A Non-Local Formulation for Higher-Order Total Variation-Based Regularization

Jan Lellmann

CIA/DAMTP, University of Cambridge

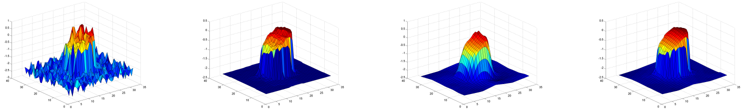
Joint work with: K. Papafitsoros, D. Spector, C. Schönlieb

Hong Kong, May 2014

Motivation – Higher-order regularization

- **Higher-order** total variation

$$\begin{aligned} \mathrm{BV}^2 &:= \{u \in W^{1,1} | \nabla u \in \mathrm{BV}(\Omega, \mathbb{R}^N)\} \\ \mathrm{TV}^2(u) &:= |D^2 u|(\Omega) \end{aligned}$$



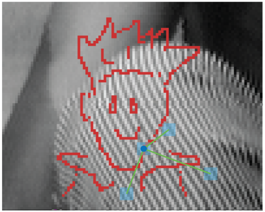
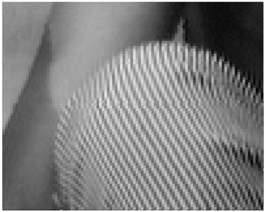
input | TV (staircase) | TV^2 (smooth) | TGV

- Not much more effort numerically
- Bounded Hessian BH [Scherzer'98] [Lysaker, Lundervold, Tai'03] [Lysaker, Tai'06] [Hinterberger, Scherzer'06] [Chan, Esedoglu, Park'07], inf-convolution [Chambolle, Lions'97], TGV [Bredies, Kunisch, Pock'09] [Setzer, Steidl, Teuber'08/'10], anisotropic variants [L., Morel, Schönlieb'13]

- Difference-based **nonlocal regularizer** [Buades, Coll, Morel'05] [Gilboa, Osher'08]:

$$J_{NL-TV_a}(u) = \int_\Omega \left(\int_\Omega |u(x) - u(y)| \sqrt{w(x,y)} dy \right) dx$$

- w can be chosen
  - *local* → approximate TV (later)
  - *non-local* → exploit self-similarity/repeating patterns in image

[Images: Gilboa, Osher: Nonlocal Operators with Applications to Image Processing, SIAM MMSIM'08]

- The weights w can have structured or far-reaching interactions
→ good for regularization of repeating structures.

Goal

How to do construct *non-local*, *higher-order* total variation?

Proposed implicit formulation

Definition

For $u \in L^1(\Omega)$, we define the **nonlocal gradient** $\mathcal{G}'(x) := \mathcal{G}'_\sigma u(x) \in \mathbb{R}^N$ and nonlocal Hessian $\mathcal{H}'(x) := \mathcal{H}'_\sigma u(x) \in \mathbb{R}^{N\times N}$ *implicitly* as the *minimizers* of

$$(\mathcal{H}'(x), \mathcal{G}'(x)) := \operatorname{argmin}_{\mathcal{G}',\mathcal{H}'} \; \frac{1}{2}\int_{\Omega-\Omega} R^2_{u,\mathcal{G}',\mathcal{H}'}(x,z)\,\sigma_x(z)dz,$$

$$R_{u,\mathcal{G}',\mathcal{H}'}(x,z) := u(x+z) - u(x) - \mathcal{G}'^\top z - \frac{1}{2} z^\top \mathcal{H}' z.$$

and the **non-local** first- and **second-order total variation**

$$TV_{NL} := \int_\Omega \|\mathcal{G}'(x)\| dx, \quad TV^2_{NL} := \int_\Omega \|\mathcal{H}'(x)\| dx.$$

- pairwise weights $\sigma_x(y)$ define the non-locality
- extension to higher orders straightforward

Computational aspects

- How to numerically solve functionals involving the non-local Gradient/Hessian

$$(\mathcal{G}'(x), \mathcal{H}'(x)) = \operatorname{argmin}_{\mathcal{G}',\mathcal{H}'} \; \frac{1}{2} \int_{\Omega-\Omega} R^2_{u,\mathcal{G}',\mathcal{H}'}(x,z)\,\sigma_x(z) dz,$$

