Sublabel-accurate convex relaxation with total generalized variation regularization

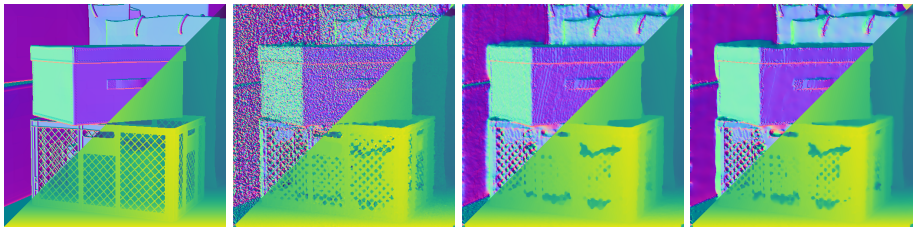
Michael Strecke[0000−0002−0322−0653] and Bastian Goldluecke[0000−0003−3427−4029]

University of Konstanz, Konstanz, Germany
firstname.lastname@uni-konstanz.de

Abstract. We propose a novel idea to introduce regularization based on second order total generalized variation (TGV) into optimization frameworks based on functional lifting. The proposed formulation extends a recent sublabel-accurate relaxation for multi-label problems and thus allows for accurate solutions using only a small number of labels, significantly improving over previous approaches towards lifting the total generalized variation. Moreover, even recent sublabel accurate methods exhibit staircasing artifacts when used in conjunction with common first order regularizers such as the total variation (TV). This becomes very obvious for example when computing derivatives of disparity maps computed with these methods to obtain normals, which immediately reveals their local flatness and yields inaccurate normal maps. We show that our approach is effective in reducing these artifacts, obtaining disparity maps with a smooth normal field in a single optimization pass.

1 Introduction

Many computer vision tasks can be formulated as continuous optimization problems over a label assignment $u : \Omega \to \Gamma$, where $\Omega \subset \mathbb{R}^d$ denotes the image domain and $\Gamma \subset \mathbb{R}^n$ the label domain. Correct solutions are characterized as the minimizers of an energy functional $E(u)$, which is designed in such a way that low-energy configurations correspond to some desired property, such as u being a smooth disparity map consistent with a given stereo pair. Thus, the energy E typically consists of a (non-convex) point-wise label cost which optimizes the fit to the observed data, and a regularization term which models interactions between neighbouring points. In order to find the optimal solutions u, approaches based on Markov random fields (MRFs) [9,4] discretize Ω as a set of nodes (e.g., pixels or superpixels) and Γ as a set of labels $\{1, \dots, \ell\}$, with the graph cut class of methods as a popular way to obtain minimizers [13]. Notably, the construction by Ishikawa [12] allows to obtain globally optimal solutions for convex interactions despite the data term being non-convex. A spatially continuous reformulation of this approach [18], based on the idea of functional lifting or the calibration method [1], reduces grid bias and memory requirements. Other related work in this context studies more general regularizers based on embeddings of the labels [14], structured label spaces [10,7] or the relationship between discrete and continuous approaches [23].



Ground truth TV [15] Huber [16] Proposed TGV^2

Fig. 1. Recent sublabel-accurate optimization algorithms [15,16] yield accurate and at first glance smooth disparity maps (bottom right corners). However, the surface normal maps (top left corners) obtained from the disparity maps by computing derivatives [22] are often still very noisy due to staircasing artifacts. The proposed approach penalizes second order derivatives based on the TGV^2-prior and yields smooth disparity and normal maps in a single optimization pass.

However, using a *discrete* label space does not allow for a faithful approximation of the underlying *continuous* model in practical problems such as image denoising or stereo matching. The above approaches [12,18] often yield results which exhibit a strong label bias degrading the result for a coarse sampling, or the discretization leads to unreasonably high demands in memory when using a finer sampling. In [15,16], Moellenhoff et al. thus substantially generalize the idea of functional lifting and derive formulations based on a fully continuous model. While their work allows for a sublabel-accurate approximation of the original (possibly nonconvex) energy, the results still exhibit some fine scale staircasing artifacts as can be seen in Fig. 1. This is particularly bad if derivatives of the result are required, such as in the above example, where normal maps are computed directly from the disparity maps [22], as this strongly emphasizes the artifacts. In practice, the normal maps become useless for subsequent tasks such as intrinsic image decomposition. Thus, [22] propose a two-pass framework, with a separate optimization pass for surface normal maps. While this achieves very good results, it is conceptually not as elegant as solving a single problem.

1.1 Contribution

In this work, we propose an empirical extension to the framework by Moellenhoff et al. [15,16], which is based on the total generalized variation (TGV) [5]. Although a mathematical validation of our approximation remains open, we demonstrate in quantitative experiments that our method obtains reasonable solutions to the optimization problem. In particular, we clearly improve over a previous approach to approximately lifting the total generalized variation [20], arriving closer to the optimal solution with less label bias due to sublabel accuracy. Furthermore, we show that our regularization effectively manages to remove staircasing artifacts produced e.g. by total variation (TV) regularization as in [15], and provides similarly accurate normals when applied to a disparity estimation task as the approach [22], which explicitly smoothes normals in a post-processing step.

