

Minimizing TGV-based Variational Models with Non-Convex Data Terms

Rene Ranftl, Thomas Pock, Horst Bischof

Institute for Computer Graphics and Vision
Graz University of Technology*
{ranftl,pock,bischof}@icg.tugraz.at

Abstract. We introduce a method to approximately minimize variational models with Total Generalized Variation regularization (TGV) and non-convex data terms. Our approach is based on a decomposition of the functional into two subproblems, which can be both solved globally optimal. Based on this decomposition we derive an iterative algorithm for the approximate minimization of the original non-convex problem. We apply the proposed algorithm to a state-of-the-art stereo model that was previously solved using coarse-to-fine warping, where we are able to show significant improvements in terms of accuracy.

Keywords: Total Generalized Variation, Optimization, Stereo

1 Introduction

Total Generalized Variation (TGV) [1], a generalization of the Total Variation (TV) regularization, has recently been successfully applied to a number of problems, like Optical Flow [2], Stereo [3] and Image Fusion [4]. Especially for Stereo and Optical Flow, TGV is arguably a better prior than the classical TV prior. For example in the second-order case, TGV does not penalize piecewise affine solutions. Such assumptions on planarity of the scene are frequently made in stereo matching (e.g. [5–7]) and also find application in optical flow estimation (e.g. [2, 8]).

However, TGV regularization currently is restricted to convex functionals (i.e. convex data terms). If the functional is non-convex, as it is the case in stereo matching, one has to rely on convex approximations to the non-convex problem, which often decreases the performance of the model. This is not the case for TV, where global solutions can be computed even in the presence of non-convex data terms, provided that the continuous label-space is discretized and some natural ordering can be imposed onto the resulting discrete label space [9]. The idea of this approach is to lift the functional to a higher dimensional space, where the resulting functional is convex. Similar results were shown by Ishikawa [10] for discrete first-order Markov Random Fields (MRF). The lifting approach [9] was later extended to a broader class of convex first-order priors such as Quadratic and Huber regularization [11].

* This work was supported by the Austrian Science Fund (project no. P22492)

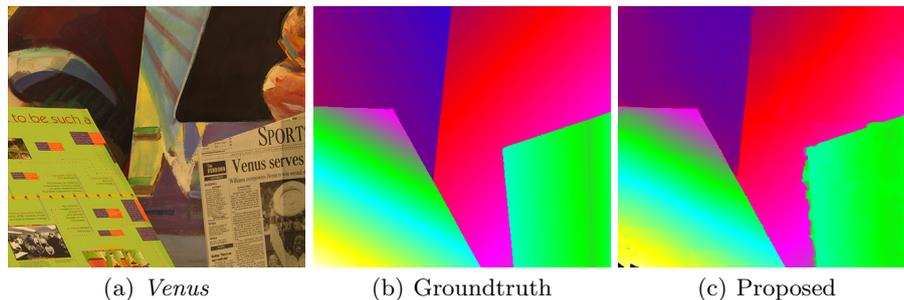


Fig. 1. Example from the Middlebury stereo dataset.

Previous work on Stereo and Optical Flow that used TGV regularization [2, 3] relied on the classical coarse-to-fine warping scheme [12] to approximately solve the original non-convex problem. The basic idea of this approach is to solve a series of convex models that arise from linearizations of the non-convex data term. In order to capture large motion or disparity ranges, respectively, this procedure has to be embedded into a coarse-to-fine framework, which is known to suffer from loss of fine details.

In the context of discrete MRFs, planarity assumptions can be enforced using a second-order prior. The resulting models can be approximately solved using a move-making strategy: The multi-label problem is reduced to a series of binary subproblems (each deciding if a node retains its label or switches to a proposed label), where each subproblem can be solved partially optimal [5]. The outcome of this approach crucially depends on the quality of the proposals in each move. Moreover, each of the subproblems is only solved partially optimal, which means that some nodes in the MRF may remain unlabeled.

Contribution. In this work we show how approximate solutions to non-convex functionals with TGV regularization can be computed. Our approach does not suffer from loss of fine details like the coarse-to-fine approaches do. The framework builds on the observation that functionals with TGV regularization and non-convex data terms can be split in two subproblems, where one is convex and the other, although non-convex, falls into the class of functionals covered by the lifting procedure described in [11] and can therefore be solved globally optimal.

In contrast to [5], where in each iteration a binary labeling problem, defined on a second-order energy, is solved, our approach solves a first-order multi-label problem in each iteration, in order to minimize the full second-order energy. This frees us from the need to specify proposals and also guarantees a complete labeling. Our splitting approach is similar to Alternating Convex Search [13], which itself falls under the broader class of Block-Relaxation methods [14].

We apply the proposed algorithm to a variational stereo model [3], which was solved using a coarse-to-fine strategy in the original formulation. By switching the optimization strategy to the herein proposed method, we are able to show

significant improvements in terms of accuracy. An exemplary result of the proposed method is shown in Figure 1. This is an example where the scene consists only of planes, which is perfectly modelled by the prior. Consequently we are able to recover high-quality disparity maps. Our evaluation shows that we obtain state-of-the-art results on the challenging KITTI stereo benchmark [15] as well as the Middlebury high-resolution benchmark [16].

