Total Cyclic Variation

Evgeny Strekalovskiy

Computer Science Department
Technical University Munich

Joint work with Daniel Cremers

Cremers, Strekalovskiy, JMIV 2012

Total Variation in Image Processing: Denoising

Input image f TV-denoised image u

$$\min_{u:\Omega\to\mathbb{R}^k} \int_\Omega (f-u)^2 \, dx + \lambda \, TV(u)$$

Rudin, Osher, Fatemi, Physica D 1992

Total Variation in Image Processing: Optical Flow

Input images f_1, f_2

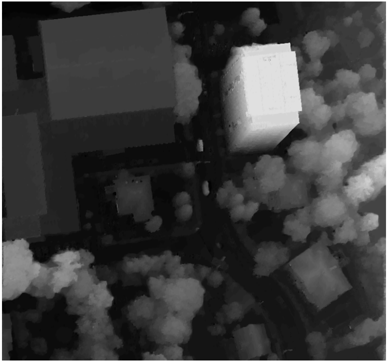
Optical flow field u

$$\min_{u:\Omega\to\mathbb{R}^k} \int_\Omega (f_1 - f_2 \circ u)^2 \, \mathrm{d}x + \lambda \, TV(u)$$

Brox et al., ECCV 2004, Zach et al., DAGM 2007

Total Variation in Image Processing: Stereo

One of two aerial images f_i



Disparity field u

$$\min_{u:\Omega\to\mathbb{R}} \int_\Omega (f_1 - f_2 \circ u)^2 \, \mathrm{d}x + \lambda \, TV(u)$$

Pock et al., ECCV 2008, SIIMS 2010

Total Variation for Scalar-Valued Functions

Total variation is defined as

$$TV(u) = \int_\Omega |Du| = \sup_{|\xi|\le 1} \int_\Omega u \operatorname{div} \xi \, \mathrm{d}x,$$

for functions $u \in L^1(\Omega, \mathbb{R})$ on a domain $\Omega \subset \mathbb{R}^m$.

Total Variation for Vector-Valued Functions

Separate directions, uncoupled (*Blomgren, Chan TIP '98*):

$$TV_S(u) := \sum_{i=1}^{k} TV(u_i) \quad = \quad \sup_{\xi_i:\Omega\to\mathbb{E}^k} \sum_{i=1}^{k} \int_\Omega u_i \operatorname{div}\xi_i \, \mathrm{d}x.$$

Separate directions, coupled (*Sapiro, Ringach TIP '96*):

$$TV_F(u) = \int_\Omega \|\nabla u\|_2 \, dx \; := \; \sup_{\xi:\Omega\to\mathbb{E}^{m\cdot k}} \sum_{i=1}^{k} \int_\Omega u_i \operatorname{div}(\xi_i) \, \mathrm{d}x.$$

Shared direction, coupled (*Goldluecke et al., SIIMS '12*):

$$TV_J(u) = \int_\Omega \|\nabla u\|_{\sigma_1} \, dx := \sup_{\xi:\Omega\to\mathbb{E}^m,\, \eta\to\mathbb{E}^k} \sum_{i=1}^{k} \int_\Omega u_i \operatorname{div}(\eta_i \xi) \, \mathrm{d}x.$$

Total Variation: Advantages

The total variation has a number of favorable properties, in particular:

- Total variation is discontinuity-preserving.
- Total variation is convex.
- Dual formulations allow for exact minimization algorithms.
- Total variation of scalar valued u equals the sum of total variations of its upper level sets (coarea formula).