2 Background and related work

2.1 Preliminaries

Our formulation is an empirical extension to the one proposed by Moellenhoff et al. [16], who propose a fully continuous model inspired by the celebrated Mumford-Shah functional [2,17],

$$\begin{aligned} E(u) = &\int_{\Omega \setminus J_u} f(x, u(x), \nabla u(x)) \,\mathrm{d}x \\ &+ \int_{J_u} d\left(x, u^-(x), u^+(x), \nu_u(x)\right) \,\mathrm{d}\mathcal{H}^{n-1}(x). \end{aligned} \tag{1}$$

The model is composed of two different integrands for the region $\Omega \setminus J_u$, where u is continuous, and the $(n-1)$-dimensional discontinuity set $J_u \subset \Omega$. The integrand $f : \Omega \times \Gamma \times \mathbb{R}^n \to [0, \infty]$ for the continuous part is a combined dataterm and regularizer, where the regularizer penalizes variations in terms of the gradient ∇u. On the discontinuity set J_u, the function $d : \Omega \times \Gamma \times \Gamma \times \mathcal{S}^{n-1} \to [0, \infty]$ penalizes jumps from u^- to u^+ in unit direction ν_u.

The energy (1) is defined for u in the space of *special functions of bounded variation* ($\mathcal{SBV}$). This is a subset of the space of functions $\mathcal{BV}(\Omega)$ of bounded variation, which are those functions $u \in L^1(\Omega; \mathbb{R})$ such that the total variation

$$\mathrm{TV}(u) = \sup \left\{ \int_\Omega u \mathrm{Div} \varphi \,\mathrm{d}x : \varphi \in C_c^1(\Omega; \mathbb{R}^n) \right\} \tag{2}$$

is finite. Functions in $\mathcal{SBV}(\Omega)$ are now exactly those $u \in \mathcal{BV}(\Omega)$ whose distributional derivative Du can be decomposed into a continuous and a jump part as required for (1),

$$Du = \nabla u \cdot \mathcal{L}^n + (u^+ - u^-)\nu_u \cdot \mathcal{H}^{n-1} \lfloor J_u, \tag{3}$$

with $\mathcal{L}^n$ denoting the n-dimensional Lebesgue measure and $\mathcal{H}^{n-1} \lfloor J_u$ the $(n-1)$-dimensional Hausdorff measure restricted to the jump set J_u.

In the above formulation, f can be nonconvex in the first two variables, and thus allows a surprisingly large class of vision problems to be represented by (1). Although this makes (1) a difficult nonconvex optimization problem, Moellenhoff et al. [16] found a sublabel-accurate formulation which employs a piecewise convex relaxation of the energy between labels. As we build upon their framework, we follow [16] and make the following simplifying assumptions on the components of the energy (1):

- The Lagrangian f in (1) is seperable into a possibly nonconvex dataterm $\rho : \Omega \times \Gamma \to \mathbb{R}$ and convex regularizer $\eta : \Omega \times \mathbb{R}^n$,

$$f(x, t, g) = \rho(x, t) + \eta(x, g). \tag{4}$$

- The isotropic jump regularizer d in (1) is induced by a concave function $\kappa : \mathbb{R}_{\geq 0} \to \mathbb{R}$:

$$d(x, u^-, u^+, \nu_u) = \kappa(\left|u^- - u^+\right|) \left\|\nu_u\right\|_2 \tag{5}$$

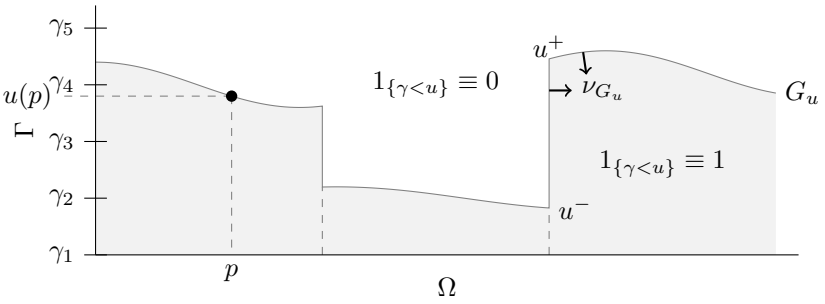


Fig. 2. The central idea behind the convex relaxation of (1) is to reformulate the problem in terms of the complete graph $G_u \subset \Omega \times \Gamma$ of $u : \Omega \to \Gamma$. Since the dimensionality of this reformulation is higher than the dimensionality of the original problem, this is often referred to as *lifting* the problem.

– The range $\Gamma = [\gamma_1, \gamma_\ell] \subset \mathbb{R}$ is a compact interval divided into $k = \ell - 1$ invervals at the boundaries $\gamma_i, i \in \{1, \dots, \ell\}$. Although most formulations work for arbitrary label intervals, we assume equidistant labels for simplicity of notation and denote the label distance by h.

2.2 Functional Lifting

The basic idea for the convex relaxation used in [16] makes use of the fact that for binary segmentation problems, the total variation (2) penalizes the length of the boundary. This was first used in [1,3,6] to derive convex relaxation of nonconvex optimization problems and applied to imaging problems by Pock et al. [18]. An illustration of this idea of reformulating the energy as the flux through the complete graph G_u of the objective u for the 1D case is shown in Fig. 2.