2 Alternating Optimization

We focus on models with second-order TGV regularization, as this is the most widely used and also the simplest instance of TGV (besides TV), i.e. we consider functionals of the form

$$\min_{u,w} \underbrace{\alpha \int_{\Omega} |Dw|_{\Gamma}}_{E_1(w|u)} + \underbrace{\int_{\Omega} |Du - w|_{\Sigma} + \lambda \int_{\Omega} \rho(u)}_{E_2(u|w)}, \quad (1)$$

where $u : \Omega \rightarrow \mathbb{R}$ and $w : \Omega \rightarrow \mathbb{R}^2$, D is the distributional derivative, which is also well defined for discontinuous functionals, and the norms are defined as $|x|_M = \langle x, Mx \rangle^{\frac{1}{2}}$, M symmetric and positive definite. The introduction of the operator M will later allow us to easily incorporate anisotropic edge-weighted diffusion into the model. Note that for $\Gamma = I$ and $\Sigma = I$, the definition reduces to the standard definition of second-order TGV [1]. We will assume throughout the rest of this paper that the data term $\rho(u)$ is non-convex. Note that an extension of this basic formulation to higher-order instances of TGV is straight-forward, as it only involves a modified version of subproblem $E_1(w|u)$.

Our main observation is as follows: It is possible to decompose problem (1) into the two subproblems $E_1(w|u)$ and $E_2(u|w)$. Let the pair (u^*, w^*) be a global minimizer of (1), then it is obvious that the relation

$$u^* = \arg \min_u E_2(u|w^*) \quad (\text{S1})$$

$$w^* = \arg \min_w E_1(w|u^*) \quad (\text{S2})$$

holds, i.e. given w^* it is possible to deduce u^* by solving a possibly simpler subproblem and vice versa. Note that (S1) is a non-convex problem, while (S2) is a convex problem, which is equivalent to a generalized vectorial TV-L1 denoising problem [17]. This observation points to an iterative scheme for finding approximate solutions to (1):

$$\begin{aligned} u^{n+1} &= \arg \min_u E_2(u|w^n) \\ w^{n+1} &= \arg \min_w E_1(w|u^{n+1}). \end{aligned} \quad (\text{A1})$$

Note that by definition we have $E(u^n, w^n) \geq E(u^{n+1}, w^n) \geq E(u^{n+1}, w^{n+1})$ and $0 \leq E(u, w) < \infty$, $\forall(u, w)$, therefore the procedure will converge in the functional value, although not necessarily to a global optimum.

The update steps in (A1) already constitute the basic iterations of the proposed algorithm for optimizing (1). It remains to show how to solve the individual subproblems in each step.

2.1 Minimizing $E_2(u|w)$

The subproblem $E_2(u|w)$ is a non-convex variational problem with a non-convex data term and a convex regularization term. It was shown by Pock et al. [11] that problems with this special structure can be solved globally optimal using the framework of calibrations. The basic idea is to lift the problem to a higher-dimensional space, where a globally optimal solution to the original problem can be computed.

Let us first introduce the general framework: In order to find a minimizer u^* of functionals of the form

$$\min_u \int_{\Omega} f(x, u(x), Du), \quad (2)$$

we can solve the auxiliary problem

$$\min_{v \in \mathcal{C}} \sup_{\phi \in \mathcal{K}} \int_{\Omega \times \mathbb{R}} \phi \cdot Dv, \quad (3)$$

where the convex sets \mathcal{C} and \mathcal{K} are given by

$$\mathcal{C} = \left\{ v \in BV(\Omega \times \mathbb{R}; [0, 1]) : \lim_{t \rightarrow -\infty} v(x, t) = 1, \quad \lim_{t \rightarrow \infty} v(x, t) = 0 \right\}$$

and

$$\mathcal{K} = \{ \phi = (\phi_x, \phi_t) \in C_0(\Omega \times \mathbb{R}; \mathbb{R}^d \times \mathbb{R}) : \phi_t(x, t) \geq f^*(x, t, \phi_x(x, t)), \forall x, t \in \Omega \times \mathbb{R} \}. \quad (4)$$

Here f^* denotes the convex conjugate of the function f . Note that the sets \mathcal{C} and \mathcal{K} are defined point-wise. The intuition behind this formulation is, that instead of minimizing u directly, one represents the energy in terms of characteristic functions of its upper level-sets v . Given a minimizer v^* the corresponding minimizer u^* can be recovered by $u^*(x) = \int_{\mathbb{R}} v^*(x, t) dt$.

