



Space-variant directional regularisation for image restoration problems

Luca Calatroni

CMAP, École Polytechnique CNRS

joint work with: M. Pragliola, A. Lanza, F. Sgallari (University of Bologna)

Computational Imaging workshop, Semester program in Computer Vision

ICERM

March 18-22 2019

Providence, USA

Introduction

Imaging inverse problems

Inverse problem formulation

Given f , seek u such that

$$f = Ku + n$$

where K (known) models blur and n noise in the data.

Task: compute reconstruction u^* . Different viewpoints...

Imaging inverse problems

Inverse problem formulation

Given f , seek u such that

$$f = Ku + n$$

where K (known) models blur and n noise in the data.

Task: compute reconstruction u^* . Different viewpoints...

Statistics

Maximise posterior:

$$u^* \in \operatorname{argmax}_u P(u; f, K)$$

Use Bayes' formula

$$u^* \in \operatorname{argmax}_u \frac{P(u)P(f, K; u)}{P(f)}$$

Optimisation

MAP estimation

$$u^* \in \operatorname{argmin}_u -\log (P(u)P(f, K; u))$$

Variational regularisation:

$$u^* \in \operatorname{argmin}_u R(u) + \mu \Phi(Ku; f)$$

Discrete/continuous framework.

PDEs

Evolution problem

$$\begin{cases} u_t = -\nabla R(u) - \mu \nabla \Phi(Ku; f) \\ u(0) = f \\ \text{b.c.} \end{cases}$$

Get u^* as the steady state
of the system above

Analogy:

- $P(u)$, $R(u)$ encode *prior* assumptions on u (**topic of this talk!**)
- $P(f, K; u)$, $\Phi(f, K; u)$ describe noise statistics

Imaging inverse problems

Inverse problem formulation

Given f , seek u such that

$$f = Ku + n$$

where K (known) models blur and n noise in the data.

Task: compute reconstruction u^* . Different viewpoints...

Statistics

Maximise posterior:

$$u^* \in \operatorname{argmax}_u P(u; f, K)$$

Use Bayes' formula

$$u^* \in \operatorname{argmax}_u \frac{P(u)P(f, K; u)}{P(f)}$$

Optimisation

MAP estimation

$$u^* \in \operatorname{argmin}_u -\log (P(u)P(f, K; u))$$

Variational regularisation:

$$u^* \in \operatorname{argmin}_u R(u) + \mu \Phi(Ku; f)$$

Discrete/continuous framework.

Tailored image regulariser?

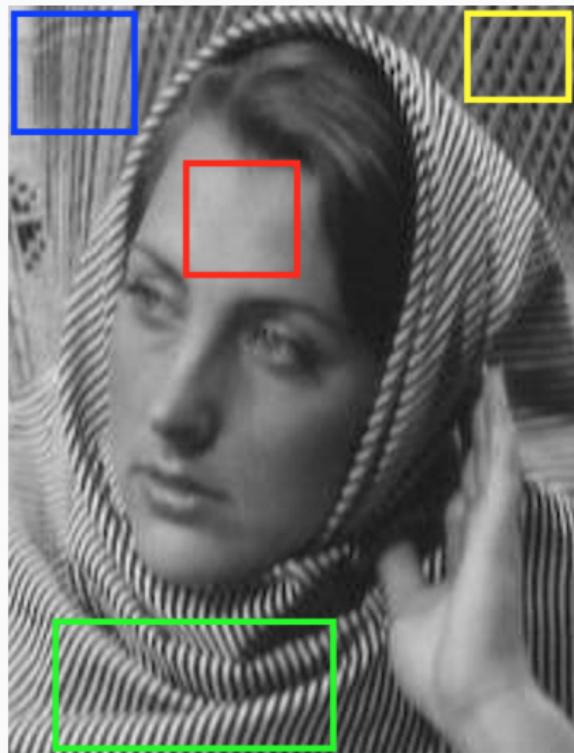


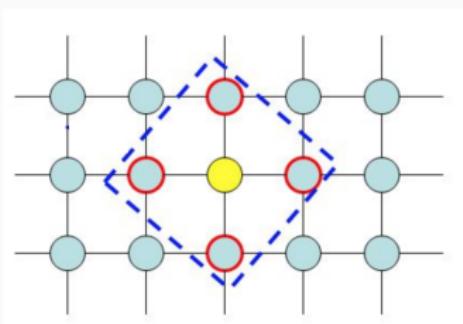
Image model adapted to local features?