Total Variation for Manifold-Valued Functions

In many applications, signals u take on values in a manifold $\mathcal{M}$:

$$u : \Omega \to \mathcal{M}.$$

The simplest non-trivial example is the unit circle

$$\mathcal{S}^1 = \left\{ x \in \mathbb{R}^2 \,\middle|\, |x| = 1 \right\}$$

Total Variation for Manifold-Valued Functions

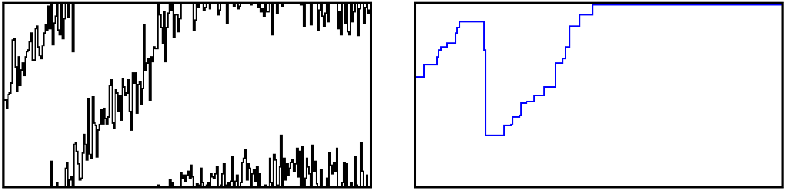
In many applications, signals u take on values in a manifold $\mathcal{M}$:

$$u : \Omega \to \mathcal{M}.$$

The simplest non-trivial example is the unit circle

$$\mathcal{S}^1 = \left\{ x \in \mathbb{R}^2 \,\middle|\, |x| = 1 \right\}$$

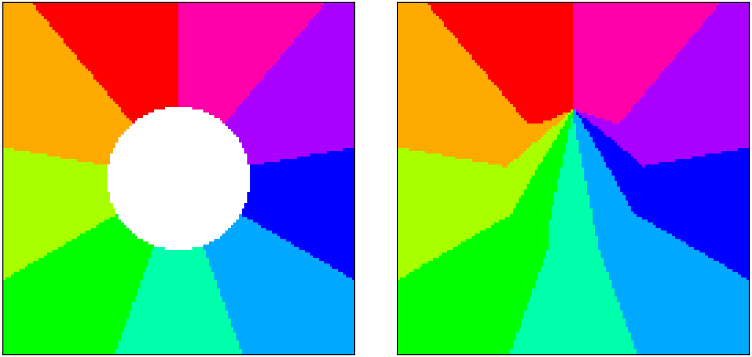
Simply representing the circle by the interval $[0, 1)$ leads to an incorrect penalization of jumps from 0 to 1:



Noisy input $f : [0, 1] \to \mathcal{S}^1$

TV-denoised

Naive Total Variation Inpainting of Cyclical Data



Inpainting problem

TV-inpainted

Simply representing the circle $\mathcal{S}^1$ by the interval $[0, 1)$ leads to a bias suppressing the transition from 0 (red) to 8/9 (pink).

Total Variation for Manifold-Valued Functions

Definition (Giaquinta, Mucci, PAMQ 2007)

Let $\mathcal{M}$ be a manifold embedded in $\mathbb{R}^N$. Then

$$\mathrm{BV}(\Omega, \mathcal{M}) := \left\{ u \in \mathrm{BV}(\Omega, \mathbb{R}^N) \,\middle|\, u(x) \in \mathcal{M} \right\},$$

$$TV_{\mathcal{M}}(u) := \underbrace{\int_{\Omega} |\nabla u| \, \mathrm{d}x}_{\text{diff. part}} + \underbrace{|D^C(u)|}_{\text{Cantor part}} + \underbrace{\int_{S_u} \mathcal{H}^1(l_x) d\mathcal{H}^{n-1}}_{\text{jump part}}$$

where l_x denotes a geodesic arc connecting $u^-(x)$ and $u^+(x)$.

Total Variation for Manifold-Valued Functions

Definition (Giaquinta, Mucci, PAMQ 2007)

Let $\mathcal{M}$ be a manifold embedded in $\mathbb{R}^N$. Then

$$\mathrm{BV}(\Omega, \mathcal{M}) := \left\{ u \in \mathrm{BV}(\Omega, \mathbb{R}^N) \,\middle|\, u(x) \in \mathcal{M} \right\},$$

$$TV_{\mathcal{M}}(u) := \underbrace{\int_{\Omega} |\nabla u| \, \mathrm{d}x}_{\text{diff. part}} + \underbrace{|D^C(u)|}_{\text{Cantor part}} + \underbrace{\int_{S_u} \mathcal{H}^1(l_x) d\mathcal{H}^{n-1}}_{\text{jump part}}$$

where l_x denotes a geodesic arc connecting $u^-(x)$ and $u^+(x)$.