This formulation is very general, the specific form of the convex regularization term only influences the set \mathcal{K} . Pock et al. [11] derived the set \mathcal{K} for Quadratic, TV, Huber and Lipschitz regularization terms. In problem $E_2(u|w)$, the regularization term is similar to TV regularization, with the difference that a constant vector is subtracted from the gradient, before the absolute value is measured. We identify $f(x, t, p) = |p(x) - w^n(x)|_{\Sigma} + \lambda\rho(x, t)$, and consequently its convex conjugate with respect to p is

$$f^*(x, t, \phi) = \begin{cases} \langle \phi_x(x, t), w^n(x) \rangle - \lambda\rho(x, t), & \text{if } |\phi_x(x, t)|_{\Sigma} \leq 1 \\ \infty, & \text{else.} \end{cases}$$

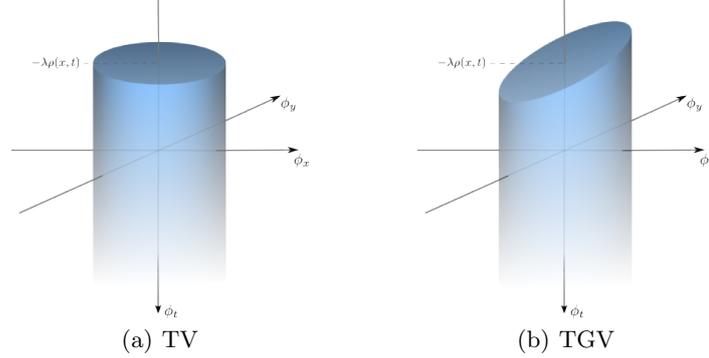


Fig. 2. The feasible set \mathcal{K} for (a) TV and (b) TGV.

The resulting set \mathcal{K} is illustrated in Figure 2(b). The feasible set for TV regularization is shown in Figure 2(a). It can be seen that for problem $E_2(u|w)$ the feasible set is slightly more complicated than in the TV case. While for TV the set is given by the interior of a cylinder with radius 1, which is bounded from below by a vertical plane centered at $(0, 0, -\lambda\rho(x, t))^T$, the set in the TGV case is bounded from below by plane that includes the point $(0, 0, -\lambda\rho(x, t))^T$ but can be arbitrarily oriented (in fact the normal of this plane is given by $[w^n, -1]^T$). This makes projection onto this set slightly harder, as a closed-form solution is no longer available.

Discretization and Optimization. In order to solve (3) it is necessary to discretize the domain $\Omega \times \mathbb{R}$ of the continuous functions v and ρ . For the sake of simplicity let us only consider the case $\Omega \subset \mathbb{R} \times \mathbb{R}$, higher-dimensional cases can be derived analogously.

We discretize on a three-dimensional grid of size $N_x \times N_y \times N_t$ with discretization steps Δx , Δy and Δt :

$$G^\Delta = \{(i\Delta x, j\Delta y, k\Delta t) : (0, 0, 0) \leq (i, j, k) < (N_x, N_y, N_t)\}. \quad (5)$$

Here the triple (i, j, k) denotes the location in the grid.

For numerical reasons we replace the vector field ϕ_x , with a rotated version $\Sigma^{\frac{1}{2}}\phi_x$, which leads to a simplification of the convex set \mathcal{K}^Δ , without changing the formulation.

The feasible sets for the discrete version of (3) are then given by

$$\mathcal{C}^\Delta = \{v^\Delta \in [0, 1]^{N_x N_y N_t} : v_{i,j,0}^\Delta = 1, v_{i,j,N_t-1}^\Delta = 0\} \quad (6)$$

and

$$\begin{aligned} \mathcal{K}^\Delta = \{ \phi^\Delta = (\phi_x^\Delta, \phi_y^\Delta, \phi_t^\Delta) \in \mathbb{R}^{3N_x N_y N_t} : \\ (\phi_t^\Delta)_{i,j,k} + \lambda(\rho)_{i,j,k} \geq \left\langle (\phi_x^\Delta, \phi_y^\Delta)_{i,j,k}^T, \Sigma^{\frac{1}{2}} w_{i,j} \right\rangle, \\ |(\phi_x^\Delta, \phi_y^\Delta)_{i,j,k}^T|_2 \leq 1, \quad \forall (i, j, k) \in G^\Delta \}. \quad (7) \end{aligned}$$

In order to discretize the differential operator D , we use forward differences with Neumann boundary conditions. Furthermore we allow Σ to vary locally, which allows us to incorporate image-driven TGV regularization similar to [3] into the framework, i.e. we define a linear operator $\nabla_{\Sigma} : \mathbb{R}^{N_x N_y N_t} \rightarrow \mathbb{R}^{3N_x N_y N_t}$, with

$$(\nabla_{\Sigma} v^{\Delta})_{i,j,k} = \begin{pmatrix} \Sigma_{i,j}^{\frac{1}{2}} & 0 \\ \Sigma_{i,j}^{\frac{1}{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} (\delta_x v^{\Delta})_{i,j,k} \\ (\delta_y v^{\Delta})_{i,j,k} \\ (\delta_t v^{\Delta})_{i,j,k} \end{pmatrix} \quad (8)$$