Statistical viewpoint: Markov Random Fields

Standard image priors are based on *stationary* Markov Random Field modelling

$$P(u) = \frac{1}{Z} \prod_{i=1}^n \exp(-\alpha V_{\mathcal{N}_i}(u)) = \frac{1}{Z} \exp\left(-\alpha \sum_{i=1}^n V_{\mathcal{N}_i}(u)\right),$$

where $\alpha > 0$, \mathcal{N}_i is the *clique* around i and $V_{\mathcal{N}_i}$ is the Gibbs' potential in \mathcal{N}_i .



Statistical viewpoint: Markov Random Fields

Standard image priors are based on *stationary* Markov Random Field modelling

$$P(u) = \frac{1}{Z} \exp \left(-\alpha \textcolor{orange}{TV}(u) \right) = \frac{1}{Z} \exp \left(-\alpha \sum_{i=1}^n \|(\nabla u)_i\|_2 \right),$$

where $\alpha > 0$ and $V_{\mathcal{N}_i}(u) = \|(\nabla u)_i\|_2$.

Statistical viewpoint: Markov Random Fields

Standard image priors are based on *stationary* Markov Random Field modelling

$$P(u) = \frac{1}{Z} \exp \left(-\alpha \textcolor{brown}{TV}(u) \right) = \frac{1}{Z} \exp \left(-\alpha \sum_{i=1}^n \|(\nabla u)_i\|_2 \right),$$

where $\alpha > 0$ and $V_{\mathcal{N}_i}(u) = \|(\nabla u)_i\|_2$.

Interpretation

$q := \|\nabla u\|_2$ is distributed **locally** as a α -half-Laplacian distribution:

$$P(q_i; \alpha) = \begin{cases} \alpha \exp(-\alpha q_i) & \text{for } q_i \geq 0, \\ 0 & \text{for } q_i < 0, \end{cases}$$

and α describes image scales.

One-parameter family describing **all** pixel features... too restrictive?

Variational and PDE approach: TV regularisation

$$TV(u) = \sum_{i=1}^n \|(\nabla u)_i\|_2$$

¹Rudin, Osher, Fatemi, '92

Variational and PDE approach: TV regularisation

$$TV(u) = \sum_{i=1}^n \|(\nabla u)_i\|_2$$

Variational model

$$\mathcal{E}(u) = TV(u) + \frac{\lambda}{2} \|u - f\|_2^2$$

PDE model (non-linear diffusion)

$$u_t = p + \lambda(f - u), \quad p \in \partial TV(u)$$

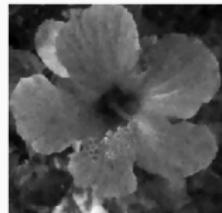
Seek u fitting the data with low TV: edges preservation & noise removal¹!



(a) f



(b) Tikhonov



(c) TV

Edge-preserving non-linear diffusion PDE...

¹Rudin, Osher, Fatemi, '92

Introduction

Describing anisotropy

Modelling anisotropy

Anisotropy operators

For all $x \in \Omega$, let $\lambda : \Omega \rightarrow \mathbb{R}^2$ be a positive vector field and let $\theta : \Omega \rightarrow [0, 2\pi)$ describe local orientation.

$$\Lambda(x) := \begin{pmatrix} \lambda_1(x) & 0 \\ 0 & \lambda_2(x) \end{pmatrix}, \quad R_\theta(x) := \begin{pmatrix} \cos \theta(x) & \sin \theta(x) \\ -\sin \theta(x) & \cos \theta(x) \end{pmatrix}.$$