Problem

- need to work with a higher dimensional embedding
- not clear how to devise a minimization algorithm

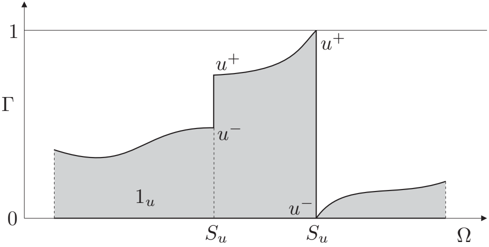
Total Cyclic Variation

Instead of an embedding, represent $\mathcal{S}^1$ as $\mathbb{R}$.

Definition

For $u \in SBV(\Omega; \mathbb{R})$ we define the *total cyclic variation* as

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



Graph of a function $u \in SBV(\Omega; \mathbb{R})$.

Total Cyclic Variation

Instead of an embedding, represent $\mathcal{S}^1$ as $\mathbb{R}$.

Definition

For $u \in SBV(\Omega; \mathbb{R})$ we define the *total cyclic variation* as

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

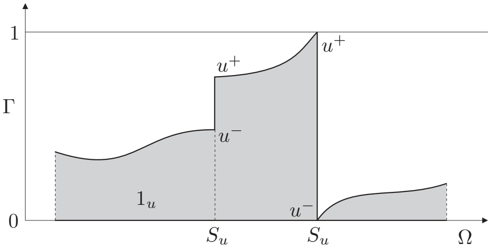
Not convex!

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

Convexification using *functional lifting* (*Alberti et al. 2003*).

Idea: Consider the graph function 1_u of u:

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



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

Convexification using *functional lifting* (*Alberti et al. 2003*).

Lemma (Functional lifting, cf. Antonin's talk)

For $u \in SBV(\Omega;\mathbb{R})$*,*

$$\int_{\Omega \setminus S_u} h(x, u(x), \nabla u(x))\,dx + \int_{S_u} d(x, u^-(x), u^+(x))\,d\mathcal{H}^{m-1}(x).$$

$$\geq \sup_{\varphi} \int_{\Omega\times\mathbb{R}} 1_u \cdot \operatorname{div}\varphi$$

with supremum taken over all $\varphi = (\varphi^x, \varphi^t) : \Omega \times \mathbb{R} \mapsto \mathbb{R}^m \times \mathbb{R}$*, s.t.*

$$\varphi^t(x,t) \geq h^*(x,t,\varphi^x(x,t)) \quad \textit{and} \quad \left| \int_t^{t'} \varphi^x(x,s)\,ds \right| \leq d(x,t,t')$$

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

Proposition (Convex representation of $TV_{\mathcal{S}^1}$)

For $u \in SBV(\Omega)$,

$$TV_{\mathcal{S}^1}(u) = \sup_{\varphi \in K} \int_{\Omega \times \mathbb{R}} 1_u \cdot \operatorname{div} \varphi$$

with the constraint set

$$K = \Big\{ \varphi : \Omega \times \mathbb{R} \mapsto \mathbb{R}^m \,\Big|\, |\varphi(x,t)| \leq 1, \\ \Big| \int_t^{t'} \varphi(x,s)\, ds \Big| \leq d_{\mathcal{S}^1}(t,t') \quad \forall x \in \Omega,\ t, t' \in \mathbb{R} \} \Big\}.$$

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

Proposition (Convex representation of $TV_{\mathcal{S}^1}$)

For $u \in SBV(\Omega)$,

$$TV_{\mathcal{S}^1}(u) = \sup_{\varphi \in K} \int_{\Omega \times \mathbb{R}} 1_u \cdot \operatorname{div} \varphi$$

with the constraint set

$$K = \Big\{ \varphi : \Omega \times \mathbb{R} \mapsto \mathbb{R}^m \;\Big|\; |\varphi(x,t)| \leq 1, \\ \Big| \int_t^{t'} \varphi(x,s)\, ds \Big| \leq d_{\mathcal{S}^1}(t,t') \quad \forall x \in \Omega,\; t,t' \in \mathbb{R} \} \Big\}.$$