and

$$(\delta_x v^{\Delta})_{i,j,k} = \begin{cases} (v_{i+1,j,k}^{\Delta} - v_{i,j,k}^{\Delta})/\Delta x & \text{if } i < N_x - 1 \\ 0 & \text{else} \end{cases} \quad (9)$$

$$(\delta_y v^{\Delta})_{i,j,k} = \begin{cases} (v_{i,j+1,k}^{\Delta} - v_{i,j,k}^{\Delta})/\Delta y & \text{if } j < N_y - 1 \\ 0 & \text{else} \end{cases} \quad (10)$$

$$(\delta_t v^{\Delta})_{i,j,k} = \begin{cases} (v_{i,j,k+1}^{\Delta} - v_{i,j,k}^{\Delta})/\Delta t & \text{if } k < N_t - 1 \\ 0 & \text{else.} \end{cases} \quad (11)$$

Note that (8) reduces to the standard discretization of a gradient operator, if $\Sigma_{i,j}^{\frac{1}{2}}$ is set to identity everywhere. On the other hand, it is possible to incorporate image-driven diffusion into the model by setting the matrix appropriately. We will later discuss the specific choice of this matrix.

The discrete version of (3) is now given by

$$\min_{v^{\Delta} \in \mathcal{C}^{\Delta}} \max_{\phi^{\Delta} \in \mathcal{K}^{\Delta}} \langle \nabla_{\Sigma} v^{\Delta}, \phi^{\Delta} \rangle \quad (12)$$

For optimization of the convex-concave saddle-point problem (12) we use the primal-dual algorithm [18]. The iterations of this algorithm are shown in Algorithm 1.

A crucial part of this algorithm are the pointwise projections $\text{Proj}_{\mathcal{K}^{\Delta}}(\cdot)$ and $\text{Proj}_{\mathcal{C}^{\Delta}}(\cdot)$ respectively. The projection of the primal variables is simple and can be carried out in closed-form:

$$(\text{Proj}_{\mathcal{C}^{\Delta}}(\hat{v}))_{i,j,k} = \begin{cases} \max\{0, \min\{1, \hat{v}_{i,j,k}\}\} & \text{if } k > 1 \\ 1 & \text{else.} \end{cases}$$

Algorithm 1. Primal-dual algorithm for solving (12)

<p>1. <i>Initialize</i> Set $(v^{\Delta})^0 \in \mathcal{C}^{\Delta}$, $(\phi^{\Delta})^0 \in \mathcal{K}^{\Delta}$, $(\bar{v})^0 = (v^{\Delta})^0$, $n = 0$ Choose time-steps $\tau, \sigma > 0$, $\tau\sigma < \frac{1}{\ \nabla_{\Sigma}\ ^2}$</p> <p>2. <i>Iterate</i></p> $\begin{cases} (\phi^{\Delta})^{n+1} \leftarrow \text{Proj}_{\mathcal{K}^{\Delta}}((\phi^{\Delta})^n + \sigma(\nabla_{\Sigma} \bar{v}^n)) \\ (v^{\Delta})^{n+1} \leftarrow \text{Proj}_{\mathcal{C}^{\Delta}}((v^{\Delta})^n - \tau(\nabla_{\Sigma}^T(\phi^{\Delta})^{n+1})) \\ \bar{v}^{n+1} \leftarrow 2(v^{\Delta})^{n+1} - (v^{\Delta})^n \end{cases}$

The projections for the dual variables $\text{Proj}_{\mathcal{K}^\Delta}(\cdot)$, although also point-wise, are more complicated. The feasible set \mathcal{K}^Δ is defined point-wise via the intersection of two convex sets. We experimented with different variants to incorporate these constraints: Lagrange multipliers, solving the projection problem in each iteration of the primal-dual algorithm using FISTA [19] (including a preconditioned variant) and finally Dykstra's Projection algorithm [20]. Our experiments show that Dykstra's algorithm provides the best performance for this type of problem and is very light-weight, we therefore resort to this variant to incorporate the dual constraints. The iterations of Dykstra's algorithm are shown in Algorithm 2, where we set $n = [w_{i,j,k}^n, -1]^T$ and $c = \lambda \rho_{i,j,k}$. In practice we run the algorithm until the distances to both convex sets ($|x^n - y^n|_2$ and $|y^n - x^{n+1}|_2$) are below a tolerance of 10^{-3} (which is typically achieved in under 10 iterations).

