

Total Variation for Cyclic Structures: Convex Relaxation and Efficient Minimization

Evgeny Strekalovski and Daniel Cremers
TU Munich, Germany

Abstract

We introduce a novel type of total variation regularizer, $TV_{\mathcal{S}^1}$, for cyclic structures such as angles or hue values. The method handles the periodicity of values in a simple and consistent way and is invariant to value shifts. The regularizer is integrated in a recent functional lifting framework which allows for arbitrary nonconvex data terms. Results are superior and more natural than with the simple total variation without special care about wrapping interval end points. In addition we propose an equivalent formulation which can be minimized with the same time and memory efficiency as the standard total variation.

1. Total Variation and Cyclic Structures

Total variation has been widely recognized as an effective and efficient means to regularize variational problems. It has been applied to a great variety of problems, including denoising [14, 17], superresolution [8], inpainting [3], multi-view reconstruction [7] and optical flow [16]. These examples all assume that the function u takes its values in a space of linear structure, e.g. \mathbb{R}^n for some $n \geq 1$.

On the contrary, values from the unit circle \mathcal{S}^1 , or equivalently the space of planar orientations, such as angles or hue values are periodic in nature. An important area for denoising of cyclical data is SAR phase unwrapping [5]. Representing \mathcal{S}^1 as just $[0, 1]$ and applying the usual TV norm leads to artifacts due to neglecting the value wrapping at the boundaries. Jumps from 0 to 1 will be penalized with a nonzero contribution to the energy whereas there is actually no jump at all with the \mathcal{S}^1 structure. Also, TV is dependent on the specific representation of \mathcal{S}^1 as $[0, 1]$, i.e. a cyclic shift of input data will not result in the corresponding cyclic shift of the output.

To the best of our knowledge, cyclic structures have not yet been accounted for in the regularization previously. Moreover, it turns out that this is by no means straightforward. In this paper we introduce a novel kind of total variation, $TV_{\mathcal{S}^1}$, for functions $u : \Omega \rightarrow \mathcal{S}^1$ with cyclic or pe-

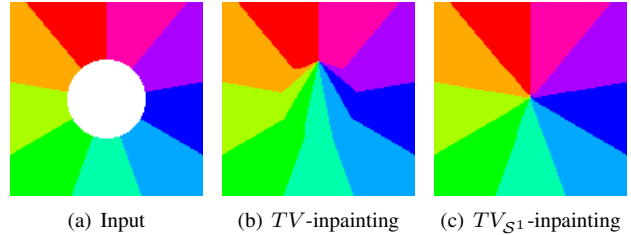


Figure 1. Inpainting of periodic hue values. While total variation regularization (*center*) does not handle wrap-arounds correctly (shrinking the interface with the highest value jump, here from magenta to red) the proposed $TV_{\mathcal{S}^1}$ formulation (*right*) is designed to provide an optimal solution for cyclic structures at no additional cost. This figure is best viewed in color.

riodic values. The regularizer provides an elegant solution to handle the wrap-around of values and has some desired mathematical properties such as invariance to cyclic shifts and isometries.

2. Definition of $TV_{\mathcal{S}^1}$

Let $\Omega \subset \mathbb{R}^n$ be an open subset, $n \geq 1$. Let

$$\mathcal{S}^1 = \{x \in \mathbb{R}^2 \mid |x| = 1\}$$

be the unit circle. Since there is a one-to-one correspondence between \mathcal{S}^1 and the unit interval $\Gamma := [0, 1)$, any function $\hat{u} : \Omega \rightarrow \mathcal{S}^1$ can be conveniently represented as a function $u : \Omega \rightarrow \Gamma$. We will use this representation throughout the paper, as it proves to be both convenient and useful for the theoretical treatment of $TV_{\mathcal{S}^1}$. As one easily checks, the distance function on \mathcal{S}^1 transfers to the distance function $d_{\mathcal{S}^1}$ on Γ given by $d_{\mathcal{S}^1}(a, b) = \min(|a - b|, 1 - |a - b|)$. For fixed a this gives a “saw-tooth” shaped curve in b (Fig. 2).

For special functions of bounded variation, $u \in SBV(\Omega)$, the distributional derivative Du of u can be decomposed as

$$Du = \nabla u \, dx + (u^+ - u^-) \nu_u \, d\mathcal{H}^{n-1} \llcorner S_u$$

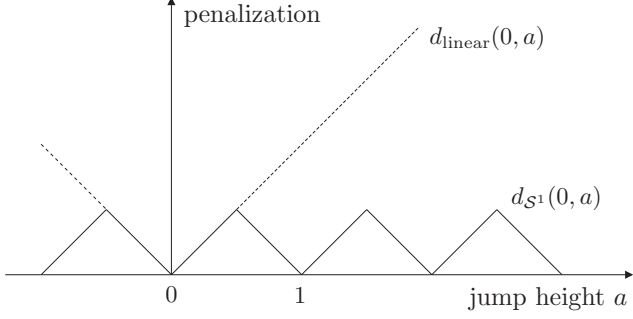


Figure 2. In contrast to the linear distance (dashed), the cyclic distance $d_{\mathcal{S}^1}$ (solid) accounts for the periodicity of \mathcal{S}^1 .

