Technische Universität Kaiserslautern

Inpainting of Cyclic Data Using First and Second Order Differences

Ronny Bergmann   Andreas Weinmann

Mathematical Image Processing and Data Analysis Group
Department of Mathematics
University of Kaiserslautern

January 15th, 2015

10th International Conference on
Energy Minimizing Methods in Computer Vision and Pattern Recognition

Hong Kong

Contents

Introduction

- employing the Rudin-Osher-Fatemi (ROF) functional [Osher, Rudin, Fatemi, 1992]

$$\sum_{i,j}(f_{i,j} - x_{i,j})^2 + \lambda \sum_{i,j} |\nabla x_{i,j}|$$

  - f noisy image
  - ∇ discrete gradient
  - $\sum_{i,j} |\nabla x_{i,j}|$ discrete total variation (TV)
  - regularization parameter $\lambda > 0$

$\Rightarrow$ edge-preserving

- several higher order variational models avoid stair caising-effect
[Chambolle, Lions, 1997; Setzer, Steidl, 2008; Bredies, Kunisch, Pock, 2010; Papafitsoros, Schönlieb, 2014]

Recently [Strekalovski, Cremers, 2013], [Lellmann et al., 2013], [Weinmann et al., 2013], [Bačák, 2013]

- TV denoising generalized to Riemannian manifolds
- several algorithms to find a minimizer x^*, e.g. PPA
- convergence for PPA on CAT(0) spaces (does not include $\mathbb{S}^1$)

Finite Differences on $\mathbb{R}$

Let $w = (w_j)_{j=1}^d \in \mathbb{R}^d \setminus \{0\}$ fulfill

$$\langle w, 1_d \rangle := \sum_{j=1}^{d} w_j = 0.$$

The finite difference operator is given by

$$\Delta(x; w) := \langle x, w \rangle, \quad x \in \mathbb{R}^d.$$

This talk: $w \in \{b_1, b_2, b_{1,1}\}$

- $b_1 = (-1, 1)$ for f_x, f_y
- $b_2 = (1, -2, 1)$ for f_{xx}, f_{yy}
- $b_{1,1} = (1, -1, -1, 1)$ for f_{xy}

Finite Differences on the Circle $\mathbb{S}^1$

Cyclic data

$$p_i \in \mathbb{S}^1 := \{q \in \mathbb{R}^2 \,:\, \|q\|_2 = 1\} \Longleftrightarrow x_i \in [-\pi, \pi)$$

Distances

- $d(x; b_1) := \arccos\langle p_1, p_2\rangle = |(x_2 - x_1)_{2\pi}| = |(\langle x, b_1\rangle)_{2\pi}|$ (arc length)

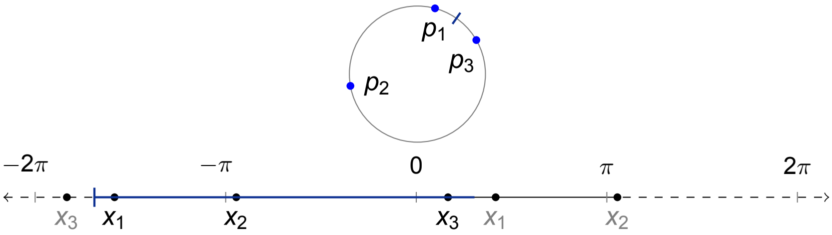
Finite Differences on the Circle $\mathbb{S}^1$

Cyclic data

$$p_i \in \mathbb{S}^1 := \{q \in \mathbb{R}^2 : \|q\|_2 = 1\} \Longleftrightarrow x_i \in [-\pi, \pi)$$

Distances

- $d(x; b_1) := \arccos\langle p_1, p_2\rangle = |(x_2 - x_1)_{2\pi}| = |(\langle x, b_1\rangle)_{2\pi}|$ (arc length)
- $d(x; b_2)$? Several unwrappings to compute $x_1 - 2x_2 + x_3$