The reformulation in terms of the characteristic function $1_{\{\gamma<u\}}$ of the subgraph $\{\gamma < u\} := \{(x, \gamma) \in \Omega \times \Gamma : \gamma < u(x)\}$ is given by

$$E(u) = F(1_{\{\gamma<u\}}) = \sup_{\varphi \in \mathcal{K}} \int_{G_u} \langle \nu_{G_u}, \varphi \rangle \, \mathrm{d}\mathcal{H}^d, \tag{6}$$

where

$$\mathcal{K} = \Big\{ (\varphi_x, \varphi_t) \in \mathcal{C}_c^1(\Omega \times \mathbb{R}; \mathbb{R}^d \times \mathbb{R}) \mid \forall x \in \Omega : \forall t, t' \in \mathbb{R} :$$

$$\varphi_t(x,t) + \rho(x,t) \geq \eta^*(x, \varphi_x(x,t)), \tag{7}$$

$$\left\| \int_t^{t'} \varphi_x(x,\tau) \, \mathrm{d}\tau \right\|_2 \leq \kappa(|t - t'|) \Big\}, \tag{8}$$

and η^* denotes the *convex conjugate* of the regularizer η. The normal ν_{G_u} in (6) is given by the distributional derivative of the characteristic function of the

subgraph $D1_{\{\gamma<u\}}$, allowing for the reformulation

$$F(1_{\{\gamma<u\}}) = \sup_{\varphi\in\mathcal{K}} \int_{\Omega\times\Gamma} \langle D1_{\{\gamma<u\}}, \varphi\rangle \,\mathrm{d}(x,\gamma), \tag{9}$$

which is then relaxed to

$$\inf_{v\in\mathcal{C}} \sup_{\varphi\in\mathcal{K}} \int_{\Omega\times\mathbb{R}} \langle Dv, \varphi\rangle \,\mathrm{d}(x,\gamma), \tag{10}$$

where

$$\begin{aligned}\mathcal{C} = \{v \in \mathcal{BV}_{\mathrm{loc}}(\Omega\times\mathbb{R};[0,1]) \mid \forall t \le \gamma_1 : v(x,t) = 1,\\ \forall t > \gamma_\ell : v(x,t) = 0, v(x,\cdot) \text{ non-increasing}\}.\end{aligned} \tag{11}$$

The key contribution of [16] is an elegant discretization of the variables v and φ, allowing for a sublabel-accurate approximation of the original energy. Their representation of the primal variable v uses coefficients for each label interval $\Gamma_i = [\gamma_i, \gamma_{i+1}]$, denoted by $\hat{v}(x,i) \in [0,1]$ to allow for continuous values in between the labels even after discretization. Writing the lifted primal variable coefficients as a vector allows for computation of the final result by just summing over the entries of that vector.

For an illustration see the point p in Fig. 2, where the discretization of the lifted variable yields:

$$\hat{v}(p,\cdot) = \sum_{i=1}^{k} e_i \hat{v}(p,i) = h\cdot[1,1,0.8,0]^T. \tag{12}$$

An intuitive derivation for this representation is also given in [15], where the vector in (12) is interpreted as a linear interpolation between labels γ_3 and γ_4, which are represented by $[1,1,0,0]^T$ and $[1,1,1,0]^T$, respectively.

Using a piecewise linear approximation for the dual variable φ_t, the authors of [16] arrive at a sublabel-accurate approximation of the original energy, enabling an implementation of the constraints (7) and (8) individually on each label interval Γ_i as orthogonal projections onto the epigraphs of η^* and ρ_i^*. Here $\rho_i = \rho + \delta_{\Gamma_i}$, which shows that this implementation computes the convex envelope of the original dataterm on each interval Γ_i. In particular, they show that for regularizers η that are support functionals of convex sets, yielding η^* as an indicator function only attaining the values 0 and ∞, the constraint (7) allows for a separation of dataterm and regularizer as implemented for the total variation in [15].

2.3 The Total Generalized Variation

The *total generalized variation* (TGV) of order k was defined in [5] and has subsequently successfully been used for example to reconstruct smooth 3D surfaces [19,8]. Generalizing from the definition (2) of the total variation, it is defined

as

$$\mathrm{TGV}_\alpha^k(u) := \sup\left\{ \int_\Omega u \mathrm{Div}^k \psi \,\mathrm{d}x \,\middle|\, \psi \in \mathcal{C}_c^k(\Omega; \mathrm{Sym}^k(\mathbb{R}^d)), \left\|\mathrm{Div}^l \psi\right\|_{2,\infty} \leq \alpha_l, l = 0, \ldots, k-1 \right\}, \tag{13}$$

with weights $\alpha = [\alpha_0, \ldots, \alpha_{k-1}]$. The space $\mathrm{Sym}^k(\mathbb{R}^d)$ denotes the symmetric k-tensors on $\mathbb{R}^d$, e.g., the space $S^{d\times d}$ of symmetric $d \times d$ matrices for $k = 2$.

We will focus on the case $k = 2$, where one can get an intuition of how TGV penalizes variations by deriving the "primal" formulation [5]

$$\mathrm{TGV}_\alpha^2(u) = \inf_{w \in \mathcal{C}^1(\Omega;\mathbb{R}^d)} \alpha_1 \left\|\nabla u - w\right\|_{2,1} + \alpha_0 \left\|\mathcal{E}(w)\right\|_{2,1}. \tag{14}$$

From (14), the TGV_α^2 penalty can be interpreted as an optimal balancing of first and second order derivative norms. Ranftl et al. [20] use this formulation to develop an approximation of the total generalized variation for nonconvex dataterms using the lifting approach [18]. They specialize the first integral in (1) to obtain

$$\min_{u,w} \overbrace{\alpha \int_\Omega \left\|Dw\right\|_2 \,\mathrm{d}x + \underbrace{\int_\Omega \left\|Du - w\right\|_2 \,\mathrm{d}x}_{E_2(u|w)} + \lambda \int_\Omega \rho(x,u) \,\mathrm{d}x}^{E_1(w|u)}, \tag{15}$$

which separates the problem into a convex subproblem E_1, optimized over w for a fixed u using standard techniques, and a nonconvex subproblem E_2, where w is assumed to be fixed and the lifted optimization [18] is applied to solve for u. In the implementation of [20], u is allowed to deviate up to half the label distance in each direction to allow for smooth surfaces.