Modelling anisotropy

Anisotropy operators

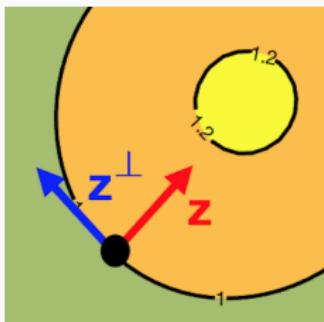
For all $x \in \Omega$, let $\lambda : \Omega \rightarrow \mathbb{R}^2$ be a positive vector field and let $\theta : \Omega \rightarrow [0, 2\pi)$ describe local orientation.

$$\Lambda(x) := \begin{pmatrix} \lambda_1(x) & 0 \\ 0 & \lambda_2(x) \end{pmatrix}, \quad R_\theta(x) := \begin{pmatrix} \cos \theta(x) & \sin \theta(x) \\ -\sin \theta(x) & \cos \theta(x) \end{pmatrix}.$$

$M_{\lambda, \theta} := \Lambda R_\theta$ is the *anisotropic metric* and $W_{\lambda, \theta} := M_{\lambda, \theta}^\top M_{\lambda, \theta}$.

Note: if $z(x) := (\cos \theta(x), \sin \theta(x))$, then:

$$M_{\lambda, \theta} \nabla u(x) = \begin{pmatrix} \lambda_1(x) \partial_{z(x)} u(x) \\ \lambda_2(x) \partial_{z(x)^\perp} u(x) \end{pmatrix} \Rightarrow \text{locally weighted gradient along } z$$



Modelling anisotropy

Anisotropy operators

For all $x \in \Omega$, let $\lambda : \Omega \rightarrow \mathbb{R}^2$ be a positive vector field and let $\theta : \Omega \rightarrow [0, 2\pi)$ describe local orientation.

$$\Lambda(x) := \begin{pmatrix} \lambda_1(x) & 0 \\ 0 & \lambda_2(x) \end{pmatrix}, \quad R_\theta(x) := \begin{pmatrix} \cos \theta(x) & \sin \theta(x) \\ -\sin \theta(x) & \cos \theta(x) \end{pmatrix}.$$

$M_{\lambda, \theta} := \Lambda R_\theta$ is the *anisotropic metric* and $W_{\lambda, \theta} := M_{\lambda, \theta}^\top M_{\lambda, \theta}$.

Note: if $z(x) := (\cos \theta(x), \sin \theta(x))$, then:

$$M_{\lambda, \theta} \nabla u(x) = \begin{pmatrix} \lambda_1(x) \partial_{z(x)} u(x) \\ \lambda_2(x) \partial_{z(x)^\perp} u(x) \end{pmatrix} \Rightarrow \text{locally weighted gradient along } z$$

Using this formalism:

- Anisotropic diffusion:

$$u_t = \operatorname{div}(W_{\lambda, \theta} \nabla u)$$

- Corresponding **directional energy**:

$$E_{\lambda, \theta}(u) := \int_{\Omega} |M_{\lambda, \theta} \nabla u|^2 dx$$

Directional regularisation for imaging: previous work

- **Statistics of natural images:** Generalised Gaussian PDFs ([Mallat, '89](#)), Laplace ([Green, '92](#)), non-stationary MRFs & optimisation ([Lanza, Morigi, Pragliola, Sgallari, '16, '18](#)), ...;
- **Directional variational regularisation:** Variable exponent ([Chen, Levine, Rao, '06](#)), DT(G)V ([Bayram, '12, Kongskov, Dong, Knudsen, '17, Parisotto, Schönlieb, Masnou, '18](#)), ...
- **Application to inverse problems:** Limited Angle Tomography ([Tovey, Benning et al., '19](#)), ...
- **PDEs:** structure tensor modelling ([Weickert, '98](#)).

$$\begin{cases} u_t = \operatorname{div}(\mathbf{D}(J_\rho(\nabla u_\sigma)))\nabla u & \text{in } \Omega \times (0, T] \\ \langle \mathbf{D}(J_\rho(\nabla u_\sigma))\nabla u, \mathbf{n} \rangle = 0 & \text{on } \partial\Omega \\ u(0, x) = f(x) & \text{in } \Omega, \end{cases}$$

where for convolution kernels K_ρ, K_σ

$$J_\rho(\nabla u_\sigma) := K_\rho * (\nabla u_\sigma \otimes \nabla u_\sigma), \quad u_\sigma := K_\sigma * u.$$

and \mathbf{D} is a smooth and symmetric diffusion tensor.

