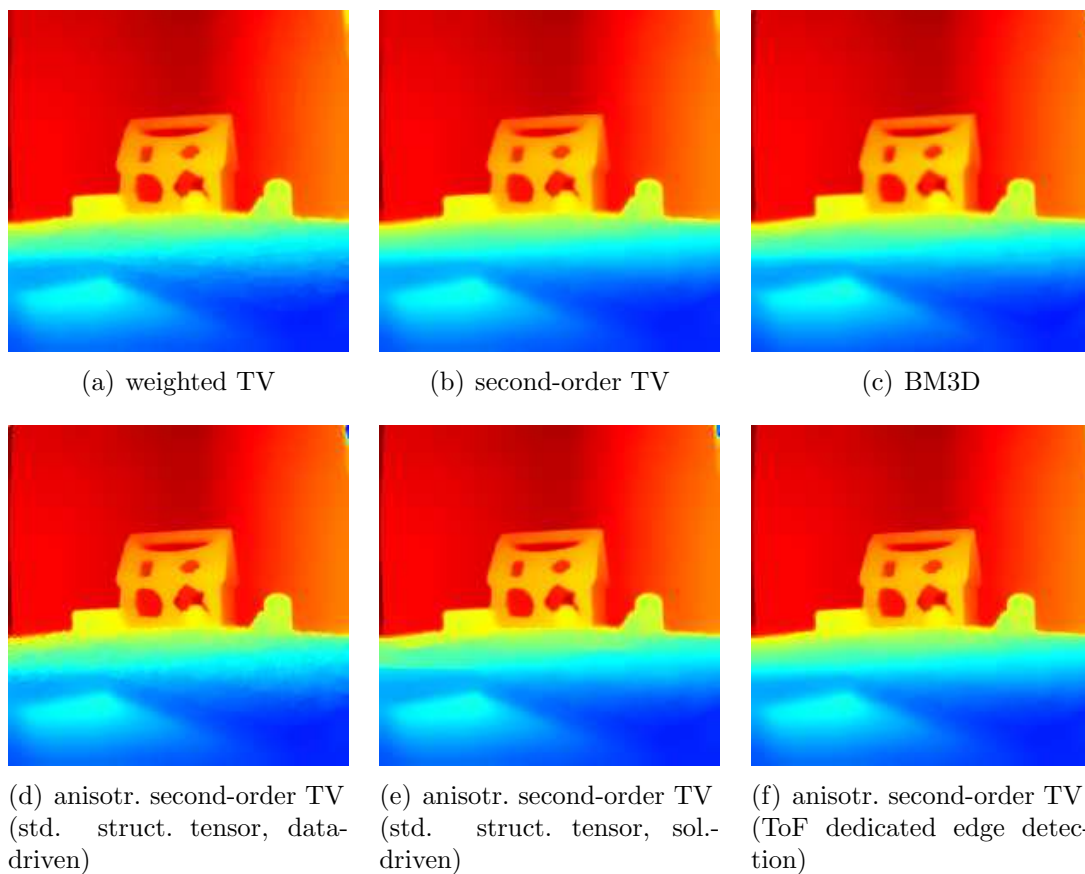


Figure 8.38: *Denoising ToF data.* Comparison of different ToF denoising approaches on the *shapes* data set. The solution-driven and the ToF specialized anisotropic second-order TV approach show the best results, while the other either show blurriness or stair-casing artifacts.

### 8.3.3 Results

The edge detection step described above has been designed to optimize the detection results on ToF data sets. In [145] we compared this approach to other standard and state-of-the-art detection approaches to show its advantages. Since the scope of this thesis lies on regularization, we focus here on the actual denoising process and the influence of the choice of the regularizer on the result. To this end, we compare the proposed approach to 1) standard first order TV with weighted data term, 2) standard second-order TV with weighted data term, (3) the BM3D denoising software, 4) data-driven and 5) solution-driven anisotropic second-order TV based on the standard structure tensor (cf. Example. 4.28). For evaluation, we use the two data sets *HCI box* (introduced in [119], available online at <http://hci.iwr.uni-heidelberg.de/Benchmarks/document/hcibox>) and *shapes* (introduced in [145]) as depicted in Fig. 8.34. In each data set, two particular regions containing piece-wise planar surfaces and edges are selected for close-ups (indicated by white rectangles in Fig. 8.34). For the *HCI box*, these regions are the stair steps in the upper part of the depth map and the ramp in the middle part. In the *shapes* depth map, we focus on the wooden objects and the toy house. Moreover, the surface of the table, which is expected to have a regular slope, is of interest.

The corresponding results for the denoising methods under considerations are shown in Figs. 8.36 and 8.38, with close-ups of the selected regions shown in Figs. 8.37 and 8.39, respectively.

On the results of the individual methods we observe the following. Firstly, *standard TV regularization* shows the well-known stair-casing effect, in particular in regions of strong noise. Besides this disadvantage, it is able to reconstruct sharp edges, but without the sufficient regularity (smoothness).

Using *isotropic second-order TV* instead avoids the stair-casing effect, but at the expense of less sharp edges.

The BM3D method, at least in its standard implementation, is not able to cope with the varying noise level. Using a smoothing strength which is suitable for most of the regions in the test images, artifacts remain in those regions with very high noise level. Increasing, however, the smoothing level leads to an over-smoothing. We remark that for the HCI box data set, the depth edges are sharply reconstructed, while for the shapes data set, the edges at the toy house appear slightly blurred.

Concerning the data-driven anisotropic second-order TV approach, we observe that despite a very strong pre-smoothing, it cannot cope with the high noise level. Artifacts from the noise are generated by the anisotropy.

Compared to the data-driven approach, the solution driven approach does not show these artifacts. However, it may produce artificial edges in the regions of high noise level, as for example for the HCI box data, where some artificial plateaus are reconstructed on the ramp. Besides this negative effect, the reconstruction of the objects in the HCI box data shows a good quality. In particular, the edge of the lowest step of the stair is reconstructed in a straight way.

Inspecting the results for the anisotropic second-order TV with edges determined

by the specialized approach presented above, we observe both edges and slopes (e.g. the ramp in the HCI box data set and the table surface in the shapes set) are reconstructed in an accurate way. Objects with different depth are clearly distinguishable (see e.g. 8.39 (f)). The only disadvantage of this approach is that the anisotropy in some regions is guided by texture edges found in the intensity data. This occurs for example at the lower part of the ramp. This effect, however, is less dominant than artifacts in the compared approaches. Concerning the shapes data set, we observe that the results of the solution-driven and the specialized approach are close-by.

Finally, we want to remark that a quantitative evaluation of the above methods based on ground truth would be advisable. For the HCI box data set, there is a virtual model available, from which the true depth data can be retrieved. The real data of the HCI box, however, are affected not only by noise, but also from various systematic errors such as artifacts from multi-path and flying pixels. These systematic errors are dominant compared to the noise error. Since the methods considered above only tackle the removal of noise, the results still contain these systematic errors. The difference to ground truth depth is strongly governed by these errors and cannot be used to measure the denoising quality. A remedy to this problem would be to generate synthetic data by means of a ToF simulator, which would allow to explicitly turn the noise component of the data on or off, while keeping the other errors fixed.

## 8.4 Fusion of ToF and Stereo Sensors

To overcome the limitations of specific imaging systems, a common strategy is to combine different types of imaging sensors. The individual outputs of the camera systems then typically are merge in a post-processing step, also referred to as *sensor fusion*. For ToF imaging an appropriate and commonly used approach is to use stereo vision as a second system. The rational behind combining ToF and stereo is, that ToF can provide useful information in regions, where the stereo cameras fail due to missing texture, while stereo can produces better results in regions of low IR reflectance. Thus, by combining both vision systems, we can exploit the individual benefits of each system. For related work on fusion of ToF and stereo data we refer to the survey [120].

In the following, we consider a variational approach with second-order anisotropic TV regularization for fusing ToF and stereo data.

### 8.4.1 Camera Setup and Available Data

We consider a camera rig with two monochrome cameras ( $1312 \times 1082$  pixel) together with a time-of-flight camera (PMD Cam Cube 3,  $200 \times 200$  pixel). It is assumed that the stereo images together the low resolution depth map from the ToF camera are recorded simultaneously. Since all cameras are fully calibrated, the stereo images can be rectified and the depth data can be warped to the view of the

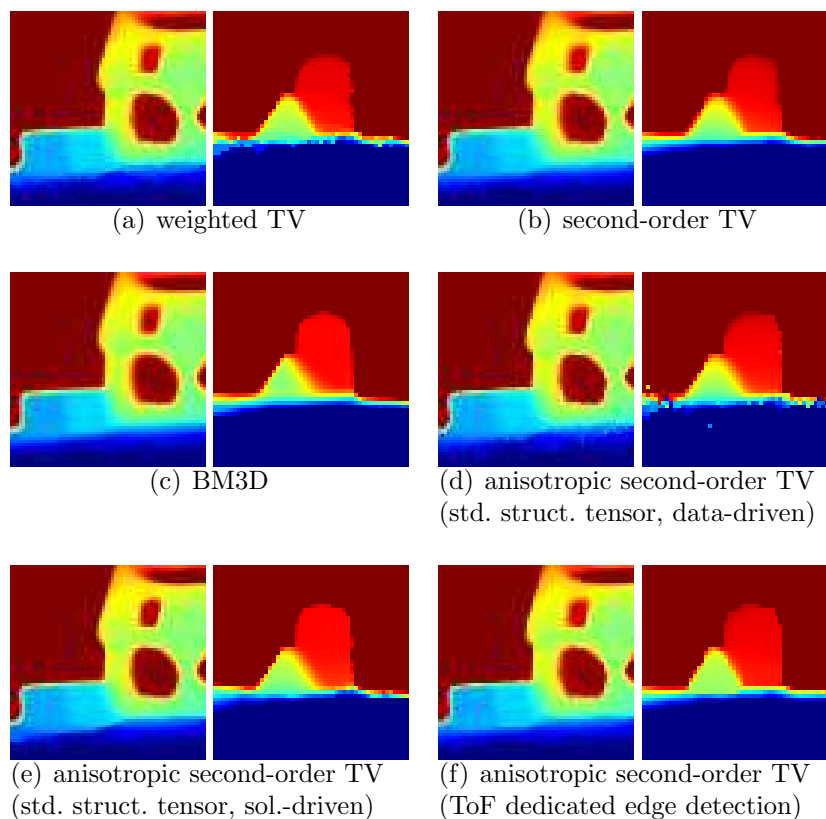


Figure 8.39: *Denosing ToF data.* Close-ups of results in Fig. 8.38, showing the toy house with two cuboids in front (left) and two other shaped blocks (left). The TV approach using a ToF specific edge detection best reproduces the affine shape of the scenes, with sharp edges and high contrast, which eases to distinguish the different objects



Figure 8.40: Camera setup for fusion of Time-of-Flight and stereo.



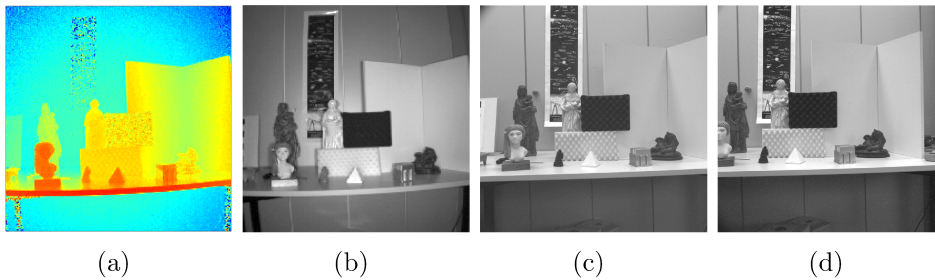


Figure 8.41: Input data for ToF stereo fusion: (a) ToF depth data ( $200 \times 200$ ), (b) ToF intensity, (c) left and (d) right stereo images  $I^L$  and  $I^R$  ( $1312 \times 1082$ , cropped for visualization).

left stereo camera. This warping also includes an upsampling of the ToF data to achieve the resolution of the left stereo image. In order to combine the results of ToF and stereo, the ToF depth data afterwards are transformed into disparities. To sum up, the data recorded by the sensors are the left and right grayscale images, denoted by  $I^L$  and  $I^R$ , the ToF intensity  $I^{\text{ToF}}$  and the disparity  $u^{\text{ToF}}$ , all in the full resolution of the stereo camera.

### 8.4.2 Approach

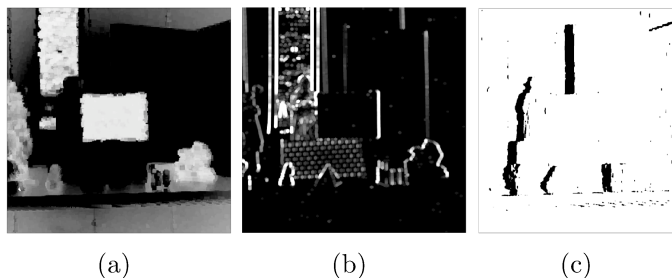


Figure 8.42: Fidelity of scene in Fig. 8.41. (a) fidelity measures  $C_{\text{ToF}}$  for ToF, (b)  $C_{\text{texture}}$  for matchable texture in the stereo data and (c)  $C_{\text{occl}}$  for occluded areas.

The task considered in this section is to obtain a high resolution displacement field or disparity  $u$  from the two stereo images  $I^L$  and  $I^R$  by taking into account the additional disparity data  $u^{\text{ToF}}$  from the ToF camera. Due to the rectification, the displacement field can be assumed to be horizontal.

We propose a variational approach with second-order TV regularization, where the data term is modeled to incorporate both the ToF and stereo data. Our approach builds on ideas presented in [118], but with a different regularization term.

In order to determine, in which image regions the ToF or stereo data can be trusted, we make use of three different fidelity measures proposed in [118]. Before describing the variational approach, we briefly recall these three measures:

1. The accuracy of the ToF depth data depends strongly on the intensity of the IR signal recorded by the camera sensor. To account for this fact, we consider the normalized intensity  $C_{\text{ToF}}$  as a fidelity measure for ToF.
2. In stereo matching, disparities can only be estimated in regions where sufficient texture is present. We use an indicator  $C_{\text{texture}}$  for texture in the left stereo image, obtained by clipping and normalizing the horizontal derivative  $\partial_x I^L$  of the image.
3. Finally, we consider occluded areas described by a binary map as  $C_{\text{occl}}$ , where occlusion is indicated by zero values. This occlusion map is found from the scene geometry using the ToF depth data.

Please note that all measures are determined with respect to the view of the left stereo camera and with the same resolution. For illustration, we depict the three fidelity measures for the *statues* data set in Fig. 8.42.

We now turn to the variational approach. We describe our approach in the continuous setting. To this end, let now  $I^L, I^R, u^{\text{ToF}} : \Omega \rightarrow \mathbb{R}$ , be defined on the image domain  $\Omega \subset \mathbb{R}^2$ . We have to determine a correspondence between the pixels of image  $I^L$  to the pixels of the image  $I^R$ . We describe this correspondence by means of a horizontal displacement field  $u : \Omega \rightarrow \mathbb{R}$ . We assume that a first approximation  $u^0$  of  $u$  is available. Here, we use the upsampled ToF data  $u^{\text{ToF}}$ , while other kinds of initialization are also feasible.

In order to refine  $u^0$ , we again make use of both the stereo data and the ToF depth map. As a consequence, we consider a variational approach with two data terms  $\mathcal{S}_{\text{stereo}}(u)$  and  $\mathcal{S}_{\text{ToF}}(u)$ . The objective function takes the general form

$$\mathcal{F}(u) := \mathcal{S}_{\text{stereo}}(u) + \mathcal{S}_{\text{ToF}}(u) + \mathcal{R}(u), \quad (8.39)$$

where  $\mathcal{S}_{\text{stereo}}(u)$  and  $\mathcal{S}_{\text{ToF}}(u)$  are two data fidelity terms for the stereo and the ToF input and  $\mathcal{R}(u)$  is the regularization term. We discuss all three terms in detail below.

In order to improve the initial displacement  $u^0$ , we search for an update  $\delta u : \Omega \rightarrow \mathbb{R}$ , such that  $u = u^0 + \delta u$ . This incremental update step is performed repeatedly to obtain a sequence  $(u^k)_k$  of refined displacements. It turns out that the incremental update is in particular crucial for the stereo data term, as we will see in the following paragraphs. We adapt the notation accordingly and consider now a stereo data term of the form  $\mathcal{S}_{\text{stereo}}(u^k, \delta)$ , where current iterate  $u^k$  and update  $\delta u$  are independent parameters. For the other data term as well as the regularization term, it suffices to evaluate them directly on  $u^k + \delta u$ , i.e. we use  $\mathcal{S}_{\text{ToF}}(u^k + \delta u)$  and  $\mathcal{R}(u^k + \delta u)$ . To conclude, we consider the sequence of optimization problems

$$\delta u^{k+1} := \arg \min_{\delta u} \mathcal{F}(u^k, \delta u), \quad k \geq 0, \quad (8.40)$$

where

$$\mathcal{F}(u^k, \delta u) := \mathcal{S}_{\text{stereo}}(u^k, \delta u) + \mathcal{S}_{\text{ToF}}(u^k + \delta u) + \mathcal{R}(u^k + \delta u), \quad (8.41)$$

and where we perform an update  $u^{k+1} := u^k + \delta u^{k+1}$  after solving each problem (8.40).

Now we turn to the choice of the individual data terms and the regularization term, which we detail in the following paragraphs.

### Stereo Data Term

The classical brightness constancy equation (BCE) [31] for stereo/optical flow, which we assume to be valid in our application, reads

$$I^L(x, y) = I^R(x + u(x, y), y). \quad (8.42)$$

Recall that we aim to recover  $u$  by a sequence of  $(u^k)_k$ , where for each  $u^k$  we search for an update  $u^k + \delta u$ . Inserting  $u = u^k + \delta u$  in (8.42), we derive

$$I^L(x, y) = I^R(x + u^k(x, y) + \delta u(x, y), y). \quad (8.43)$$

Linearizing  $I^R(x, y)$  by Taylor expansion around  $x + u^k(x, y)$ , we find

$$I^R(x + u^k(x, y) + \delta u(x, y), y) \approx I^R(x + u^k(x, y)) + \delta u(x, y) \frac{\partial}{\partial x} I^R(x + u^k(x, y), y), \quad (8.44)$$

which is justified for small  $\delta u$  [131]. Inserting (8.44) into (8.43) gives the linearized brightness constancy equation

$$I^L(x, y) \approx I^R(x + u^k(x, y), y) + \delta u(x, y) \frac{\partial}{\partial x} I^R(x + u^k(x, y), y). \quad (8.45)$$

The first fidelity term  $\mathcal{S}_{\text{stereo}}$  in (8.41) is to penalize derivations from (8.45). In order to be robust against outliers, we use the  $L^1$ -norm, i.e.

$$\int_{\Omega} \left| I^R(x + u^k(x, y), y) + \delta u(x, y) \frac{\partial}{\partial x} I^R(x + u^k(x, y), y) - I^L(x, y) \right| dx dy. \quad (8.46)$$

Please note that the fidelity term is discriminative only in regions, where matchable textures are present. We exploit the fidelity measure  $C_{\text{texture}}$  to identify regions with sufficiently matchable textures. Moreover, the stereo fidelity term is not reasonable, if occlusions occur. Fidelity measure  $C_{\text{occl}}$  indicates areas with occlusions. To respect both issues, consider a weighting function  $w_1(x, y) := C_{\text{texture}}(x, y) \cdot C_{\text{occl}}(x, y)$  for the stereo data term. Combining (8.46) with this weighting function, we end up with the data term

$$\begin{aligned} \mathcal{S}_{\text{stereo}}(u^k, \delta u) := \\ \int_{\Omega} w_1(x, y) \left| (\partial_x I^R(x + u^k(x, y), y) \delta u + I^R(x + u^k(x, y), y) - I^L(x, y)) \right| dx dy. \end{aligned} \quad (8.47)$$

### ToF Data Term

As already mentioned before, we assume that the ToF data have been upsampled to the resolution of the stereo images, warped to match to the view of the left camera and transferred into disparities  $u^{\text{ToF}}$ .

We can make use of the ToF data in regions where no sufficiently matchable structures for stereo are available, taking into account the fidelity  $C_{\text{ToF}}$  for ToF. To this end, we define a second weighting term  $w_2$  by

$$w_2(x, y) := (1 - w_1(x, y)) \cdot C_{\text{ToF}}(x, y), \quad (8.48)$$

where  $w_1$  is the weighting function for the stereo data as defined above. The second fidelity term  $\mathcal{S}_{\text{ToF}}$  in (8.41) penalizes the deviation of a given  $u = u^k + \delta u$  from the data  $u^{\text{ToF}}$ . Again, to be robust against possible outliers, we use the  $L^1$ -norm, weighted with the above function  $w_2(x)$ :

$$\mathcal{S}_{\text{ToF}}(u^k + \delta u) := \int_{\Omega} w_2(x, y) |u^k(x, y) + \delta u(x, y) - u^{\text{ToF}}(x, y)| \, dx \, dy. \quad (8.49)$$

**Remark 8.4.** *As already mentioned, ToF data suffer from several systematic errors. By fusion ToF and stereo data, we are able to cope with most of them. There is, however, one systematic error which strongly affects some regions in the image domain, namely the multi-path error. The multi-path error is a systematic error, where depth is incorrectly estimated due to multiple reflections of the IR signal at different objects, such that the optical path is prolonged (cf. [92] for details), and consequently the depth is falsified by a large shift. This error does not affect only isolated pixels, but larger regions in the image. Due to the high impact of this error on the depth data, the ToF data should be discarded in these regions, which could be done by adapting the weight  $w_2$ . However, we up to now have not studied a criterion to automatically detect regions with multi-path errors.*

### Regularization Term

As for the ToF data term, for regularization it makes sense to consider the full displacement field  $u = u^k + \delta u$  instead of penalizing  $\delta u$ . Thus, in the following we specify  $\mathcal{R}(u) = \mathcal{R}(u^k + \delta u)$ .