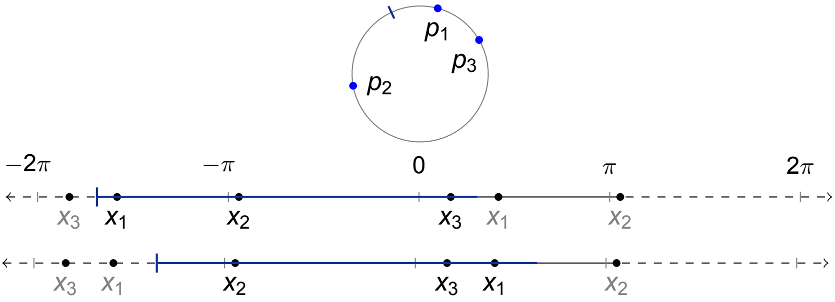
Finite Differences on the Circle $\mathbb{S}^1$

Cyclic data

$$p_i \in \mathbb{S}^1 := \{q \in \mathbb{R}^2 : \|q\|_2 = 1\} \Longleftrightarrow x_i \in [-\pi, \pi)$$

Distances

- $d(x; b_1) := \arccos\langle p_1, p_2\rangle = |(x_2 - x_1)_{2\pi}| = |(\langle x, b_1\rangle)_{2\pi}|$ (arc length)
- $d(x; b_2)$? Several unwrappings to compute $x_1 - 2x_2 + x_3$



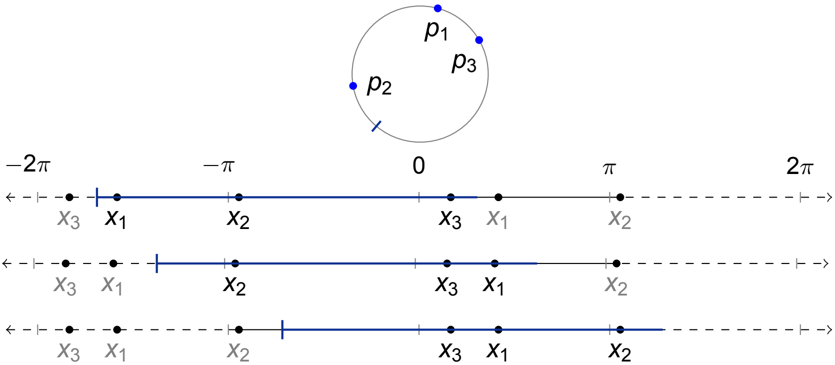
Finite Differences on the Circle $\mathbb{S}^1$

Cyclic data

$$p_i \in \mathbb{S}^1 := \{q \in \mathbb{R}^2 \,:\, \|q\|_2 = 1\} \iff x_i \in [-\pi, \pi)$$

Distances

- $d(x; b_1) := \arccos\langle p_1, p_2\rangle = |(x_2 - x_1)_{2\pi}| = |(\langle x, b_1\rangle)_{2\pi}|$ (arc length)
- $d(x; b_2)$? Several unwrappings to compute $x_1 - 2x_2 + x_3$



Absolute Finite Differences of Cyclic Data

The cyclic absolute finite difference w.r.t. w

$$d(x; w) := \min_{\mu\in\mathbb{R}} \left|\Delta\big([x + \mu 1_d]_{2\pi}; w\big)\right|, \quad x \in [-\pi, \pi)^d$$

- $[x]_{2\pi}$: element-wise mod 2π, except for $x_i = (2k + 1)\pi$: take both $\pm\pi$
- $d(x; w)$ is shift invariant
- notation: $d_1 := d(\cdot\,; b_1)$, $d_2 := d(\cdot\,; b_2)$ and $d_{1,1} := d(\cdot\,; b_{1,1})$
- minimization not necessary for $w \in \{b_1, b_2, b_{1,1}\}$: it holds

$$d(x; w) = \left|(\Delta(x, w))_{2\pi}\right|$$

this does not hold e.g. for $w = b_3 = (1, -3, 3, -1)$

Second Order TV for Cyclic Data

For given data $f = (f_i)_{i=1}^N \in (\mathbb{S}^1)^N$, $\alpha, \beta \geq 0$, compute

$$\operatorname*{arg\,min}_{x\in[-\pi,\pi)^N} J(x), \quad J(x) := \sum_{i=1}^{N} d_1(f_i, x_i)^2 + \alpha\,\mathrm{TV}_1(x) + \beta\,\mathrm{TV}_2(x)$$

where

$$\mathrm{TV}_1(x) = \sum_{i=1}^{N-1} d_1(x_i, x_{i+1}), \quad \mathrm{TV}_2(x) = \sum_{i=2}^{N-1} d_2(x_{i-1}, x_i, x_{i+1}).$$