into a differentiable part and a jump part. Here, S_u is the $(n-1)$ -dimensional jump set of u where the values jump from u^- to u^+ , ν_u is the normal to S_u from the u^- to the u^+ side, and ∇u is the gradient of u outside the jump set S_u (Fig. 3). $\mathcal{H}^{n-1} \llcorner S_u$ denotes the $(n-1)$ -dimensional Hausdorff measure restricted to the set S_u . We refer to [2] for a comprehensive introduction to functions of bounded variation.

Definition 2.1. For $u : \Omega \rightarrow \Gamma$ we define the *total cyclic variation* as

$$TV_{\mathcal{S}^1}(u) = \int_{\Omega \setminus S_u} |\nabla u| + \int_{S_u} d_{\mathcal{S}^1}(u^-, u^+). \quad (2.1)$$

A straightforward generalization is the *weighted* $TV_{\mathcal{S}^1}$ where the contribution to $TV_{\mathcal{S}^1}$ at each point $x \in \Omega$ is weighted by a function $g : \Omega \rightarrow \mathbb{R}_{\geq 0}$:

$$TV_{\mathcal{S}^1}^g(u) = \int_{\Omega \setminus S_u} g |\nabla u| + \int_{S_u} g d_{\mathcal{S}^1}(u^-, u^+).$$

Writing $|u^- - u^+|$ instead of $d_{\mathcal{S}^1}(u^-, u^+)$ would give just the usual TV seminorm. Note that $TV_{\mathcal{S}^1}(u)$ is not convex in u because of the nonconvex distance $d_{\mathcal{S}^1}$. In the next section, we will introduce a convex representation for efficiently minimizing functionals involving $TV_{\mathcal{S}^1}(u)$.

2.1. Convex Representation of $TV_{\mathcal{S}^1}$

In a series of papers [1, 12, 9, 10, 11] it was shown that for functionals of the form

$$F(u) = \int_{\Omega \setminus S_u} h(x, u(x), \nabla u(x)) dx + \int_{S_u} d(s, u^-(s), u^+(s)) ds \quad (2.2)$$

a tight convex relaxation can be found using the method of cartesian currents [6] and that this relaxation can be employed to minimize these functionals. The idea is to restate

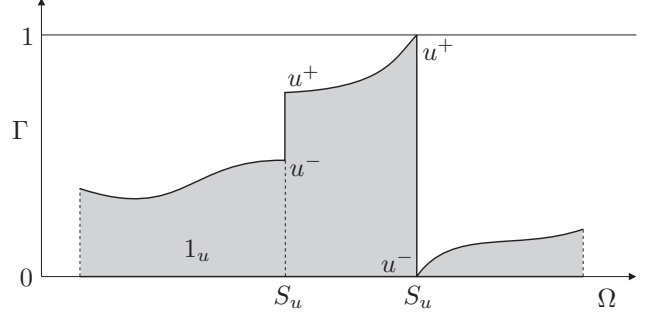


Figure 3. Graph of a function $u \in SBV(\Omega)$. The jump set S_u consists here of two points, the second jump being just a wrap-around and not an actual jump in \mathcal{S}^1 sense. The function 1_u is defined as 1 in the shaded area under the graph and 0 otherwise.

$F(u)$ in terms of the graph function of u (Fig. 3):

$$1_u : \Omega \times \Gamma \rightarrow \mathbb{R}, \quad 1_u(x, t) = \begin{cases} 1, & \text{if } t < u(x) \\ 0, & \text{else} \end{cases}$$

Although the functional $F(u)$ is in general highly non-convex, it turns out [1, Lemma 3.7] that in case where $h : \Omega \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$ is convex in its third argument (i.e. in ∇u) it can be recast as

$$F(u) = \sup_{\varphi \in K} \int_{\Omega \times \Gamma} \varphi \cdot \nabla 1_u =: \mathcal{F}(1_u) \quad (2.3)$$

with the convex set

$$K = \left\{ \varphi = (\varphi^x, \varphi^t) : \Omega \times \Gamma \mapsto \mathbb{R}^n \times \mathbb{R} \mid \begin{aligned} &\varphi^t(x, t) \geq h^*(x, t, \varphi^x(x, t)), \\ &\left| \int_{t_1}^{t_2} \varphi^x(x, t) dt \right| \leq d(x, t_1, t_2) \end{aligned} \right\}. \quad (2.4)$$

Here, $h^*(x, t, q)$ is the Legendre-Fenchel conjugate of $h(x, t, p)$ with respect to p . For an introduction to convex analysis we refer to [13]. The main advantage of formulation (2.3) is that \mathcal{F} is convex in 1_u . In general, (2.3) holds with \geq , and equality holds for a given u if there is a vector field $\varphi \in K$ satisfying certain properties. Since (2.3) is fundamental for our discussion, a proof of the equality for $F = TV_{\mathcal{S}^1}$ is presented in the appendix.