Algorithm 2. Algorithm for projecting onto the set \mathcal{K}

<p>1. <i>Initialize</i> Set $n = 0$, $x^0 = \phi_{i,j,k}$, $p^0 = 0$, $q^0 = 0$</p> <p>2. <i>Iterate</i></p> $\begin{cases} y^n & \leftarrow \frac{(x^n + p^n)}{\max\{1, x^n + p^n _2\}} \\ p^{n+1} & \leftarrow p^n + x^n - y^n \\ x^{n+1} & \leftarrow \begin{cases} y^n + q^n & \text{if } \langle y^n + q^n, n \rangle \leq c \\ y^n + q^n - \frac{\langle y^n + q^n, n \rangle - c}{\langle n, n \rangle} n & \text{else} \end{cases} \\ q^{n+1} & \leftarrow q^n + y^n - x^{n+1} \end{cases}$

2.2 Minimizing $E_1(w|u)$

The subproblem $E_1(w|u)$ is a non-smooth convex optimization problem, which can be solved using standard techniques. We will show how to cast this problem in a saddle-point formulation and again apply the primal-dual algorithm [18].

The optimization problem reads

$$\min_w \int_{\Omega} |Du^{n+1} - w|_{\Sigma} + \alpha \int_{\Omega} |Dw|_{\Gamma}, \quad (13)$$

where u^{n+1} is given by the last solution of problem $E_2(u|w)$. Note that this problem corresponds to denoising the gradients of u^{n+1} .

Using the definition $\text{div}_M z = \text{div}(M^{\frac{1}{2}} z)$, the equivalent saddle-point formulation is given by:

$$\min_w \sup_{\substack{\|p\|_{\infty} \leq 1 \\ \|q\|_{\infty} \leq 1}} - \int_{\Omega} u^{n+1} \text{div}_{\Sigma} p \, dx - \int_{\Omega} \langle w, \Sigma^{\frac{1}{2}} p + \alpha \text{div}_{\Gamma} q \rangle \, dx. \quad (14)$$

Discretization of (14) follows analogously to the lifted problem: The two-dimensional grid is given by

$$\hat{\mathcal{K}}^\Delta = \{(i\Delta x, j\Delta y) : (0, 0) \leq (i, j) < (N_x, N_y)\}, \quad (15)$$

Algorithm 3. Primal-dual algorithm for solving (17)

1. <i>Initialize</i>	
Set $(u^\Delta)^0 = u^{n+1}, (w^\Delta)^0 = \nabla_\Sigma u^{n+1}, (\bar{u})^0 = (u^\Delta)^0, (\bar{w})^0 = (w^\Delta)^0$	
Set $((p^\Delta)^0)_{i,j} = (0, 0)^T, (q^\Delta)_{i,j} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, n = 0$	
Choose time-steps $\tau, \sigma > 0, \tau\sigma < \frac{1}{\ A\ ^2}$, where $A = \begin{pmatrix} \nabla_\Sigma & -I \\ 0 & \mathcal{D}_\Gamma \end{pmatrix}$	
2. <i>Iterate</i>	
$\begin{cases} (p^\Delta)^{n+1} \\ (q^\Delta)^{n+1} \\ (u^\Delta)^{n+1} \\ (w^\Delta)^{n+1} \\ \bar{u}^{n+1} \end{cases}$	$\leftarrow \begin{cases} \text{Proj}_{\ p\ _\infty \leq 1}((p^\Delta)^n + \sigma(\nabla_\Sigma \bar{u}^n - \text{diag}(\Sigma_{i,j}^{\frac{1}{2}})\bar{w}^n)) \\ \text{Proj}_{\ q\ _\infty \leq 1}((q^\Delta)^n + \sigma(\mathcal{D}_\Gamma \bar{w}^n)) \\ \text{Proj}_{\mathcal{B}}((u^\Delta)^n - \tau \nabla_\Sigma^T (p^\Delta)^{n+1}) \\ (u^\Delta)^n - \tau(\mathcal{D}_\Gamma^T (q^\Delta)^{n+1} - \text{diag}(\Sigma_{i,j}^{\frac{1}{2}})(p^\Delta)^{n+1}) \\ 2(u^\Delta)^{n+1} - (u^\Delta)^n, \quad \bar{w}^{n+1} \leftarrow 2(w^\Delta)^{n+1} - (w^\Delta)^n \end{cases}$

where the tuple (i, j) again denotes a location in the grid, which also coincides with the spatial coordinates of the lifted problem. The discrete saddle-point problem can be written as

$$\min_{w^\Delta} \max_{\substack{\|p^\Delta\|_\infty \leq 1 \\ \|q^\Delta\|_\infty \leq 1}} \langle \nabla_\Sigma u^{n+1}, p^\Delta \rangle - \langle \text{diag}(\Sigma^{\frac{1}{2}})w^\Delta, p^\Delta \rangle + \alpha \langle \mathcal{D}_\Gamma w^\Delta, q^\Delta \rangle, \quad (16)$$

where the discrete differential operators ∇_Σ and \mathcal{D}_Γ are again based on forward differences with Neumann boundary conditions, i.e. we have

$$(\nabla_\Sigma u^\Delta)_{i,j} = \Sigma_{i,j}^{\frac{1}{2}} \begin{pmatrix} (\delta_x u^\Delta)_{i,j} \\ (\delta_y u^\Delta)_{i,j} \end{pmatrix} \quad (\mathcal{D}_\Gamma w^\Delta)_{i,j} = \Gamma_{i,j}^{\frac{1}{2}} \begin{pmatrix} (\delta_x w_1^\Delta)_{i,j} & (\delta_y w_2^\Delta)_{i,j} \\ (\delta_y w_1^\Delta)_{i,j} & (\delta_x w_2^\Delta)_{i,j} \end{pmatrix}.$$

In practice direct usage of (13) for the estimation of the second-order part w may be problematic if the discretization step Δt for the solution of the lifted problem was chosen too coarsely. In this case discretization artifacts are propagated from the lifted problem to problem (13), which may deteriorate the estimation of the second-order part, since in the context of this subproblem such artifacts are merely additional edges.