Similar: 2D model for data $f_{i,j} \in (\mathbb{S}^1)^{N,M}$

- vertical and horizontal first and second order differences
- mixed second order difference on each 2×2 submatrix of $x \in [-\pi, \pi)^{N,M} \Rightarrow$ additional term $\gamma\,\mathrm{TV}_{1,1}(x)$

Models for Inpainting

- image domain: $\Omega_0 \subset \mathbb{N}^2$ with $\Omega \subset \Omega_0$ unknown
- data $f_{i,j}$, $(i,j) \in \overline{\Omega} = \Omega_0 \backslash \Omega$
- noiseless case

$$\operatorname*{arg\,min}_{x \in [-\pi,\pi)^{N,M}} \alpha \operatorname{TV}_1^\Omega(x) + \beta \operatorname{TV}_2^\Omega(x) + \gamma \operatorname{TV}_{1,1}^\Omega(x),$$
$$\text{subject to} \quad x_{i,j} = f_{i,j} \quad \text{for all} \quad (i,j) \in \overline{\Omega}$$

- noisy case

$$\operatorname*{arg\,min}_{x \in [-\pi,\pi)^{N,M}} \sum_{(i,j) \in \overline{\Omega}} d_1(f_{i,j}, x_{i,j})^2 + \alpha \operatorname{TV}_1(x) + \beta \operatorname{TV}_2(x) + \gamma \operatorname{TV}_{1,1}(x)$$

- iterative initialization: set $x_{i,j}$, $(i,j) \in \Omega$, by solving $d(x;w) = 0$, whenever all other data items are known [Almeida, Figueiredo, 2013]

$\Rightarrow$ "cold start initialization"

Proximal Mappings on $\mathbb{S}^1$

[Rockafellar, 1976; Ferreira, Oliveira, 2002]

$$\operatorname{prox}_{\lambda\varphi}(g) = \operatorname*{arg\,min}_{x\in[-\pi,\pi)^d} \frac{1}{2}\sum_{i=1}^{d} d_1(g_i, x_i)^2 + \lambda\varphi(x), \quad \lambda > 0$$

Theorem: Proximal Mapping I [B., Laus, Steidl, Weinmann, 2014]

The unique minimizer x^* of $\operatorname{prox}_{\lambda d_1(f,\,\cdot\,)^2}(g)$ is

$$x^* = \left(\frac{g+\lambda f}{1+\lambda} + \frac{\lambda}{1+\lambda} 2\pi\, v\right)_{2\pi}, \quad v = \begin{cases} 0 & \text{for } |g-f| \leq \pi, \\ \operatorname{sgn}(g-f) & \text{for } |g-f| > \pi. \end{cases}$$

Proximal Mappings on $\mathbb{S}^1$

[Rockafellar, 1976; Ferreira, Oliveira, 2002]

$$\operatorname{prox}_{\lambda\varphi}(g) = \operatorname*{arg\,min}_{x\in[-\pi,\pi)^d} \frac{1}{2}\sum_{i=1}^{d} d_1(g_i,x_i)^2 + \lambda\varphi(x), \quad \lambda > 0$$

Theorem: Proximal Mapping I [B., Laus, Steidl, Weinmann, 2014]

The unique minimizer x^* of $\operatorname{prox}_{\lambda d_1(f,\cdot)^2}(g)$ is

$$x^* = \left(\frac{g+\lambda f}{1+\lambda} + \frac{\lambda}{1+\lambda}2\pi v\right)_{2\pi}, \quad v = \begin{cases} 0 & \text{for } |g-f| \leq \pi, \\ \operatorname{sgn}(g-f) & \text{for } |g-f| > \pi. \end{cases}$$