\+ convex in 1_u

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

Proposition (Convex representation of $TV_{\mathcal{S}^1}$)

For $u \in SBV(\Omega)$,

$$TV_{\mathcal{S}^1}(u) = \sup_{\varphi \in K} \int_{\Omega \times \mathbb{R}} 1_u \cdot \operatorname{div} \varphi$$

with the constraint set

$$K = \Big\{ \varphi : \Omega \times \mathbb{R} \mapsto \mathbb{R}^m \,\Big|\, |\varphi(x,t)| \le 1, \\ \Big| \int_t^{t'} \varphi(x,s)\, ds \Big| \le d_{\mathcal{S}^1}(t,t') \quad \forall x \in \Omega,\; t,t' \in \mathbb{R} \} \Big\}.$$

+ convex in 1_u

− quadratic complexity after discretization

From Quadratic to Linear Complexity

The quadratic number of constraints is equivalent to a linear number of constraints.

Proposition (Constraint equivalence)

$$\Big| \int_t^{t'} \varphi(x,s)\, ds \Big| \le d_{\mathcal{S}^1}(t,t') \quad \forall\; t, t' \in \mathbb{R}$$

$$\Longleftrightarrow$$

$$\varphi(x,\cdot) \text{ 1-}\textit{periodic},$$

$$|\varphi(x,t)| \le 1 \quad \textit{for a.e. } t \in [0,1),$$

$$\int_0^1 \varphi(x,s)\, ds = 0.$$

From Quadratic to Linear Complexity

As a consequence, we obtain the following representation of $TV_{\mathcal{S}^1}$.

Proposition (Dual representation)

For $u \in SBV(\Omega)$ it holds

$$TV_{\mathcal{S}^1}(u) = \sup_{\varphi \in K} \int_\Omega \int_0^{u(x)} \operatorname{div} \varphi(x,s)\, ds\, dx,$$

$$K = \Big\{ \varphi : \Omega \times \mathbb{R} \to \mathbb{R}^m \;\Big|\; \varphi(x,\cdot) \text{ 1-\textit{periodic}}, \\ |\varphi(x,t)| \le 1, \quad \int_0^1 \varphi(x,s)\, ds = 0 \quad \forall x \in \Omega,\ t \in [0,1) \Big\}.$$

From Quadratic to Linear Complexity

As a consequence, we obtain the following representation of $TV_{\mathcal{S}^1}$.

Proposition (Dual representation)

For $u \in SBV(\Omega)$ it holds

$$TV_{\mathcal{S}^1}(u) = \sup_{\varphi \in K} \int_\Omega \int_0^{u(x)} \operatorname{div} \varphi(x,s)\, ds\, dx,$$

$$K = \Big\{ \varphi : \Omega \times \mathbb{R} \to \mathbb{R}^m \;\Big|\; \varphi(x,\cdot) \text{ 1-periodic}, \quad |\varphi(x,t)| \le 1, \quad \int_0^1 \varphi(x,s)\, ds = 0 \quad \forall x \in \Omega,\ t \in [0,1) \Big\}.$$

- \+ convex in 1_u
- \+ linear complexity after discretization

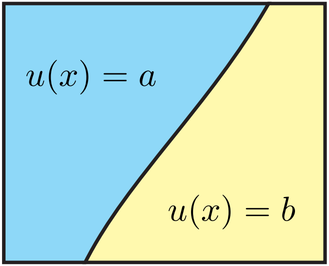
Properties of $TV_{\mathcal{S}^1}$

Proposition (Consistency with $d_{\mathcal{S}^1}$)

Let $u = a\chi_A + b\chi_{\bar{A}}$ *for some* $a, b \in \mathbb{R}$ *and* $A \subset \Omega$*. Then*