Regularized non-local L1-TV2

$$\inf_u \|u - f\|_{L^1} + \lambda \int_\Omega \|\mathcal{H}'(x)\| dx$$

$$\text{s.t.} \quad (\mathcal{G}'(x), \mathcal{H}'(x)) = \operatorname{argmin}_{\mathcal{G}',\mathcal{H}'} \; \frac{1}{2} \int_{\Omega-\Omega} R^2_{u,\mathcal{G}',\mathcal{H}'}(x,z)\,\sigma_x(z) dz,$$

$$R_{u,\mathcal{G}',\mathcal{H}'}(x,z) := u(x+z) - u(x) - \mathcal{G}'^\top z - \frac{1}{2} z^\top \mathcal{H}' z.$$

- *Convex* (SOCP) objective with *linear* constraints $\Rightarrow$ solvable!
- Free to choose σ_x – *local* or *non-local*

Overview

Localization

Analysis

Question

Can the non-local TV^2 be **localized**, i.e.,

$$\mathrm{TV}^2_{NL} \to \mathrm{TV}^2$$

in some sense for a suitable sequence of σ_x?

Related work – first order non-local TV

- *Another look at Sobolev spaces* [Bourgain, Brezis, Mironescu'01]: $\Omega \subseteq \mathbb{R}^N$ open bounded with Lipschitz boundary, $\rho_n(z) = \hat{\rho}_n(\|z\|) \geqslant 0$ radially symmetric localization kernel, $\int_{\mathbb{R}^N} \rho_n(z) dz = 1$, converging to δ_0 as $n \rightarrow \infty$.

$$\lim_{n\rightarrow\infty} \int_\Omega \int_{\Omega-\Omega} \frac{|u(x+z) - u(x)|}{\|z\|} \hat{\rho}_n(\|z\|) dz dx \ = K_{1,N} TV(u).$$

- This **characterizes** BV!
- Also: Sobolev spaces $W^{1,p}$, generalization to non-radial ρ_ε [Ponce'04] and arbitrary open domains [Spector'11].
- Compare to [Gilboa, Osher'08]

$$J_{NL-TV_a}(u) = \int_\Omega \left(\int_{\Omega-\Omega} |u(x+z) - u(x)| \sqrt{w(x, x+z)} dz \right) dx$$

Related work – first-order non-local gradients

- We can define a full **explicit nonlocal gradient** at a point x: [Gilboa, Osher'08], [Du, Gunzburger, Lehoucq, Zhou'13], [Mengesha, Spector'13]

$$\mathcal{G}_\rho u(x) := N \int_{\Omega-\Omega} \frac{u(x+z)-u(x)}{\|z\|} \frac{z}{\|z\|} \rho(z) dz$$

  $\rho \geqslant 0$ radial, $u : \Omega \subseteq \mathbb{R}^N \to \mathbb{R}$ scalar.

- Then $\mathcal{G}_{\rho_n} u \mathcal{L}^N \to Du$ strictly as measures and

$$u \in \mathrm{BV}(\Omega) \quad \Longleftrightarrow \quad \liminf_{n\to\infty} \|\mathcal{G}_{\rho_n}\|_{L^1(\Omega)} < +\infty.$$

- Relates to "gradient-based" formulation in [Gilboa, Osher'08]

Proposed model – second-order non-local TV

Implicit model

$$(\mathcal{G}'(x), \mathcal{H}'(x)) = \operatorname{argmin}_{\mathcal{G}',\mathcal{H}'} \ \frac{1}{2} \int_{\Omega-\Omega} R^2_{u,\mathcal{G}',\mathcal{H}'}(x,z)\,\sigma_x(z)dz$$

Explicit model

$$\mathcal{H}_\rho u(x) = C \int_{\mathbb{R}^N} \frac{u(x+z) - 2u(x) + u(x-z)}{\|z\|^2} \left(\frac{zz^\top}{\|z\|^2} - \frac{1}{N+2} I \right) \rho(z)dz$$

- Reconstructs the full Hessian from all directional second derivatives

Connection to implicit model

- Restrict implicit formulation to spheres with radius h:

$$(\mathcal{G}'_h u(x), \mathcal{H}'_h u(x)) := \arg\min_{\mathcal{G}',\mathcal{H}'} \frac{1}{2} \int_{h\mathcal{S}^{N-1}} R^2_{u,\mathcal{G}',\mathcal{H}'}(x,z)dz$$