Theorem: Proximal Mapping II [B., Laus, Steidl, Weinmann, 2014]

The minimizers of $\operatorname{prox}_{\lambda d(\cdot;w)}(g)$, $w \in \{b_1, b_2, b_{1,1}\}$, are given by

$$x^* = \left(g - \operatorname{sgn}([\langle g,w\rangle]_{2\pi}) \min\left\{\lambda, \frac{|(\langle g,w\rangle)_{2\pi}|}{\|w\|_2^2}\right\} w\right)_{2\pi}.$$

For $|(\langle g,w\rangle)_{2\pi}| = \pi$ there are two minimizers, otherwise it is unique.

The Cyclic Proximal Point Algorithm

Find $\arg\min_x \varphi(x)$, $\varphi\colon \mathbb{R}^N \to \mathbb{R}$, convex, proper, lsc, by Picard iteration:

[Moreau, 1965; Rockafellar, 1976]

$$x^{(k)} = \operatorname{prox}_{\lambda\varphi}(x^{(k-1)}), \quad k > 0$$

$\Rightarrow$ fast evaluation of $\operatorname{prox}_{\lambda\varphi}$ needed

For $\varphi = \sum_{i=1}^{c} \varphi_i$ use Cyclic Proximal Point Algorithm (CPPA) [Bertsekas, 2011]

$$x^{(k+\frac{i+1}{c})} = \operatorname{prox}_{\lambda_k\varphi_i}(x^{(k+\frac{i}{c})}), \quad i = 0,\ldots,c-1,\ k > 0$$

converges to a minimizer if $\{\lambda_k\} \in \ell_2(\mathbb{Z})\setminus\ell_1(\mathbb{Z})$.

For our model $J\colon (\mathbb{S}^1)^N \to \mathbb{R}$ we can prove convergence if additionally

- data f_i *locally* dense enough
- $\alpha, \beta(,\gamma)$ are sufficiently small

An Efficient Splitting for the CPPA

- data $\displaystyle \frac{1}{2}\sum_{i=1}^{N} d_1(f_i, x_i)^2 =: J_1(x)$

  proximal mapping I (simultaneously elementwise)

- first order differences

$$\alpha \,\mathrm{TV}_1(x) = \alpha \sum_{i=1}^{N-1} d_1(x_i, x_{i+1})$$

- second order differences

$$\beta \,\mathrm{TV}_2(x) = \beta \sum_{i=2}^{N-1} d_2(x_{i-1}, x_i, x_{i+1})$$

An Efficient Splitting for the CPPA

- data $\frac{1}{2}\sum_{i=1}^{N} d_1(f_i, x_i)^2 =: J_1(x)$

  proximal mapping I (simultaneously elementwise)

- first order differences

$$\alpha\,\mathrm{TV}_1(x) = \sum_{l=0}^{1} \alpha \sum_{i=1}^{\lfloor \frac{N-1}{2} \rfloor} d_1(x_{2i-1+l}, x_{2i-l}) =: \sum_{l=0}^{1} J_{2+l}(x)$$

- second order differences

$$\beta\,\mathrm{TV}_2(x) = \beta \sum_{i=2}^{N-1} d_2(x_{i-1}, x_i, x_{i+1})$$

An Efficient Splitting for the CPPA

- data $\displaystyle\frac{1}{2}\sum_{i=1}^{N} d_1(f_i, x_i)^2 =: J_1(x)$

  proximal mapping I (simultaneously elementwise)
- first order differences

$$\alpha\,\mathrm{TV}_1(x) = \sum_{l=0}^{1} \alpha \sum_{i=1}^{\lfloor\frac{N-1}{2}\rfloor} d_1(x_{2i-1+l}, x_{2i-l}) =: \sum_{l=0}^{1} J_{2+l}(x)$$

  inner sum: distinct data $\Rightarrow$ proximal mapping II with $w = b_1$
- second order differences

$$\beta\,\mathrm{TV}_2(x) = \beta \sum_{i=2}^{N-1} d_2(x_{i-1}, x_i, x_{i+1})$$

An Efficient Splitting for the CPPA

- data $\displaystyle\frac{1}{2}\sum_{i=1}^{N} d_1(f_i, x_i)^2 =: J_1(x)$

  proximal mapping I (simultaneously elementwise)