For the task of fusing ToF and stereo data we choose the same regularizer  $\mathcal{R}(u)$  as for denoising ToF data (cf. Section 8.3), namely second-order TV, which is anisotropic in the first order term and adaptive in the regularization strength for the second-order term. Different to the approach for ToF denoising, we gather the required edge information from the left stereo image instead of the ToF intensity and depth.

To this end, let  $\chi_e : \Omega \rightarrow [0, 1]$  be an edge indicator function and  $v_e(x) : \Omega \rightarrow S^1$  be a vector field providing the edge normals. In our application, we use edge data obtained by a similar approach as described in Section 8.3.2 (structure tensor

combined with Canny’s edge detection approach), but applied on the left stereo image.

Based on the two functions  $\chi_e$  and  $v_e$ , we define  $\mathcal{R}(u)$  as described in Example 4.26 in Section 4.7.

Please note that this regularization term differs from the one proposed in [118], where adaptivity of second-order TV was achieved by adapting the discrete differential operator near edge structures to avoid smoothing across edges. We compare both regularizers in the experimental section below.

### Numerical Aspects

Concerning numerical treatment of the variational problem (8.40), two issues are to be mentioned.

Firstly, we remark that due to the dense and quite accurate initialization by the local approach a classical coarse-to-fine strategy can be omitted. Instead, the update  $u^k \rightarrow u^k + \delta u^k$  is performed iteratively for several steps on the finest scale.

Secondly, in order to ease the numerical treatment, the objective function is modified by introducing a variable splitting, as it is standard in the literature, see e.g. [45, 173, 182]. Introducing auxiliary variables  $\tilde{u}$  and  $\bar{u}$  and penalizing their difference to  $u = u^k + \delta u$  with additional quadratic cost terms, we derive

$$\min_{\delta u, \tilde{u}, \bar{u}} \mathcal{S}_{\text{stereo}}(u^k, \delta u) + \mathcal{S}_{\text{ToF}}(\bar{u}) + \mathcal{R}(\tilde{u}) + \frac{1}{2\lambda_1} \|u^k + \delta u - \tilde{u}\|_2^2 + \frac{1}{2\lambda_2} \|\bar{u} - \tilde{u}\|_2 \quad (8.50)$$

as an approximation of problem (8.40). With  $\lambda_1, \lambda_2 > 0$  tending to zero, both problems become equivalent. In practice, however, we use fixed small values  $\lambda_1 = \lambda_2 = 0.01$ . Typically (8.50) is minimized by alternately solving the three subproblems

$$\min_{\delta u} \mathcal{S}_{\text{stereo}}(u^k, \delta u) + \frac{1}{2\lambda_1} \|u^k + \delta u - \tilde{u}\|_2^2, \quad (8.51)$$

$$\min_{\bar{u}} \mathcal{S}_{\text{ToF}}(\bar{u}) + \frac{1}{2\lambda_2} \|\bar{u} - \tilde{u}\|_2^2, \quad (8.52)$$

$$\min_{\tilde{u}} \mathcal{R}(\tilde{u}) + \frac{1}{2\lambda_1} \|\tilde{u} - u^k + \delta u\|_2^2 + \frac{1}{2\lambda_2} \|\tilde{u} - \bar{u}\|_2^2. \quad (8.53)$$

For the subproblems (8.51) and (8.52), closed forms can be provided using the pseudo-inverse of  $(\frac{\partial}{\partial x} I^R(x + u^k(x), y))$  for the first and utilizing soft thresholding for the second one. We solve subproblem (8.53) by applying the primal-dual algorithm proposed by Chambolle & Pock [36], see Section 7.1.

### 8.4.3 Experiments

In this experimental section we follow two aims. Firstly, we demonstrate the competitiveness of the approach presented above to two state-of-the-art methods, namely the local approach in [118] and the semi-global matching (SGM) proposed by Hirschmüller [75]. In addition, we compare to the closely related global approach from [118], in which second-order adaptive TV based on edge masks is used. Secondly,

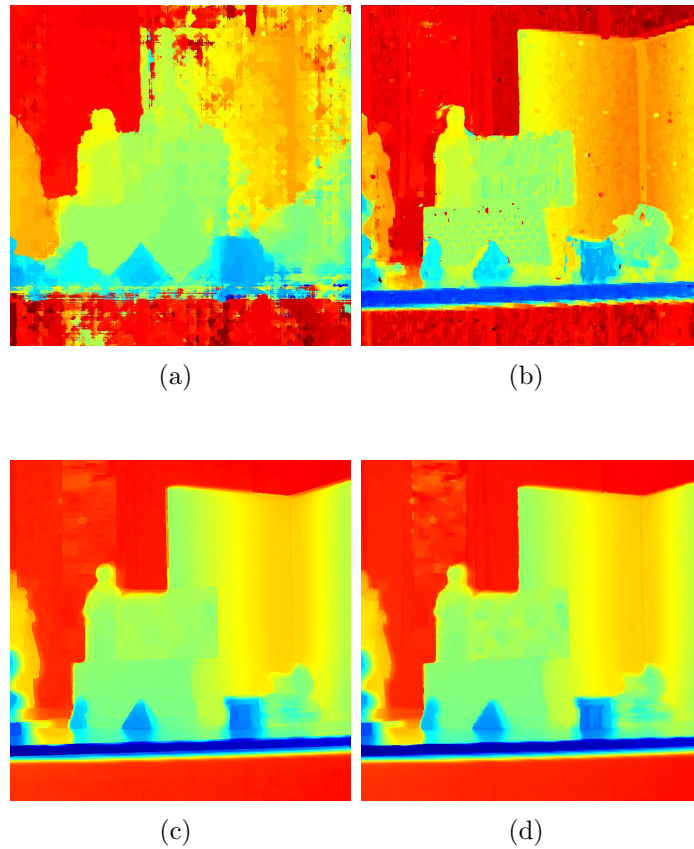
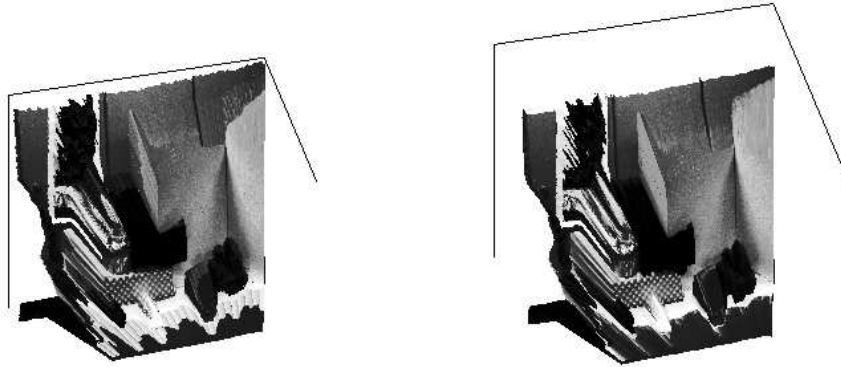


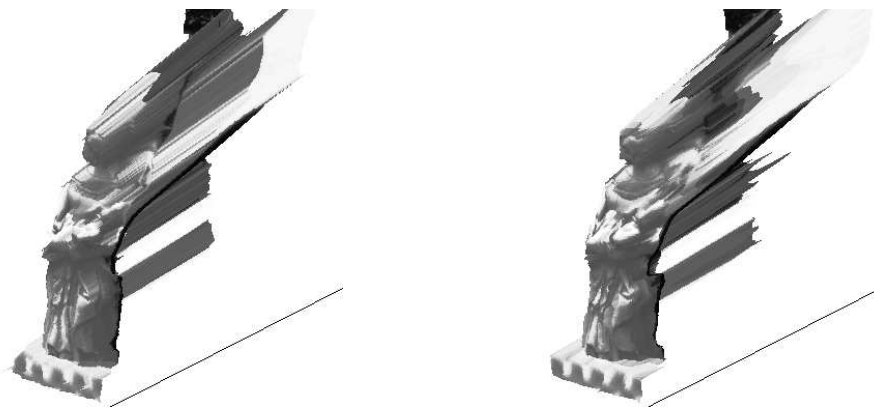
Figure 8.43: [Best viewed in color.] Comparison of different approaches: (a) SGM stereo [75], (b) local method from [118], (c) variational approach from [118] and (d) the proposed approach. We observe that the global approaches, due to the underlying regularization, provide much smoother and more homogeneous results than SGM and the local methods. This comes with the cost of larger computational effort.



(a) variational approach from [118]

(b) described method

Figure 8.44: 3D surface of the *statues* scene reconstructed from the smoothed depth map. Differences between the method [118] and the one describe above are hardly visible (showing the need for a quantitative comparison).



(a) variational approach from [118]

(b) described method

Figure 8.45: Close-up of the 3D surface of the right statue of Virgin Maria in Fig. 8.44. Only slight differences between the method [118] and the one describe above are visible, e.g. the shape of the head of Maria.

we investigate the influence of the choice of regularizer on the quality of the reconstruction. We utilize two different data sets, the *statues* data set (cf. Fig. 8.41) and the *HCI box*<sup>1</sup> (cf. Fig. 8.46), which already has been considered in Section 8.3 for the task of denoising ToF data.

The results depicted in Fig. 8.43 give an impression on how the variational methods, the one proposed in [118] and the one proposed above, compare to the local and the semi-global method. Due to the fact, that variational methods provide a global minimum and due to the smoothing property of the regularizer, the results of the latter are much more regular than those of the local approach and SGM. Comparing our previous approach from [118] with the one proposed above, only minor differences show up in the disparity maps as well as in the back-projected 3D data (see Figs. 8.44 and 8.45), respectively.

Evaluating the different methods on the second data set, the *HCI box*, see Fig. 8.47, we come to the same conclusion. The results of the TV-based variational methods are more regular than those of the simple upsampling, local method and SGM, while difference inbetween this class are noticeable only after close inspection. Main differences are occurring at the ramp, the edge of the lowest stair step, the border between side walls and backside of the box and the polystyrene objects.

Obviously the difference between the variational approaches are too small to compare/rank them just based on visual inspection. Instead, we resort to a quantitative evaluation on the second data set, the *HCI box*, for which ground truth data is available. We recall that, when we considered the task of denoising ToF data in Section 8.3, we refrained of using the ground truth data for quantitative evaluation because strong systematic errors are present, which were not removed by the denoising approaches. Here, however, the situation is different. By complementing the ToF camera with a stereo imaging system, we are able to cope with various errors in the ToF data except for the already mentioned the multi-path error (cf. Remark 8.4). After excluding regions affected by this multi-path error, which for the HCI box are mainly the right and bottom inner wall, the ground truth data are suitable to evaluate our results.

Table 8.9 shows the error to the ground truth data for various methods including simple upsampling of the ToF data, the local fusion approach from [118], semi-global matching [75], a variational approach applied to the ToF data only (i.e. using only the data term  $\mathcal{S}_{\text{ToF}}$ ) as well as variational approach with the coupled ToF and stereo terms using several variants of TV regularization. The quantitative evaluation shows that the variational fusion approaches produce a considerably smaller error than the competing methods. Concerning the kind of regularization used, one observation is that both the steps of introducing adaptivity and of introducing second-order, lead to a decrease of the error. Finally, this quantitative comparison shows that the proposed approach with anisotropic second-order TV provides the best result.

---

<sup>1</sup>available online at <http://hci.iwr.uni-heidelberg.de/Benchmarks/document/hcibox>



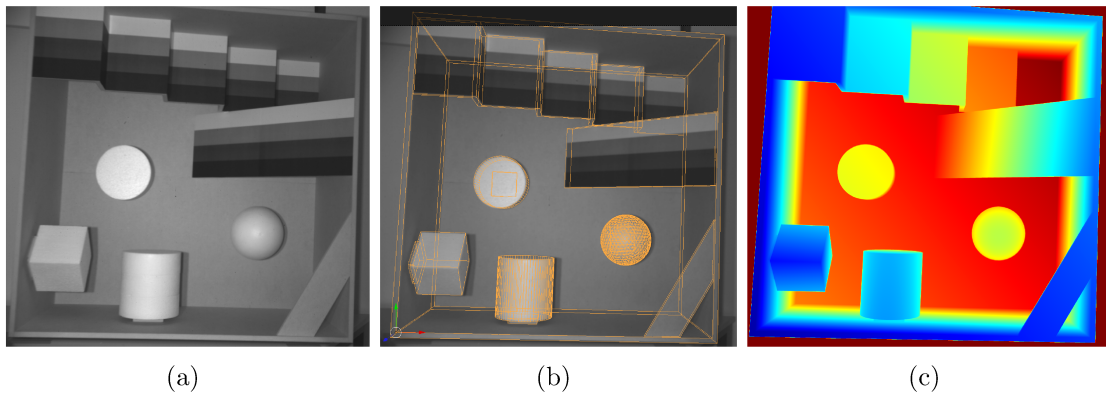


Figure 8.46: *HCI box*. (a) grayscale image, (b) overlay with the ground truth 3D mesh projected onto the image plane, (c) ground truth depth

Data	Method	Regularization	Error
ToF-data	upsampling	—	2.755
ToF-data	global method	standard second-order TV	2.916
ToF-data	global method	adaptive second-order TV	2.731
Stereo	SGM [75]	penalty on disparity differences	2.924
Fusion	local method [118]	—	2.783
Fusion	global method	standard TV	2.543
Fusion	global method	anisotropic first-order TV	2.485
Fusion	global method	second-order TV	2.526
Fusion	global method	adaptive second-order TV [118]	2.563
Fusion	global method	anisotropic second-order TV as described above	<b>2.451</b>

Table 8.9: Summary of GT evaluation (disparities) on regions without multi-path errors. The proposed approach with anisotropic second-order TV regularization gives the best result. Please note that in particular the result of anisotropic second-order TV regularization improves compared to anisotropic first-order TV.

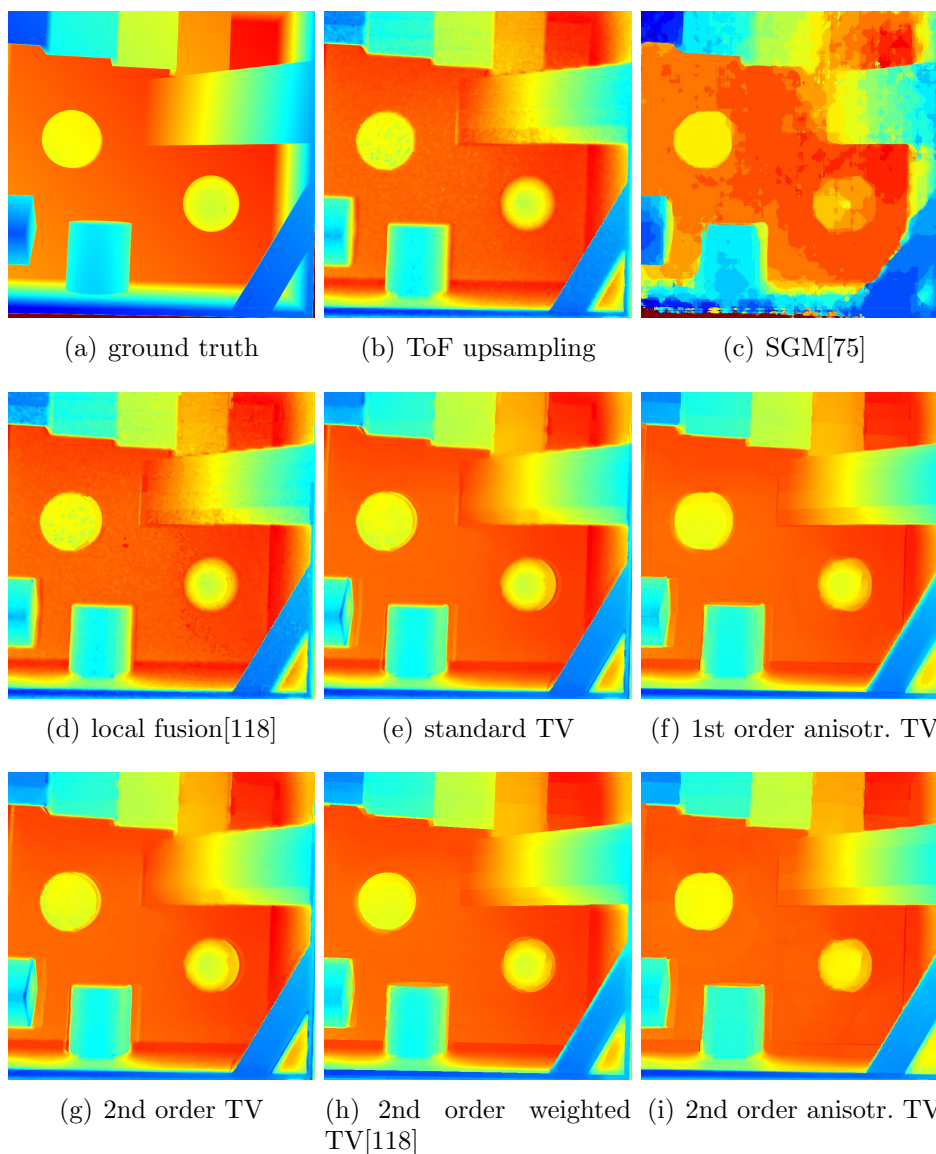


Figure 8.47: Reconstructed disparities for the HCI box data set. Different state-of-the-art approaches and a variational approach with different regularizers are visually compared to ground truth (top left). The variational approaches using adaptive second-order TV provide the best results. For a quantitative comparison, we refer to Table 8.9.

## 8.5 Upsampling of Color Images

### 8.5.1 Problem Description

In this section, we consider the task of *image upsampling*, also referred to as *zooming in*. The problem statement is as follows. We assume an equi-distant two-dimensional grid with nodes  $x_{i,j} \in \mathbb{R}^2$  and indices  $i = 1, \dots, N_x$  and  $j = 1, \dots, N_y$ , which covers a rectangular image domain  $\Omega \subset \mathbb{R}^2$ . Assigned to each grid node  $x_{i,j}$  we consider some discrete image data  $z_{m,i,j}$ ,  $m = 1, \dots, M$ , where  $M$  denotes the number of color channels of the image. The aim is to determine a vector valued function  $u : \Omega \rightarrow \mathbb{R}^M$ ,  $u(x) = (u_1(x), \dots, u_m(x))^\top$ , which has a certain regularity and interpolates the discrete data  $z_{m,i,j}$ . In a mathematically strict sense, this means that  $u_m(x_{i,j}) = z_{m,i,j}$  for each  $(m, i, j)$ . We will, however consider a slightly generalized concept of interpolation constraints, which acknowledges the fact that camera sensor perform a certain local averaging of the recorded image. This concept will be described in detail in the next section.

We remark that our approach can be easily adapted to the general case of image interpolation, where the domain  $\Omega$ , on which the image is to be interpolated and the pixel grid of the input image might not be congruent. The latter for example is the case when an image has to be rotated by an arbitrary angle and where the color values of the at the rotated pixel grid are determined by interpolation. Moreover, our restriction to two-dimensional data is only for the simplicity of notation. An adaptation to higher-dimensional data is straightforward.

### 8.5.2 Interpolation Constraints

Let

$$\Omega := \left(\frac{1}{2}, N_x + \frac{1}{2}\right) \times \left(\frac{1}{2}, N_y + \frac{1}{2}\right), \quad (8.54)$$

where  $N_x, N_y \in \mathbb{N}$ . The domain is partitioned into cells ('pixels')

$$Q_{i,j} := \left(i - \frac{1}{2}, i + \frac{1}{2}\right) \times \left(j - \frac{1}{2}, j + \frac{1}{2}\right) \quad (8.55)$$

for indices  $i = 1, \dots, N_x$  and  $j = 1, \dots, N_y$ . By  $(x_{i,j}, y_{i,j})$  we denote the center of the cell  $Q_{i,j}$ .

Let  $z \in \mathbb{R}^{M \times N_x \times N_y}$  denote the discrete data of a given image, where  $(i, j), i \in \{1, \dots, N_x\}, j \in \{1, \dots, N_y\}$  denotes the pixel index, and  $m \in \{1, \dots, M\}$  denotes the channel index. We assume that the data  $u^0$  have been sampled from an (unknown) function  $u_{\text{true}} : \Omega \rightarrow \mathbb{R}^M$ . We denote the Cartesian product of the indices  $(i, j, m)$  by  $\mathcal{I}$ .

The task is to find a function  $u : \Omega \rightarrow \mathbb{R}^M$  which satisfies interpolation constraints based on the discrete data  $z$ . Interpolation in the classical sense is achieved if  $u$  fulfills  $u_m(x_{i,j}, y_{i,j}) = z_{m,i,j}$ .