- Consistent numerical schemes ([Fehrenbach, Mirebeau, '14](#), ...).

Directional regularisation for imaging: previous work

- **Statistics of natural images:** Generalised Gaussian PDFs ([Mallat, '89](#)), Laplace ([Green, '92](#)), non-stationary MRFs & optimisation ([Lanza, Morigi, Pragliola, Sgallari, '16, '18](#)), ...;
- **Directional variational regularisation:** Variable exponent ([Chen, Levine, Rao, '06](#)), DT(G)V ([Bayram, '12, Kongskov, Dong, Knudsen, '17, Parisotto, Schönlieb, Masnou, '18](#)), ...
- **Application to inverse problems:** Limited Angle Tomography ([Tovey, Benning et al., '19](#)), ...
- **PDEs:** structure tensor modelling ([Weickert, '98](#)).

$$\begin{cases} u_t = \operatorname{div}(\mathbf{D}(J_\rho(\nabla u_\sigma)))\nabla u & \text{in } \Omega \times (0, T] \\ \langle \mathbf{D}(J_\rho(\nabla u_\sigma))\nabla u, \mathbf{n} \rangle = 0 & \text{on } \partial\Omega \\ u(0, x) = f(x) & \text{in } \Omega, \end{cases}$$

where for convolution kernels K_ρ, K_σ

$$J_\rho(\nabla u_\sigma) := K_\rho * (\nabla u_\sigma \otimes \nabla u_\sigma), \quad u_\sigma := K_\sigma * u.$$

and \mathbf{D} is a smooth and symmetric diffusion tensor.

- Consistent numerical schemes ([Fehrenbach, Mirebeau, '14, ...](#)).

Common problem

Tailored regularisation adapted to local image orientation/structures?

Introduction

Variational space-variant regularisation

Space-variant and directional regularisation models

Discrete formulation.

Let Ω be the image domain with $|\Omega| = n$.

$$\text{TV}(u) = \sum_{i=1}^n \|(\nabla u)_i\|_2$$

Space-variant and directional regularisation models

Discrete formulation.

Let Ω be the image domain with $|\Omega| = n$.

$$\text{TV}_p(u) = \sum_{i=1}^n \|(\nabla u)_i\|_2^p,$$

Enforcing sparsity:

- $p = 2$: Tikhonov regularisation ([Tikhonov, Arsenin, '77](#)).
- $p = 1$: Total Variation ([Rudin, Osher, Fatemi, '92](#)).
- $0 < p < 1$: Non-convex regularisation ([Hintermüller, Wu, Valkonen, '13, '15, Nikolova, Ng, Tam, '10](#)).

Space-variant and directional regularisation models

Discrete formulation.

Let Ω be the image domain with $|\Omega| = n$.

$$\text{TV}_p^{\text{sv}}(u) = \sum_{i=1}^n \|(\nabla u)_i\|_2^{p_i}$$

Space-variant modelling:

- $1 \leq p_i \leq 2$: Convex, space-variant regularisation (Blomgren, Chan, Mulet, Wong, '97, Chen, Levine, Rao, '06).
- $p_i \in (0, 2]$: Non-convex space-variant regularisation (Lanza, Morigi, Pragliola, Sgallari, '16, '18).

Space-variant and directional regularisation models

Discrete formulation.

Let Ω be the image domain with $|\Omega| = n$.

$$\text{DTV}_p(u) = \sum_{i=1}^n \|M_{\lambda_i, \theta_i}(\nabla u)_i\|_2^p, \quad \theta_i \in [0, 2\pi)$$

where $M_{\lambda_i, \theta_i} = \Lambda_i R_{\theta_i}$ as before.

Directional modelling:

- $p = 2$: Anisotropic diffusion ([Weickert, '98](#)).
- $p = 1$: Directional Total (Generalised) Variation for dominant direction $\theta_i \equiv \bar{\theta}$ ([Kongskov, Dong, Knudsen, '17](#)) and inverse problems ([Tovey, Benning et al., '19](#)).