- first order differences

$$\alpha\,\mathrm{TV}_1(x) = \sum_{l=0}^{1}\alpha\sum_{i=1}^{\lfloor\frac{N-1}{2}\rfloor} d_1(x_{2i-1+l}, x_{2i-l}) =: \sum_{l=0}^{1} J_{2+l}(x)$$

  inner sum: distinct data $\Rightarrow$ proximal mapping II with $w = b_1$

- second order differences

$$\beta\,\mathrm{TV}_2(x) = \sum_{l=0}^{2}\beta\sum_{i=1}^{\lfloor\frac{N-1}{3}\rfloor} d_2(x_{3i-2+l}, x_{3i-1+l}, x_{3i+l}) =: \sum_{l=0}^{2} J_{4+l}(x)$$

  inner sum: distinct data $\Rightarrow$ proximal mapping II with $w = b_2$

An Efficient Splitting for the CPPA

- data $\frac{1}{2}\sum_{i=1}^{N} d_1(f_i, x_i)^2 =: J_1(x)$

  proximal mapping I (simultaneously elementwise)
- first order differences

$$\alpha\,\mathrm{TV}_1(x) = \sum_{l=0}^{1} \alpha \sum_{i=1}^{\lfloor \frac{N-1}{2} \rfloor} d_1(x_{2i-1+l}, x_{2i-l}) =: \sum_{l=0}^{1} J_{2+l}(x)$$

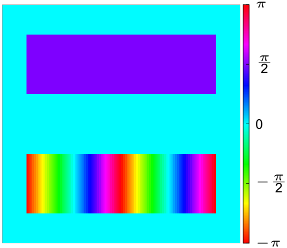
  inner sum: distinct data $\Rightarrow$ proximal mapping II with $w = b_1$
- second order differences

$$\beta\,\mathrm{TV}_2(x) = \sum_{l=0}^{2} \beta \sum_{i=1}^{\lfloor \frac{N-1}{3} \rfloor} d_2(x_{3i-2+l}, x_{3i-1+l}, x_{3i+l}) =: \sum_{l=0}^{2} J_{4+l}(x)$$

  inner sum: distinct data $\Rightarrow$ proximal mapping II with $w = b_2$

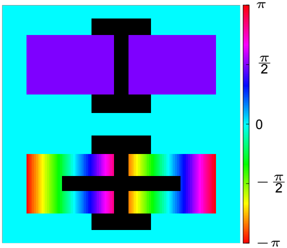
$\Rightarrow$ $J(x) = \sum_{l=1}^{6} J_l(x)$ $\Rightarrow$ cycle length $c = 6$ (2D: $c = 15$)

Example: Inpainting



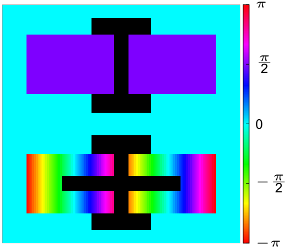
Original image.

Example: Inpainting

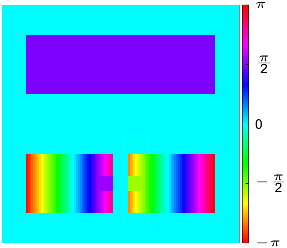


Original image (lost area in black).

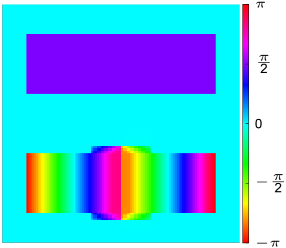
Example: Inpainting



Original image (lost area in black).