Besides the classical interpolation constraint, we consider an alternative concept for image interpolation proposed by [70, 111], which accounts for the fact that the cells of an image sensor have a certain extend, so that the intensity is averaged over the cell area. This concept can be briefly characterized as follows. Let  $G$  be a kernel function defined on  $\mathbb{R}^2$  and compactly supported in  $[-\frac{1}{2}, \frac{1}{2}]^2$ . We consider the convolution of  $G$  with some function  $u$  representing an arbitrary image. This convolution models the integration process on the sensor. We denote this convolution, which is performed channel-wise, by  $G * u$ . Sampling the resulting function  $G * u$  at pixel positions  $(i, j)$  results in a discrete data set  $Z := (z_{m,i,j})_{m,i,j}$  with

$$z_{m,i,j} := (G * u_m)(x_{i,j}, y_{i,j}), \quad \text{for } (i, j, m) \in \mathcal{I}. \quad (8.56)$$

Examples for kernel functions typically used in the literature are listed in [139]. The case of classical interpolation is retrieved by setting

$$G(x, y) = \delta(x)\delta(y), \quad (8.57)$$

where  $\delta(x)$  is the delta distribution. We rewrite (8.56) as follows: Let

$$G_{i,j} := G(\cdot - (x_{i,j}, y_{i,j})), \quad (8.58)$$

then for each  $(i, j, m) \in \mathcal{I}$

$$z_{m,i,j} = \langle G_{i,j}, u_m \rangle_{L^2(\Omega)}, \quad (8.59)$$

where

$$\langle u, v \rangle_{L^2(\Omega)} = \int_{\Omega} u \cdot v \quad (8.60)$$

is the  $L^2$  inner product.

We say that a function  $u : \Omega \rightarrow \mathbb{R}^M$ ,  $u = (u_1, \dots, u_M)^\top$  satisfies the interpolation constraints for some discrete data  $Z = (z_{m,i,j})$ , if

$$\langle G_{i,j}, u_m \rangle_{L^2(\Omega)} = z_{m,i,j}. \quad (8.61)$$

The set of functions satisfying the interpolation constraints for data  $Z$  is denoted by  $\mathcal{U}_{Z,G}$ .

### 8.5.3 Related work

Since the problem of image interpolation can be considered one of the fundamental tasks in image processing, it has been intensively studied in the literature. For a detailed insight into this topic, we refer to the two surveys [113, 163], which cover the topic approximately up to the year 2000. For more recent work, we refer to the papers [41, 79, 83, 88] and the discussion of related work therein.

In the following, we will derive an approach using partial differential equations (PDE) to solve the task of image processing. PDE based methods have also been focus of previous work, see e.g. [19, 139, 140, 166]. The difference between the approaches in [139, 140, 166] and ours are the particular PDEs used for filtering: [139, 140, 166] use *anisotropic diffusion*, whereas we will utilize a PDE introduced in Section 5.3, which generalizes the *mean curvature flow* (MCF).

### 8.5.4 Approach

The simplest model one can think of for image interpolation is the nearest neighbor interpolation, where the color of each node of the high resolution grid is set to the color of the nearest node of the low resolution grid, as considered in the following example.

**Example 8.5.** *We consider for  $G$  the two-dimensional  $\delta$  distribution, i.e.,*

$$G(x, y) = \delta(x)\delta(y). \quad (8.62)$$

*Then  $z_{m,i,j} = u_m((x_{i,j}, y_{i,j}))$ . The component-wise nearest neighbor interpolation provides a function  $u^0 : \Omega \rightarrow \mathbb{R}^M$ , such that*

$$u_m^0|_{Q_{i,j}} = z_{m,i,j}, \quad (i, j, m) \in \mathcal{I} \quad (8.63)$$

*is piece-wise constant on each cell  $Q_{i,j}$ .*

*We can interpret the interpolation error to the true function  $u_{\text{true}}$  as a displacement error as introduced in Section 8.2. Recall that we model a displacement error in terms of a displacement field  $\Phi : \Omega \rightarrow \Omega$ , such that a distorted function  $u^0(x, y)$  is derived from the original function  $u$  by  $u^0(x, y) = u(\phi(x, y))$  (cf. (8.8)). Considering the displacement field*

$$\phi(x, y)|_{Q_{i,j}} := (x_{i,j}, y_{i,j}), \quad (8.64)$$

*we find that  $u^0 = u \circ \phi$  is satisfied in the case of nearest neighbor interpolation.  $\diamond$*

The above example motivates the following approach for interpolating images.

1. Perform an initialization step, where a function  $u^0 \in \mathcal{U}_{Z,G}$  satisfying the interpolation constraints is determined by some basic interpolation approach, e.g. nearest neighbor interpolation or interpolation by cubic splines.
2. Interpret the difference between  $u^0$  and the original function  $u_{\text{true}}$  as a displacement error.
3. Adapt the evolution equation (8.27), which is motivated for smoothing images with displacement errors, so that the solution  $u(t)$  for initial value  $u(0) = u^0 \in \mathcal{U}_{Z,G}$  stays in the set  $\mathcal{U}_{Z,G}$ .

In the remainder of this section we show the adaptation to guarantee  $u(t) \in \mathcal{U}_{Z,G}$ .

Now let  $u^0 \in \mathcal{U}_{Z,G}$  be arbitrary. The nearest neighbor interpolation in Example 8.5 motivates the assumption that, for a sampled function  $u$ , there exists  $\phi$  such that  $u^0 = u \circ \phi$ . Consequently, we can interpret  $u^0$  as a function with sampling errors. Recalling the approach of displacement regularization presented in Section 8.2 we consider the functional (8.24) restricted to the set  $\mathcal{U}_{Z,G}$  in order to reconstruct  $u$  from given data  $u^0$ . In turn, we restrict the flow equation (8.27) to  $\mathcal{U}_{Z,G}$ :

$$\partial_t u = P_{\mathcal{U}_{0,G}} \left( (\nabla u^\top \nabla u + \varepsilon \text{Id}) \nabla \cdot \left( \frac{\nabla u}{\|\nabla u\|_2} \right) \right), \quad (8.65)$$

where

$$P_{\mathcal{U}_{0,G}}(v) := v - \|G\|_{L^2(\mathbb{R}^2)}^{-2} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \langle G_{i,j}, v \rangle_{L^2(\Omega)} G_{i,j} \quad (8.66)$$

is applied on each component separately. Please note that the assumption  $u^0 \in \mathcal{U}_{Z,G}$  together with  $\partial_t u \in \mathcal{U}_{0,G}$  asserts that the solution  $u(t)$  stays in  $\mathcal{U}_{Z,G}$  for all  $t \geq 0$ .

Since the PDE (8.65) comprises a projection, for a the numerical solution a time-explicit scheme with sufficiently small step size  $\Delta t$  is required.

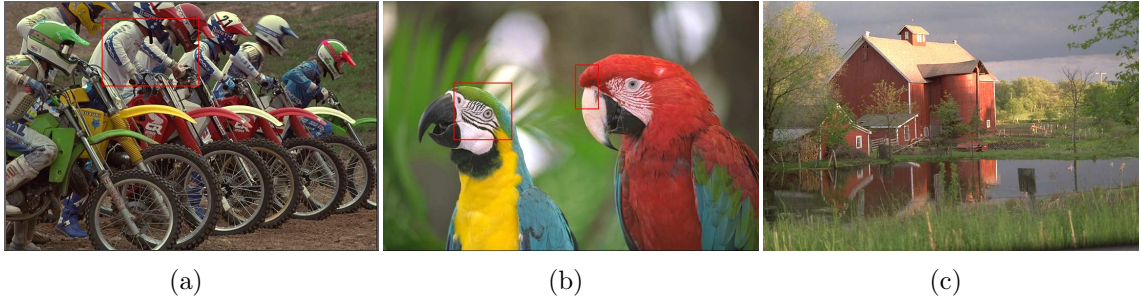


Figure 8.48: Test images *motocross* (a), *parrots* (b) and *house* (c) (taken from the Kodak image collection, see e.g. <http://r0k.us/graphics/kodak/>). Each test image is available in a low and a high resolution version with a factor of four between both resolution. The red boxes in the first two images indicate regions, which will be used to qualitatively evaluate the proposed approach, see Figs. 8.49 to 8.51.

### 8.5.5 Experiments

We compare our method, which numerically solves (8.65), to two standard interpolation methods, namely nearest neighbor and cubic interpolation, as well as to the state-of-the-art interpolation methods proposed by Tschumperlé & Deriche [167] and by Roussos & Maragos [140]. The method of Tschumperlé & Deriche is implemented in the GREYCstoration software (see <http://cimg.sourceforge.net/greycstoration/>), test results for the method of Roussos & Maragos are available from the web page <http://cvsp.cs.ntua.gr/~tassos/PDEinterp/ssvm07res/>.

In our method, the kernel function has to be chosen appropriately. We use

$$G(x, y) := \frac{1}{\int_{[-1/2, 1/2]^2} g_\sigma(x, y) dx dy} \chi_{[-1/2, 1/2]^2} g_\sigma(x, y), \quad (8.67)$$

where  $g_\sigma$  is the two-dimensional isotropic Gaussian kernel with standard deviation  $\sigma$ .

The initial data  $u^0$  are obtained from a sinc-interpolation satisfying the interpolation constraints introduced in Section 8.5.4.

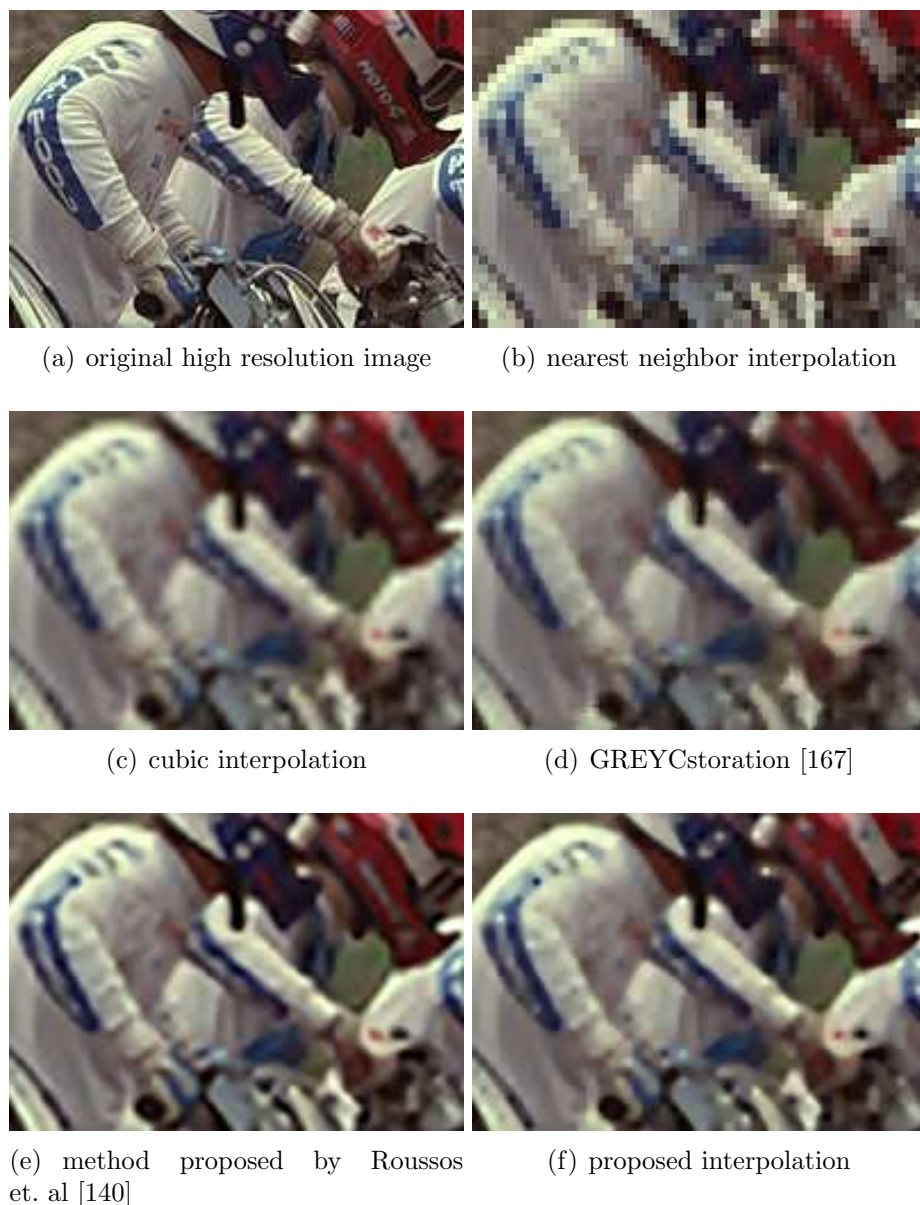


Figure 8.49: [Best viewed in color.] Detail of the first test image after upsampling by a factor of four using standard and state-of-the-art methods. As can be expected, the nearest neighbor interpolation produces blocky results with high contrast. Cubic interpolation as well as GREYCstoration [167] produce blurry results. The method by Roussos et al. [140] and the proposed method provide the best results. Differences between the latter are merely visible on this test image, but will be discussed based on close-ups of the second test image *parrots*, cf. Figures 8.50 and 8.51.



Figure 8.50: [Best viewed in color.] Top row: detail of an edge in the original high resolution image (a) and interpolated images, using GREYCstoration with interpolation constraints (b), Roussos' method (c), and the proposed method (d). Bottom row: results of subsequently applying the Sobel operator to the interpolated images. The proposed method provides the sharpest and most regular reconstruction of the edge.



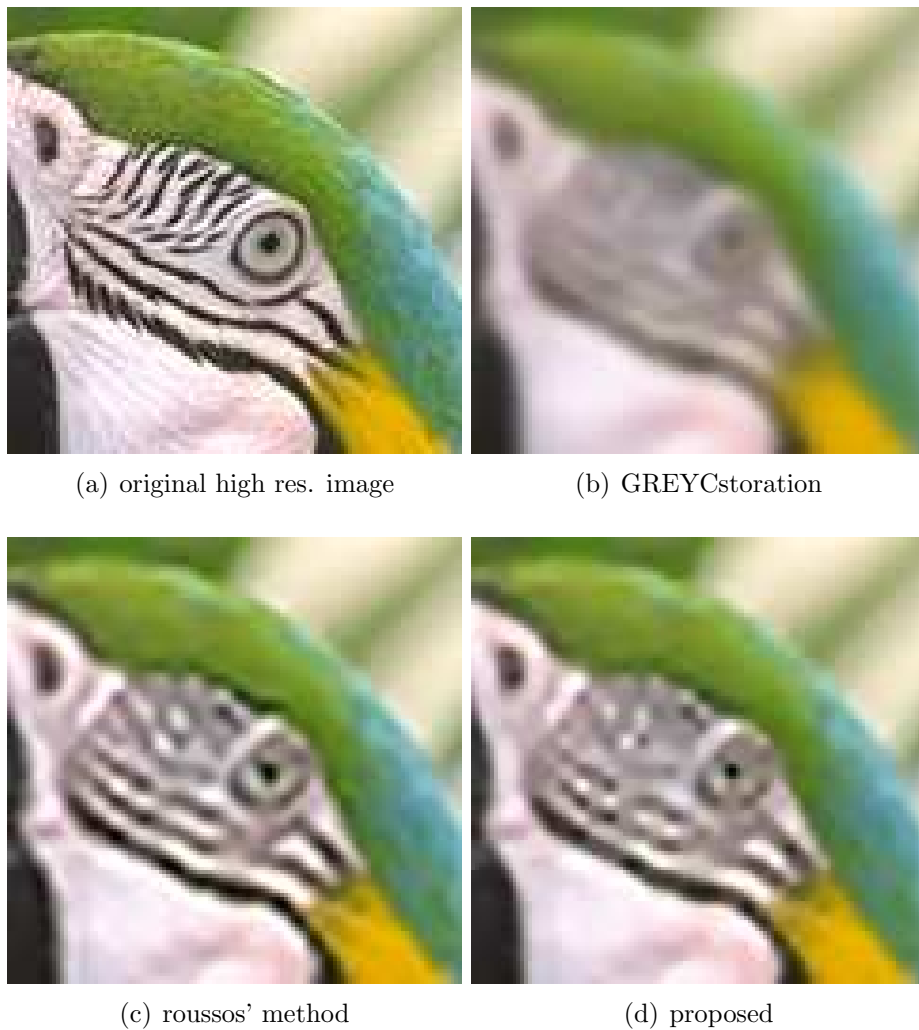


Figure 8.51: [Best viewed in color.] A texture detail of the original (a) and interpolated images using GREYCstoration (b), Roussos' method (c) and the proposed method (d). While GREYCstoration produces a blurry result, Roussos' method and the proposed approach provide satisfactory results. Differences can be observed at the black and white stripes and the parrot's eye. The proposed method reconstructs the striped region less smooth than Roussos' method, which on the other hand seems to enhance one specific orientation, that is the one of the three longer black stripes.

For evaluating the different methods, we use the three test images shown in Fig. 8.48. For each image, a low and a high resolution version, denoted by  $u^0$  and  $u_{\text{true}}$  respectively, are available, where the low resolution image is obtained from the high resolution image via low-pass filtering (convolution with a bi-cubic spline) and down-sampling by a factor of four. The test images were obtained from <http://cvsp.cs.ntua.gr/~tassos/PDEinterp/ssvm07res/>.

The methods mentioned above are used to upsample the low resolution image by a factor of four.

Our method is applied with 100 time steps,  $\Delta t = 0.03$ ,  $\varepsilon = 0.05$  and  $\sigma^2 = 20$  for the first and 100 time steps,  $\Delta t = 0.05$ ,  $\varepsilon = 0.01$  and  $\sigma^2 = 20$  for the second test image, respectively. Except if noted otherwise, we use the PDE (8.65), which corresponds to the choice  $p = 1$  in (5.22).

For GREYCstoration (version 2.9) we use the option '-resize' together with the aimed size of the high resolution image and parameters '-anchor true', '-iter 3' and '-dt 10'. For the remaining parameters the default values are used. The results of Roussos' method were obtained from the web page mentioned above.

Before we turn to the results of the mentioned methods, we remark that presenting the results side by side in their full image size requires a downsampling, so that differences between the individual approaches are no longer visible. We therefore compare the methods based on specific image regions displayed in high resolution.

Let us consider the results of upsampling the first test image (see Fig. 8.48 (a)). For this image we use one specific region of the image to highlight the differences between the methods only details of the resulting images, see Fig. 8.49. The results with nearest neighbor and cubic interpolation are shown in Fig. 8.49, (b) and (c), respectively. Both results are unsatisfactory and confirm, what is well known from the literature, that by nearest neighbor interpolation the upsampled images look blocky and cubic interpolation produces blurry images. The result of GREYCstoration with interpolation constraints (Fig. 8.49 (d)) also appears blurry, but compared to cubic interpolation better reconstruct the edges in the image. The method proposed by Roussos & Maragos as well as our method (see Fig. 8.49 (e) and (f)) produce sharp and well reconstructed edges.

Concerning our proposed method, we like to remark that this method can be slightly varied. This modification is based on the energy functional (5.22) from Section 5.3 with the choice  $A(u) := ((\nabla u^\top \nabla u)^{\frac{1}{2}})^\dagger$  and the corresponding PDE (5.38). With this modification, the results are slightly smoother than with PDE (8.27), but the difference is merely visible. For this reason the results of the former are not depicted here.

In order to further investigate the differences between the PDE based methods, we zoom into two regions of the second test image (see Fig. 8.48 (b)), one region containing an edge, see Fig. 8.50 and one region with texture, see Fig. 8.51.

Fig. 8.50 shows the edge region after applying the methods proposed by Tschumperlé with interpolation constraints (b), Roussos (c) and our method (d). For comparison we have plotted also the detail of the original image (a). One can see that

by Tschumperlé's method the edges appear blurry and irregular. This seems to be an effect of the interpolation constraints, because when Tschumperlé's method is applied without constraints, strong anisotropic diffusion orthogonal to the image gradient enhances the edges. By the method of Roussos the edge is reconstructed in a sharp way, but overshoots appear. Our method is also able to reconstruct the edge sharply but with little overshoots. Concerning the gray mark at the parrot's beak, we observe that Tschumperlé's method reconstructs the shape of the mark better than the other methods do.

The differences in the behavior of the methods can also be recognized when applying the Sobel operator to the interpolated images, see Fig. 8.50, bottom row. The thickness of the edges in the result of the Sobel operator indicates the blurriness of the reconstructed edge. We see that the proposed method produces sharper edges than the method by Roussos and more regular edges than the method by Tschumperlé. The overshoots introduced by Roussos' method can also be observed in the outcome of the Sobel operator. They are far stronger than the overshoots produced by our method.

Now we investigate the effect of the interpolation methods on textures. Fig. 8.51 (a) shows a textured region of the original image. The results of the methods proposed by Tschumperlé (with interpolation constraints) and Roussos are given in Fig. 8.51 (b) and (c), respectively. The result of the proposed method is shown in Fig. 8.51 (d). One observes a certain blurriness in the results by Tschumperlé's method. As for the result before, we point out that incorporating the interpolation constraints seems to have a strong effect on the result. When applying GREYC-storation without imposing constraints, the results are much more influenced by the anisotropic diffusion and the edges and the texture are accentuated. In the result of the interpolation method proposed by Roussos, we see a strong effect of the anisotropic diffusion on the texture, so that the result is more visually appealing than the other results. Nevertheless, a comparison with the original image shows that original and reconstructed texture differ significantly. In particular the orientations of the short stripes in the face of the parrot are different. Note that the anisotropic diffusion induced by the direction of the texture also affects the pupil of the parrot. On the result of our method we remark that the reconstruction of the texture is quite conservative, i.e., we stay near the initial guess. The blockyness is slightly reduced by the evolution process. Taking a look at the eye of the parrot, the relation of our method to mean curvature flow can be observed: The pupil is reconstructed as a perfectly circular shape.

We also provide a quantitative error measure to compare the proposed method with standard and state-of-the art interpolation methods.

To this end we choose the three different test images shown in Figure 8.48 and seven different interpolation methods: nearest neighbor interpolation, cubic interpolation, Lanczos interpolation, the methods proposed by Tschumperlé [167] and by Roussos [140] and the proposed method. For the latter we also apply the modified version with  $p = 1/2$ .