3 Lifting the Total Generalized Variation

In our approach to lifting the total generalized variation, we make use of the fact that the framework of [16] allows for a *label-wise* optimization of the problem. Since at discontinuities, TGV approximates the total variation (cf. [5]), we set $\kappa(a) = a$ in (1) and focus on the formulation for η as TGV_α^2 in the following. The main idea of our approach is motivated by the fact that the only difference between the definitions of TV (2) and TGV (13) is a more constrained set of dual variables in TGV. Based on this observation, we set η^* in the constraint set (7) to the indicator function

$$\eta^*(\varphi_x) = \delta_{\{-\mathrm{Div}_x \psi \,|\, \|\psi\| \leq \alpha_0, \|\mathrm{Div}_x \psi\| \leq \alpha_1\}}(\varphi_x). \tag{16}$$

This way, we perform TGV regularization on each label interval individually, just as it has been done for TV in [15]. The constraints can be implemented

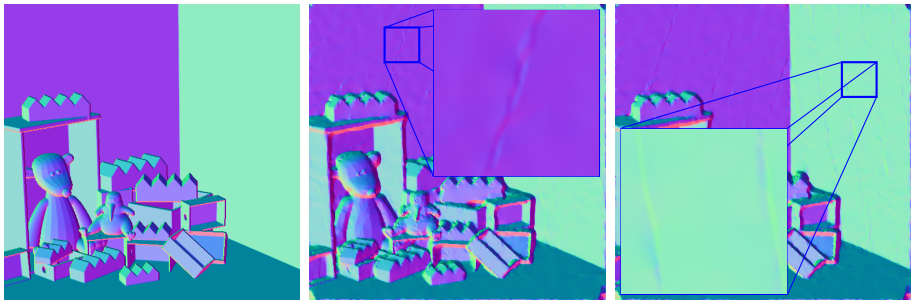


Fig. 3. A naive application of label-wise TGV leads to artifacts in the normal field at disparity label transitions. **Left:** Ground truth normal map, **Center:** One can see label artifacts ("kinks" in the wall on the left) for $\ell = 4$ labels, **Right:** With $\ell = 8$ labels, there are more artifacts at different locations for the $k = 7$ label boundaries.

in the publicly available framework `prost`[1] presented in [15] by using Lagrange multipliers w to enforce equality of φ_x to the negative divergence of some vector field ψ. One can see this by rewriting

$$\begin{aligned} \mathrm{TGV}_{\alpha}^{2} &= \sup_{\|\varphi_x\| \leq \alpha_1, \|\psi\| \leq \alpha_0} \inf_{w} \{\langle \nabla u - w, \varphi_x \rangle + \langle \mathcal{E} w, \psi \rangle\} \\ &= \sup_{\|\varphi_x\| \leq \alpha_1, \|\psi\| \leq \alpha_0} \inf_{w} \{\langle \nabla u, \varphi_x \rangle + \langle w, -\varphi_x - \mathrm{Div} \psi \rangle\}. \end{aligned} \tag{17}$$

The lifted implementation uses the same piecewise constant discretization of the spatial variables as explained in [16], while for the dual variable φ_t (which has not changed in our formulation) a piecewise linear discretization is chosen to allow for piecewise convex sublabel-accurate approximation of the original energy.

However, if we just penalize label-wise with TGV, the solution exhibits artifacts which indicate that our formulation is not yet entirely correct. This can be observed in Fig. 3, where we display the resulting normal maps from this naive implementation for light field depth estimation task (cf. section 4 for more details).

One can obtain an intuition for why these artifacts appear when observing a 1D closeup of the artifacts as shown in Fig. 4 (Center). One can see that especially the lower label is oversmoothed by the TGV prior towards the label boundary (constant 1.0) on the left. This leads to an irregularity in the final result retrieved as the sum over the labels. The desired result is shown in Fig. 4 (Left) and illustrates that the plane with constant slope would be best approximated by both labels allowing a sharp transition to the label boundary at a single pixel.

The theoretical reason behind this failure is that when we just perform label-wise regularization, we assume every change of label is a jump discontinuity, for which we do not penalize the second order derivative. Obviously, as the above example shows, this assumption does not always hold true. Thus, we need

[1] https://github.com/tum-vision/prost

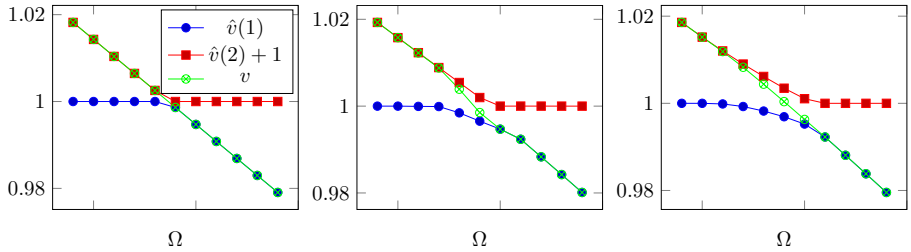


Fig. 4. Left: Synthetic 1D example for two label indicator functions (red and blue), where the exact result is a line with constant slope (green). **Center:** Label-wise TGV tends to oversmooth the coefficients $\hat{v}$ according to the second order component at the transition to constant maximal or minimal value of the respective label, leading to artifacts when computing the result v as the sum of the coefficients. **Right:** Our additional prior penalizing differences in the second order gradients of neighboring labels yields a consistent result.

an additional penalty for the second order gradient in the jump penalization. However, the basic framework developed in [1,3,6] used in [16] does not allow for this to be implemented, since the jump penalty in (1) may only depend on the height of the jump and its direction.