Proposition

With this definition, explicit ($\mathcal{H}_\rho$) and implicit ($\mathcal{H}'_\sigma$) model are related:

$$\mathcal{H}_\rho u(x) = \int_0^\infty \mathcal{H}'_h u(x) \tau_\rho(h) dh,$$

where

$$\tau_\rho(h) := \int_{h\mathcal{S}^{N-1}} \rho(x) d\mathcal{H}^{N-1}(x).$$

Main results – Localization in the regular case

Explicit model

$$\mathcal{H}_\rho u(x) = C \int_{\mathbb{R}^N} \frac{u(x+z) - 2u(x) + u(x-z)}{\|z\|^2} \left(\frac{zz^\top}{\|z\|^2} - \frac{1}{N+2} I \right) \rho(z) dz$$

Theorem (regular case)

Let $1 \leq p < \infty$. Then for every $u \in W^{2,p}(\mathbb{R}^N)$ we have that

$$\mathcal{H}_{\rho_n} u \to \nabla^2 u \quad in\ L^p(\mathbb{R}^N, \mathbb{R}^{N \times N}) \quad as\ n \to \infty.$$

- Steps in proof:
  - show uniform convergence for C_c^2 functions using mean value theorem
  - show manually that $\mathcal{H}_\rho$ is in fact an L^p function for $u \in W^{2,p}$
  - use density of $C_c^2(\mathbb{R}^N)$ in $W^{2,p}(\mathbb{R}^N)$

Main results – Localization for BV2

Theorem (Localization for BV2 functions)

Let $u \in \mathrm{BV}^2(\mathbb{R}^N)$. *Then*

$$\mathcal{H}_{\rho_n} u \mathcal{L}^N \to D^2 u \quad \textit{weakly}^*,$$

i.e., for every $\phi \in C_0(\mathbb{R}^N, \mathbb{R}^{N\times N})$

$$\lim_{n\to\infty} \int_{\mathbb{R}^N} \mathcal{H}_{\rho_n} u(x) \cdot \phi(x) dx = \int_{\mathbb{R}^N} \phi(x)\, dD^2 u.$$

- For $N = 1$: strict convergence as measures, i.e., additionally

$$|\mathcal{H}_{\rho_n} u \mathcal{L}^N|(\mathbb{R}) \to |D^2 u|(\mathbb{R})$$

Main results – Characterization of BV2

Theorem (Characterization of higher-order Sobolev and BV spaces)

Let $u \in L^p(\mathbb{R}^N)$ *for some* $1 < p < \infty$. *Then*

$$u \in W^{2,p}(\mathbb{R}^N) \iff \liminf_{n \to \infty} \int_{\mathbb{R}^N} |\mathcal{H}_{\rho_n} u(x)|^p dx < \infty.$$

Let now $u \in L^1(\mathbb{R}^N)$. *Then*

$$u \in \mathrm{BV}^2(\mathbb{R}^N) \iff \liminf_{n \to \infty} \int_{\mathbb{R}^N} |\mathcal{H}_{\rho_n} u(x)| dx < \infty.$$

Overview

Numerical Results

Applications

Explicit model

$$\mathcal{H}_\rho u(x) = C \int_{\mathbb{R}^N} \frac{u(x+z) - 2u(x) + u(x-z)}{\|z\|^2} \left(\frac{zz^\top}{\|z\|^2} - \frac{1}{N+2} I \right) \rho(z) dz$$

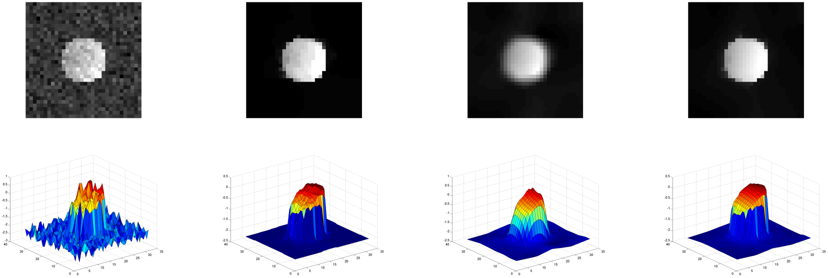
Implicit model

$$(\mathcal{G}'(x), \mathcal{H}'(x)) = \operatorname{argmin}_{\mathcal{G}', \mathcal{H}'} \; \frac{1}{2} \int_{\Omega - \Omega} R^2_{u, \mathcal{G}', \mathcal{H}'}(x, z) \, \sigma_x(z) dz$$

- Copes naturally with bounded/irregular domains, easier numerically
- How to choose the weights?