$BV(u - u_{\text{true}})$	Motocross	Parrots	House
Nearest n.	127.24	40.67	69.53
Cubic	114.66	33.46	64.38
Lanczos	119.10	34.99	66.15
Tschumperlé	120.45	37.35	67.89
Roussos	109.04	31.60	62.74
proposed $p = 1/2$	108.89	31.58	<b>62.68</b>
proposed $p = 1$	<b>108.04</b>	<b>31.54</b>	62.72

Table 8.10: Difference between the interpolated image and the high-resolution counterpart, measured in the  $BV$ -norm for the different test images and interpolation methods.

We measure the difference between an interpolated image  $u$  and the corresponding high-resolution image  $u_{\text{true}}$  with respect to the (discrete)  $BV$ -norm,

$$\|v\|_{BV} := \sum_{i,j} |v_{i,j}| + \sum_{i,j} \sqrt{(v_{i+1,j} - v_{i,j})^2 + (v_{i,j+1} - v_{i,j})^2}, \quad (8.68)$$

see Table 8.10. We have chosen the  $BV$ -norm for the following reasons:

- Using an  $L^p$ -norm of  $u - u_{\text{true}}$  is inappropriate, since this error measure is biased by the interpolation constraint.
- The above discussion shows that the methods act differently mainly in regions, where the absolute value of  $\nabla u_{\text{true}}$  is high, i.e., edges and textures. A suitable error measure should therefore incorporate the gradient.
- The  $BV$ -norm is known to be a suitable norm for image processing purposes.

Table 8.10 shows that the proposed method achieves optimal interpolation errors with respect to the  $BV$ -norm. Only Roussos' method is capable of providing comparable error levels.

## 8.6 Joint Estimation of Scene and Egomotion

In the following, we recall an approach presented in [16, 17], which jointly estimates scene and egomotion from a sequence recorded by monocular camera mounted in a vehicle. We assume that the observed scene is static and that the vehicle drives towards or through the given scene. Moreover, due to a possibly high speed of the vehicle, we assume that large displacements of objects in the subsequent frames occur.

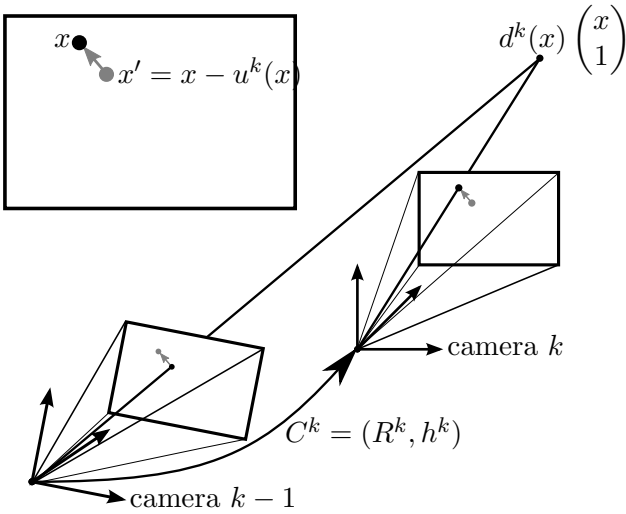


Figure 8.52: Optical flow  $u^k$  induced by given camera motion  $C^k = (R^k, h^k)$  and scene depth  $d^k$ .

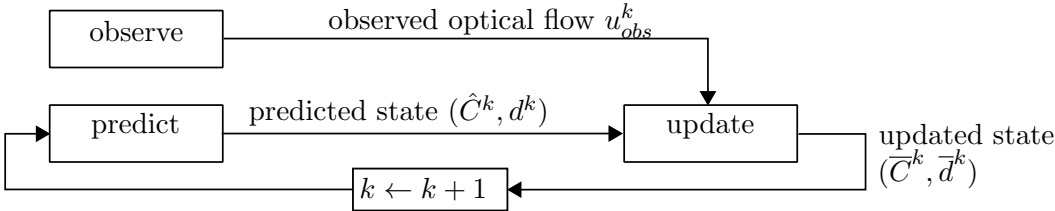


Figure 8.53: Illustration of the prediction-update loop. In each iteration  $k$  the optical flow  $u_{\text{obs}}^k$  between  $I^{k-1}$  and  $I^k$  is observed. The update step takes into account the prediction  $(\hat{C}^k, \hat{d}^k)$  as well as the observation  $u_{\text{obs}}^k$  to determine a final estimate  $(\bar{C}^k, \bar{d}^k)$ .

### 8.6.1 Task

Given an image sequence  $I_k, k = 1, \dots, K$  from a moving camera. We assume that the camera moves with approximately constant speed, but with possibly curved trajectory and, as mentioned before, that the scene is static.

The task is to estimate egomotion in terms of camera translation  $h^k$  and rotation  $R^k$  from frame  $k - 1$  to  $k$  and scene depth  $d^k$  with respect to the frame  $k$ .

### 8.6.2 Related Work

There are four features of the approach described below, based on which our method can be distinguished from related work:

1. We perform a joint estimation of scene (depth) and egomotion (camera translation and rotation).
2. We consider a monocular approach (using an image sequence from a single moving camera).
3. our approach provides a *dense* estimation (to be distinguished from *sparse* approaches in the literature).
4. it is an *online* approach, i.e. it handles the image sequence in one run through the whole sequence, only processing the current frame and its predecessor at the same time (to be distinguished from *batch* approaches).

Various approaches have been proposed which tackle only one of the tasks of egomotion or scene estimation.

Concerning approaches that estimate egomotion only, we exemplarily refer to [85, 126] for methods using sparse features and to [46] for an approach with dense correspondences.

For the task of scene estimation, the most common strategy is the well studied stereo approach [62, 75, 179]. Note that the stereo approach by Geiger [62] is especially designed for stereo data sets as they appear in automotive applications. Stereo approaches can easily be adapted to estimate scene flow based on sequences [133, 169, 174].

The dense reconstruction of the scene depth based on arbitrary multiple views as been studied e.g. in [66, 123, 160]. These methods, however, require information about the camera motion to be estimated in a pre-processing step.

Among the methods jointly estimating egomotion and scene, there exist various based on sparse tracked features. The most commonly used techniques here are factorization methods (cf. [161]) and bundle adjustment (cf. [165]). The latter in contrast to online approaches process the whole sequence at once. Two recently local and efficient approaches are proposed in [86, 116]. The approach by Lin et al. [107] used only two frames to estimated both camera poses and estimation of the 3D scene by a point cloud.

Methods most related to our approach, since they estimate jointly and densely egomotion and scene, are the approaches by Weishaupt et al. [176] (with restrictions on the camera rotation), Irani et al. [78] (using reference planes for image alignment), Newcombe et al. [124], Bagnato et al. [11] (omni-directional camera) and Engel et al. [52] (semi-dense disparity maps and egomotion). On the method proposed by Newcombe we remark that it relies on the concept of a reference frame, based to which the other frames are processed. As a consequence the method is better suited for the reconstruction of a fixed scene observed from multiple positions with various views, as in the context of a camera moving fastly through a scene, which would require a frequent change of the reference plane.

Finally, we also want to mention learning based methods. Exemplarily, we refer to the offline learning approached by Lee et al. [89], Saxena et al. [144] and Liu et al. [108], and the online learning approach by Hadsell et al. [72].

### 8.6.3 A Model for Optical Flow

In the following, we assume that for two consecutive frames  $I^{k-1}$  and  $I^k$  the camera motion  $C^k$  in terms of a rotation  $R^k$  and translation  $h^k$  is known as well as the scene depth  $d^k$ . Recall that we also assume the scene to be static. Our goal is to derive a model  $u(R^k, h^k, d^k)$  for the optical flow induced by this camera movement and scene depth. Fig. 8.52 illustrates the setting.

For the sake of notation, we drop the indices  $k$  and  $k-1$  and refer to the variables of the frame  $k$  by  $C, R, h$  and  $d$ . Variables concerning the frame  $k-1$  will be denoted with a primes. We refer to the frame  $k-1$  as the first and to frame  $k$  as the second frame.

Let us now consider a fixed 3D point with coordinates  $X' = (X'_1, X'_2, X'_3)^\top$  with respect to the coordinate system of the first camera position, i.e. the coordinate system with the origin being the camera position and vector  $(0, 0, 1)^\top$  being its viewing direction.

In the coordinate system of the second frame, this point takes the coordinates

$$X = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = R^\top (X' - h). \quad (8.69)$$

If this point is visible in both frames, we can compare the 2D coordinates of this point projected to the two image planes. E.g. for the second camera this projection is given as

$$P_C(X) = \frac{1}{X_3} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}. \quad (8.70)$$

Using normalized coordinates  $x$  together with their corresponding depth  $d(x)$ , we can back-project  $x$  along the viewing ray  $(x, 1)^\top$  and  $x'$  analogously along  $(x', 1)^\top$  to find  $X$  and  $X'$ :

$$X = d(x) \begin{pmatrix} x \\ 1 \end{pmatrix}, \quad X' = d'(x') \begin{pmatrix} x' \\ 1 \end{pmatrix}. \quad (8.71)$$

Combining (8.69) and (8.71) yields

$$d(x) \begin{pmatrix} x \\ 1 \end{pmatrix} = R^\top \left( d'(x') \begin{pmatrix} x' \\ 1 \end{pmatrix} - h \right). \quad (8.72)$$

Now, we define the displacement  $u(x)$  of points  $x$  and  $x'$  with respect to the second camera frame by

$$u(x) = x - x'. \quad (8.73)$$

Using (8.72) and the camera projection  $P_C$  with respect to the second camera (cf. (8.70)), we obtain the following formula for  $u(x)$ ,

$$u(x; R, h, d) = x - P_C \left( d(x)R \begin{pmatrix} x \\ 1 \end{pmatrix} + h \right). \quad (8.74)$$

Please note that  $u(x; R, h, d)$  describing the mapping from point  $x'$  in the first frame to  $x$  in the second frame is defined as a function for  $x$ . Consequently, we consider optical flow to only those points, which are visible in the second frame. Assuming that the camera viewing direction coincides with the moving direction, which is true in automotive scenarios, and that the rotation component of the camera motion is small, any such point is also visible in the first frame. For this reason it is more suitable to use the flow (8.74) than the flow defined in coordinates  $x'$ .

In the approach considered below we will compare the model flow  $u(X; R, h, d)$  with the flow observed from the image sequence  $(I^k)_k$ . Thus, information of the two input frames is only used indirectly in our variational approach. An alternative would be to directly compare the first frame to the image obtained by warping the second frame according to the flow  $u(x; R, h, d)$ , i.e. to compare  $I^{k-1}(x - u(x))$  with  $I^k$  (typically after linearization), as it is for example proposed in [11, 52, 176]. The reason for choosing the former alternative is, that we will consider a stochastic approach, in which we can make use of a measure of uncertainty for the observed optical flow.

**Remark 8.6** (Epipolar geometry). *When searching for matching points  $x$  and  $x'$  in two views of a static scene, the epipolar geometry (cf. e.g. [73]) is of importance. Before describing this issue, let us first define the epipoles for each of the two views as follows. Let  $X_{C_1}$  and  $X'_{C_2}$  be the 3D coordinates of the camera positions the two frames are taken from, both with respect to the coordinate system of the other camera. We denote the projection of  $X'_{C_2}$  onto the image plane for the first camera position by  $e'$  and the projection of  $X_{C_1}$  onto the image plane for the second camera position by  $e$ . As a fact, for static scenes, if we consider a 3D point  $X'$  matching  $x$  in the second frame, then its projection onto the first frame always has to lie on an epipolar line, i.e. a line between  $x'$  and  $e'$  in the first frame, and analogously for points  $X$  matching  $x'$ . In our approach, an epipolar line can be retrieved by keeping  $x$  and  $C = (R, h)$  in (8.74) fixed and allowing  $d$  to vary between 0 and  $+\infty$ . Then,  $x' = x - u(x; R, h, d)$  lies on an epipolar line. For this reason, our approach respects the epipolar geometry of the camera setup.*



### 8.6.4 Approach

The proposed approach follows an online strategy, by which the image sequence is processed in one sweep. This strategy is in view of allowing a future on-board implementation, since it in principle allows an infinitely long image sequence. Our approach follows the structure of a recursive Bayesian filter, which is briefly described in the next remark.

**Remark 8.7** (Recursive Bayesian Filtering). *A Recursive Bayesian Filter [6] deals with a sequence of observations  $O^k, k = 0, \dots, K$ , which are linked with latent state variables  $S^k, k = 0, \dots, K$  via a hidden Markov model [51].*

*The Markov assumption is that every state variable  $S^k$  is conditionally dependent only on its predecessor  $S^{k-1}$ , while an observation variable  $O^k$  is only conditionally dependent on the corresponding state variable  $S^k$ . Thus, the system is fully described by the two conditional probability densities  $p(O^k|S^k)$  and  $p(S^k|S^{k-1})$ .*

*The aim of a recursive Bayesian filtering is to find an approximation of the conditional probability density  $p(S^k|O^k)$ , which then can be used to infer e.g. the most likely hidden state  $S^{k_0}$  from observations  $\{O^k, k = 0, \dots, k_0\}$ .*

*The main steps of a recursive Bayesian filtering are a **prediction step**, where*

$$p(S^k|O^{k-1}, \dots, O^0) = \int p(S^k|S^{k-1})p(S^{k-1}|O^{k-1}, \dots, O^0) dS^{k-1} \quad (8.75)$$

*is calculated, and an **update step**, in which the conditional probability density*

$$p(S^k|O^k, \dots, O^0) \propto p(O^k|S^k)p(S^k|O^{k-1}, \dots, O^0) \quad (8.76)$$

*is derived. Both steps are subsequently applied in an iteration loop  $k = 1, \dots, K$ .*

*A widely used specialization of recursive Bayesian filtering is the Kalman filter [81, 177].*

In our approach we assume that the observations  $O^k, k = 0, \dots, K$  are given as the optical flow  $u_{\text{obs}}^k$  between frames  $I^{k-1}$  and  $I^k$ , estimated by means of a standard optical flow algorithm. We refer to  $u_{\text{obs}}^k$  as the *observed* optical flow. Moreover, we assume that this flow can be provided in terms of a Gaussian probability density function, i.e.

$$u_{\text{obs}}^k(x) \propto \mathcal{N}(\mu_{\text{obs}}(x), \Sigma_{\text{obs}}(x)) \quad (8.77)$$

for some  $\mu_{\text{obs}}(x) \in \mathbb{R}^2$  and  $\Sigma_{\text{obs}}(x) \in \mathbb{R}^{2 \times 2}$  is symmetric and positive definite. In particular,  $\Sigma_{\text{obs}}(x)$  serves as an accuracy measure in our approach.

The state variables  $S^k, k = 0, \dots, K$  comprise the camera motion  $C^k = (h^k, R^k)$  and the scene depth  $d^k$ . We refer to Fig. 8.53 for an illustration.

As already indicated, our approach follows the scheme of a recursive Bayesian filter. It consists of an iteration  $k = 1, \dots, K$ , in which a prediction and update step are performed. To ease the discrimination between variables and parameters used in either the prediction or the update step, we use the 'hat' symbol for variables in the prediction step and the 'bar' symbol in the update step, e.g.  $\hat{x}$  and  $\bar{x}$ , respectively.

In the prediction step, we determine a prediction of  $S^k$  in terms of a conditional Gaussian probability density  $p_{\text{pre}}(S^k|S^{k-1})$  with mean  $\hat{\mu}_{S^k}$  and covariance matrix  $\hat{\Sigma}_{S^k}$ :

$$p_{\text{pre}}(S^k|S^{k-1}) = \mathcal{N}(\hat{\mu}_{S^k}, \hat{\Sigma}_{S^k}). \quad (8.78)$$

We will describe the details of this step later on.

In the update step, we exploit the probability density

$$p(S^k|O^k, S^{k-1}) = p(O^k|S^k)p_{\text{prior}}(S^k|S^{k-1}), \quad (8.79)$$

where  $p_{\text{prior}}(S^k|S^{k-1})$  incorporates the prediction  $p_{\text{pre}}(S^k|S^{k-1})$  as a temporal prior as well as a spatial prior  $p_{\text{spatial}}(S^k)$  for the depth  $d^k$ :

$$p_{\text{prior}}(S^k|S^{k-1}) = p_{\text{pre}}(S^k|S^{k-1})p_{\text{spatial}}(S^k). \quad (8.80)$$

The probability density  $p(O^k|S^k)$ , which describes the dependency of the observation to the state variable, will be non-Gaussian. As a consequence,  $p(S^k|O^k, S^{k-1})$  will also be non-Gaussian. We find the actual estimate  $\bar{S}^k$  for  $S^k$  by determining one of the modes of  $p(S^k|O^k, S^{k-1})$ . Performing a Laplacian approximation [164] at the point  $\bar{S}^k$  yields

$$p(S^k|O^k, S^{k-1}) \propto \mathcal{N}(\bar{\mu}_{S^k}, \bar{\Sigma}_{S^k}) \quad (8.81)$$

with mean  $\bar{\mu}_{S^k}$  and covariance matrix  $\bar{\Sigma}_{S^k}$ .

We now detail the choice for each probability density mentioned so far.

**Observation** As observation of  $O^k$  in the  $k$ -th step we consider the empirical optical flow  $u_{\text{obs}}^k$  between frame  $I^{k-1}$  and frame  $I^k$ . As already mentioned, we apply a local Lukas-Kanade approach to obtain this optical flow. In the first iteration, where no image pair  $(I^{k-1}, I^k)$  is available, we set the observation to a zero flow.

In addition to estimating the flow  $u_{\text{obs}}^k$ , we exploit the structure tensor used for determining  $u_{\text{obs}}^k$  to derive also a measure for the certainty of  $u_{\text{obs}}^k$ . We interpret  $u_{\text{obs}}^k$  as the mean of a Gaussian distribution and its certainty as the precision matrix  $(\Sigma_{u_{\text{obs}}})^{-1}$ .

In order to compare the model optical flow  $u(X; R^k, h^k, d^k)$  from (8.74) with the observed one, we evaluate its Gaussian distribution at point  $u(x; R^k, h^k, d^k)$  and obtain

$$\begin{aligned} p(O^k|S^k) &= \frac{1}{C} \sum_x \exp \left( -\frac{1}{2} (u(x; R^k, h^k, d^k) - \mu_{\text{obs}}^k)^\top (\Sigma_{\text{obs}}^k)^{-1} (u(x; R^k, h^k, d^k) - \hat{u}(x)) \right) \end{aligned} \quad (8.82)$$

with a fixed normalization constant  $C > 0$ .

**Spatial prior** Concerning the spatial prior for the depth map  $d^k$  we resort to a classical quadratic regularization of the gradient of  $d^k$ . To this end, we set

$$p_{\text{spatial}}(S^k) = \frac{1}{C} \sum_x \exp\left(-\frac{1}{2\sigma_{\text{spatial}}^2} \|\nabla d(x)\|^2\right) \quad (8.83)$$

with a fixed standard deviation  $\sigma_{\text{spatial}} > 0$  and a normalization constant  $C > 0$ . Experiments show that this choice leads to a robust estimation of the scene depth. We note that for the ego-motion parameters  $C^k = (h^k, R^k)$  the only regularization will be a temporal one.

**Temporal prior** For the temporal prior, we consider the distribution  $p_{\text{pre}}(S^k|S^{k-1})$  defined as in (8.78). Note that  $p_{\text{pre}}(S^k|S^{k-1})$  in general is a joint distribution for the variables  $d^k$  and  $C^k = (h^k, R^k)$ . In our approach we make the assumption that depth  $d^k$  and camera motion  $C^k$  are independent from each other. Consequently,  $\hat{\Sigma}_{S^k}$  has a block structure of the form

$$\hat{\Sigma}_{S^k} = \begin{pmatrix} \hat{\Sigma}_{d^k} & 0 \\ 0 & \hat{\Sigma}_{C^k} \end{pmatrix}. \quad (8.84)$$

To derive the parameters  $\hat{\mu}_{S^k}, \hat{\Sigma}_{S^k}$  is part of the prediction step detailed below.

**Prediction step** In the prediction step, we determine the probability  $p_{\text{pre}}(S^k|S^{k-1})$ , which we assume to be Gaussian. Moreover, we assume  $d^k$  and  $C^k$  to be independent from each other, i.e.  $p_{\text{pre}}(S^k|S^{k-1}) = p_{\text{pre}}(C^k|S^{k-1}) p_{\text{pre}}(d^k|S^{k-1})$ . With these assumptions, it suffices to determine the mean  $\hat{\mu}_{C^k}$  and covariance  $\hat{\Sigma}_{C^k}$  of  $p_{\text{pre}}(C^k|S^{k-1})$  and the mean  $\hat{\mu}_{d^k}$  and covariance  $\hat{\Sigma}_{d^k}$  of  $p_{\text{pre}}(d^k|S^{k-1})$ . On the later as a simplification we assume an independent distribution in each image location, so that  $\hat{\Sigma}_{d^k}$  becomes a diagonal matrix.

For the first iterate of the Kalman filter, the prediction is set using initial values  $\hat{\mu}_{d^0} = 0, \hat{\mu}_{R^0} = \text{Id}$  and  $\hat{\mu}_{h^0} \neq 0$  to a constant value in the range of a reasonable camera speed. For the covariance matrices we use diagonal matrices of the form  $\sigma^2 \text{Id}$  with large constants  $\sigma > 0$  for each random variable. The subsequent predictions ( $k \geq 1$ ) are based on the result of the previous update step.