Unfortunately, current attempts at solving this problem in a principled theoretical way, by deriving a formulation based on [1,6] for a TGV-based prior, have not led to a satisfying implementation so far. Nevertheless, we managed to implement an empirically accurate solution inspired by the observations in Fig. 4 (Left). We observe that in most points, both label coefficients have a constant slope (either zero, or the slope of the resulting plane), yielding zero second order derivatives. Nonzero second order derivatives only appear at a single point for both coefficients, where $\hat{v}(1)$ changes from the slope of the result to constant 1.0 and $\hat{v}(2)$ changes from constant 0.0 to the slope of the result. Thus, we need to penalize different second order derivatives across neighboring labels. Since enforcing this relation using a hard constraint would make the lifting obsolete as the dual variables would be the same across all labels, we enforce a soft constraint $\|\hat{\psi}(x,i+1)-\hat{\psi}(x,i)\| \leq \alpha_0$, effectively penalizing $\alpha_0 \|\mathcal{E}\hat{w}(x,i+1)-\mathcal{E}\hat{w}(x,i)\|$. Thus, we impose the same penalty on these differences as on the second order part itself. Fig. 4 (Right) shows how this approach successfully removes the artifacts in the 1D closeup. Due to the additional penalty, both labels receive a similar amount of second order smoothing, which cancel each other out, yielding a consistent surface structure after retrieving the final result. Fig. 5 shows the results for the normal maps, which are now almost free of discretization artifacts.

4 Experiments

4.1 ROF Denoising

In order to verify the quality of our lifted approximation to the total generalized variation experimentally, we first apply our formulation to the convex ROF

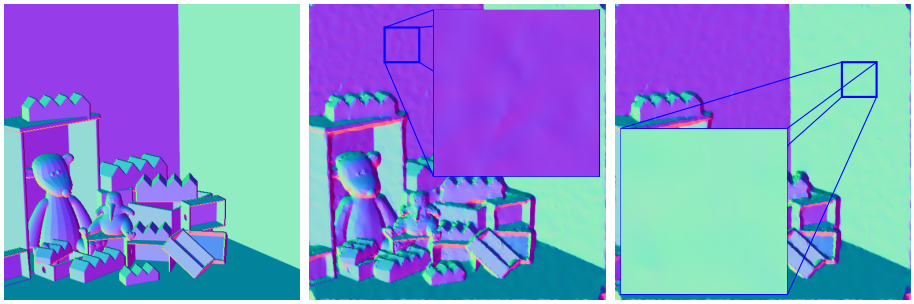


Fig. 5. Our proposed method correctly penalizes different second order derivatives across different labels and thus can remove the label wise artifacts from the naive approach in figure 3. **Left:** Ground truth normal map, **Center and Right:** Results of the proposed lifted implementation with $\ell = 4$ and $\ell = 8$ labels, respectively.

denoising problem [21] where the dataterm is designed for Gaussian noise,

$$\rho(x,t) = (t - f(x))^2 \tag{18}$$

for a grayscale input image $f : \Omega \to \mathbb{R}$. For this convex energy, the convex envelope of the energy used for direct optimization as well as for the case $\ell = 2$ is the same as for lifted versions with higher numbers of labels. In Fig. 6, we compare our method to quadratic and Huber regularization as enabled by [16] (see [18] for derivations of the epigraphical projections required), and to our implementation of the baseline TGV^2 lifting approach [20]. The energies for Huber and quadratic regularization are computed in a straightforward way from their definitions using the resulting u. For the TGV energies, we first solve subproblem $E_2(w|u)$ in (15) for w as an instance of ROF given u, and then compute the final energy using the resulting pair u and w. One can see that even for small numbers of labels, our approach manages to achieve similar energies as the direct optimization, the differences to the direct energy only being slightly larger than for Huber and quadratic regularization demonstrated in [16]. The baseline approach [20], however, despite allowing values in between the labels, still has a strong label bias. Since the parameters in (15) are different from the original TGV parameters in (14), we compute them as $\lambda = \frac{1}{\alpha_1}$ and $\alpha = \frac{\alpha_0}{\alpha_1}$ to have a fair comparison.

4.2 Robust Truncated Quadratic Denoising

In practice, the simple ROF denoising term [21] is only of limited use because the model does not account for possible outliers which are often present in a captured image (i.e., salt-and-pepper noise). For this kind of noise, a truncated version of the dataterm can achieve better results, penalizing outliers with a constant value,

$$\rho(x,t) = \frac{\alpha}{2} \min\{(t - f(x))^2, \nu\}. \tag{19}$$

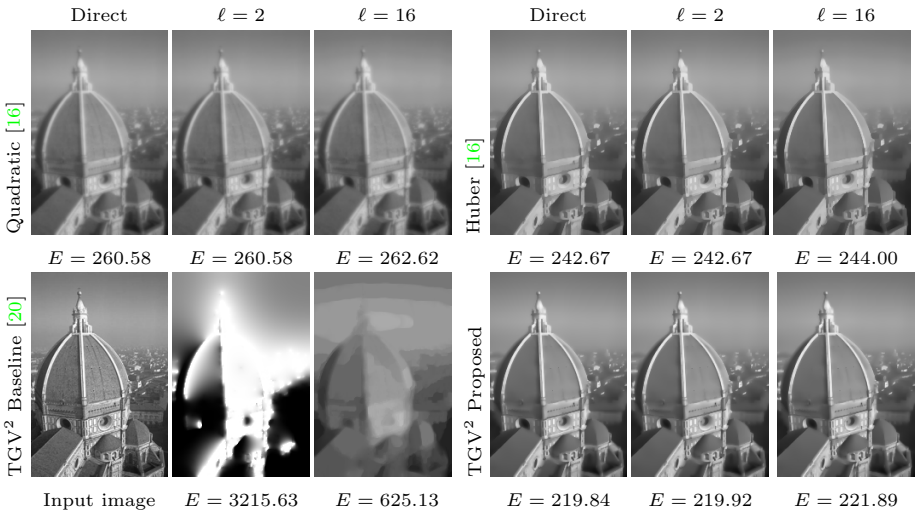


Fig. 6. Results of different regularizers for convex ROF-denoising with different numbers of labels. The proposed method recovers the original energy with only slightly larger margins than for Huber and Quadratic regularization which have been demonstrated in [16], while the baseline method [20] shows strong label artifacts. Note: the direct TGV^2 optimization result displayed for TGV^2 proposed is the same as for the baseline approach [20], where we show the input image instead.