For $TV_{\mathcal{S}^1}$, we have $d = d_{\mathcal{S}^1}$ and $h(x, u(x), \nabla u(x)) = |\nabla u|$ in (2.2). The Legendre-Fenchel dual of h is given by

$$h^*(x, t, q) = \sup_p p \cdot q - h(x, t, p) = \begin{cases} 0, & \text{if } |q| \leq 1 \\ \infty, & \text{else.} \end{cases}$$

Thus, the first constraint $\varphi^t(x, t) \geq h^*(x, t, \varphi^x(x, t))$ in (2.4) is equivalent to $\varphi^t \geq 0$ and $|\varphi^x| \leq 1$. The supremum over $\varphi^t \geq 0$ in (2.3) can be easily eliminated since

1_u is nonincreasing in the t -direction, i.e. $\partial_t 1_u \leq 0$ and $\varphi^t \partial_t 1_u \leq 0$ with 0 being the maximum possible value. So only the constraint $|\varphi^x| \leq 1$ remains.

Next, the integral constraint in (2.4), using this formulation, must be satisfied for all pairs $0 \leq t_1 < t_2 < 1$. In practice, the space of cyclical values Γ needs to be discretized at N levels $\frac{0}{N}, \dots, \frac{N-1}{N}$ for some $N \geq 1$. In the discretization, this means that the number of constraints in (2.4) grows quadratically with the number of levels. Enforcing them quickly becomes unfeasible with respect to both runtime and memory consumption even for a moderate number of levels. Surprisingly, for $d = d_{S^1}$ there is an equivalent formulation of this constraint which leads to a much more efficient implementation.

2.2. From Quadratic to Linear Complexity

The main contribution of this paper is to show that one can efficiently minimize functionals involving the TV_{S^1} regularizer. This is based on the following equivalent formulation of the integral constraint:

Proposition 2.2 (Constraint equivalence).

$$\left| \int_{t_1}^{t_2} \varphi^x(x, t) dt \right| \leq d_{S^1}(t_1, t_2) \quad \text{for } t_1, t_2 \in \Gamma \quad (2.5)$$

\iff

$$|\varphi^x(x, t)| \leq 1 \quad \text{for } t \in \Gamma, \quad (2.6)$$

$$\int_{\Gamma} \varphi^x(x, t) dt = 0 \quad (2.7)$$

Proof. “ \Rightarrow ”: From (2.5) we have

$$\left| \int_{t_1}^{t_2} \varphi^x(x, t) dt \right| \leq d_{S^1}(t_1, t_2) \leq t_2 - t_1.$$

Dividing both sides by $t_2 - t_1$ for $t_1 < t_2$ and letting $t_1, t_2 \rightarrow t$ for some $t \in \Gamma$ we obtain (2.6). Furthermore, setting $t_1 = 0$ and $t_2 = 1$ yields

$$\left| \int_{\Gamma} \varphi^x(x, t) dt \right| \leq d_{S^1}(0, 1) = 0$$

and therefore (2.7).

“ \Leftarrow ”: With (2.6) we have

$$\left| \int_{t_1}^{t_2} \varphi^x \right| \leq \int_{t_1}^{t_2} |\varphi^x| \leq \int_{t_1}^{t_2} 1 = t_2 - t_1$$

and using (2.7) also

$$\begin{aligned} \left| \int_{t_1}^{t_2} \varphi^x \right| &= \left| \int_0^1 \varphi^x - \int_{t_1}^{t_2} \varphi^x \right| = \left| \int_0^{t_1} \varphi^x + \int_{t_2}^1 \varphi^x \right| \\ &\leq \int_0^{t_1} |\varphi^x| + \int_{t_2}^1 |\varphi^x| \leq \int_0^{t_1} 1 + \int_{t_2}^1 1 \\ &= 1 - (t_2 - t_1). \end{aligned}$$

Combining these two estimates we get

$$\left| \int_{t_1}^{t_2} \varphi^x \right| \leq \min(|t_2 - t_1|, 1 - |t_2 - t_1|) = d_{S^1}(t_1, t_2). \quad \square$$

Effectively, this reduces the memory and time complexity of enforcing the constraint from quadratic to just linear. Thus, regularization of general data terms with TV_{S^1} becomes just as inexpensive as the usual TV regularization [11] where the number of constraints also grows linearly with the number of levels. Based on the proposition above, integrating by parts in (2.3) we obtain a useful dual formulation of TV_{S^1} which does not require to know the jump set S_u explicitly:

Proposition 2.3. For $u : \Omega \rightarrow \Gamma$ it holds

$$TV_{S^1}(u) = \sup_{\varphi^x \in \hat{K}} \int_{\Omega} \int_0^{u(x)} \operatorname{div} \varphi^x(x, t) dt dx \quad (2.8)$$

with

$$\hat{K} = \left\{ \varphi^x \in C_c^1(\Omega \times \Gamma, \mathbb{R}^n) \mid \begin{aligned} &|\varphi^x(x, t)| \leq 1, \int_{\Gamma} \varphi^x(x, s) ds = 0 \end{aligned} \right\}. \quad (2.9)$$

Note that for the formulation (2.8) the vector fields φ^x in (2.9) must be smooth and with compact support with regard to the x variable. One can in fact restrict (2.4) to these vector fields because of their L^1 -density in the set (2.4). This can be easily seen by applying a cut-off operator together with a smoothing convolution, which is a standard procedure in the PDE literature.

3. Properties of TV_{S^1}

We will now investigate some mathematical properties of TV_{S^1} . First of all, TV_{S^1} has the desired property that, whenever u attains two different values $a, b \in \Gamma$, TV_{S^1} penalizes the value change by the d_{S^1} distance of a and b multiplied with the length of the interface:

Proposition 3.1 (Consistency with d_{S^1}). Let $u = a\chi_A + b\chi_{\bar{A}}$. Then

$$TV_{S^1}(u) = d_{S^1}(a, b) \operatorname{Per}(A)$$

where $\operatorname{Per}(A) = TV(\chi_A)$ is the perimeter of A .

Proof. See appendix. \square

Thus, TV_{S^1} as defined above is indeed the total cyclic variation of u . As a corollary we obtain $TV_{S^1}(u) = 0$ for any constant $u : \Omega \rightarrow \Gamma$.

Since the values in S^1 are periodic in nature one expects that $TV_{S^1}(u)$ does not change if the values of u are cyclically shifted everywhere by the same amount. This is indeed the case as stated in the following proposition. This

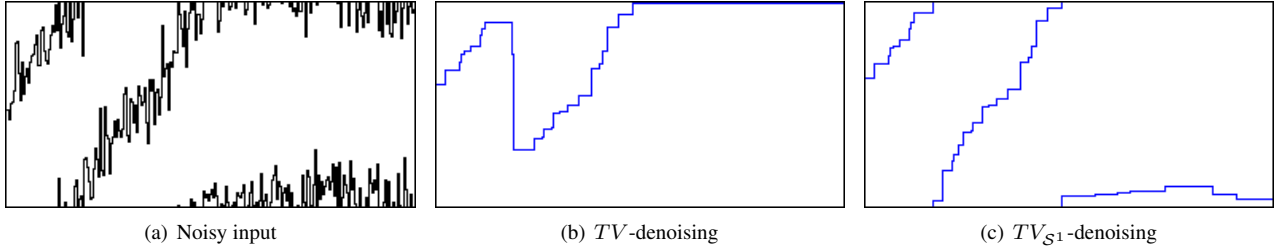


Figure 4. Denoising of angle values. TV_{S^1} (right) reconstructs the monotonous and the constant signal part correctly despite the strong noise (Gaussian 10%). TV (center) falsely penalizes wrap-arounds driving the values towards the middle to reduce jump height. The reconstruction lacks monotonicity and is falsely equal to 1 in the constant part to avoid the jump from 1 to 0.05 in the middle.

means that TV_{S^1} is independent of a concrete representation of S^1 as the unit interval $[0, 1]$, i.e. it does not matter which point of S^1 is taken to be 0. This property is not shared by TV regularization (Fig. 5). For $\alpha \in \mathbb{R}$ we define the shift operator $T_\alpha : \Gamma \rightarrow \Gamma$ by $T_\alpha(t) = t + \alpha \pmod{1}$.

Proposition 3.2 (Cyclic shift invariance). *Let $u : \Omega \rightarrow \Gamma$ and $\alpha \in \mathbb{R}$. Then*

$$TV_{S^1}(T_\alpha \circ u) = TV_{S^1}(u).$$

Proof. See appendix. \square

Finally, TV_{S^1} is invariant with respect to rigid transformations of the domain Ω .

Proposition 3.3 (Rigid transformation invariance). *Let $\Omega' \subset \mathbb{R}^m$ and $A : \Omega' \rightarrow \Omega$ be an isometry. Then for any u ,*

$$TV_{S^1}(u \circ A) = TV_{S^1}(u).$$

Proof. Follows from representation (2.8) by the integral transformation formula and the invariance of the set \tilde{K} in (2.9) with respect to linear isometries. \square