To cope with this problem, we modify (16) to allow u^{n+1} to slightly vary in a neighborhood of half the discretization step Δt of the lifted problem:

$$\min_{w^\Delta, u^\Delta \in \mathcal{B}} \max_{\substack{\|p^\Delta\|_\infty \leq 1 \\ \|q^\Delta\|_\infty \leq 1}} \langle \nabla_\Sigma u^\Delta - \text{diag}(\Sigma_{i,j}^{\frac{1}{2}})w^\Delta, p^\Delta \rangle + \alpha \langle \mathcal{D}_\Gamma w^\Delta, q^\Delta \rangle, \quad (17)$$

where $\mathcal{B} = \{u^\Delta \in \mathbb{R}^{N_x N_y} : |(u^\Delta)_{i,j} - (u^{n+1})_{i,j}| \leq \Delta t/2\}$.

The iterations for optimizing (17) are shown in Algorithm 3. As before, we again have to perform projections onto convex sets in each iteration of the algorithm. The projections of the dual variables are given by $(\text{Proj}_{\|r\|_\infty \leq 1}(r))_{i,j} = \frac{r_{i,j}}{\max\{1, |r_{i,j}|\}} \cdot \frac{1}{2}$. For the primal variables u , the projection onto \mathcal{B} can be computed by clamping $(u)_{i,j}$ in the interval $[(u^{n+1})_{i,j} - \frac{\Delta t}{2}, (u^{n+1})_{i,j} + \frac{\Delta t}{2}]$.

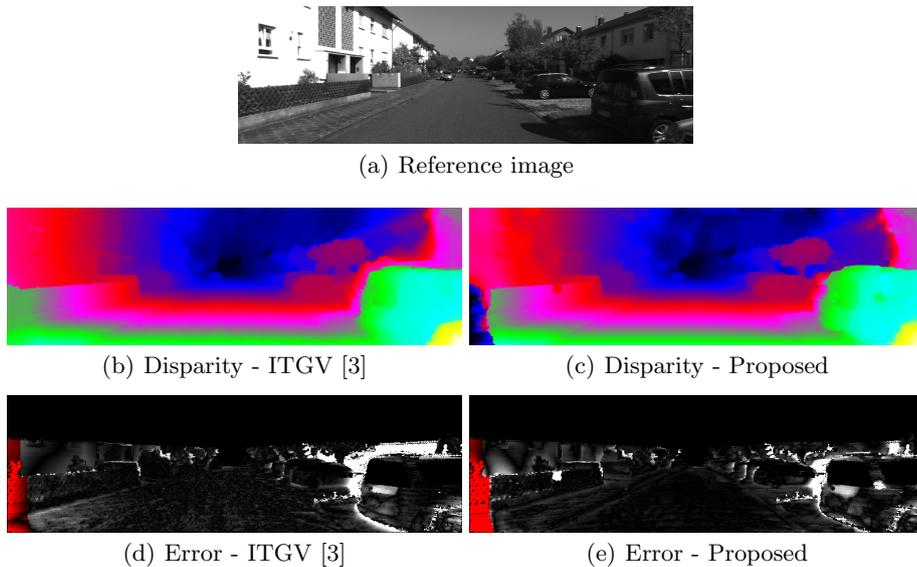


Fig. 3. Example from the KITTI benchmark for (b) ITGV [3] and (c) the proposed algorithm. The corresponding error maps are shown in (d) and (e). Occluded pixels are marked red in the error maps.

3 Application to Stereo

We show the effectiveness of our optimization approach on the application of stereo matching. We use the variational model that was proposed in [3] as basis for our experiments. This model introduced an image-driven TGV regularizer and based the matching term on the Census Transform. The original formulation used a warping procedure together with a coarse-to-fine scheme for the optimization. Such approaches are commonly used in variational stereo and optical flow, but are known to suffer from loss of detail due to the downsampling procedure.