Fig. 7 shows that our TGV-based lifted regularization achieves good results already for $\ell = 10$ labels, similarly as quadratic and Huber regularization with the approach in [16]. In contrast, the baseline TGV approach [20] exhibits a strong label bias for $\ell = 10$ and severe oversmoothing for $\ell = 20$. The energies displayed in Fig. 7 show a corresponding substantial improvement of our approach compared to the baseline TGV relaxation [20].

4.3 Light Field Disparity Estimation

We evaluate our formulation on the light field disparity estimation task with a special focus on the quality of surface normals. Since the behaviour of our algorithm across different numbers of labels was already evaluated in the previous experiments, we focus on a comparison of TV, Huber, and the proposed TGV^2 regularization with $\ell = 8$ labels and compare it to the approach by Strecke et al. [22], which explicitly smoothes the normal map derived from sublabel-accurate TV-based optimization of disparity using [15] in a post-processing step. This method is still among the top-ranked algorithms for surface normal quality on the benchmark [11]. To have a fair comparison of how the regularization influences the result, we use the same dataterm used in [22] and refer to their work for details on how it is constructed and implemented. Figs. 8 and 9 demonstrate that we manage to achieve similar results as [22], denoted as `OFSY 330DNR` in the figures, without the need of computing the surface normals explicitly for

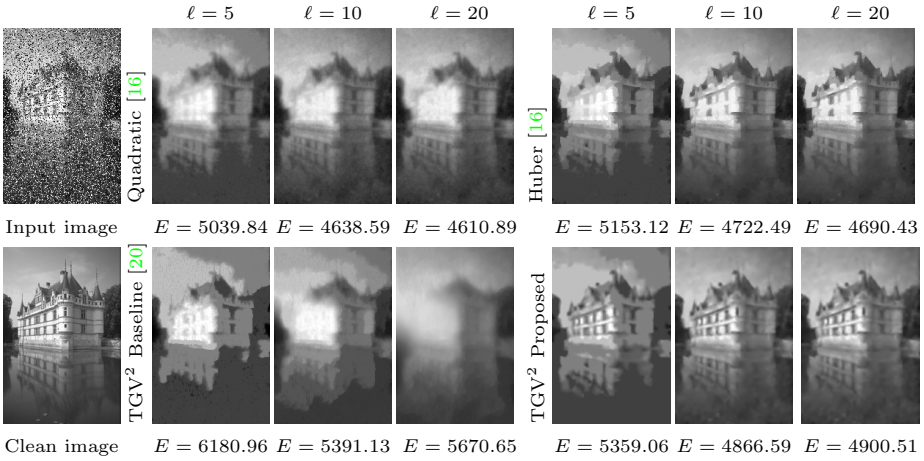


Fig. 7. Results of different regularizers for nonconvex truncated ROF denoising with different numbers of labels. The quadratic regularizer oversmooths details and Huber exhibits similar piecewise constant artifacts at the total variation. The proposed method retains details and results in smooth transitions, while the baseline method [20] does exhibit label artifacts and severe oversmoothing for the same parameter setting.

optimization or running two optimization passes. One can further see that our approach yields significantly better results in this respect than TV or Huber regularization.

Besides the *Median Angular Error (MAE)* for planar and non-planar surfaces in Fig. 8, Fig. 9 shows that the quality of our obtained disparity maps, measured by the percentage of pixels deviating more than 0.07 from the ground truth. This so-called *BadPix(0.07)* measure is similar to the appoach [22].

5 Conclusions

Our approach extends the recent functional lifting approaches [15,16] to obtain experimentally accurate solutions using a TGV-based prior. We show in our experiments that our formulation significantly improves over the previous approach to lifting the total generalized variation [20], which is based on an earlier functional lifting approach [18] and decouples the regularizer from the data term. When applied to the task of light field disparity estimation, our TGV^2-based approach outperforms other regularizers implemented in [15,16] in terms of the quality of surface normals. In a single pass optimization, we manage to achieve similar results as [22], who explicitly smooth the normal map in a post-processing step after sublabel-accurate optimization using [15].

Unfortunately, our method is only empirical so far and based on heuristic observations - while results are convincing, it currently lacks theoretical verification of correctness. However, we hope that the insights presented in this paper can inspire future studies on a more principled approach towards deriving our constraint sets, as our experiments indicate that they seem to work in practice.