For the camera motion, assuming an almost constant camera speed, we set  $\hat{\mu}_{C^k} := \bar{\mu}_{C^{k-1}}$ , where  $\bar{\mu}_{C^{k-1}}$  is the mean the distribution of  $C^{k-1}$  obtained from the previous update step. For the covariance matrix for  $C^k$ , we set  $\hat{\Sigma}_{C^k} := \bar{\Sigma}_{C^{k-1}} + \Sigma_C$ , where  $\bar{\Sigma}_{C^{k-1}}$  is again known from the previous update step and  $\Sigma_C$  is a diagonal matrix with diagonal entries  $\sigma_h$  and  $\sigma_R$  for the translational and rotational part, respectively, accounting for some uncertainty introduced in the transition from step  $k-1$  to  $k$ . The predicted distribution of  $C^k$  then is given as

$$p(C^k, C^{k-1}) = \mathcal{N}_{\mathcal{M}^C}(\hat{\mu}_{C^k}, \hat{\Sigma}_{C^k}), \quad (8.85)$$

where  $\mathcal{N}_{\mathcal{M}^C}(\mu, \Sigma)$  is the approximate normal distribution with mean  $\mu$  and covariance matrix  $\Sigma$  on the manifold  $\mathcal{M}^C = SE(3)$  (cf. [17] for details).

In order to find  $\hat{\mu}_{d^k}$ , we will warp the map  $\bar{d}^{k-1}$  from the view of the current step to the view of that camera position, which is mostly likely to occur in the next step. After having set the prediction for  $C^k$  the most likely position is obtained by translating and rotating the camera position by

$$\hat{C}^k := \hat{\mu}_{\hat{C}^k} = \bar{\mu}_{C^{k-1}} \quad (8.86)$$

$$\Leftrightarrow (\hat{h}^k, \hat{R}^k) = (\bar{h}^{k-1}, \bar{R}^{k-1}). \quad (8.87)$$

The warping process is performed as follows. In order to determine  $\hat{d}^k$  at position  $x$  in frame  $I^k$ , we use (8.73) and (8.74) together with  $\hat{C}^k, d^{k-1}$  to find a point  $x'$  in frame  $I^{k-1}$  corresponding to  $x$  in the sense that their two reprojections provide the same point in the 3D scene. We find from (8.73), (8.74) and (8.87) that

$$x' \approx x - u(x; \hat{R}^k, \hat{h}^k, d^{k-1}(x)) = P_C(d^{k-1}(x) \bar{R}^{k-1} \begin{pmatrix} x \\ 1 \end{pmatrix} + \bar{h}^{k-1}). \quad (8.88)$$

By bilinear interpolation of  $d^{k-1}$  at point  $x'$  we find  $d'(x')$ . We then transport the depth  $d'(x')$  to the camera view of the  $k$ -th frame by setting

$$\hat{d}^k(x) := (\bar{R}^{k-1})^\top \left( d'(x') \begin{pmatrix} x' \\ 1 \end{pmatrix} + \bar{h}^{k-1} \right) \quad (8.89)$$

We then choose  $\hat{d}$  to be the mean of the prediction  $p_{\text{pre}}(d^k | S^{k-1})$ .

Finally, the covariance  $\hat{\Sigma}_{d^k}$  is determined as follows. For simplicity, we assume that the covariance matrix is diagonal with entries  $\hat{\sigma}_{d^k}^2(x)$ . In order to derive  $\hat{\sigma}_{d^k}(x)$  for  $\hat{d}(x)$ , we first extract the diagonal entries  $\bar{\sigma}_{d^{k-1}(x)}$  from  $\bar{\Sigma}_{S^{k-1}}$  (from the last update step) and then transport them analogously to  $d^{k-1}(x)$  to the next time step. To account for the uncertainty of this transformation, the resulting variance is increased by adding an additional value  $\sigma_d^2$ ,

$$(\hat{\sigma}^{k-1}(x))^2 := (\bar{\sigma}^{k-1}(x'))^2 + \sigma_d^2. \quad (8.90)$$

With the above setting, the probability density  $p_{\text{pre}}(S^k | S^{k-1})$  is fully determined.

**Update step** In the update step, we determine a mode of  $p(S^k | O^k, S^{k-1})$ . To this end, we solve the non-convex optimization problem

$$\bar{S}^k := \arg \min_{S^k \in X} \mathcal{F}(S^k), \quad \mathcal{F}(S^k) := -\ln p(S^k | O^k, S^{k-1}), \quad (8.91)$$

where  $S^k = (d^k, C^k) \in X = \mathbb{R}_{\geq 0}^n \times SE(3)$  with  $SE(3)$  being the special Euclidean group of  $\mathbb{R}^n$  and  $\mathbb{R}_{\geq 0}^n$  is the set of  $n$ -dimensional vectors with non-negative entries. The components  $\bar{d}^k$  and  $\bar{C}^k$  of  $\bar{S}^k$  are the actual estimates for  $d^k$  and  $C^k$ . Moreover, we determine  $\bar{\Sigma}_{d^k}$  and  $\bar{\Sigma}_{C^k}$ , which are used in the next prediction step.

This concludes the overview over the different steps of our approach.

Please observe that in our prediction step, we estimate the probability density  $p(S^k|O^{k-1}, \dots, O^0)$  only approximately. The reason is that computational efficiency prohibits to perform the integration in (8.75), which is over all possible states  $S^{k-1}$ , to a high accuracy. Instead, we retrieve an approximation of  $p(S^k|O^{k-1}, \dots, O^0)$  by assuming that  $p(S^k|O^{k-1}, S^{k-1})$  is Gaussian (with diagonal covariance matrix for the depth part) and obtaining the mean and covariance matrix by warping the mean and covariance matrix of  $p(S^k|O^k)$  to the next step as described above. Due to this inexact integration, our approach in a strict sense is not recursive Bayesian filter as in Remark 8.7.

**Remark 8.8** (Regularization). *For the task of determining the optical flow from a pair/sequence of images (monocular or stereo) a common strategy is to take into account for the epipolar geometry, i.e. use a-priori information about which lines in the images corresponding matches have to lie on (cf. Remark 8.6). Our approach implicitly makes use of the epipolar geometry by using the optical flow model (8.74). This can be seen by assuming the camera parameters  $C = (R, h)$  in (8.74) to be fixed and allowing  $d(x)$  to vary. Then, the resulting point  $x' := x - u(x; R, h, d(x))$  corresponding to  $x$  moves exactly according to the epipolar geometry. Therefore, in our approach the optical flow is implicitly regularized with respect to the epipolar geometry by its dependence on  $d$ , which is in turn regularized by its spatial and temporal prior. Moreover, we stress that it is more suitable to have a regularization on a physical quantity like the depth rather than on an artificial variable like the optical flow.*

**Remark 8.9** (Non-convexity of  $\mathcal{F}$  in the update step). *We observe that the functional  $\mathcal{F}(S)$  in (8.91) to be minimized during the update step is non-convex. This convexity is due to both the data term and the manifold constraint on the camera parameters  $C \in SE(3)$ . In the data term both the product of depth  $d(x)$  with camera rotation  $R$  and the camera projection  $P_C$  are causing non-convexity. It raises the question if one of the strategies discussed in Chapters 5 and 6 can be applied in this context. Firstly, in view of a relaxation of the functional, we have to remark that a relaxation technique would also have to relax the manifold constraint, which is not an option in this case. Secondly, we observe that a fixed point approach would decouple the joint optimization with respect to  $d, R, h$ , leading to an alternating minimization. In our previous work ([16] with an alternating minimization and [17] with a joint minimization) we have observed that the joint approach shows a higher reconstruction quality. Moreover, concerning the non-convexity of the data term due to the camera projection, we remark that especially this projection enables us to estimate depth from a monocular image sequence, which would not be possible e.g. with a orthogonal projection. We expect a fixed point approach to weaken the influence of the data term. Finally, the non-convexity due to the manifold constraint cannot be resolved by a fixed point approach. These issues led us to directly tackle the non-convex optimization.*

### 8.6.5 Numerics

Concerning the numerical implementation, some issues arise, which are briefly discussed in the following. For a detailed presentation of the implementation of our approach we refer to [17].

**Multiscale approach** In the implementation of our approach we choose a Lukas-Kanade approach [110] to estimate the optical flow  $u_{\text{obs}}$  in terms of mean  $\mu_{\text{obs}}$  and covariance matrix  $\Sigma_{\text{obs}}$ . The Lukas-Kanade approach relies on a linearized brightness constancy assumption which assumes small displacements. This limitation is commonly resolved by considering a multi-scale approach, see e.g. [31]. Therefore, we apply the above approach to a cascade of downsampled problems, starting with the coarsest, and propagating the results to the next finer level until finally reaching the full resolution of the input images.

**Joint optimization on Manifold** Compared to earlier work [16], we propose in [17] to jointly optimize the objective function (8.91) with respect to the unknowns  $d$  and  $C$ . To this end, some effort has been made to adapt the previous method to work on the manifold  $\mathcal{M} := \mathbb{R}^n \times SE(3)$ . For details we refer to Section 7.3 and [17].

**Newton method, proximal point algorithm** We apply a Newton-like method to minimize the functional  $\mathcal{F}(S)$  in (8.91), where we determine the search direction by  $-B\nabla_T\mathcal{F}(S)$ , where  $\nabla_T\mathcal{F}(S)$  is the gradient of  $\mathcal{F}(S)$  in the tangent space  $T_S\mathcal{M}$  of  $\mathcal{M}$  at  $S$  (cf. Section 7.3). Moreover,  $B$  is a positive definite bilinear form on the tangent space, which incorporates second order information on  $\mathcal{F}$  as in the case of the classical Newton approach. In regions, where  $\mathcal{F}(S)$  is convex, the bilinear form  $B$  is chosen similar as in the classical Newton approach (inverse of the Hessian at point  $S$ ), except that on the manifold  $\mathcal{M}$  we exploit the Levi-Civita connection instead of the classical Hessian of  $S$ . The resulting bilinear form  $B$  then is positive definite, a property which guarantees that  $-B\nabla_T\mathcal{F}(S)$  is a descent direction.

In regions, where  $\mathcal{F}(S)$  is concave, care has to be taken to assert the positive definiteness of  $B$ . To this end, we propose to use a proximal point method [137] in concave regions, i.e. we add an proximity term to  $\mathcal{F}(S)$  so that the objective function becomes convex. The required modifications are described in detail in [17, Section 4].

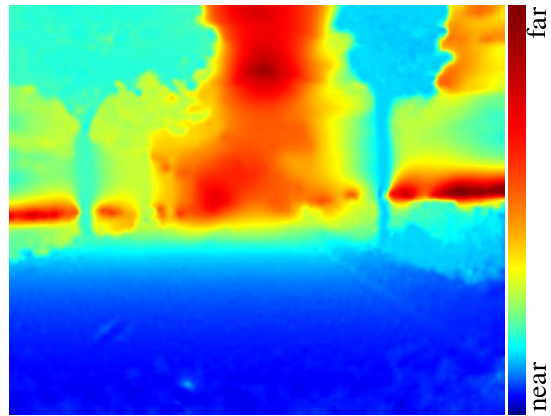
We combine this Newton-like method with a line search following Wolfe's Rule [27].

### 8.6.6 Experiments

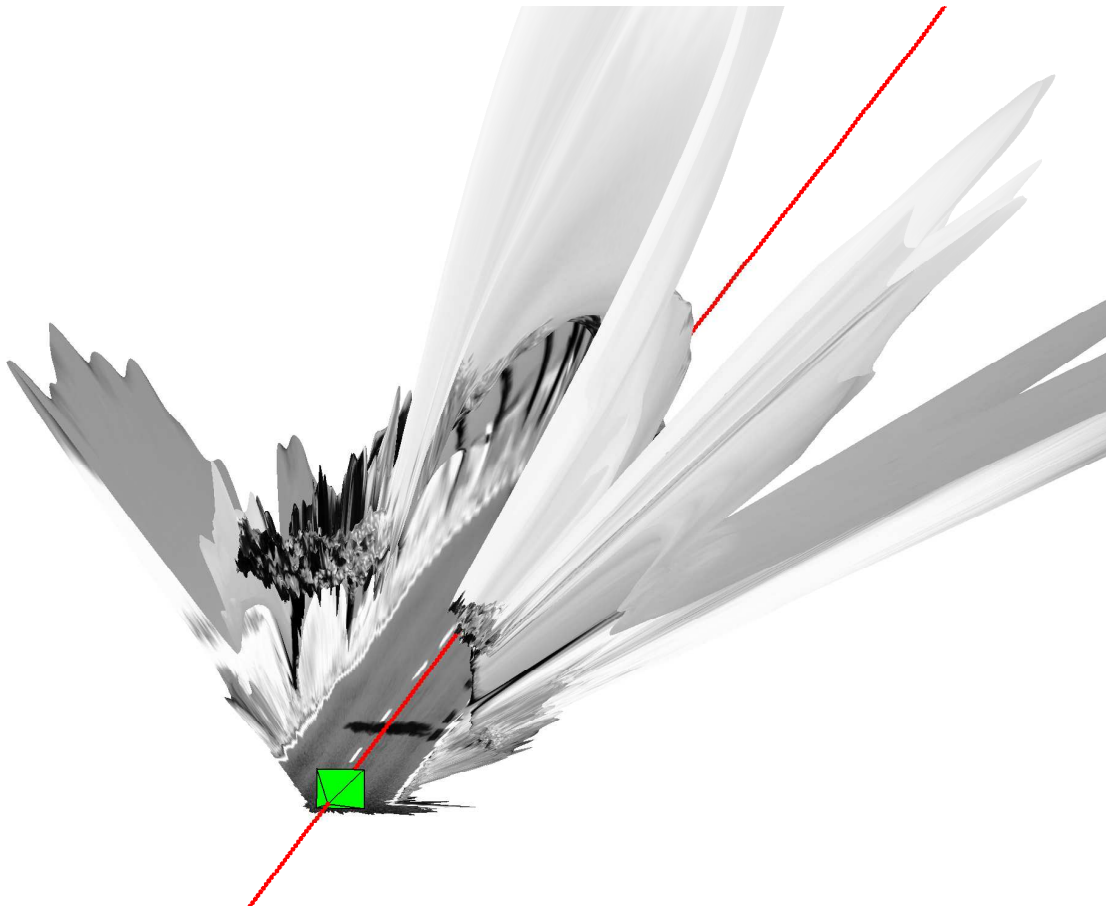
We evaluated our approach on two different data sets. The first one is a data set of sequences recorded by a camera system mounted in a driving car [114]. We refer to



(a) input frame



(b) estimated depth map

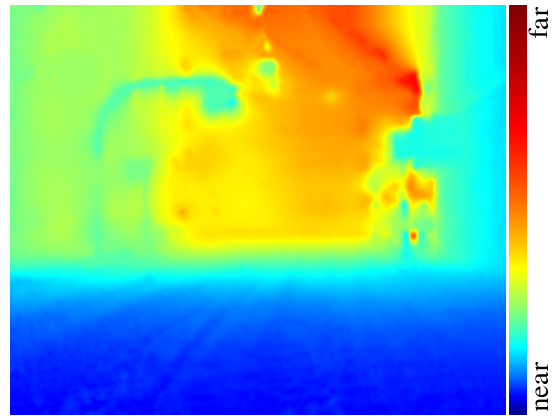


(c) reconstructed scene and camera track

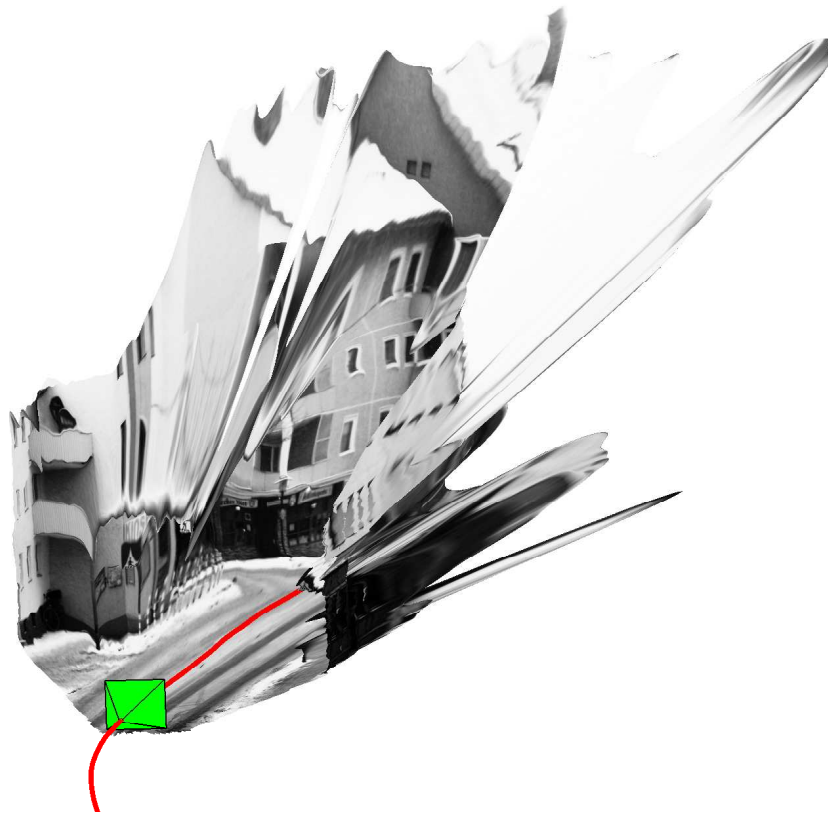
Figure 8.54: *Avenue* sequence: joint estimation of scene and egomotion. The depth map corresponding to the view of one input frame (a) is depicted in (b). The camera trajectory and the projected 3D scene is shown in (c).



(a) input frame



(b) estimated depth map



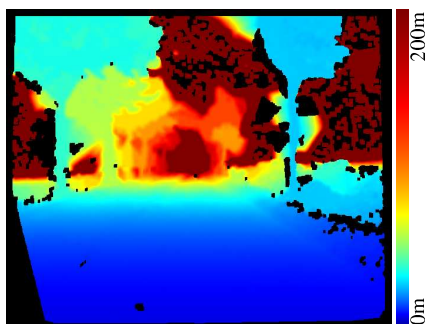
(c) reconstructed scene and camera track

Figure 8.55: *City* sequence: joint estimation of scene and egomotion. The depth map corresponding to the view of one input frame (a) is depicted in (b). The camera trajectory and the reprojected 3D scene is shown in (c).

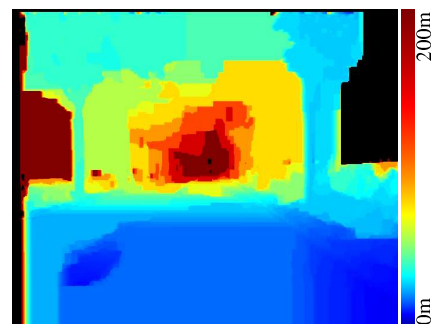




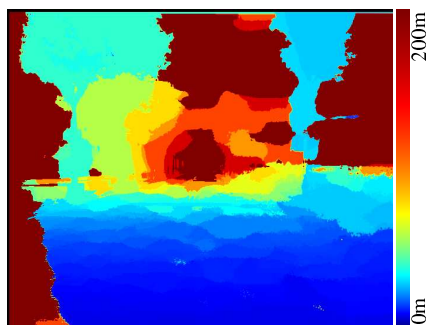
(a) frame



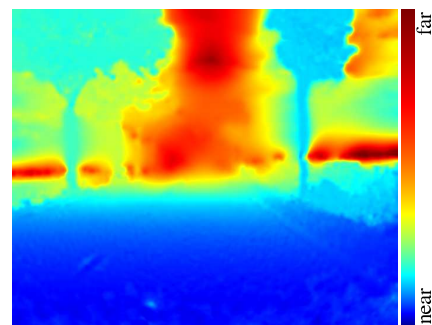
(b) depth estimated by Geiger et al. [62]



(c) depth estimated by Szeliski et al. [162]



(d) depth estimated by Rhemann et al. [134]



(e) depth estimated by our approach

Figure 8.56: Comparison of the proposed approach to three state-of-the-art stereo approaches. The proposed approach provides a dense and smooth depth map. It compares favorably with the stereo approaches, except for the region around the epipole, where, due to nature of the problem, accurate estimation of depth is not possible. Please note that the depth and egomotion from monocular sequences can only be estimated up to a global scale. Thus, to allow a comparison to the results of the stereo approaches, the scale has been determined by a one-parameter fit to the stereo data.

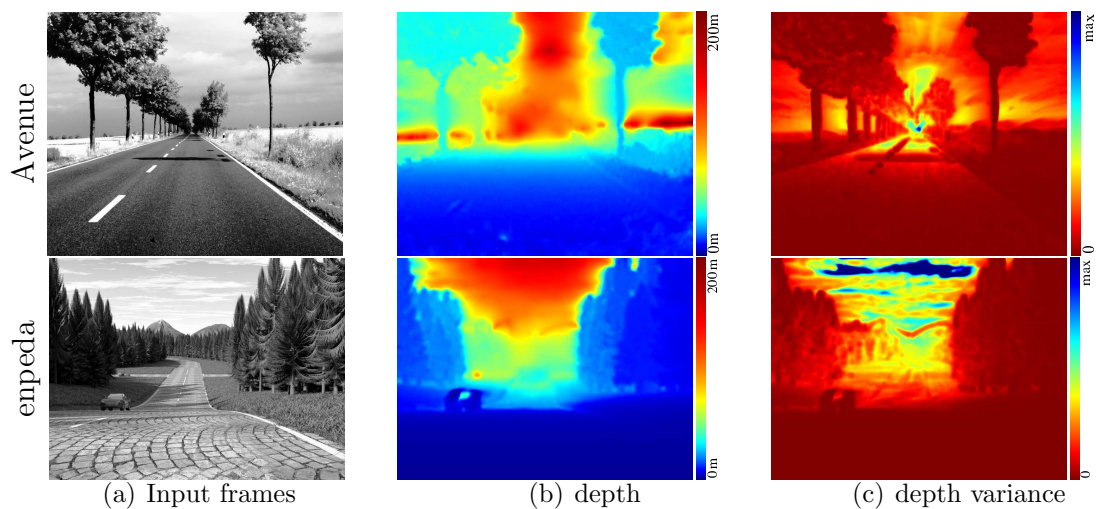


Figure 8.57: Augmenting the estimated depth map  $\bar{d}^k$  with its variance  $\bar{\sigma}^k$  (diagonal entries of the covariance matrix  $\bar{\Sigma}^k$ ). We show two exemplary frames. In particular, our approach is aware of the uncertainty of the estimation near the epipole.