Let us briefly explain, how the model [3] is realized in our framework: The matching term is based on the ternary Census transform [21]. We denote the ternary Census transform of the image I by $C(I)$. Then the matching cost for disparity t is given by the Hamming distance [22] between the ternary Census transforms of the warped matching image I_L and the reference image I_R , i.e.:

$$\hat{\rho}(x, t) = \Delta(C(I_L(x + [t, 0]^T)), C(I_R(x))), \quad \Delta(p, q) = \sum_{p_i \neq q_i} 1. \quad (18)$$

In order to cope with small calibration errors and to improve robustness with respect to the discretization, we employ a similar strategy to the Birchfeld-Tomasi dissimilarity measure [23], i.e. we sample the cost in a neighborhood of x and assign the minimum value as the final data term:

$$\rho(x, t) = \min\{\hat{\rho}(x, t), \hat{\rho}(x + a, t), \hat{\rho}(x - a, t), \hat{\rho}(x + b, t), \hat{\rho}(x - b, t)\}, \quad (19)$$

where the offset vectors a and b are given by $a = [\frac{\Delta x}{2}, 0]^T$ and $b = [0, \frac{\Delta y}{2}]^T$.

Image-driven regularization can be realized by setting $\Gamma_{i,j}^{\frac{1}{2}} = I$ and $\Sigma_{i,j}^{\frac{1}{2}} = \exp(-\gamma |\nabla I_L|_{i,j}^\beta) nn^T + n^\perp n^{\perp T}$, where $n = (\frac{\nabla I_L}{|\nabla I_L|})_{i,j}$ and $\gamma, \beta > 0$.

Evaluation. We focus the evaluation on qualitative results on the task of stereo estimation, instead of direct comparisons of final energies. A meaningful and fair comparison in terms of final energies between our approach and the baseline [3] is hardly possible, since the results of the baseline heavily depend on the parameters of the coarse-to-fine strategy.

We first compare the proposed approach to the baseline algorithm using the KITTI stereo benchmark [15]. This benchmark consists of 195 test images and 194 training images captured from an automotive platform. Groundtruth data is given in the form of semi-dense disparity maps that were captured using a laser sensor. We used the groundtruth that is provided with the training images to tune the parameters of the model. For all experiments the discretization of the disparity range was fixed to $\Delta t = 1px$ and Algorithm (A1) was run for 10 iterations. The optimization of each subproblem was run for 2000 iterations.

Figure 3 shows an example from the test set and compares the proposed approach (Figure 3 (c)) to the baseline approach (Figure 3 (b)). We observe that the proposed approach preserves more fine details and is better at handling large disparities than the coarse-to-fine approach. This higher accuracy is also reflected in the average number of bad pixels on the test set (Table 1), where the proposed approach currently ranks second, while the baseline is ranked on the 7th place. Note, that the method shows a slightly worse inpainting capability, when compared to the baseline, in areas, where there is no overlap between the input images, which results in a slightly higher error if those regions are considered in the evaluation (Avg-All in Table 1).

Our second evaluation uses a subset of 9 images from the Middlebury high-resolution benchmark [16] (*Teddy, Cones, Lamp2, Cloth3, Aloe, Art, Dolls, Baby3, Rocks2*). Exemplary results for this benchmark are shown in Figure 4. The average scores for different error thresholds and a comparison to state-of-the-art methods is shown in Table 2. We again observe that the proposed method is competitive to the other methods.

Table 1. Results on the KITTI-Benchmark. Columns *Out-Noc* and *Out-All* show the average percentage of pixels with an error larger than 3px in non-occluded and all regions, respectively. Columns *Avg-Noc* and *Avg-All* show the mean absolute errors.

Rank	Method	Out-Noc	Out-All	Avg-Noc	Avg-All	Runtime
1	PCBP [7]	4.13 %	5.45 %	0.9 px	1.2 px	5 min
2	Proposed	5.05 %	6.91 %	1.0 px	1.6 px	6 min
3	iSGM [24]	5.16 %	7.19 %	1.2 px	2.1 px	8s
7	ITGV [3]	6.31 %	7.40 %	1.3 px	1.5 px	7s

Table 2. Error on the Middlebury high-resolution benchmark.

Method	> 2 pixels	> 3 pixels	> 4 pixels	> 5 pixels
PCBP [7]	2.8 %	2.4 %	2.1 %	2.0 %
Proposed	4.4 %	3.1 %	2.5 %	2.2 %
ELAS [25]	4.7 %	3.9 %	3.5 %	3.2 %
OCV-SGBM [26]	5.9 %	5.5 %	5.3 %	5.2 %

4 Conclusion

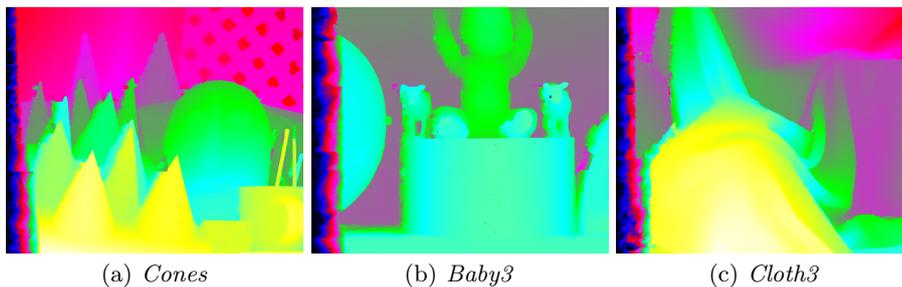
We presented an approach to approximately solve variational models with Total Generalized Variation regularization and non-convex data terms. Our approach alternates between solving a non-convex subproblem that can be solved globally optimal using functional lifting, and solving a convex subproblem.