Combine (possibly non-convex) space-variance
AND
directional modelling?

A flexible directional & space-variant regularisation

A flexible directional & space-variant regularisation

Statistical motivation

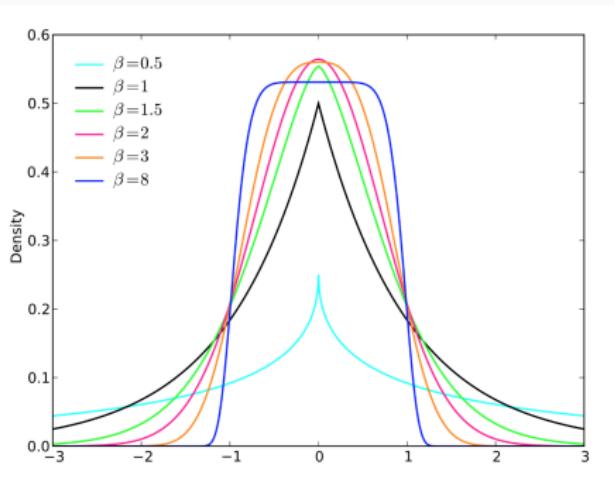
Bivariate Generalised Gaussian Distribution (BGGD) prior

Idea: model locally the joint distribution of $(\nabla u)_i$ in a more flexible way:

$$P((\nabla u)_i; p_i, \Sigma_i) = \frac{1}{2\pi|\Sigma_i|^{1/2}} \frac{p_i}{\Gamma(2/p_i) 2^{2/p_i}} \exp\left(-\frac{1}{2}((\nabla u)_i^T \Sigma_i^{-1} (\nabla u)_i)^{p_i/2}\right),$$

where:

- Γ is the Gamma function in \mathbb{R} ;
- Σ_i are gradient covariance matrices.



One-dimensional GGD with *shape parameter* $\beta = 2/p$.

Bivariate Generalised Gaussian Distribution (BGGD) prior

Idea: model locally the joint distribution of $(\nabla u)_i$ in a more flexible way:

$$P((\nabla u)_i; p_i, \Sigma_i) = \frac{1}{2\pi|\Sigma_i|^{1/2}} \frac{p_i}{\Gamma(2/p_i) 2^{2/p_i}} \exp\left(-\frac{1}{2}((\nabla u)_i^T \Sigma_i^{-1} (\nabla u)_i)^{p_i/2}\right),$$

where:

- Γ is the Gamma function in \mathbb{R} ;
- Σ_i are gradient covariance matrices.

Gaussian case

$p_i \equiv 2$: standard bivariate Gaussian distribution with pixel-wise covariance matrix Σ_i .

Image prior $P(u) = \prod_{i=1}^n P((\nabla u)_i; p_i, \Sigma_i).$

Via MAP estimate, derive the **variational space-variant, directional regulariser...**

A new space-variant, directional TV regulariser

By defining \mathbf{R}_{θ_i} and Λ_i from $\Sigma_i = \mathbf{R}_{\theta_i}^T \Lambda_i^2 \mathbf{R}_{\theta_i}$, find :

$$\text{DTV}_p^{\text{sv}}(u) := \sum_{i=1}^n \|\Lambda_i \mathbf{R}_{\theta_i}(\nabla u)_i\|_2^{p_i}, \quad p_i \in (0, 2], \quad \theta_i \in [0, 2\pi)$$

DTV_p^{sv}-L₂ image restoration model (LC, Lanza, Pragliola, Sgallari, '18)

We aim to solve

$$\min_u \left\{ \text{DTV}_p^{\text{sv}}(u) + \frac{\mu}{2} \|Ku - f\|_2^2 \right\}, \quad \mu > 0,$$

for Gaussian image reconstruction.

Highly **flexible**, more degrees of freedom to describe
natural images!