4. Regularization with TV_{S^1}

We consider an arbitrary, possibly nonconvex pointwise data term $\int_\Omega \varrho(x, u(x))$ and apply TV_{S^1} regularization to obtain the overall optimization problem

$$\inf_u \int_\Omega \varrho(x, u(x)) + TV_{S^1}(u).$$

Using the representation (2.3) this can be rewritten as

$$\inf_u \sup_{\varphi \in K} \int_{\Omega \times \Gamma} \varphi \cdot \nabla 1_u \quad (4.1)$$

with

$$K = \left\{ \varphi = (\varphi^x, \varphi^t) : \Omega \times \Gamma \rightarrow \mathbb{R}^n \times \mathbb{R} \mid \varphi^t \geq -\varrho, |\varphi^x| \leq 1, \int_\Gamma \varphi^x = 0 \right\}. \quad (4.2)$$

For the weighted variant $TV_{S^1}^g$ the second constraint $|\varphi^x| \leq 1$ must be replaced by $|\varphi^x| \leq g$. The set K is derived similarly as in section 2.1 using the equivalent constraint formulation of proposition 2.2. Since $\{1_u \mid u \in SBV(\Omega)\}$ is not convex we relax the problem (4.1) to

$$\inf_{v \in C} \sup_{\varphi \in K} \int_{\Omega \times \Gamma} \varphi \cdot \nabla v \quad (4.3)$$

with the convex set

$$C = \left\{ v : \Omega \times \Gamma \rightarrow \mathbb{R} \mid v(x, t) \in [0, 1], v(x, 0) = 1, v(x, 1) = 0 \right\}. \quad (4.4)$$

The condition on φ^t in (4.2) ensures that the supremum in (4.3) is finite only if v is nonincreasing in the t -direction. An optimal solution u of the original problem (4.1) can be recovered from a binary solution v of (4.3) by simple thresholding, $u(x) = \sup\{t \mid v(x, t) > \frac{1}{2}\}$. For nonbinary v this still produces an approximate solution u . Though this u is not necessarily optimal, an energy bound holds:

$$|F(u) - F(u^*)| \leq |\mathcal{F}(1_u) - \mathcal{F}(v)| \quad (4.5)$$

with a true solution u^* of (4.1) and \mathcal{F} in (2.3). This follows from the simple observation $\mathcal{F}(v) \leq \mathcal{F}(1_{u^*}) = F(u^*) \leq F(u) = \mathcal{F}(1_u)$, where the inequalities hold because v is optimal for \mathcal{F} and u^* is optimal for F . In all our experiments the deviation from the optimal solution was less than 5%. This is an overestimation of the true deviation obtained by formula (4.5).

4.1. Efficient Primal-Dual Minimization

We solve (4.3) by the general algorithm of [10]. This is a fast primal-dual algorithm which consists essentially in a gradient descent in v and a gradient ascent in φ , with an orthogonal reprojecton onto the sets C and K . Since K includes the non-local constraint $\int_\Gamma \varphi^x = 0$ there is no simple formula for the projection. Therefore, instead of a