We demonstrated the benefit of our approach on a variational stereo model that was previously solved using coarse-to-fine warping. Experiments on the challenging KITTI stereo benchmark show that this alternating minimization algorithm is able to significantly increase the performance of the model and consequently provides state-of-the-art results.

For future work, we plan to extend our approach to non-convex variants of TGV (i.e. truncated potentials). While we expect such regularization terms to be stronger priors and a splitting is in principle still possible, the problem is much harder to solve, because both of the resulting subproblems are non-convex.

References

1. Bredies, K., Kunisch, K., Pock, T.: Total Generalized Variation. *SIAM J. Img. Sci.* **3**(3) (2010) 492–526
2. Werlberger, M.: Convex Approaches for High Performance Video Processing. PhD thesis, Institute for Computer Graphics and Vision, Graz University of Technology, Graz, Austria (2012)
3. Ranftl, R., Gehrig, S., Pock, T., Bischof, H.: Pushing the Limits of Stereo Using Variational Stereo Estimation. In: *Proc. Intelligent Vehicles Symposium.* (2012)

**Fig. 4.** Example results from the Middlebury stereo benchmark.

4. Pock, T., Zebedin, L., Bischof, H.: TGV-Fusion. Maurer Festschrift, LNCS 6570 (2011) 245–258
5. Woodford, O., Torr, P., Reid, I., Fitzgibbon, A.: Global stereo reconstruction under second order smoothness priors. In: Proc. CVPR. (2008) 1–8
6. Bleyer, M., Rhemann, C., Rother, C.: Patchmatch Stereo - Stereo Matching with Slanted Support Windows. In: Proc. BMVC. (2011) 14.1–14.11
7. Yamaguchi, K., Hazan, T., McAllester, D., Urtasun, R.: Continuous markov random fields for robust stereo estimation. In: Proc. ECCV. (2012) 45–58
8. Trobin, W., Pock, T., Cremers, D., Bischof, H.: An unbiased second-order prior for high-accuracy motion estimation. In: Proc. DAGM. (2008) 396–405
9. Pock, T., Schoenemann, T., Graber, G., Bischof, H., Cremers, D.: A convex formulation of continuous multi-label problems. In: Proc. ECCV. (2008) 792–805
10. Ishikawa, H.: Exact optimization for markov random fields with convex priors. PAMI **25** (2003) 1333–1336
11. Pock, T., Cremers, D., Bischof, H., Chambolle, A.: Global solutions of variational models with convex regularization. SIAM J. Img. Sci. **3**(4) (2010) 1122–1145
12. Brox, T., Bruhn, A., Papenbergh, N., Weickert, J.: High accuracy optical flow estimation based on a theory for warping. In: Proc. ECCV. (2004) 25–36
13. Gorski, J., Pfeuffer, F., Klamroth, K.: Biconvex sets and optimization with biconvex functions: a survey and extensions. Mathematical Methods of Operations Research **66** (2007) 373–407
14. de Leeuw, J.: Block-relaxation algorithms in statistics. Technical report, Dept. of Statistics, UCLA (1994)
15. Geiger, A., Lenz, P., Urtasun, R.: Are we ready for autonomous driving? the kitti vision benchmark suite. In: Proc. CVPR. (2012) 3354–3361
16. Scharstein, D., Szeliski, R., Zabih, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. In: IEEE Workshop on Stereo and Multi-Baseline Vision (SMBV). (2001) 131–140
17. Nikolova, M.: A variational approach to remove outliers and impulse noise. JMIV **20**(1-2) (2004) 99–120
18. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. JMIV **40** (2011) 120–145
19. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Img. Sci. **2**(1) (2009) 183–202
20. Boyle, J.P., Dykstra, R.L.: A method for finding projections onto the intersection of convex sets in Hilbert spaces. Lecture Notes in Statistics **37** (1986) 28–47
21. Zabih, R., Li, J.W.: Non-parametric local transforms for computing visual correspondence. In: Proc. ECCV. (1994) 151–158
22. Hamming, R.W.: Error detecting and error correcting codes. BELL SYSTEM TECHNICAL JOURNAL **29**(2) (1950) 147–160
23. Birchfield, S., Tomasi, C.: A pixel dissimilarity measure that is insensitive to image sampling. PAMI **20**(4) (1998) 401–406
24. Hermann, S., Klette, R.: Iterative semi-global matching for robust driver assistance systems. In: Proc. ACCV. (2012)
25. Geiger, A., Roser, M., Urtasun, R.: Efficient large-scale stereo matching. In: Proc. ACCV. (2010) 25–38
26. Hirschmüller, H.: Accurate and efficient stereo processing by semi-global matching and mutual information. In: Proc. CVPR. (2005) 807–814