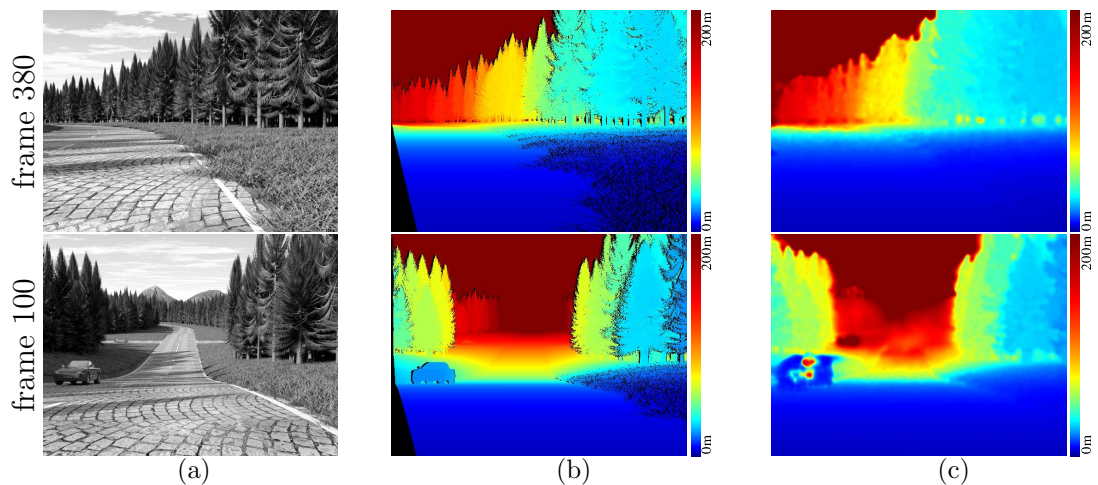


Figure 8.58: Synthetic sequence *enpeda-2-2*. (a) Image frames 100 and 380, (b) ground truth depth maps (black pixels indicate missing depth information), (c) estimated depth map with global scale adapted to ground truth.

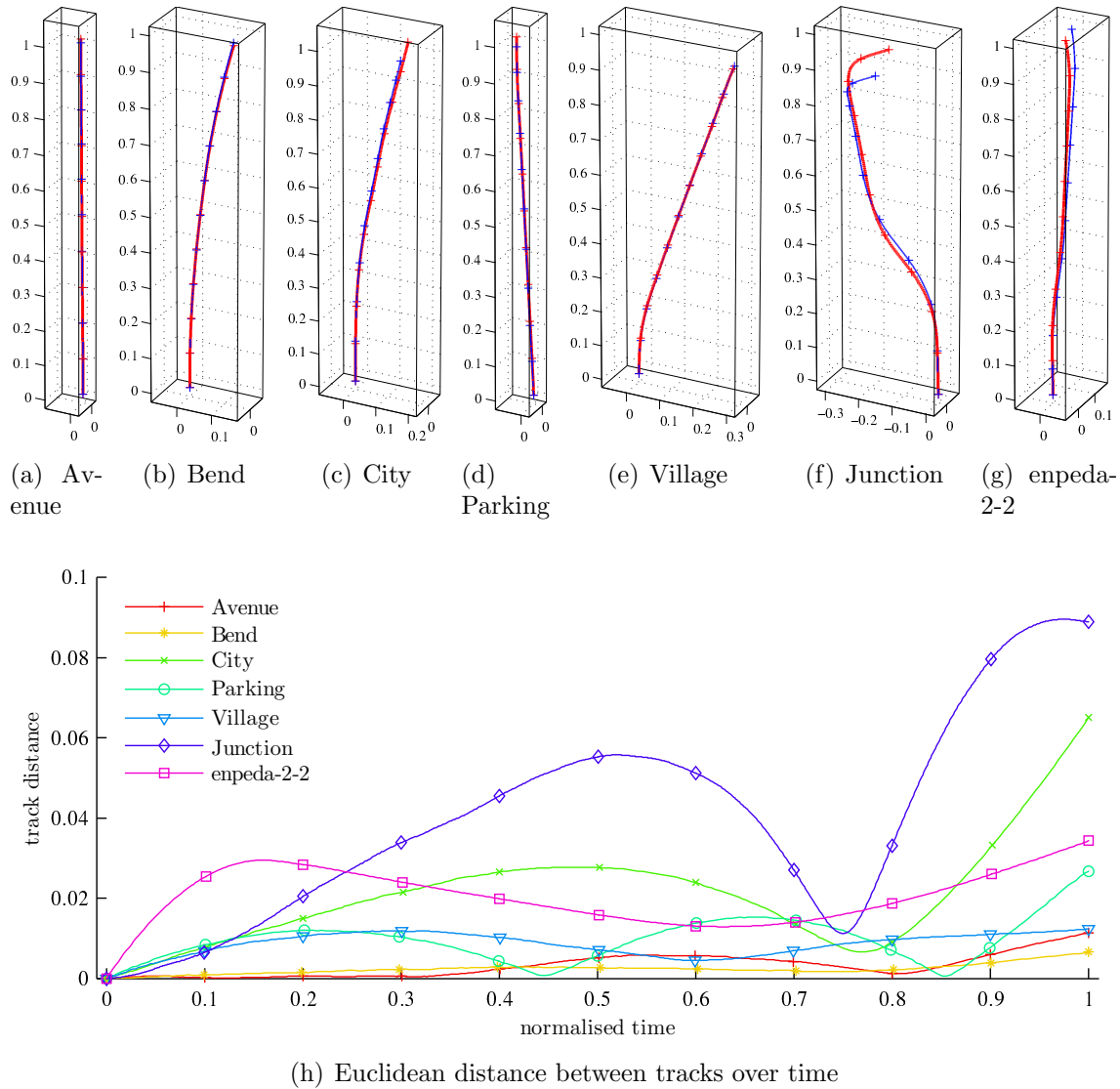


Figure 8.59: Comparison of the camera trajectory estimated by the proposed approach to the result of the Voodoo Tracker software. (a) - (g) visual comparison. The shape of the estimated trajectories matches well to those of the Voodoo tracker. However, we observe some integration errors, which aggregate over time and lead to a different speed, in which the camera moves along the trajectory (as can be seen from the markers, which indicate a selection of camera positions), which might be due to an over-smoothing of the camera translation  $\bar{h}^k$ . The misfit for the *Junction* sequence is due to the deceleration and acceleration of the car, which violates our assumptions of a fairly constant camera speed. (h) error between the estimated camera positions relative to the length (start-to-end-point) of the trajectories.

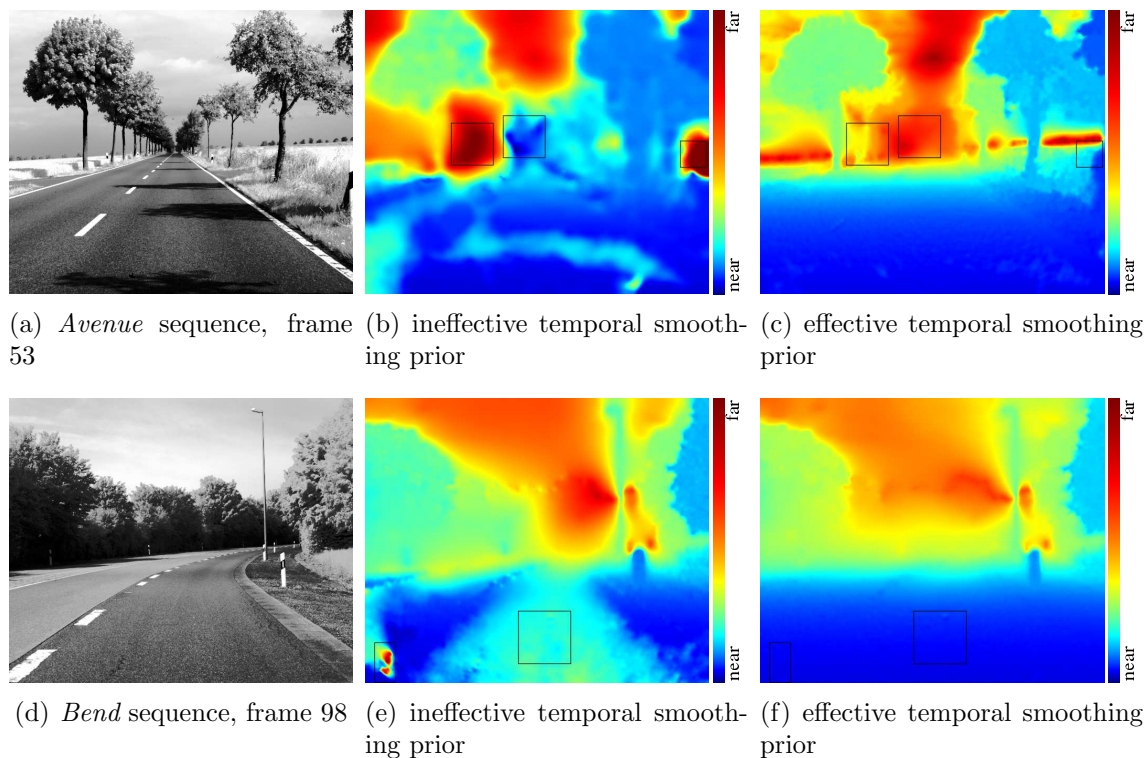


Figure 8.60: Comparison of our approach without (middle) and with temporal regularization (based on prediction). Without temporal regularization the depth is inaccurately estimated in various regions. Problems with wrong depth estimates occur in particular near the epipole (centered boxes in the top row) and in low-textured regions with (marked regions in the bottom row). Evaluating the full sequence (see movie at the website [http://hci.iwr.uni-heidelberg.de/VSFM/ICCV11/videos/Avenue\\_depthmap\\_without\\_and\\_with\\_prior.avi](http://hci.iwr.uni-heidelberg.de/VSFM/ICCV11/videos/Avenue_depthmap_without_and_with_prior.avi)) shows that these errors are highly oscillating, leading to a severe flickering in the sequence.

it as the HCI data set. These sequences are actually stereo sequences, from which for our approach we use only the image from the left camera. Most of the selected scene were almost static scenes up to some minor exceptions like shaking of trees, pedestrians, snow fall, etc. It turned out that our approach is robust against such small movements. The selected set of sequences is available online <sup>2</sup>. All considered sequences and the corresponding estimations of depth and egomotion can be found on the web page <http://hci.iwr.uni-heidelberg.de/VSFM/>.

The second data set is a synthetic image sequence taken from the *enpeda* project (cf. <http://ccv.wordpress.fos.auckland.ac.nz/eisats/>). The sequence is created with computer graphics tools and is augmented by ground truth information such as the true scene depth, which allows us to evaluate the exact error of our approach.

Let us start with the first data set. We present two exemplary results for the *Avenue* and *City* sequence in Figs. 8.54 and 8.55. Each figure depicts the estimated depth for one particular frame as well as the observed grayscale image backprojected to 3D by means of the estimated depth and augmented by the estimated camera trajectory. The estimated camera position corresponding to the particular image frame is indicated by a green icon. We remark that in contrast to stereo, in the monocular setup scene and egomotion can only be determined up to an unknown global scale. Experiments show, that this global scale due to temporal regularization stay constant over time after being fixed by the initialization.

In order to give an impression on the quality of the results, we compare our approach to three state-of-the-art stereo approaches proposed by Rhemann et al. [134], Szeliski et al. [162] and Geiger et al. [62], using the stereo image sequences from the HCI data set. Due to the known baseline of the stereo cameras, these methods typically are calibrated to the true global scale. To allow the comparison of our results to the stereo results, we perform a least-squares fit of the depth data to retrieve the unknown global scale. The comparison (cf. Fig. 8.56) then shows, that our approach produces a dense and smooth depth map, which except from regions near the epipole, where due to the ill-posedness of the problem an estimation of depth per se is difficult, our monocular approach is competitive to stereo approaches. Moreover, since we follow a probabilistic approach, our results are equipped with a certainty measure for the depth (obtained from the covariance matrix  $\bar{\Sigma}_{d^k}$  of the depth estimate  $\bar{d}^k$ ), which allows us to identify regions with uncertain depth estimates due to closeness to the epipole or absence of matchable structures in the images. (cf. Fig. 8.57).

To further evaluate our approach, we consider also the synthetic *enpeda* data set. Fig. 8.58 shows the results of our approach for two distinct image frames together with the ground truth depth data from the data set. Please note that in the second frame two moving cars (one oncoming, one approaching the crossing from left) are present, which violates our assumption of a static scene. As a consequence, the

---

<sup>2</sup>[http://hci.iwr.uni-heidelberg.de/Benchmarks/document/Real\\_Static\\_Scene\\_Image\\_Sequences\\_Dataset](http://hci.iwr.uni-heidelberg.de/Benchmarks/document/Real_Static_Scene_Image_Sequences_Dataset)

depth of the cars is incorrectly estimated. The violation of our assumption results, however, only in a local distortion and has no effect on the estimated depth of the static parts (and the estimated egomotion, see below), which demonstrates the robustness of our approach to such violations. Considering the static parts of the scene, the results of our approach are close to the ground truth depth except for a region around the epipole in the images (left to the image center for the first and in the image center for the second frame), where on principle a precise estimate of the depth due to vanishing information in the flow is impossible. As mentioned before, we account for this problem by also estimating the certainty of the depth by means of a covariance matrix see Fig. 8.57.

In order to evaluate our approach with respect to the estimated egomotion, we compare it to the results of the *Voodoo Camera Tracker*<sup>3</sup>, a batch approach, which based on the full sequence uses sparse feature matches to determine the egomotion. We remark that we retrieve the camera trajectory in our approach by integrating the individual camera translations  $h^k$  over time. A comparison of the camera trajectories found by our approach and the batch approach is depicted in Fig. 8.59. We observe that the camera trajectory in terms of congruency is quite well estimated by our approach, except for the *Junction* sequence. In the latter sequence, the car slows down before taking a right turn and accelerates afterwards. This violates our assumption that the camera speed is constant. As a consequence, deceleration and acceleration are underestimated resulting in some error in the camera trajectory. Moreover, for most of the test sequences, we observe some error in the exact localization of the camera positions (cross markers) on the camera trajectory, which might be due to a certain temporal smoothing of the egomotion.

Finally, we also want to stress the importance of the temporal prior obtained from the recursive approach with a prediction step. To this end we compare our approach to a variant without temporal smoothing, i.e. where we omit the term  $p_{\text{pre}}$ . As a consequence, the depth is only estimated based on two consecutive image frames. The results of both are depicted in Fig. 8.60. The approach without temporal filtering shows a far worse results with several regions of severely mis-estimated depth. The problematic regions appear not only near the epipole, (centered boxes in the top row), but also in the foreground (boxes in the bottom row) and due to missing temporal regularization show a high variability in time (see also the corresponding movie available from the web page <http://hci.iwr.uni-heidelberg.de/VFSM/>).

---

<sup>3</sup><http://www.digilab.uni-hannover.de/docs/manual.html>, v1.2.0b

# Chapter 9

## Conclusion

In this thesis, we studied variational approaches for solving inverse problems in image processing and computer vision. Hereby, a major focus was on approaches where the regularization term adapts to the specific inverse problem at hand.

We presented a general approach for adaptive TV regularization. This concept allows for various models for penalizing the gradient of the unknown signal, e.g. isotropic TV regularization with locally varying regularization parameter and anisotropic TV regularization. Moreover, by introducing arbitrary differential operators, it covers also models of higher-order TV. For this general model we showed existence and uniqueness of a minimizer of the objective function. We stress that our concept embraces various adaptive regularization approaches proposed in the literature.

Based on our general model, we considered two classes of adaptivity. The first one is the class of data-driven adaptive methods, where the adaptivity is steered solely by the (possibly pre-processed) input data. The second class comprises solution-driven adaptive variants, where the adaptivity is defined as an explicit mapping from the solution space to some parameter space. The latter approach has the advantage, that adaptivity can be steered by the reconstructed signal and thus better promotes the reconstruction of important data structures, while being less sensitive to distortions in the input data.

Both classes differ in the nature of the underlying minimization problem. While the data-driven approaches in general lead to convex minimization problems, the solution-driven approaches, if tackled in a direct way, in general results in non-convex minimization problems. In particular for the models of adaptivity discussed in this thesis, we obtain non-convex functionals. It has to be stressed that for non-convex minimization problem in contrast to convex problems challenges arise both in theory and in the numerical treatment.

The standard way to resolve these issues is to apply a relaxation technique. In Chapter 5 we have considered specific relaxation techniques for functionals defined on  $W^{1,p}(\Omega)$  and  $BV(\Omega)$ . We considered two applications, namely displacement regularization and image interpolation in Sections 8.2 and 8.5, respectively, where a solution-driven adaptive weighting of the *data term* lead to non-convex optimization

problems. For these applications we showed that a relaxation strategy is beneficial and lead to satisfactory results.

For solution-driven adaptive *regularization*, we observed that the non-convex dependency of the regularization term on the unknown in theory can be tackled by relaxation, but that due to missing analytic formulas for the relaxed functionals it is unclear how to numerically solve the relaxed minimization problem.

We therefore introduced an alternative strategy to resolve the problem of non-convexity using the concept of a fixed point problem. The core of this strategy is (i) to consider the operator  $T(v)$  mapping a function  $v$  which determines the adaptivity of the regularizer to the solution of the convex optimization problem (with this  $v$  being fixed), and (ii) to search for a fixed point of  $T$ .

Such fixed point approaches already have been considered in the literature, but then usually as ad hoc approaches without providing mathematical theory to support the approach.

For our strategy we were able provide theory both in the continuous and discrete setting. In the *continuous* setting we have proved existence of a fixed point under mild assumptions. In the *discrete* setting, we identified the equivalence to quasi-variational inequality problems (QVIPs). This equivalence enabled us to use *existence* theory for QVIPs from the literature.

For showing *uniqueness*, however, the situation was different due to the missing strong monotonicity of the operator occurring in the QVIPs occurring in our approach, which in general leads to an ambiguity in one component of the solution of the QVIP. Strong monotonicity, however, is a prerequisite for applying standard theory for uniqueness from the literature. This issue made it necessary to adapt standard theory to our special case. By exploiting the coupling between the fixed point and the solution of the QVIP, we were able to show uniqueness of the *fixed point of  $T$*  under sufficient conditions (guaranteeing a contraction). In addition, we provided a convergent algorithm for solving the considered fixed point problem.

We finally discussed various applications from image processing and computer vision. In these applications we utilized the theory and numerical methods presented in this thesis. The results of our applications demonstrated that adaptive regularization in general provides reconstructions of better quality than comparable non-adaptive approaches. In several applications, the solution-driven adaptive regularization produced results, which were significantly better than the results from the data-driven counterparts. This improvement is due to the fact that the regularization term is able to better adapt to structures present in the unknown signal, while being less sensitive to distortions of the input data. In particular, this shows that it is worthwhile to solve the more challenging fixed point problem.

We identified some drawbacks of our approaches. Firstly, since they are build on the total variation semi-norm, our approaches oversmooth textured regions in the results. Secondly, we require strictly convex data terms in the solution-driven approaches. This is due to the fact that the adaptivity is determined based on a single function  $v$ . Currently we are not aware of a meaningful generalization to



---

sets of  $v$ . Finally, our experiments show that, while solution-driven approach in general improve the results during their outer iteration, there are cases where the data-driven variant with optimally chosen parameters provides even better results. We are currently not aware of a criterion to identify these cases.

One plan for future work is to tackle the problem of oversmoothing of textures. In this regard, generalizing our concept to either non-local total variation models or structure/texture decomposition models as discussed in Chapter 2 seems to be promising.