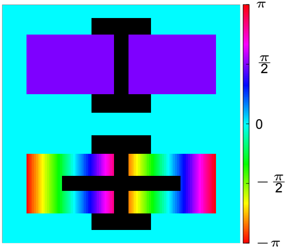


First order real-valued model.

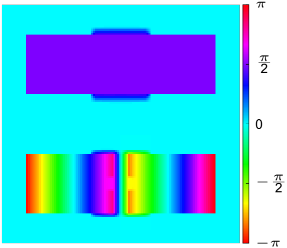


Our model, $\alpha = 2$, $\beta = \gamma = 0$.

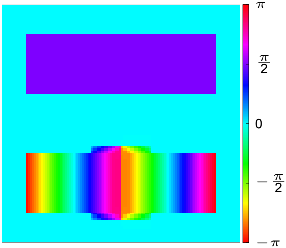
Example: Inpainting



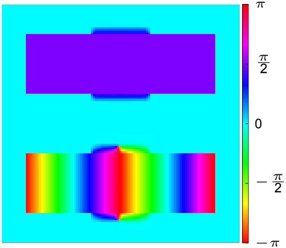
Original image (lost area in black).



First & second order real-valued model.

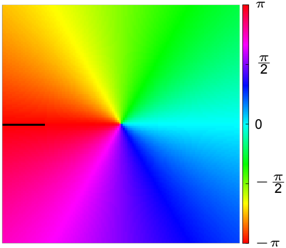


Our model, $\alpha = 2$, $\beta = \gamma = 0$.

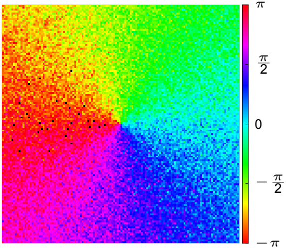


Our model, $\alpha = 2$, $\beta = \gamma = 1$.

Example: Combined Inpainting and Denoising

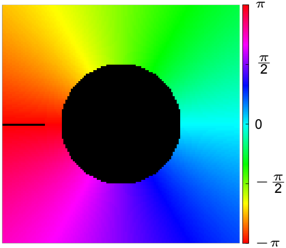


Original image.

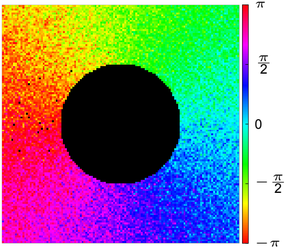


Noisy image, $\sigma = 0.3$.

Example: Combined Inpainting and Denoising

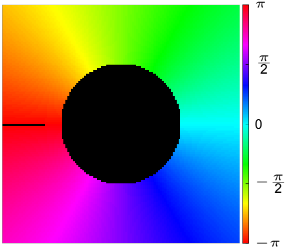


Original image (lost area in black).

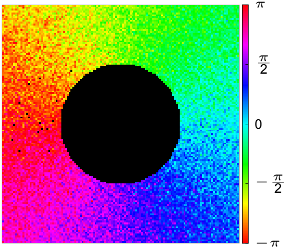


Noisy image, $\sigma = 0.3$.

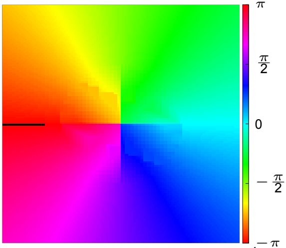
Example: Combined Inpainting and Denoising



Original image (lost area in black).

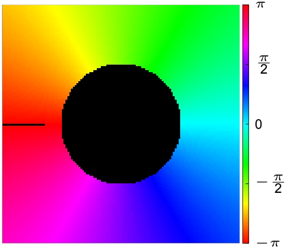


Noisy image, $\sigma = 0.3$.

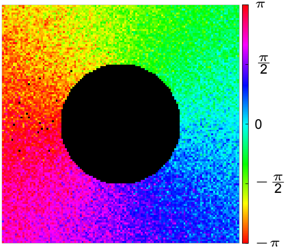


Noiseless model, $\alpha = \frac{1}{8}$.