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

where $Per(A) = TV(\chi_A)$ *is the perimeter of* A*.*

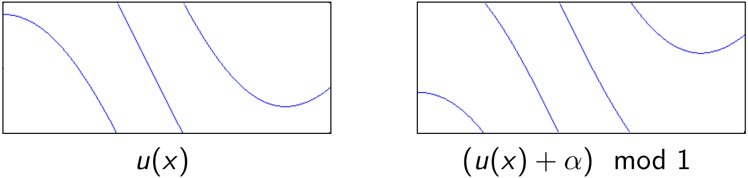


Properties of $TV_{\mathcal{S}^1}$

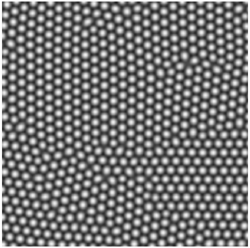
Proposition (Cyclic shift invariance)

Let $u : \Omega \to [0,1)$ *and* T_α *a cyclic shift by* $\alpha \in \mathbb{R}$. *Then*

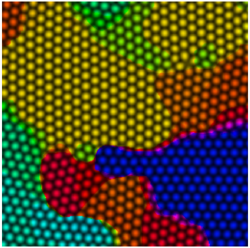
$$TV_{\mathcal{S}^1}(T_\alpha \circ u) = TV_{\mathcal{S}^1}(u).$$



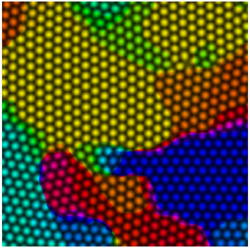
Properties of $TV_{\mathcal{S}^1}$: Cyclic shift invariance



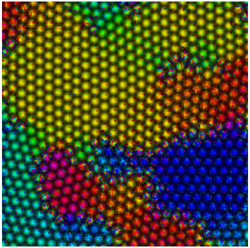
Input



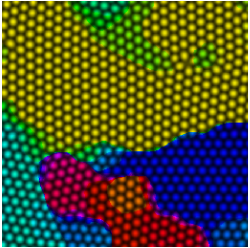
TV



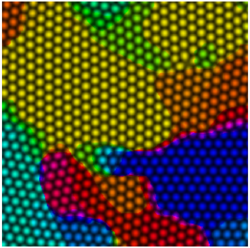
$TV_{\mathcal{S}^1}$



Data term



TV, shifted data



$TV_{\mathcal{S}^1}$, shifted data

Properties of $TV_{\mathcal{S}^1}$

Proposition (Representation invariance)

For a $u : \Omega \to \mathbb{R}$ *and an integer function* $k : \Omega \to \mathbb{Z}$ *define* $u_k(x) := u(x) + k(x)$. *Then*

$$TV_{\mathcal{S}^1}(u_k) = TV_{\mathcal{S}^1}(u).$$

Properties of $TV_{\mathcal{S}^1}$

Proposition (Representation invariance)

For a $u : \Omega \to \mathbb{R}$ *and an integer function* $k : \Omega \to \mathbb{Z}$ *define* $u_k(x) := u(x) + k(x)$. *Then*

$$TV_{\mathcal{S}^1}(u_k) = TV_{\mathcal{S}^1}(u).$$

Proposition (Lower-semicontinuity)

$TV_{\mathcal{S}^1}$ *is lower-semicontinuous with respect to pointwise a.e. convergence in the* $d_{\mathcal{S}^1}$ *metric, i.e. it holds*

$$TV_{\mathcal{S}^1}(u) \leq \liminf_{n\to\infty} TV_{\mathcal{S}^1}(u_n)$$

for $u_n : \Omega \to \mathbb{R}$, $n \geq 1$, *with* $d_{\mathcal{S}^1}(u_n(x), u(x)) \to 0$ *for a.e.* $x \in \Omega$.