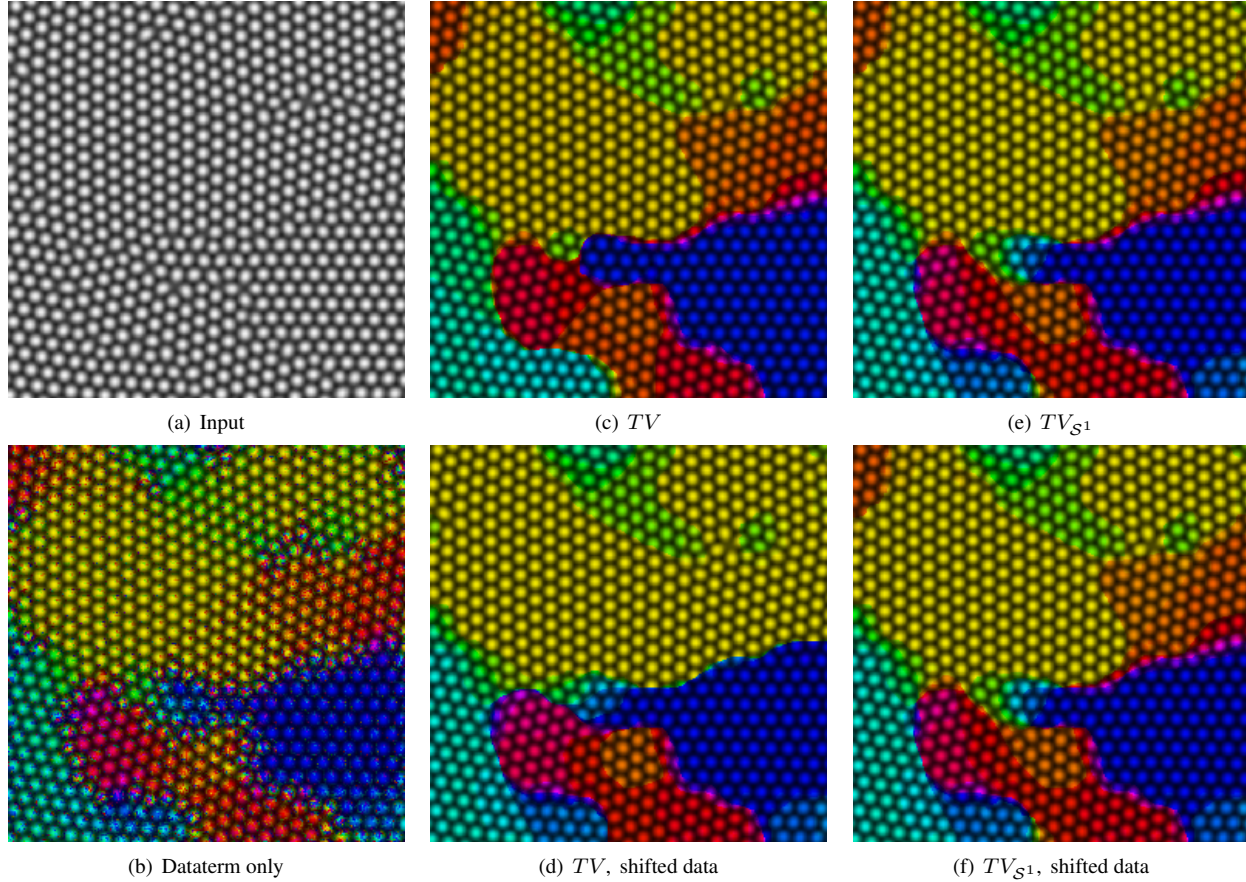


Figure 5. Segmentation of regions with homogeneous atom lattice *orientation*, coloring different orientations by different hues. This example demonstrates the cyclic shift invariance of TV_{S^1} in proposition 3.2. The results in (e) for the original data term $\varrho(x, t)$ and in (f) for the cyclically shifted one, $\varrho(x, t - 0.5 \bmod 1)$, are nearly indistinguishable using TV_{S^1} (coloring in (f) is shifted accordingly for comparison). In contrast, using TV the solution may be different for different zero orientations (c–d). This figure is best viewed in color.

direct projection we implement this integral constraint by Lagrange multipliers. For this, we add the term

$$\inf_{q: \Omega \rightarrow \mathbb{R}^n} \int_{\Omega} q(x) \cdot \left(\int_{\Gamma} \varphi^x(x, t) dt \right) dx \quad (4.6)$$

to the energy (4.3). The overall optimization problem becomes

$$\inf_{v \in C, q} \sup_{\varphi \in K} \int_{\Omega \times \Gamma} \varphi \cdot \nabla v + \int_{\Omega \times \Gamma} q \cdot \varphi^x \quad (4.7)$$

with no constraints on the Lagrange multiplier $q: \Omega \rightarrow \mathbb{R}^n$, C as in (4.4) and K without the integral constraint:

$$K = \left\{ \varphi : \Omega \times \Gamma \rightarrow \mathbb{R}^n \times \mathbb{R} \mid \varphi^t \geq -\varrho, |\varphi^x| \leq 1 \right\}.$$

The equations for minimization of (4.7) can now be obtained in a straightforward way from the general algorithm in [10].

5. Experiments

We now test the potential of TV_{S^1} regularization on several experiments. To evaluate the results we compare TV_{S^1} with TV , since no regularizers for S^1 valued functions have yet been investigated and TV_{S^1} is supposed to behave like TV in regions where u is smooth. With a parallel CUDA implementation on NVIDIA GTX 480 a typical runtime for 256×256 images using 64 levels for Γ is about 30 seconds for TV_{S^1} . TV requires about 12 seconds. For illustration, the TV_{S^1} runtime without using our efficient formulation in proposition 2.2 is more than 2 hours even for only 32 levels, and still about 40 minutes for 16 levels.

The analogon of the famous ROF denoising model of Rudin, Osher and Fatemi [14] for the cyclic case is

$$\inf_u \frac{\lambda}{2} \int_{\Omega} d_{S^1}(u(x) - f(x))^2 + TV_{S^1}(u). \quad (5.1)$$

We apply this to a one dimensional example in Fig. 4. The signal f consisting of a monotonous and a constant part has

been degraded by adding 10% Gaussian noise producing numerous wrap-arounds. The results show that only TV_{S^1} is able to reconstruct the signal. Thus cyclic wrap-arounds are handled correctly by TV_{S^1} . In contrast, for TV there is no choice of λ leading to a reconstruction without heavy fluctuations or large displacements. This example clearly demonstrates the advantage of TV_{S^1} over noncyclic regularizers.

Fig. 1 shows an inpainting experiment for the periodic hue values of the HSV color space. Values $0/9, \dots, 8/9$ are circularly arranged starting from the top left red region with value 0 and ending with the top right magenta region with value $8/9$. We use model (5.1) with $\lambda = 0$ within the inpainting circle area and $\lambda = \infty$ otherwise. As Fig. 1 shows, the image prior for cyclic values in regions with no data available is more natural with TV_{S^1} than with TV .