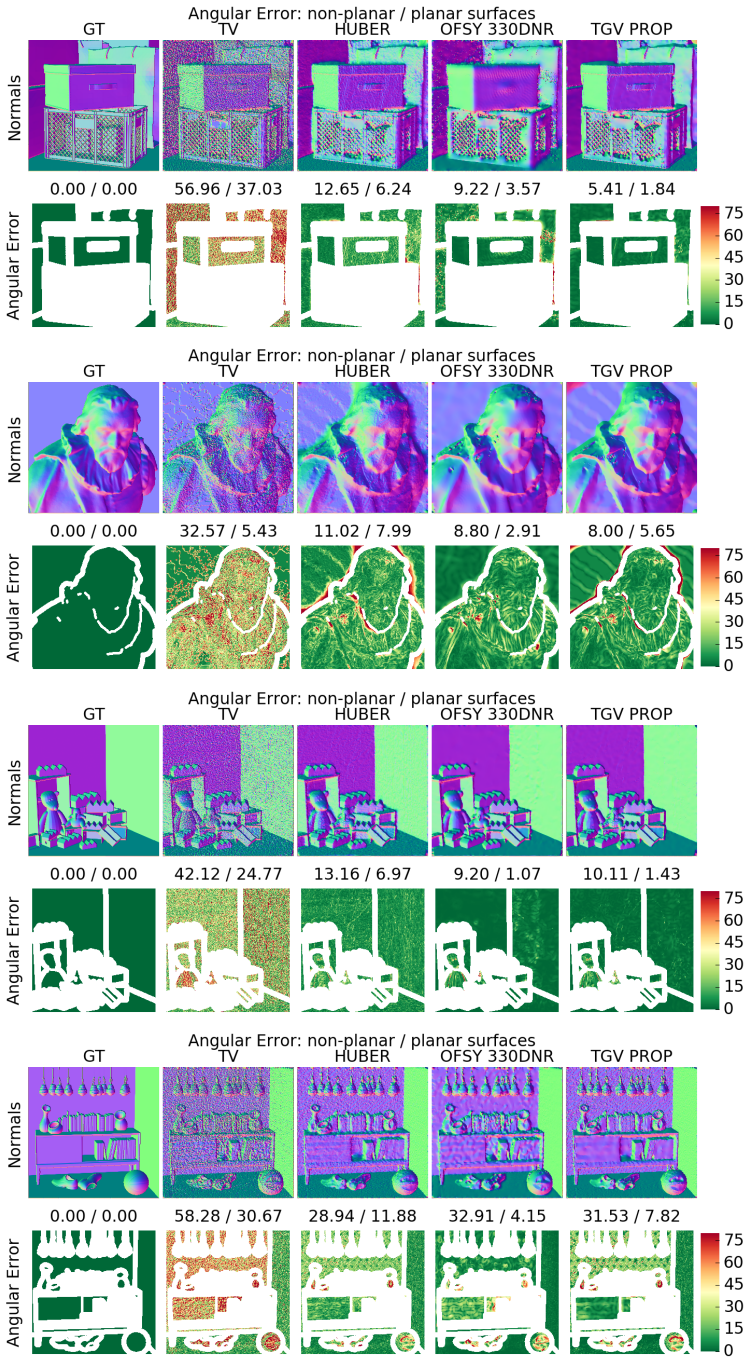


Fig. 8. Normal maps and *Median Angular Errors* obtained on the benchmark training dataset [11] for different regularizers and the method in [22].

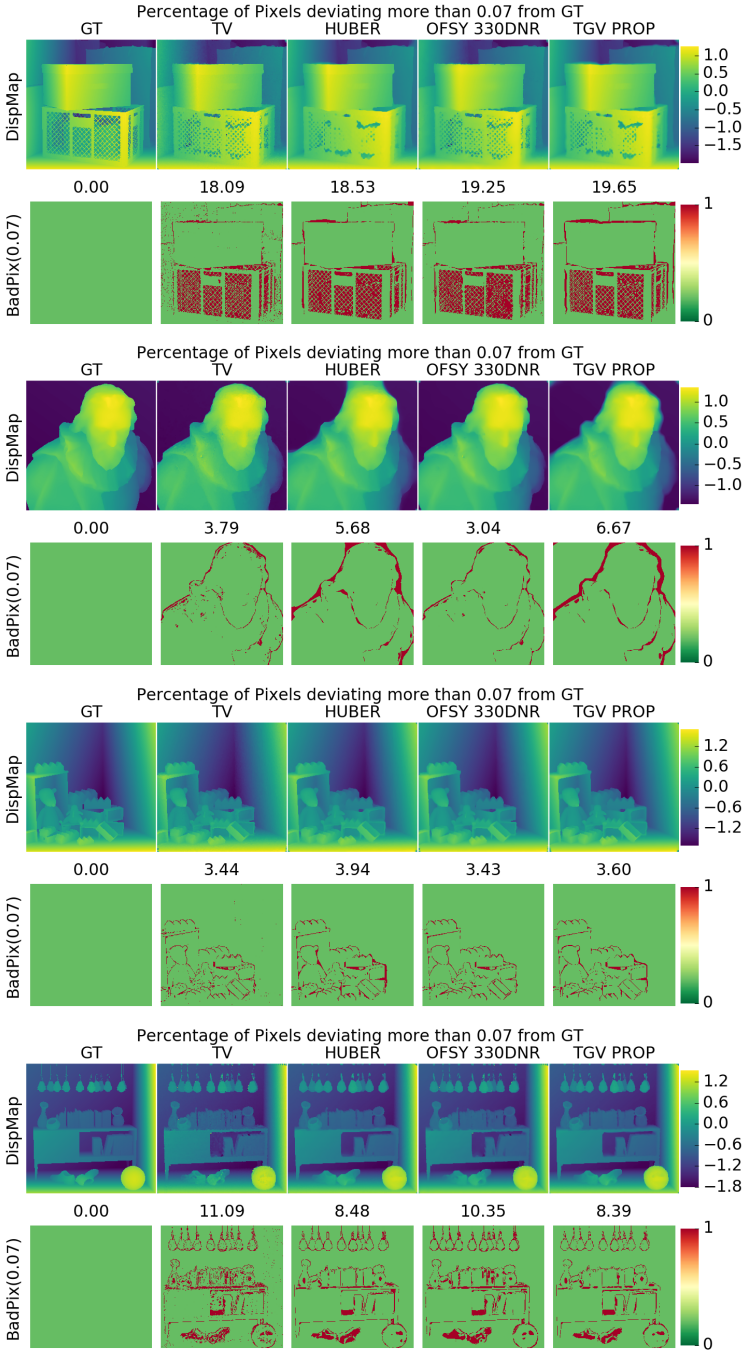


Fig. 9. Disparity maps and *BadPix0.07* errors obtained on the benchmark training dataset [11] for different regularizers and the method [22].