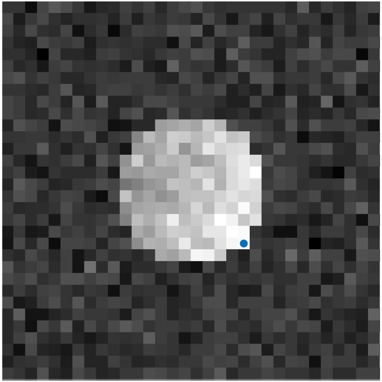
Application: higher-order structure elements

- Jumps and higher-order regularization clash

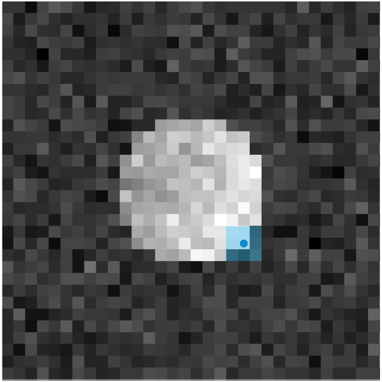


input | TV (staircase) | TV^2 (smooth) | TGV

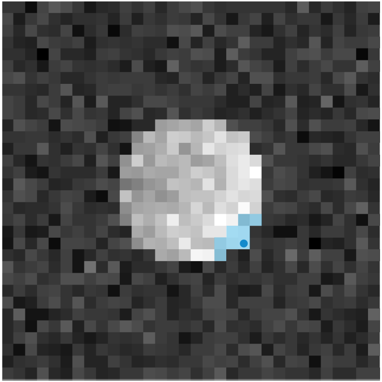
Locally adaptive discretization



Locally adaptive discretization



Locally adaptive discretization



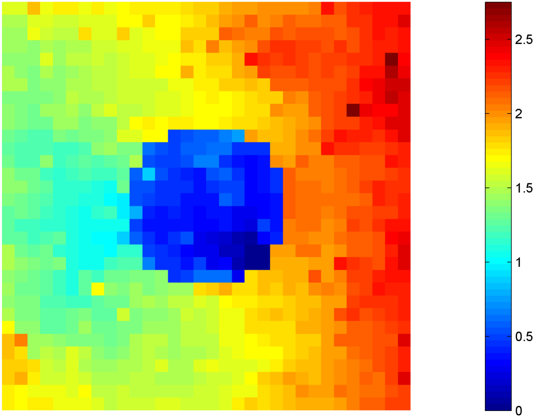
Application: higher-order structure elements

- Solve weighted Eikonal equation

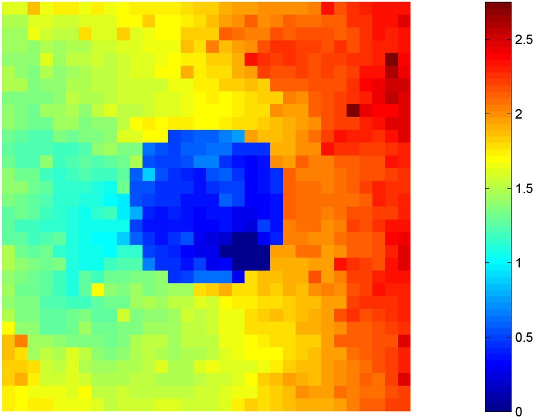
$$\|\nabla c_x(y)\| = \varphi(\nabla I(y)), \quad c_x(x) = 0,$$

  using Fast Marching Method.
- $\varphi(g) = \|g\|^2 + \alpha$ works, but others are possible
- $c_x(y)$ cost to move from x to y through the *gradient* image
  - high if x and y are separated by strong edge
  - can go around outliers/structural noise
- "Amoeba" filters [Lerallut, Decencière, Meyer'05/'07], Active Contours [Welk'11]