Many material properties such as deformation behavior can be deduced knowing the structures at atomic scale. In particular one is interested in the segmentation of grains, i.e. regions with homogeneous *orientation* of the atom lattice. To this end, when smooth transitions between neighboring regions are allowed, Boerdgen et al. [4] used TV regularization with a nonconvex data term. Since orientations are cyclic, the total *cyclic* variation TV_{S^1} is a more natural regularizer for this problem. Fig. 5 shows a comparison of segmentations using TV_{S^1} and TV , applied to an image obtained by the phase field crystal simulation model [15]. This experiment shows the solution behavior if one uses different representations of the orientation space S^1 as $[0, 1)$. The two solutions with TV_{S^1} are essentially identical due to the cyclic shift invariance in proposition 3.2. On the contrary, with TV they may differ and important structures may not be recognized in the segmentation (Fig. 5, (b)).

6. Conclusion

We have introduced a novel kind of total variation for cyclic structures. The regularizer penalizes the value jumps with the S^1 distance instead of linearly. A convex formulation is obtained through a recent theory for general functionals using cartesian currents. We provide an equivalent formulation of the constraint sets which allows to minimize functionals with TV_{S^1} regularization and arbitrary data terms as efficiently as the usual total variation. Experiments on a variety of imaging problems show the clear advantage of TV_{S^1} in the correct handling of value wrap-arounds as opposed to noncyclic regularizers such as TV .

References

[1] G. Alberti, G. Bouchitté, and G. D. Maso. The Calibration Method for the Mumford-Shah Functional and Free-discontinuity Problems. *Calc. Var. Partial Differential Equations*, (16(3)):299–333, 2003. 2, 7

[2] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford University Press, USA, 2000. 2

[3] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher. Simultaneous Structure and Texture Image Inpainting. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 2:707, 2003. 1

[4] M. Boerdgen, B. Berkels, M. Rumpf, and D. Cremers. Convex Relaxation for Grain Segmentation at Atomic Scale. *Vision, Modeling, and Visualization*, 2010. 6

[5] D. C. Ghiglia and M. D. Pritt. *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*. John Wiley & Sons, 1 edition, 1998. 1

[6] M. Giaquinta, G. Modica, and J. Souček. *Cartesian Currents in the Calculus of Variations II: Variational Integrals*. Springer Verlag, 1 edition, 1998. 2

[7] K. Kolev, M. Klodt, T. Brox, and D. Cremers. Continuous Global Optimization in Multiview 3D Reconstruction. *International Journal of Computer Vision*, 84(1):80–96, August 2009. 1

[8] A. Marquina and S. Osher. Image Super-resolution by TV Regularization and Bregman Iteration. *Journal of Scientific Computing*, (37(3)):367–382, 2008. 1

[9] T. Pock, A. Chambolle, H. Bischof, and D. Cremers. A Convex Relaxation Approach for Computing Minimal Partitions. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Miami, Florida, 2009. 2

[10] T. Pock, D. Cremers, H. Bischof, and A. Chambolle. An Algorithm for Minimizing the Piecewise Smooth Mumford-Shah Functional. In *IEEE International Conference on Computer Vision (ICCV)*, Kyoto, Japan, 2009. 2, 4, 5

[11] T. Pock, D. Cremers, H. Bischof, and A. Chambolle. Global Solutions of Variational Models with Convex Regularization. *SIAM Journal on Imaging Sciences*. To appear, 2010. 2, 3

[12] T. Pock, T. Schoenemann, G. Graber, H. Bischof, and D. Cremers. A Convex Formulation of Continuous Multi-Label Problems. In *European Conference on Computer Vision (ECCV)*, Marseille, France, October 2008. 2

[13] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1996. 2

[14] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear Total Variation Based Noise Removal Algorithms. *Phys. D*, 60(1-4):259–268, 1992. 1, 5

[15] P. Stefanovic, M. Haataja, and N. Provatas. Phase Field Crystal Study of Deformation and Plasticity in Nanocrystalline Materials. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, 80(4):046107+, 2009. 6

[16] F. Steinbruecker, T. Pock, and D. Cremers. Large Displacement Optical Flow Computation without Warping. In *IEEE International Conference on Computer Vision (ICCV)*, Kyoto, Japan, 2009. 1

[17] C. Vogel and M. Oman. Fast, Robust Total Variation-based Reconstruction of Noisy, Blurred Images. *IEEE Transactions on Image Processing*, (7(6)):813–824, 1998. 1