## Nomenclature

### measures

$\int f d\mathcal{L}$	Lebesgue integral of $f$	
$Du,  Du $	signed/ positive Radon measures	p. 19
$Du_s$	singular part of $Du$	—“—

### function spaces

$C^k(\Omega, Y)$	$k$ -times continuously differentiable functions defined on $\Omega$ with values in $Y$	
$C_c^k(\Omega, Y)$	functions in $C^k(\Omega, Y)$ with compact support	
$L^p(\Omega)$	Space of Lebesgue-integrable functions we omit $\Omega$ if it is clear from the context	
$W^{m,p}(\Omega)$	Sobolev functions with weak derivatives up to order $m$ in $L^p$	
$BV(\Omega)$	space of functions of bounded variations	p. 13
$BV^l(\Omega)$	space of functions of bounded variations of order $l$	p. 34

### (semi-)norms and metrics

$\ \cdot\ _2$	Euclidean norm in $\mathbb{R}^n$	
$\ \cdot\ _{L(\Omega)}$	norm of $L^p(\Omega)$	
TV	total variation semi-norm	p. 12
TGV	total generalized variation semi-norm	—“—
$\text{dist}(x, A)$	distance from point $x$ to set $A$ .	p. 81

### functionals

$\mathcal{F}$	functional to be minimized	
$\mathcal{G}$	Fenchel dual of $\mathcal{F}$	p. 28
$\mathcal{F}_{ROF}$	ROF functional	p. 12
$\mathcal{S}$	data term	
$\mathcal{R}$	regularization term	
$\mathcal{F}_\diamond$	relaxed functional in $W^{1,p}(\Omega)$	p. 65
$\mathcal{F}_{\diamond BV}$	relaxed functional in $BV(\Omega)$	—“—

### differential operators

$\nabla^l$	differential operator including all orders equal to $l$	p. 33
$\text{div}^l$	adjoint operator to $\nabla^l$	—“—
$\overline{\nabla}$	differential operator including all orders equal up to $l$	—“—
$\overline{\text{div}}$	adjoint operator to $\overline{\nabla}$	—“—
$\overline{\text{TV}}^l$	total variation semi-norm $\overline{\text{TV}}^l(u) = \sum_{k=1}^l \text{TV}^k(u)$	—“—
$L$	discrete divergence operator	p. 16

### operators

$Q(x, v)$	linear operator defining adaptivity	p. 77
A	application dependent operator	p. 38
M	linear operator in QVIP	p. 75
$T(v)$ ,	solution operator to $\min_u \mathcal{F}(u, v)$	p. 75
$\hat{T}$	solution operator to VI	p. 91

**constants**

$C_b, C_A$	constants of A for lower/upper bounds	p. 45
$C_a,$	constant of A for coercivity	p. 46
$C_{min}$	radius of a ball included in $\mathcal{D}^{loc}$	p. 41
$C_{max}$	radius of a ball including in $\mathcal{D}^{loc}$	—  —
$C_l$	Lipschitz constant of $v \rightarrow A(x, v)$	p. 79
$\gamma$	Lipschitz constant of $M \circ \hat{T}$	p. 91
$\eta$	variation rate	p. 91
$\mu$	Lipschitz constant of $g$	p. 91
$\nu$	monotonicity constant of $g$	—  —

**functions**

$f : X \rightrightarrows Y$	set-valued function	p. 89
$\chi_e$	indicator function for edge locations	p. 50
$v_e$	vector field of edge normals	—  —
$\chi_{sd}$	indicator function for locations of slope discontinuities	p. 56

**sets**

$\overline{\mathbb{R}}$	extended real values $\mathbb{R} \cup \{+\infty\}$	
$\Omega$	$d$ -dimensional open bounded domain with Lipschitz domain	
$\partial\Omega$	boundary of $\Omega$	p. 48
$level_{\leq t}(f)$	sublevel set of $f$	p. 26
$level_{\geq t}(f)$	superlevel set of $f$	—  —
$\mathcal{K}$	convex set	p. 25
$\mathcal{D}$	(convex) constraint set for dual variable $p$	p. 35
$\mathcal{C}$	(convex) constraint set for dual variable $v$	p. 88
$\mathcal{D}^{loc}$	local constraint set for variable $p_i$	p. 36
$\mathring{B}_r(x)$	open ball of radius $r$ centered at $x$	p. 126
$B_r(x)$	closed ball of radius $r$ centered at $x$	
$Rect(r, \alpha, \beta)$	rectangle	p. 37
$Ell(r, \alpha, \beta)$	ellipse	p. 58
$Par(r_1, r_2, \alpha)$	parallelogram	p. 59

**differential geometry**

Exp	exponential map	p. 111
$SE(3)(n)$	special Euclidean group of $\mathbb{R}^n$	p. 110

**statistic**

$\mathcal{N}(\mu, \Sigma)$	Gaussian distribution of mean $\mu$ and covariance matrix $\Sigma$	p. 11
----------------------------	--	-------

**weak convergence**

$\rightharpoonup$	weak convergence in Sobolev spaces	p. 26
$\overset{*}{\rightharpoonup}$	weak * convergence in $BV$	—“—
$\overset{L^2}{\rightharpoonup}$	weak convergence $L^2(\Omega)$	p. 87
$\overset{X}{\rightharpoonup}$	joint weak/weakstar convergence in $L^2(\Omega) \cap BV(\Omega)$	p. 39

**Bayesian filtering**

$O$	observation variable	p. 201
$S$	state variable	—“—

## Figure Acknowledgements

We acknowledge the content of the following figures:

- Fig. 2.3            standard test image
- Fig. 4.5            courtesy of Florian Becker
- Fig. 4.6            courtesy of Florian Becker
- Fig. 8.1            standard test images
- Fig. 8.29          the mountain image was taken by the author,  
the other images are standard test images
- Fig. 8.33          image taken by the author
- Fig. 8.34          courtesy of Stephan Meister, Rahul Nair and Henrik Schäfer,  
the HCI box is taken from [Db1]
- Figs. 8.36 – 8.39 in collaboration with Henrik Schäfer and Christoph S. Garbe
- Figs. 8.40 – 8.42 courtesy of the authors of [118, 120]
- Fig. 8.43          in collaboration with the authors of [118, 120]
- Fig. 8.46          courtesy of the authors of [118, 120], taken from [Db1]
- Fig. 8.47          in collaboration with the authors of [118, 120]
- Fig. 8.48          images taken from the Kodak image data base [Db2]
- Figs. 8.52, 8.53   courtesy of Florian Becker
- Figs. 8.54, 8.55   input: [Db3], depth and egomotion: in cooperation with the  
authors of [16, 17]
- Figs. 8.56, 8.57   input images: [Db3,Db4], other images in cooperation with  
the authors of [16, 17]
- Fig. 8.58          input image: [Db4], depth maps: in cooperation with the  
authors of [16, 17]
- Fig. 8.59          in cooperation with the authors of [16, 17]
- Fig. 8.60          input: [Db4], depth maps: in cooperation with the  
authors of [16, 17]

All figures not explicitly listed above were created by the author.

Referenced data bases:

- [Db1]    HCI box  
[http://hci.iwr.uni-heidelberg.de//Benchmarks/  
document/hcibox/](http://hci.iwr.uni-heidelberg.de//Benchmarks/document/hcibox/)
- [Db2]    Kodak image database  
<http://r0k.us/graphics/kodak/>
- [Db3]    Static Scenes in an automotive context  
[http://hci.iwr.uni-heidelberg.de//Benchmarks/  
document/Real\\_Static\\_Scene\\_Image\\_Sequences\\_  
Dataset/](http://hci.iwr.uni-heidelberg.de//Benchmarks/document/Real_Static_Scene_Image_Sequences_Dataset/)
- [Db4]    Test sequence from the enpeda project  
[www.mi.auckland.ac.nz/enpeda/](http://www.mi.auckland.ac.nz/enpeda/)



# Acknowledgements

First of all, I would like to thank Prof. Dr. Christoph Schnörr for giving me the opportunity to work at the Heidelberg Collaboratory Image Processing (HCI), for his support throughout my time at the HCI and for being mentor for my habilitation. Moreover, I would like also to thank Prof. Dr. Bernd Jähne and Prof. Dr. Rolf Rannacher for being mentors. I thank Prof. Dr. Fred Hamprecht and Prof. Dr. Björn Ommer from the directorate of the HCI for supporting my work.

I thank Dr. Wolfgang Niehsen from Bosch for initiating the collaboration on computer vision tasks in the automotive environment. I am thankful to our secretaries Karin Kruljac, Dr. Christan Schmidt, Barbara Werner and Evelyn Wilhelm for their administrative support and our (former and current) system administrators Dr. Ole Hansen, Dr. Nils Krahn and Dominik Spangenberg for maintaining the IT infrastructure and helping out with any software or hardware problem.

I would like to thank Dr. Florian Becker, Johannes Berger, Andreea Denițiu, Dr. Rahul Nair, Dr. Stefania Petra, and Dr. Henrik Schäfer for proof-reading this thesis. I thank PD Dr. Christoph Garbe and the (former) PhD students involved in the project *Algorithms for Low Cost Depth Imaging*, namely Jens-Malte Gottfried, Dr. Stephan Meister, Dr. Rahul Nair and Dr. Henrik Schäfer for their collaboration and continued discussions on Time-of-Flight imaging as well as for collecting Time-of-Flight data. Moreover, I thank Katrin Honauer, Dr. Daniel Kondermann and Dr. Stephan Meister for their support on the automotive data sets.

I thank all the (current and former) HCI postdocs, the members of the *Image and Pattern Analysis group* and the research training group *Probabilistic Graphical Models and Applications in Image Analysis* for fruitful discussions on a broad field of research topics.

Finally, I thank my family, especially Bettina, Philipp, Pia, Dieter, Gisela, Erwin and Hedi for supporting me throughout the time.





# Bibliography

- [1] T. Aach, C. Mota, I. Stuke, M. Mühlich, and E. Barth. Analysis of superimposed oriented patterns. *Image Processing, IEEE Transactions on*, 15(12):3690–3700, 2006.
- [2] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2008.
- [3] R. Agarwal and D. O’Regan. Fixed-point theory for weakly sequentially upper-semicontinuous maps with applications to differential inclusions. *Nonlinear Oscillations*, 5(3), 2002.
- [4] H. W. Alt. *Linear functional analysis. An application oriented introduction.* (*Lineare Funktionalanalysis. Eine anwendungsorientierte Einführung.*). Springer, 2006.
- [5] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*, volume 254. Clarendon Press Oxford, 2000.
- [6] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *Signal Processing, IEEE Transactions on*, 50(2):174–188, 2002.
- [7] F. Aström, G. Baravdish, and M. Felsberg. A tensor variational formulation of gradient energy total variation. In X.-C. Tai, E. Bae, T. Chan, and M. Lysaker, editors, *Energy Minimization Methods in Computer Vision and Pattern Recognition*, volume 8932 of *Lecture Notes in Computer Science*, pages 307–320. Springer, 2015.
- [8] J.-F. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle. Image decomposition application to SAR images. In *Scale Space Methods in Computer Vision*, pages 297–312. Springer, 2003.
- [9] J.-F. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle. Image decomposition into a bounded variation component and an oscillating component. *Journal of Mathematical Imaging and Vision*, 22(1):71–88, 2005.

- [10] J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher. Structure-texture image decomposition – modeling, algorithms, and parameter selection. *Int. J. Comp. Vision*, 67(1):111–136, 2006.
- [11] L. Bagnato, P. Frossard, and P. Vandergheynst. A variational framework for structure from motion in omnidirectional image sequences. *J. Math. Imaging Vis.*, 41(3):182–193, Nov. 2011.
- [12] C. Baiocchi and A.-C. Capelo. Variational and quasivariational inequalities: Applications to free boundary problems. *JOHN WILEY & SONS*, 1984.
- [13] H. Bauschke and P. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Canadian Math. Society. Springer, 2011.
- [14] I. Bayram and M. E. Kamasak. A directional total variation. In *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, pages 265–269. IEEE, 2012.
- [15] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009.
- [16] F. Becker, F. Lenzen, J. H. Kappes, and C. Schnörr. Variational recursive joint estimation of dense scene structure and camera motion from monocular high speed traffic sequences. In D. Metaxas, L. Quan, A. Sanfeliu, and L. Van Gool, editors, *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 1692–1699. IEEE, 2011.
- [17] F. Becker, F. Lenzen, J. H. Kappes, and C. Schnörr. Variational recursive joint estimation of dense scene structure and camera motion from monocular high speed traffic sequences. *International Journal of Computer Vision*, 105(3):269–297, 2013.
- [18] F. Becker, S. Petra, and C. Schnörr. Optical flow. *Handbook of Mathematical Methods in Imaging*, 2014.
- [19] A. Belahmidi and F. Guichard. A partial differential equation approach to image zoom. In *Image Processing, 2004. ICIP'04. 2004 International Conference on*, volume 1, pages 649–652. IEEE, 2004.
- [20] A. Bensoussan and J.-L. Lions. *Applications des inéquations variationnelles en contrôle stochastique*. Dunod, 1978.
- [21] M. Bergounioux. On Poincaré-Wirtinger inequalities in spaces of functions of bounded variation. *Control and Cybernetics*, 40(4):921–930, 2011.

- 
- [22] B. Berkels, M. Burger, M. Droske, O. Nemitz, and M. Rumpf. Cartoon extraction based on anisotropic image classification. In *Vision, Modeling, and Visualization Proceedings*, pages 293–300, 2006.
- [23] J. Besag, J. York, and A. Mollié. Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43(1):1–20, 1991.
- [24] P. Blomgren and T. F. Chan. Color TV: total variation methods for restoration of vector-valued images. *Image Processing, IEEE Transactions on*, 7(3):304–309, 1998.
- [25] P. Blomgren, T. F. Chan, and P. Mulet. Extensions to total variation denoising. In *Optical Science, Engineering and Instrumentation'97*, pages 367–375. International Society for Optics and Photonics, 1997.
- [26] E. M. Bollt, R. Chartrand, S. Esedoglu, P. Schultz, and K. R. Vixie. Graduated adaptive image denoising: local compromise between total variation and isotropic diffusion. *Advances in Computational Mathematics*, 31(1-3):61–85, 2009.
- [27] J. F. Bonnans, J. C. Gilbert, C. Lemaréchal, and C. Sagastizábal. *Numerical Optimization*. Springer Verlag, 2003.
- [28] K. Bredies, K. Kunisch, and T. Pock. Total Generalized Variation. *SIAM J. Imaging Sciences*, 3(3):492–526, 2010.
- [29] X. Bresson and T. F. Chan. Fast dual minimization of the vectorial total variation norm and applications to color image processing. *Inverse Problems and Imaging*, 2(4):455–484, 2008.
- [30] H. Brezis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, volume 5. North Holland, 1973.
- [31] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In T. Pajdla and J. Matas, editors, *Computer Vision – ECCV 2004*, volume 3024 of *LNCS*, pages 25–36, Prague, Czech Republic, 5 2004. Springer.
- [32] T. Brox, O. Kleinschmidt, and D. Cremers. Efficient nonlocal means for denoising of textural patterns. *Image Processing, IEEE Transactions on*, 17(7):1083–1092, July 2008.
- [33] J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1986.
- [34] A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imaging Vision*, 20(1–2):89–97, 2004.

- [35] A. Chambolle and P.-L. Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76(2):167–188, 1997.
- [36] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.
- [37] D. Chan and J.-S. Pang. The generalized quasi-variational inequality problem. *Math. Operat. Res.*, 7(2):211–222, 1982.
- [38] T. Chan, A. Marquina, and P. Mulet. High-order total variation-based image restoration. *SIAM Journal on Scientific Computing*, 22(2):503–516, 2000.
- [39] T. F. Chan, S. H. Kang, and J. Shen. Total variation denoising and enhancement of color images based on the CB and HSV color models. *Journal of Visual Communication and Image Representation*, 12(4):422 – 435, 2001.
- [40] K. Chang, T. Qin, W. Xu, and Z. Tang. Reconstruction of multi-view compressed imaging using weighted total variation. *Multimedia Systems*, 20(4):363–378, 2014.
- [41] C.-H. Chen, C.-M. Kuo, T.-K. Yao, and S.-H. Hsieh. Anisotropic probabilistic neural network for image interpolation. *Journal of Mathematical Imaging and Vision*, pages 1–11, 2013.
- [42] Q. Chen, P. Montesinos, Q. S. Sun, P. A. Heng, et al. Adaptive total variation denoising based on difference curvature. *Image Vision Comput.*, 28(3):298–306, 2010.
- [43] Y. Chen and T. Wunderli. Adaptive total variation for image restoration in bv space. *Journal of mathematical analysis and applications*, 272(1):117–137, 2002.
- [44] L.-J. Chu. Unified approaches to nonlinear optimization. *Optimization*, 46(1):25–60, 1999.
- [45] C. Clason, B. Jin, and K. Kunisch. A duality-based splitting method for L1-TV image restoration with automatic regularization parameter choice. *SIAM J. Sci. Comp.*, pages 1484–1505, 2010.
- [46] A. Comport, E. Malis, and P. Rives. Accurate Quadri-focal Tracking for Robust 3D Visual Odometry. In *IEEE International Conference on Robotics and Automation, ICRA '07*, 2007.
- [47] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *Image Processing, IEEE Transactions on*, 16(8):2080–2095, 2007.

- 
- [48] B. Dacorogna. *Weak continuity and weak lower semicontinuity of non-linear functionals*, volume 922. Springer-Verlag Berlin, 1982.
- [49] B. Dacorogna. *Direct methods in the calculus of variations*, volume 78. Springer, 2008.
- [50] P. Elbau, M. Grasmair, F. Lenzen, and O. Scherzer. Evolution by non-convex functionals. *Numerical Functional Analysis and Optimization*, 31:489–517, 2010.
- [51] R. Elliot, L. Aggoun, and J. Moore. *Hidden Markov Models*. Springer, 1995.
- [52] J. Engel, J. Sturm, and D. Cremers. Semi-dense visual odometry for a monocular camera. In *Computer Vision (ICCV), 2013 IEEE International Conference on*, pages 1449–1456, 2013.
- [53] S. Esedoglu and S. J. Osher. Decomposition of images by the anisotropic Rudin-Osher-Fatemi model. *Communications on pure and applied mathematics*, 57(12):1609–1626, 2004.
- [54] V. Estellers, S. Soato, and X. Bresson. Adaptive regularization with the structure tensor. *Image Processing, IEEE Transactions on*, 24(6):1777–1790, 2015.
- [55] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions*, volume 5. CRC press, 1992.
- [56] J. M. Fadili and G. Peyré. Total variation projection with first order schemes. *Image Processing, IEEE Transactions on*, 20(3):657–669, 2011.
- [57] W. Förstner and E. Gülch. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In *Proc. ISPRS conf. on fast processing of photogrammetric data*, pages 281–305, 1987.
- [58] M. Frank, M. Plaue, K. Rapp, U. Köthe, Jähne, and H. F.A. Theoretical and experimental error analysis of continuous-wave time-of-flight range cameras. *Optical Engineering*, 48(1):13602, 2009.
- [59] S. Fu and C. Zhang. Adaptive non-convex total variation regularisation for image restoration. *Electronics Letters*, 46(13):907–908, 2010.
- [60] D. Gabay. Minimizing a differentiable function over a differential manifold. *Journal of Optimization Theory and Applications*, 37(2):177–219, 1982.
- [61] D. Garcia. Robust smoothing of gridded data in one and higher dimensions with missing values. *Computational statistics & data analysis*, 54(4):1167–1178, 2010.

- [62] A. Geiger, M. Roser, and R. Urtasun. Efficient large-scale stereo matching. In *Asian Conference on Computer Vision*, Queenstown, New Zealand, November 2010.
- [63] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 6(6):721–741, 1984.
- [64] G. Gilboa, J. Darbon, S. Osher, and T. Chan. Nonlocal convex functionals for image regularization. Technical Report 06-57, UCLA CAM report, 2006.
- [65] G. Gilboa and S. Osher. Nonlocal operators with applications to image processing. *Multiscale Model. Simul.*, 7(3):1005–1028, 2008.
- [66] G. Graber, T. Pock, and H. Bischof. Online 3D reconstruction using Convex Optimization. In *1st Workshop on Live Dense Reconstruction From Moving Cameras, ICCV 2011*, pages 708–711, 2011.
- [67] M. Grasmair. Locally adaptive total variation regularization. In *proceedings SSVM '09*, volume 5567, pages 331–342, 2009.
- [68] M. Grasmair and F. Lenzen. Anisotropic total variation filtering. *Applied Mathematics & Optimization*, 62:323–339, 2010.
- [69] M. Grasmair, F. Lenzen, A. Obereder, O. Scherzer, and M. Fuchs. A non-convex PDE scale space. In R. Kimmel, N. A. Sochen, and J. Weickert, editors, *Proceedings of 5th International Conference on Scale Space and PDE Methods in Computer Vision 2005*, volume 3459 of *LNCS*, pages 303–315. Springer, 2005.
- [70] F. Guichard and F. Malgouyres. Total variation based interpolation. In *Proceedings of the European signal processing conference*, volume 3, pages 1741–1744, 1998.
- [71] J. Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton university bulletin*, 13(49-52):28, 1902.
- [72] R. Hadsell, P. Sermanet, J. Ben, A. Erkan, M. Scoffier, K. Kavukcuoglu, U. Muller, and Y. LeCun. Learning long-range vision for autonomous off-road driving. *J. Field Robot.*, 26:120–144, Feb. 2009.
- [73] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge Univ. Press, 2000.
- [74] C. Himmelberg. Fixed points of compact multifunctions. *Journal of Mathematical Analysis and Applications*, 38(1):205–207, 1972.