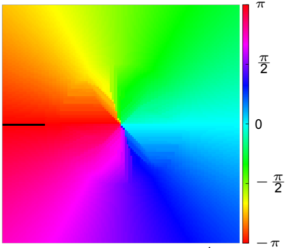
Example: Combined Inpainting and Denoising



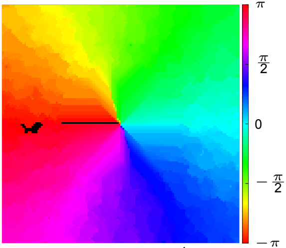
Original image (lost area in black).



Noisy image, $\sigma = 0.3$.

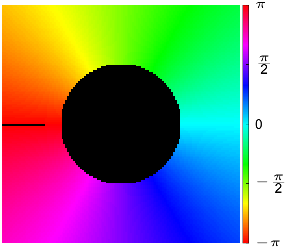


Noiseless model, $\alpha = \frac{1}{8}$ (& diags).

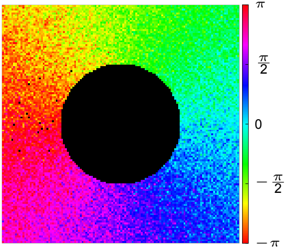


Noisy model, $\alpha = \frac{1}{8}$ (& diags).

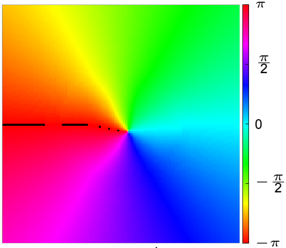
Example: Combined Inpainting and Denoising



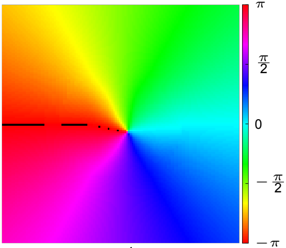
Original image (lost area in black).



Noisy image, $\sigma = 0.3$.



Noiseless model, $\alpha = \frac{1}{8}$ (& diags), $\beta = \gamma = \frac{1}{4}$.



Noisy model, $\alpha = \frac{1}{8}$ (& diags), $\beta = \gamma = \frac{1}{4}$.

Conclusion

We have

- defined a model for second order differences of cyclic data
- derived a first and second order TV type functional
- extended the model to inpainting
- employed CPPA in order to minimize the functional
- simultaneous inpainting and denoising

Future work

- combined cyclic and linear vector space data (submitted)
- extension to other manifolds

Literature

[1] M. Bačák. Computing medians and means in Hadamard spaces. *SIAM J. Optim.*, 2014.

[2] R. Bergmann, F. Laus, G. Steidl, and A. Weinmann. Second order differences of cyclic data and applications in variational denoising. *SIAM J. Imaging Sci.*, 2014.

[3] R. Bergmann and A. Weinmann. *A second order TV-type approach for inpainting and denoising higher dimensional combined cyclic and vector space data*. *Preprint, ArXiv, submitted*, 2015.

[4] E. Strekalovskiy and D. Cremers. Total cyclic variation and generalizations. *J. Math. Imaging Vis.*, 2013.

[5] A. Weinmann, L. Demaret, and M. Storath. Total variation regularization for manifold-valued data. *SIAM J. Imaging Sci.*, 2014.

Literature

[1] M. Bačák. Computing medians and means in Hadamard spaces. *SIAM J. Optim.*, 2014.

[2] R. Bergmann, F. Laus, G. Steidl, and A. Weinmann. Second order differences of cyclic data and applications in variational denoising. *SIAM J. Imaging Sci.*, 2014.

[3] R. Bergmann and A. Weinmann. *A second order TV-type approach for inpainting and denoising higher dimensional combined cyclic and vector space data*. *Preprint, ArXiv, submitted*, 2015.

[4] E. Strekalovskiy and D. Cremers. Total cyclic variation and generalizations. *J. Math. Imaging Vis.*, 2013.

[5] A. Weinmann, L. Demaret, and M. Storath. Total variation regularization for manifold-valued data. *SIAM J. Imaging Sci.*, 2014.

Thank you for your attention.