Optimization Problem

Regularization of an arbitrary data term $\varrho(x, u(x))$:

$$\inf_{u:\Omega\to\mathbb{R}} E(u) = \int_\Omega \varrho(x, u(x))\, dx + \lambda\, TV_{\mathcal{S}^1}(u). \tag{1}$$

Optimization Problem

Regularization of an arbitrary data term $\varrho(x, u(x))$:

$$\inf_{u:\Omega\to\mathbb{R}} E(u) = \int_\Omega \varrho(x, u(x))\, dx \,+\, \lambda\, TV_{\mathcal{S}^1}(u). \tag{1}$$

Theorem (Existence of minimizers)

For $\varrho(x, \cdot)$ lower-semicontinuous with respect to $d_{\mathcal{S}^1}$ convergence, the problem (1) admits a minimizer $u : \Omega \to \mathbb{R}$ with $TV_{\mathcal{S}^1}(u) < \infty$.

Optimization problem: Algorithm

- Setting $v = 1_u$ and relaxing the constraint $v(x,t) \in \{0,1\}$ to $v(x,t) \in [0,1]$, the overall optimization problem is *convex*.
- Discretization of the range $[0,1)$ of u into $n \geq 1$ levels $\frac{0}{n}, \frac{1}{n}, \ldots, \frac{n-1}{n}$.
- Lagrange multipliers q for $\int_0^1 \varphi(x,s)\,ds = 0$
- Primal-dual algorithm (*Chambolle, Pock, JMIV 2011*), GPU parallelization

Optimization problem: Algorithm

Dual ascent in φ: for $0 \leq i < n$, pointwise in $x \in \Omega$:

$$\varphi_i^{x,k+1} = \pi_{|\cdot| \leq \frac{\lambda}{n}} \left(\varphi_i^{x,k} + \tfrac{\sigma}{3} \left(\nabla \bar{v}_i^k - \bar{q}_i^k \right) \right),$$

$$\varphi_i^{t,k+1} = \max \left(\varphi_i^{t,k} + \tfrac{\sigma}{2}\, \partial_t^+ \bar{v}_i^k \,,\ -\varrho_i \right).$$

Primal descent in v, q: for $1 \leq i < n$, pointwise in $x \in \Omega$:

$$v_i^{k+1} = \pi_{[0,1]} \left(v_i^k - \tfrac{\tau}{6} \left(\operatorname{div}_x \varphi_i^{x,k+1} + \partial_t^- \varphi_i^{t,k+1} \right) \right),$$

$$q^{k+1} = q^k + \tfrac{\tau}{n} \sum_{j=0}^{n-1} \varphi_j^{x,k+1}.$$

Extrapolation step: for $1 \leq i < n$, pointwise in $x \in \Omega$:

$$\bar{v}_i^{k+1} = 2 v_i^{k+1} - v_i^k,$$

$$\bar{q}^{k+1} = 2 q^{k+1} - q^k.$$

Optimization problem: Algorithm

Dual ascent in φ: for $0 \le i < n$, pointwise in $x \in \Omega$:

$$\varphi_i^{x,k+1} = \pi_{|\cdot| \le \frac{\lambda}{n}} \left(\varphi_i^{x,k} + \tfrac{\sigma}{3} \left(\nabla \bar{v}_i^k - \bar{q}_i^k \right) \right),$$

$$\varphi_i^{t,k+1} = \max \left(\varphi_i^{t,k} + \tfrac{\sigma}{2} \, \partial_t^+ \bar{v}_i^k \, , \, -\varrho_i \right).$$

Primal descent in v, q: for $1 \le i < n$, pointwise in $x \in \Omega$:

$$v_i^{k+1} = \pi_{[0,1]} \left(v_i^k - \tfrac{\tau}{6} \left(\operatorname{div}_x \varphi_i^{x,k+1} + \partial_t^- \varphi_i^{t,k+1} \right) \right),$$

$$q^{k+1} = q^k + \tfrac{\tau}{n} \sum_{j=0}^{n-1} \varphi_j^{x,k+1}.$$

256×256 images, $n = 64$ levels:
10 sec for $TV_{\mathcal{S}^1}$, 5 sec for TV

Extrapolation step: for $1 \le i < n$, pointwise in $x \in \Omega$:

$$\bar{v}_i^{k+1} = 2 v_i^{k+1} - v_i^k,$$

$$\bar{q}^{k+1} = 2 q^{k+1} - q^k.$$

Optimization problem: Optimality

Proposition (Discrete $TV_{\mathcal{S}^1}$ is NP-hard)

For a discrete range with $n \geq 3$ levels, where $n = 3z$ for some $z \in \mathbb{Z}$, the regularization with $TV_{\mathcal{S}^1}$ is NP-hard.