- 
- [75] H. Hirschmüller. Stereo processing by semiglobal matching and mutual information. *TPAMI*, 30(2):328–341, 2008.
- [76] J. Hu and H. Wang. Adaptive total variation based on feature scale. *Trans. Eng. Comput. Technol. vV2*, pages 245–248, 2005.
- [77] K. Huper and J. Trunpf. Newton-like methods for numerical optimization on manifolds. In *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Eighth Asilomar Conference on*, volume 1, pages 136–139. IEEE, 2004.
- [78] M. Irani, P. Anandan, and M. Cohen. Direct recovery of planar-parallax from multiple frames. *TPAMI*, 24(11):1528–1534, 2002.
- [79] S. Jang, H. Nam, Y. J. Lee, B. Jeong, R. Lee, and J. Yoon. Data-adapted moving least squares method for 3-d image interpolation. *Physics in medicine and biology*, 58(23):8401, 2013.
- [80] M. Jung and M. Kang. Simultaneous cartoon and texture image restoration with higher-order regularization. *SIAM Journal on Imaging Sciences*, 8(1):721–756, 2015.
- [81] R. E. Kalman et al. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45, 1960.
- [82] S. Keller, F. Lauze, and M. Nielsen. A total variation motion adaptive deinterlacing scheme. In R. Kimmel, N. Sochen, and J. Weickert, editors, *Scale Space and PDE Methods in Computer Vision*, volume 3459 of *Lecture Notes in Computer Science*, pages 408–418. Springer, 2005.
- [83] K. I. Kim and Y. Kwon. Single-image super-resolution using sparse regression and natural image prior. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 32(6):1127–1133, 2010.
- [84] S. Kindermann, S. Osher, and P. Jones. Deblurring and denoising of images by nonlocal functionals. *Multiscale Model. Simul.*, 4(4):1091–1115, 2005.
- [85] G. Klein and D. Murray. Parallel tracking and mapping for small AR workspaces. In *Proc. Sixth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'07)*, 2007.
- [86] K. Konolige and M. Agrawal. FrameSLAM: From Bundle Adjustment to Real-Time Visual Mapping. *IEEE Trans. Robotics*, 24(5):1066–1077, 2008.
- [87] U. Köthe. Edge and junction detection with an improved structure tensor. In *Proc. of the 25th DAGM Symposium on Pattern Recognition*, LNCS. DAGM, 2003.

- [88] H. Lakshman, W.-Q. Lim, H. Schwarz, D. Marpe, G. Kutyniok, and T. Wiegand. Image interpolation using shearlet based iterative refinement. *arXiv preprint arXiv:1308.1126*, 2013.
- [89] D. C. Lee, M. Hebert, and T. Kanade. Geometric reasoning for single image structure recovery. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, June 2009.
- [90] S. Lefkimmiatis, A. Roussos, P. Maragos, and M. Unser. Structure tensor total variation. *SIAM J. Imaging Sci*, 2015.
- [91] S. Lefkimmiatis, A. Roussos, M. Unser, and P. Maragos. Convex generalizations of total variation based on the structure tensor with applications to inverse problems. In A. Kuijper, K. Bredies, T. Pock, and H. Bischof, editors, *Scale Space and Variational Methods in Computer Vision*, volume 7893 of *Lecture Notes in Computer Science*, pages 48–60. Springer, 2013.
- [92] D. Lefloch, R. Nair, F. Lenzen, H. Schäfer, L. Streeter, M. J. Cree, R. Koch, and A. Kolb. Technical foundation and calibration methods for time-of-flight cameras. In M. Grzegorzec, C. Theobalt, R. Koch, and A. Kolb, editors, *Time-of-Flight and Depth Imaging: Sensors, Algorithms, and Applications*, volume 8200 of *Lecture Notes in Computer Science*, pages 3–24. Springer, 2013.
- [93] J. Lellmann, F. Becker, and C. Schnörr. Convex optimization for multi-class image labeling with a novel family of total variation based regularizers. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 646–653. IEEE, 2009.
- [94] J. Lellmann, J.-M. Morel, and C.-B. Schönlieb. Anisotropic third-order regularization for sparse digital elevation models. In A. Kuijper, K. Bredies, T. Pock, and H. Bischof, editors, *Scale Space and Variational Methods in Computer Vision*, volume 7893 of *Lecture Notes in Computer Science*, pages 161–173. Springer Berlin Heidelberg, 2013.
- [95] J. Lellmann, K. Papafitsoros, C.-B. Schönlieb, and D. Spector. Analysis and application of a non-local hessian. *arXiv preprint arXiv:1410.8825*, 2014.
- [96] J. Lellmann and C. Schnörr. Continuous multiclass labeling approaches and algorithms. *SIAM Journal on Imaging Sciences*, 4(4):1049–1096, 2011.
- [97] F. Lenzen. *Statistical Regularization and Denoising*. PhD thesis, University of Innsbruck, Innsbruck, 2006.
- [98] F. Lenzen, F. Becker, and J. Lellmann. Adaptive second-order total variation: An approach aware of slope discontinuities. In A. Kuijper, K. Bredies, T. Pock, and H. Bischof, editors, *Scale Space and Variational Methods in Computer Vision*, volume 7893 of *Lecture Notes in Computer Science*, pages 61–73. Springer Berlin Heidelberg, 2013.



- 
- [99] F. Lenzen, F. Becker, J. Lellmann, S. Petra, and C. Schnörr. Variational image denoising with adaptive constraint sets. In *3rd Int. Conference on Scale Space and Variational Methods in Computer Vision*, pages 206–217. Springer, 2012.
- [100] F. Lenzen, F. Becker, J. Lellmann, S. Petra, and C. Schnörr. A class of quasi-variational inequalities for adaptive image denoising and decomposition. *Computational Optimization and Applications*, 54(2):371–398, 2013.
- [101] F. Lenzen and J. Berger. Solution-driven adaptive total variation regularization. In J.-F. Aujol, M. Nikolova, and N. Papadakis, editors, *Proceedings of SSVM 2015*, volume 9087 of *LNCS*, pages 203–215, 2015.
- [102] F. Lenzen, K. I. Kim, H. Schäfer, R. Nair, S. Meister, F. Becker, C. S. Garbe, and C. Theobalt. Denoising strategies for time-of-flight data. In M. Grzegorzek, C. Theobalt, R. Koch, and A. Kolb, editors, *Time-of-Flight and Depth Imaging: Sensors, Algorithms, and Applications*, volume 8200 of *Lecture Notes in Computer Science*, pages 25–45. Springer, 2013.
- [103] F. Lenzen, J. Lellmann, F. Becker, and C. Schnörr. Solving quasi-variational inequalities for image restoration with adaptive constraint sets. *SIAM Journal on Imaging Sciences (SIIMS)*, 7(4):2139–2174, 2014.
- [104] F. Lenzen and O. Scherzer. A geometric PDE for interpolation of m-channel data. In X.-C. Tai, K. Mørken, M. Lysaker, and K.-A. Lie, editors, *Scale Space and Variational Methods in Computer Vision*, volume 5567 of *Lecture Notes in Computer Science*, pages 413–425. Springer, 2009.
- [105] F. Lenzen and O. Scherzer. Partial differential equations for zooming, deinterlacing and dejittering. *International Journal of Computer Vision*, 92(2):162–176, 2011.
- [106] F. Li, M. Ng, and C. Shen. Multiplicative noise removal with spatially varying regularization parameters. *SIAM Journal on Imaging Sciences*, 3(1):1–20, 2010.
- [107] W.-Y. Lin, L.-F. Cheong, P. Tan, G. Dong, and S. Liu. Simultaneous Camera Pose and Correspondence Estimation with Motion Coherence. *International Journal of Computer Vision*, pages 1–17, 2011.
- [108] B. Liu, S. Gould, and D. Koller. Single image depth estimation from predicted semantic labels. In *Computer Vision and Pattern Recognition, CVPR 2010 IEEE Conference on*, pages 1253–1260, 2010.
- [109] X. Liu and L. Huang. An efficient algorithm for adaptive total variation based image decomposition and restoration. *International Journal of Applied Mathematics and Computer Science*, 24(2):405–415, 2014.

- [110] D. B. Lucas and T. Kanade. An Iterative Image Registration Technique with an Application to Stereo Vision. In *DARPA*, pages 121–130, 1981.
- [111] F. Malgouyres and F. Guichard. Edge direction preserving image zooming: a mathematical and numerical analysis. *SIAM Journal on Numerical Analysis*, 39(1):1–37, 2001.
- [112] C. Mantegazza. *Lecture notes on mean curvature flow*, volume 290. Springer Science & Business Media, 2011.
- [113] E. Meijering. A chronology of interpolation: from ancient astronomy to modern signal and image processing. *Proceedings of the IEEE*, 90(3):319–342, 2002.
- [114] S. Meister, D. Kondermann, and B. Jähne. An outdoor stereo camera system for the generation of real-world benchmark datasets with ground truth. *SPIE Optical Engineering*, 51(2), 2012.
- [115] Y. Meyer. *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, volume 22 of *Univ. Lect. Series*. Amer. Math. Soc., 2001.
- [116] E. Mouragnona, M. Lhuilliera, M. Dhomea, F. Dekeyserb, and P. Sayd. Generic and real-time structure from motion using local bundle adjustment. *Image Vis. Comp.*, 27(8):1178–1193, 2009.
- [117] M. Mühlich and T. Aach. Analysis of multiple orientations. *Image Processing, IEEE Transactions on*, 18(7):1424–1437, July 2009.
- [118] R. Nair, F. Lenzen, S. Meister, H. Schäfer, C. S. Garbe, and D. Kondermann. High accuracy ToF and stereo sensor fusion at interactive rates. In *Computer Vision–ECCV 2012. Workshops and Demonstrations*, pages 1–11. Springer Berlin Heidelberg, 2012.
- [119] R. Nair, S. Meister, M. Lambers, M. Balda, H. Hoffmann, A. Kolb, D. Kondermann, and B. Jähne. Ground truth for evaluating time of flight imaging. In M. Grzegorzec, C. Theobalt, R. Koch, and A. Kolb, editors, *Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications*, volume 8200 of *Lecture Notes in Computer Science*, pages 52–74. Springer, 2013.
- [120] R. Nair, K. Ruhl, F. Lenzen, S. Meister, H. Schäfer, C. S. Garbe, M. Eisemann, M. Magnor, and D. Kondermann. A survey on time-of-flight stereo fusion. In M. Grzegorzec, C. Theobalt, R. Koch, and A. Kolb, editors, *Time-of-Flight and Depth Imaging: Sensors, Algorithms, and Applications*, volume 8200 of *Lecture Notes in Computer Science*, pages 105–127. Springer, 2013.
- [121] M. Z. Nashed. *Generalized Inverses and Applications: proceedings of an advanced seminar*, volume 1973. Academic Pr, 1976.

- 
- [122] Y. Nesterov and L. Scrimali. Solving strongly monotone variational and quasi-variational inequalities. *Discrete Contin. Dynam. Systems*, 31(4), 2011.
- [123] R. A. Newcombe and A. J. Davison. Live dense reconstruction with a single moving camera. In *Computer Vision and Pattern Recognition, CVPR 2010. IEEE Computer Society Conference on*, 2010.
- [124] R. A. Newcombe, S. J. Lovegrove, and A. J. Davison. DTAM: Dense Tracking and Mapping in Real-Time. In *2011 IEEE International Conference on Computer Vision (ICCV)*, pages 2320–2327, 2011.
- [125] C. Nieuwenhuis and D. Cremers. Spatially varying color distributions for interactive multilabel segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(5):1234–1247, May 2013.
- [126] D. Nister, O. Naroditsky, and J. Bergen. Visual odometry. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 1, pages 652 – 659, June 2004.
- [127] M. A. Noor and W. Oettli. On general nonlinear complementarity problems and quasi-equilibria. *Le Matematiche*, 49:313–331, 1994.
- [128] J. P. Oliveira, J. M. Bioucas-Dias, and M. A. Figueiredo. Adaptive total variation image deblurring: a majorization–minimization approach. *Signal Processing*, 89(9):1683–1693, 2009.
- [129] K. Papafitsoros and C.-B. Schönlieb. A combined first and second order variational approach for image reconstruction. *Journal of mathematical imaging and vision*, 48(2):308–338, 2014.
- [130] K. Papafitsoros, C.-B. Schönlieb, and B. Sengul. Combined first and second order total variation inpainting using split Bregman. *Image Processing On Line*, 2013:112–136, 2013.
- [131] N. Papenberg, A. Bruhn, T. Brox, S. Didas, and J. Weickert. Highly accurate optic flow computation with theoretically justified warping. *Int. J. Comput. Vision*, 67(2):141–158, 2006.
- [132] P. Perona, T. Shiota, and J. Malik. Anisotropic diffusion. In *Geometry-driven diffusion in computer vision*, pages 73–92. Springer, 1994.
- [133] C. Rabe, T. Müller, A. Wedel, and U. Franke. Dense, Robust, and Accurate Motion Field Estimation from Stereo Image Sequences in Real-Time. In K. Daniilidis, P. Maragos, and N. Paragios, editors, *Proceedings of the 11th European Conference on Computer Vision*, volume 6314 of *Lecture Notes in Computer Science*, pages 582–595. Springer, September 2010.

- [134] C. Rhemann, A. Hosni, M. Bleyer, C. Rother, and M. Gelautz. Fast Cost-Volume Filtering for Visual Correspondence and Beyond. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, pages 3017–3024, 2011.
- [135] W. Ring and B. Wirth. Optimization methods on Riemannian manifolds and their application to shape space. *SIAM Journal on Optimization*, 22(2):596–627, 2012.
- [136] R. Rockafellar and R. J.-B. Wets. *Variational Analysis*. Springer, 2nd edition, 2004.
- [137] R. T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization*, 14(5):877–898, 1976.
- [138] R. Rojas and P. Rodriguez. Spatially adaptive total variation image denoising under salt and pepper noise. In *Proceedings of the European Signal Processing Conference (EUSIPCO'11)*, pages 314–318, 2011.
- [139] A. Roussos and P. Maragos. Vector-valued image interpolation by an anisotropic diffusion-projection PDE. In *Scale Space and Variational Methods in Computer Vision*, pages 104–115. Springer, 2007.
- [140] A. Roussos and P. Maragos. Reversible interpolation of vectorial images by an anisotropic diffusion-projection PDE. *Int. J. Comput. Vision*, 84(2):130–145, 2009.
- [141] Y. Ruan, H. Fang, and Q. Chen. Semiblind image deconvolution with spatially adaptive total variation regularization. *Mathematical Problems in Engineering*, 2014.
- [142] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
- [143] J. Santner, M. Unger, T. Pock, C. Leistner, A. Saffari, and H. Bischof. Interactive texture segmentation using random forests and total variation. In *BMVC*, pages 1–12, 2009.
- [144] A. Saxena, S. H. Chung, and A. Y. Ng. 3-D Depth Reconstruction from a Single Still Image. *Int. J. Comput. Vision*, 76:53–69, 2008.
- [145] H. Schäfer, F. Lenzen, and C. S. Garbe. Depth and intensity based edge detection in time-of-flight images. In *3DV-Conference, 2013 International Conference on*, pages 111–118, 2013.
- [146] H. Schäfer, F. Lenzen, and C. S. Garbe. Model based scattering correction in time-of-flight cameras. *Optics Express*, 22:29835–29846, 2014.

- 
- [147] O. Scherzer. Denoising with higher order derivatives of bounded variation and an application to parameter estimation. *Computing*, 60:1–27, 1998.
- [148] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational methods in imaging*, volume 167 of *Applied Mathematical Sciences*. Springer, 2009.
- [149] O. Scherzer and J. Weickert. Relations between regularization and diffusion filtering. *Journal of Mathematical Imaging and Vision*, 12(1):43–63, 2000.
- [150] M. Schmidt. *Analysis, Modeling and Dynamic Optimization of 3D Time-of-Flight Imaging Systems*. PhD thesis, IWR, Fakultät für Physik und Astronomie, Univ. Heidelberg, 2011.
- [151] U. Schmidt, K. Schelten, and S. Roth. Bayesian deblurring with integrated noise estimation. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 2625–2632. IEEE, 2011.
- [152] S. Setzer, G. Steidl, and T. Teuber. Restoration of images with rotated shapes. *Numerical Algorithms*, 48:49–66, 2008.
- [153] S. Setzer, G. Steidl, and T. Teuber. Infimal convolution regularizations with discrete l1-type functionals. *Comm. Math. Sci.*, 9:797–872, 2011.
- [154] G. Steidl and T. Teuber. Anisotropic smoothing using double orientations. In *Proceedings of the 2nd International Conference on Scale Space and Variational Methods in Computer Vision*, pages 477–489, 2009.
- [155] G. Steidl and T. Teuber. Removing multiplicative noise by Douglas-Rachford splitting methods. *Journal of Mathematical Imaging and Vision*, 36(2):168–184, 2010.
- [156] D. Strong and T. Chan. Edge-preserving and scale-dependent properties of total variation regularization. *Inverse problems*, 19(6):S165, 2003.
- [157] D. M. Strong. *Adaptive Total Variation Minimizing Image Restoration*. PhD thesis, UCLA, 1997.
- [158] D. M. Strong, P. Blomgren, and T. F. Chan. Spatially adaptive local-feature-driven total variation minimizing image restoration. In *Optical Science, Engineering and Instrumentation'97*, volume 3167, pages 222–233. International Society for Optics and Photonics, 1997.
- [159] D. M. Strong and T. F. Chan. Spatially and scale adaptive total variation based regularization and anisotropic diffusion in image processing. In *Diffusion in Image Processing, UCLA Math Department CAM Report*, 1996.

- [160] J. Stühmer, S. Gumhold, and D. Cremers. Parallel Generalized Thresholding Scheme for Live Dense Geometry from a Handheld Camera. In A. Doucet, N. De Freitas, and N. Gordon, editors, *CVGPU*, 2010.
- [161] P. Sturm and B. Triggs. A factorization based algorithm for multi-image projective structure and motion. In *Computer Vision – ECCV 1996*, pages 709–720. Springer, 1996.
- [162] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother. A Comparative Study of Energy Minimization Methods for Markov Random Fields with Smoothness-Based Priors. *TPAMI*, 30:1068–1080, June 2008.
- [163] P. Thévenaz, T. Blu, and M. Unser. Image interpolation and resampling, 2000.
- [164] L. Tierney and J. B. Kadane. Accurate Approximations for Posterior Moments and Marginal Densities. *Journal of the American Statistical Association*, 81(393):82–86, 1986.
- [165] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon. *Bundle Adjustment – A Modern Synthesis*, volume 1883. Springer-Verlag, 2000.
- [166] D. Tschumperlé. Fast anisotropic smoothing of multi-valued images using curvature-preserving PDE’s. *International Journal of Computer Vision*, 68(1):65–82, 2006.
- [167] D. Tschumperle and R. Deriche. Vector-valued image regularization with PDEs: A common framework for different applications. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(4):506–517, 2005.
- [168] C. Udriste. *Convex functions and optimization methods on Riemannian manifolds*, volume 297. Springer, 1994.
- [169] L. Valgaerts, A. Bruhn, H. Zimmer, J. Weickert, C. Stoll, and C. Theobalt. Joint estimation of motion, structure and geometry from stereo sequences. In *Computer Vision – ECCV 2010*, Computer Vision – ECCV 2010, pages 568–581. Springer, 2010.
- [170] X. Wang, Y. Liu, H. Zhang, and L. Fang. A total variation model based on edge adaptive guiding function for remote sensing image de-noising. *International Journal of Applied Earth Observation and Geoinformation*, 34:89–95, 2015.
- [171] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: from error visibility to structural similarity. *Image Processing, IEEE Transactions on*, 13(4):600–612, 2004.

- 
- [172] A. Wedel, D. Cremers, T. Pock, and H. Bischof. Structure- and motion-adaptive regularization for high accuracy optic flow. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 1663–1668, Sept 2009.
- [173] A. Wedel, T. Pock, C. Zach, D. Cremers, and H. Bischof. An improved algorithm for TV-L1 optical flow. In *Proc. Dagstuhl Motion Workshop*, LNCS. Springer, 2008.
- [174] A. Wedel, C. Rabe, T. Vaudrey, T. Brox, U. Franke, and D. Cremers. Efficient dense scene flow from sparse or dense stereo data. In *Computer Vision – ECCV 2008*, volume 3021 of *LNCS*, 2008.
- [175] J. Weickert. *Anisotropic Diffusion in Image Processing*. Teubner Verlag, 1998.
- [176] A. Weishaupt, L. Bagnato, and P. Vanderghenst. Fast Structure from Motion for Planar Image Sequences. In *Proc. EUSIPCO*, 2010.
- [177] G. Welch and G. Bishop. An introduction to the Kalman filter, 1995.
- [178] M. Werlberger, W. Trobin, T. Pock, A. Wedel, D. Cremers, and H. Bischof. Anisotropic Huber-L1 optical flow. In *BMVC*, pages 1–11, 2009.
- [179] K. Yamaguchi, T. Hazan, D. McAllester, and R. Urtasun. Continuous Markov Random Fields for Robust Stereo Estimation. In *Computer Vision – ECCV 2012*, 2012.
- [180] L. Yan, H. Fang, and S. Zhong. Blind image deconvolution with spatially adaptive total variation regularization. *Optics letters*, 37(14):2778–2780, 2012.
- [181] Y. Yang. Globally convergent optimization algorithms on riemannian manifolds: Uniform framework for unconstrained and constrained optimization. *Journal of Optimization Theory and Applications*, 132(2):245–265, 2007.
- [182] C. Zach, T. Pock, and H. Bischof. A duality based approach for realtime TV-L1 optical flow. In *Pattern Recognition*, pages 214–223. Springer, 2007.
- [183] J. Zhang, R. Lai, and C.-C. Kuo. Latent fingerprint segmentation with adaptive total variation model. In *Biometrics (ICB), 2012 5th IAPR International Conference on*, pages 189–195, 2012.
- [184] Y. Zhao, J. G. Liu, B. Zhang, W. Hong, and Y.-R. Wu. Adaptive total variation regularization based SAR image despeckling and despeckling evaluation index. *Geoscience and Remote Sensing, IEEE Transactions on*, 53(5):2765–2774, 2015.
- [185] H. Zheng and O. Hellwich. An edge-driven total variation approach to image deblurring and denoising. In *Innovative Computing, Information and Control, 2006. ICICIC '06. First International Conference on*, volume 2, pages 705–710, 2006.

- [186] H. Zheng and O. Hellwich. Adaptive data-driven regularization for variational image restoration in the BV space. In *VISAPP (1)*, pages 53–60, 2007.