Application: higher-order structure elements



Application: higher-order structure elements



Application: higher-order structure elements

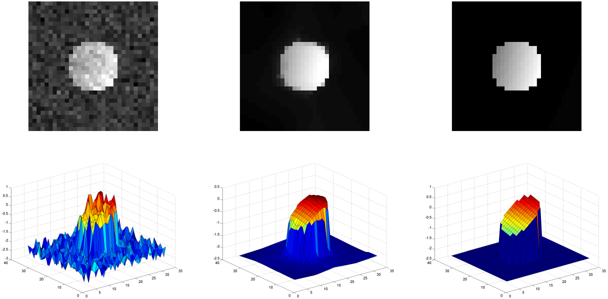
Implicit model

$$(\mathcal{G}'(x), \mathcal{H}'(x)) = \operatorname{argmin}_{\mathcal{G}',\mathcal{H}'} \ \frac{1}{2} \int_{\Omega-\Omega} R^2_{u,\mathcal{G}',\mathcal{H}'}(x,z)\, \sigma_x(z) dz$$

- Sort neighbours $y^1, y^2, \ldots$ so that $c_x(y^1) \leqslant c_x(y^2) \leqslant \ldots$, and use $M \in \mathbb{N}$ points y^i with shortest distance to x:

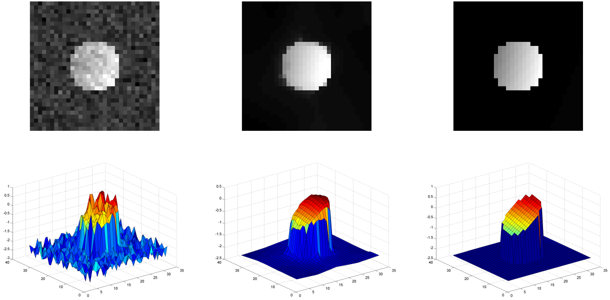
$$\sigma_x(y^i) = \begin{cases} 1/c_x^2(y^i), & i \leqslant M, \\ 0, & i > M. \end{cases}$$

Results

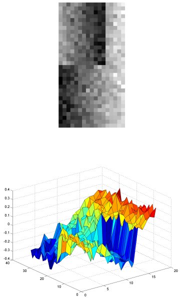


Why?

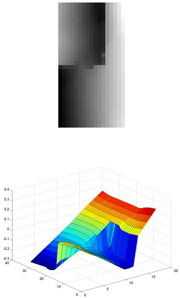
- regularisation weight can be set relatively large
- outliers are ignored by the construction of the distances
- we can have true piecewise affine regularization



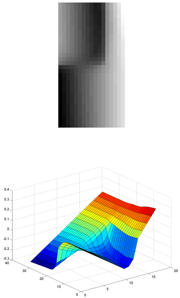
More examples



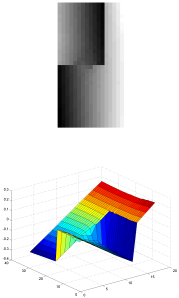
input



TGV-L^1 low



TGV-L^1 high



non-local L1

Adaptive iterative regularization

input

$n = 1$

$n = 2$

$n = 3$

$n = 4$

$n = 5$

More examples and iterating

More examples and iterating

A Non-Local Formulation for Higher-Order Total Variation-Based Regularization

Jan Lellmann

CIA/DAMTP, University of Cambridge

Joint work with: K. Papafitsoros, D. Spector, C. Schönlieb

Hong Kong, May 2014

Non-local integration by parts

- Non-local second-order divergence for functions $\phi : \mathbb{R}^N \to \mathbb{R}^{N\times N}$:

$$\mathcal{D}^2_\rho \phi(x) = C \int_{\mathbb{R}^N} \frac{\phi(x+z) - 2\phi(x) + \phi(x-z)}{\|z\|^2} \cdot \left(\frac{zz^\top}{\|z\|^2} - \frac{1}{N+2} I \right) \rho(z) dz. \tag{1}$$

Theorem (Second order non-local integration by parts)

Assume $u \in L^1(\mathbb{R}^N)$, $\frac{|d^2 u(x,y)|}{|x-y|^2}\rho_n(x-y) \in L^1(\mathbb{R}^N \times \mathbb{R}^N)$ *and let* $\phi \in C^2_c(\mathbb{R}^N, \mathbb{R}^{N\times N})$. *Then*

$$\int_{\mathbb{R}^N} \mathcal{H}_\rho u(x) \cdot \phi(x) dx = \int_{\mathbb{R}^N} u(x) \mathcal{D}^2_\rho \phi(x) dx.$$