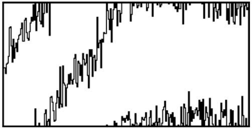
Optimization problem: Optimality

Proposition (Discrete $TV_{\mathcal{S}^1}$ is NP-hard)

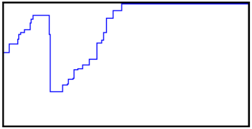
For a discrete range with $n \geq 3$ levels, where $n = 3z$ for some $z \in \mathbb{Z}$, the regularization with $TV_{\mathcal{S}^1}$ is NP-hard.

- In general, no optimal solutions can be expected.
- At most 5% away from the global optimum in practice

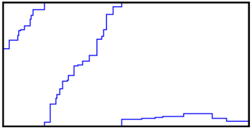
Experimental Comparison: Denoising



Noisy input

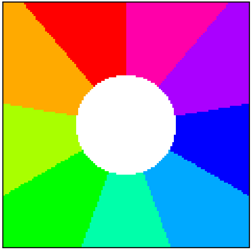


TV-denoised

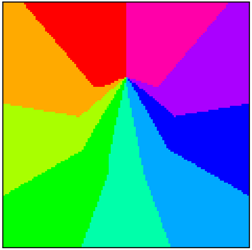


$TV_{\mathcal{S}^1}$-denoised

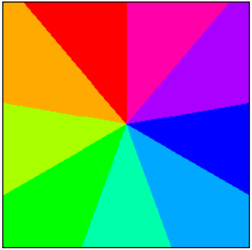
Experimental Comparison: Inpainting



Input

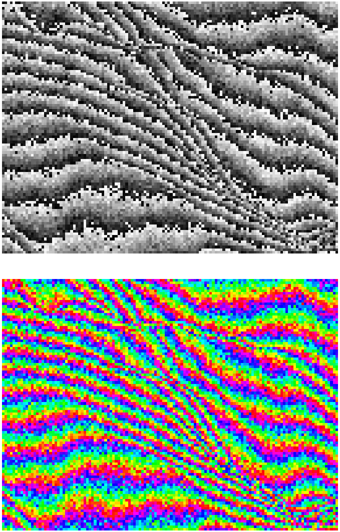


TV-inpainting

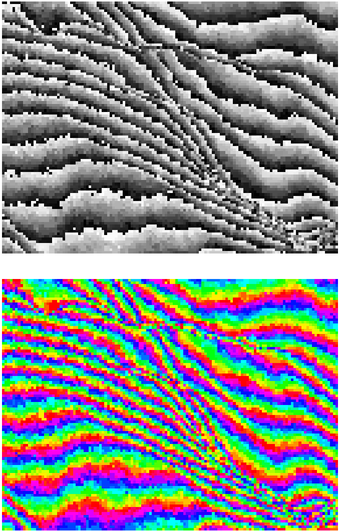


$TV_{\mathcal{S}^1}$-inpainting

Application: SAR Image Denoising



Noisy phase



$TV_{\mathcal{S}^1}$-denoising

Extension: General cyclic regularizers

Quadratic $\mathcal{S}^1$-regularization:

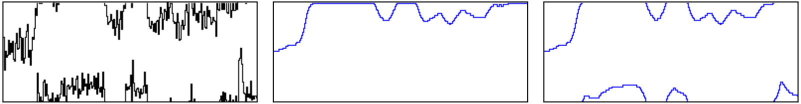
$$R(u) = \lambda \int_\Omega |\nabla u|^2 \, dx \; + \; \int_{S_u} \left. \begin{cases} 0 & \text{if } d_{\mathcal{S}^1}(u^-, u^+) = 0 \\ \infty & \text{else} \end{cases} \right\} d\,\mathcal{H}^{m-1}$$

Extension: General cyclic regularizers

Quadratic $\mathcal{S}^1$-regularization:

$$R(u) = \lambda \int_{\Omega} |\nabla u|^2 \, dx \; + \; \int_{S_u} \left\{ \begin{matrix} 0 & \text{if } d_{\mathcal{S}^1}(u^-, u^+) = 0 \\ \infty & \text{else} \end{matrix} \right\} d\,\mathcal{H}^{m-1}$$

Smoothing of cyclical signals:



Noisy input

quadratic reg.
cyclic data term

quadratic $\mathcal{S}^1$-reg.
cyclic data term

Extension: General cyclic regularizers

Huber $\mathcal{S}^1$-regularization:

$$R(u) = \lambda \int_{\Omega \setminus S_u} h_\varepsilon(\nabla u)\, dx \;+\; \lambda \int_{S_u} d_{\mathcal{S}^1}(u^-, u^+)\, d\,\mathcal{H}^{m-1}$$

with

$$h_\varepsilon(p) := \begin{cases} \frac{|p|^2}{2\varepsilon} & \text{if } |p| \le \varepsilon, \\ |p| - \frac{\varepsilon}{2} & \text{else} \end{cases} \tag{2}$$

Reduces oversmoothing (staircasing) of $TV_{\mathcal{S}^1}$.

Extension: General cyclic regularizers

Truncated Linear $\mathcal{S}^1$-regularization

$$R(u) = \lambda \int_{\Omega \setminus S_u} |\nabla u| \, dx \; + \; \int_{S_u} \min\left(\nu, \, \lambda \, d_{\mathcal{S}^1}(u^-, u^+)\right) d\,\mathcal{H}^{m-1}$$

Extension: General cyclic regularizers

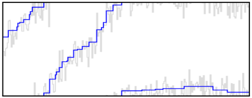
Truncated Linear $\mathcal{S}^1$-regularization

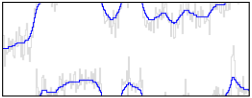
$$R(u) = \lambda \int_{\Omega \setminus S_u} |\nabla u| \, dx \; + \; \int_{S_u} \min\left(\nu, \, \lambda \, d_{\mathcal{S}^1}(u^-, u^+)\right) d\,\mathcal{H}^{m-1}$$

Mumford-Shah $\mathcal{S}^1$-regularization:

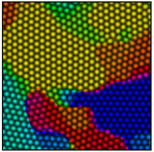
$$R(u) = \lambda \int_{\Omega \setminus S_u} |\nabla u|^2 \, dx \; + \; \int_{S_u} \left\{ \begin{matrix} 0 & \text{if } d_{\mathcal{S}^1}(u^-, u^+) = 0 \\ \nu & \text{else} \end{matrix} \right\} d\,\mathcal{H}^{m-1}$$

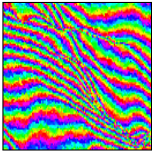
Conclusion





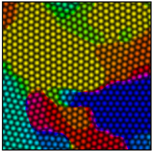
- Novel kind of total variation, $TV_{\mathcal{S}^1}$, for cyclic structures
- Useful invariance and lower-semicontinuity properties
- Same complexity as usual TV, arbitrary data terms allowed
- Clear advantage in handling value wrap-arounds as opposed to noncyclic regularizers





Conclusion

- Novel kind of total variation, $TV_{\mathcal{S}^1}$, for cyclic structures
- Useful invariance and lower-semicontinuity properties
- Same complexity as usual TV, arbitrary data terms allowed
- Clear advantage in handling value wrap-arounds as opposed to noncyclic regularizers



Thank you!