Acknowledgements

This work was supported by the ERC Starting Grant "Light Field Imaging and Analysis" (LIA 336978, FP7-2014) and the SFB Transregio 161 "Quantitative Methods for Visual Computing".

References

1. Alberti, G., Bouchitté, G., Dal Maso, G.: The calibration method for the Mumford-Shah functional and free-discontinuity problems. Calc. Var. Partial Differential Equations **16**(3), 299–333 (2003) 1, 4, 8
2. Blake, A., Zisserman, A.: Visual Reconstruction. MIT Press (1987) 3
3. Bouchitté, G.: Recent Convexity Arguments in the Calculus of Variations (1998), (lecture notes from the 3rd Int. Summer School on the Calculus of Variations, Pisa) 4, 8
4. Boykov, Y., Veksler, O., Zabih, R.: Markov Random Fields with Efficient Approximations. In: Proc. International Conference on Computer Vision and Pattern Recognition. pp. 648–655. Santa Barbara, California (1998) 1
5. Bredies, K., Kunisch, K., Pock, T.: Total generalized variation. SIAM Journal on Imaging Sciences **3**(3), 492–526 (2010) 2, 5, 6
6. Chambolle, A.: Convex representation for lower semicontinuous envelopes of functionals in L^1. J. Convex Anal. **8**(1), 149–170 (2001) 4, 8
7. Cremers, D., Strekalovskiy, E.: Total cyclic variation and generalizations. Journal of Mathematical Imaging and Vision **47**(3), 258–277 (November 2012) 1
8. Ferstl, D., Reinbacher, C., Ranftl, R., Rüther, M., Bischof, H.: Image guided depth upsampling using anisotropic total generalized variation. In: Proc. International Conference on Computer Vision. pp. 993–1000. IEEE (2013) 5
9. Geman, S., Geman, D.: Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence **6**, 721–741 (1984) 1
10. Goldluecke, B., Strekalovskiy, E., Cremers, D.: Tight convex relaxations for vector-valued labeling. SIAM Journal on Imaging Sciences **6**(3), 1626–1664 (2013) 1
11. Honauer, K., Johannsen, O., Kondermann, D., Goldluecke, B.: A dataset and evaluation methodology for depth estimation on 4d light fields. In: Asian Conf. on Computer Vision (2016) 10, 12, 13
12. Ishikawa, H.: Exact optimization for Markov random fields with convex prio rs. IEEE Trans. Pattern Analysis and Machine Intelligence **25**(10), 1333–1336 (2003) 1, 2
13. Kolmogorov, V., Zabih, R.: What Energy Functions can be minimized via Graph Cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence **26**(2), 147–159 (2004) 1
14. Lellmann, J., Becker, F., Schnörr, C.: Convex Optimization for Multi-Class Image Labeling with a Novel Family of Total Variation Based Regularizers. In: IEEE International Conference on Computer Vision (ICCV) (2009) 1
15. Moellenhoff, T., Laude, E., Moeller, M., Lellmann, J., Cremers, D.: Sublabel-accurate relaxation of nonconvex energies. In: Proc. International Conference on Computer Vision and Pattern Recognition (2016) 2, 5, 6, 7, 10, 11
16. Möllenhoff, T., Cremers, D.: Sublabel-accurate discretization of nonconvex free-discontinuity problems. In: International Conference on Computer Vision (ICCV). Venice, Italy (October 2017) 2, 3, 4, 5, 6, 7, 8, 9, 10, 11

17. Mumford, D., Shah, J.: Optimal approximations by piecewise smooth functions and associated variational problems. Communications on Pure and Applied Mathematics **42**, 577–685 (1989) 3
18. Pock, T., Cremers, D., Bischof, H., Chambolle, A.: Global Solutions of Variational Models with Convex Regularization. SIAM Journal on Imaging Sciences (2010) 1, 2, 4, 6, 9, 11
19. Ranftl, R., Gehrig, S., Pock, T., Bischof, H.: Pushing the limits of stereo using variational stereo estimation. In: IEEE Intelligent Vehicles Symposium (2012) 5
20. Ranftl, R., Pock, T., Bischof, H.: Minimizing TGV-based variational models with non-convex data terms. In: International Conference on Scale Space and Variational Methods in Computer Vision. pp. 282–293. Springer (2013) 2, 6, 9, 10, 11
21. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica D **60**, 259–268 (1992) 9
22. Strecke, M., Alperovich, A., Goldluecke, B.: Accurate depth and normal maps from occlusion-aware focal stack symmetry. In: Proc. International Conference on Computer Vision and Pattern Recognition (2017) 2, 10, 11, 12, 13
23. Zach, C., Hne, C., Pollefeys, M.: What is optimized in convex relaxations for multilabel problems: Connecting discrete and continuously inspired map inference. IEEE Transactions on Pattern Analysis and Machine Intelligence **36**(1), 157–170 (Jan 2014) 1