A. Appendix

Proof of (2.3) with (2.4) for $F = TV_{S^1}$. First we will prove that, for any feasible φ , the integral on the right hand side of (2.3) is less than or equal to $F(u) = TV_{S^1}(u)$. For a general vector field $\varphi : \Omega \times \Gamma \rightarrow \mathbb{R}^n \times \mathbb{R}$, $\varphi = (\varphi^x, \varphi^t)$ with $\varphi^t \equiv 0$ lemma 2.8 of [1] states that

$$\int_{\Omega \times \Gamma} \varphi \cdot \nabla 1_u = \int_{\Omega \setminus S_u} \varphi^x(x, u) \nabla u + \int_{S_u} \left(\int_{u^-}^{u^+} \varphi^x \right) \cdot \nu_u$$

This is less than or equal to

$$TV_{S^1}(u) = \int_{\Omega \setminus S_u} |\nabla u| + \int_{S_u} d_{S^1}(u^-, u^+)$$

if

$$\begin{aligned} \varphi^x(x, u) \nabla u &\leq |\nabla u| && \text{for } x \in \Omega \setminus S_u, \\ \left(\int_{u^-}^{u^+} \varphi^x \right) \cdot \nu_u &\leq d_{S^1}(u^-, u^+) && \text{for } x \in S_u. \end{aligned} \quad (\text{A.1})$$

These constraints are fulfilled, if

$$|\varphi^x(x, t)| \leq 1, \quad \left| \int_{t_1}^{t_2} \varphi^x(x, t) dt \right| \leq d_{S^1}(t_1, t_2) \quad (\text{A.2})$$

for all $x \in \Omega$, $t, t_1, t_2 \in \Gamma$. This proves (2.3) with \geq instead of equality. To show equality, we must construct a φ^x with (A.2) such that equalities hold in (A.1). For $x \in S_u$ set

$$\begin{aligned} \varphi^x(x, t) &= \nu_u(x) d_{S^1}(u^-, u^+) \\ &\quad \left(\frac{\chi_{[u^-, u^+]}(t)}{u^+ - u^-} - \frac{\chi_{[0, 1] \setminus [u^-, u^+]}(t)}{1 - (u^+ - u^-)} \right) \end{aligned}$$

and for $x \in \Omega \setminus S_u$ set $\varphi^x(x, t) = \frac{\nabla u(x)}{|\nabla u(x)|} (t - \frac{1}{2})$ if $\nabla u(x) \neq 0$ and $\varphi^x(x, t) = 0$ otherwise. One easily checks that φ^x satisfies the required conditions. The integral inequality is most conveniently checked using proposition 2.2. \square

Proof of proposition 3.1. We have $u = a$ inside A and $u = b$ outside A . The jump set S_u is therefore just the boundary $S_u = \partial A$ of A . Furthermore, $\nabla u \equiv 0$ since u is constant in every connected component of $\Omega \setminus S_u$. Thus, the definition (2.1) of TV_{S^1} yields

$$TV_{S^1}(u) = \int_{S_u} d_{S^1}(a, b) = d_{S^1}(a, b) \text{Per}(A). \quad \square$$

Proof of proposition 3.2. Without loss of generality, let $\alpha \in [0, 1)$. By proposition 2.3,

$$TV_{S^1}(T_\alpha \circ u) = \sup_{\varphi \in \hat{K}} \int_{\Omega} \int_0^{T_\alpha(u(x))} \text{div} \varphi(x, t) dt dx.$$

We have $T_\alpha(u(x)) = u(x) + \alpha - k$ with $k = 0$ or $k = 1$. Writing $p(t) = \text{div} \varphi(x, t \bmod 1)$ for brevity,

$$\begin{aligned} \int_0^{T_\alpha(u(x))} p(t) &= \int_0^{u(x)+\alpha-k} p(t) \\ &= \int_{\alpha-k}^{u(x)+\alpha-k} p(t) + \int_{-k}^{\alpha-k} p(t) - \int_{-k}^0 p(t) \\ &= \int_0^{u(x)} p(t + \alpha - k) + \int_0^\alpha p(t - k) - \int_0^k p(t - k) \\ &= \int_0^{u(x)} \text{div} \varphi_\alpha + \int_0^\alpha \text{div} \varphi - \int_0^k \text{div} \varphi \end{aligned}$$

with $\varphi_\alpha(x, t) := \varphi(x, t + \alpha \bmod 1)$. The last integral vanishes trivially for $k = 0$, and also for $k = 1$ pulling out the div operator and using $\int_\Gamma \varphi = 0$. Hence, we obtain

$$\int_{\Omega} \int_0^{T_\alpha(u(x))} \text{div} \varphi = \int_{\Omega} \int_0^{u(x)} \text{div} \varphi_\alpha + \int_0^\alpha \int_{\Omega} \text{div} \varphi.$$

The last integral is zero by the Gaussian integral formula since φ has zero values at the boundary. It is easy to check that $\varphi_\alpha \in \hat{K}$ if $\varphi \in \hat{K}$ and that $(\varphi_{-\alpha})_\alpha = \varphi$, hence $\{\varphi_\alpha \mid \varphi \in \hat{K}\} = \hat{K}$. Taking the sup over $\varphi \in \hat{K}$ on both sides we obtain the required identity. \square