**A flexible directional & space-variant
regularisation**

Automated parameter estimation

ML approach for parameter estimation

For any pixel i , Σ_i is s.p.d.:

$$\Sigma_i = \begin{bmatrix} \sigma_1 & \sigma_3 \\ \sigma_3 & \sigma_2 \end{bmatrix} \quad \text{with} \quad \begin{cases} \sigma_1 > 0 \\ |\Sigma| = \sigma_1\sigma_2 - \sigma_3^2 > 0 \end{cases}$$

Four-parameter per-pixel: $\sigma_1, \sigma_2, \sigma_3$ and p .

²see Sharifi, Leon-Garcia, '95, Song, '06, Pascal, Bombrun, Tourneret, Berthoumieu, '13

ML approach for parameter estimation

For any pixel i , Σ_i is s.p.d.:

$$\Sigma_i = \begin{bmatrix} \sigma_1 & \sigma_3 \\ \sigma_3 & \sigma_2 \end{bmatrix} \quad \text{with} \quad \begin{cases} \sigma_1 > 0 \\ |\Sigma| = \sigma_1\sigma_2 - \sigma_3^2 > 0 \end{cases}$$

Four-parameter per-pixel: $\sigma_1, \sigma_2, \sigma_3$ and p .

Maximum likelihood² approach from collection of N samples around i : reformulation as a constrained problem using polar coordinates (ϱ, ϕ) in the plane $\sigma_1 - \sigma_3$ for Σ_i .

$$(p^*, \phi^*, \varrho^*) \in \operatorname{argmin}_{\bar{\mathcal{C}}} \left(\mathcal{F}(p, \phi, \varrho) := N \log \left(\Gamma \left(\frac{2}{p} + 1 \right) \frac{\pi}{\sqrt{1-\varrho^2}} \left(\frac{p}{2N} \right)^{2/p} \right) + \frac{2N}{p} \log \frac{p}{4N} + \frac{2N}{p} \log \left[\sum_{j=1}^N ((1 + \varrho \cos \phi)(\nabla u)_{j,1}^2 + (1 - \varrho \cos \phi)(\nabla u)_{j,2}^2 - 2\varrho \sin \phi (\nabla u)_{j,1}(\nabla u)_{j,2})^{p/2} \right] \right).$$

where $\bar{\mathcal{C}} := \{(p, \phi, \varrho) : p \in [\bar{\varepsilon}, \bar{p}], \phi \in [0, 2\pi], \varrho \in [0, 1 - \epsilon]\}$, after pre-processing.

²see Sharifi, Leon-Garcia, '95, Song, '06, Pascal, Bombrun, Tourneret, Berthoumieu, '13

ML parameter selection results

TEST: single BGGD with fixed parameters (p, ϕ, ϱ) .

- **Unbiased estimator** + empirical variance and RMSE going to 0 as $N \rightarrow +\infty$.

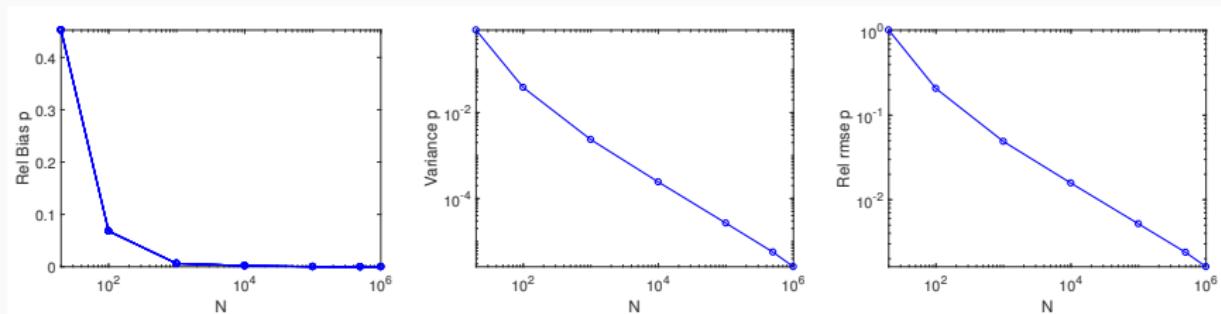


Figure 2: Relative bias, empirical variance and RMSE for estimated p^* . Comparison with synthetic BGGD. Analogous plots for ϕ^* and ϱ^* .

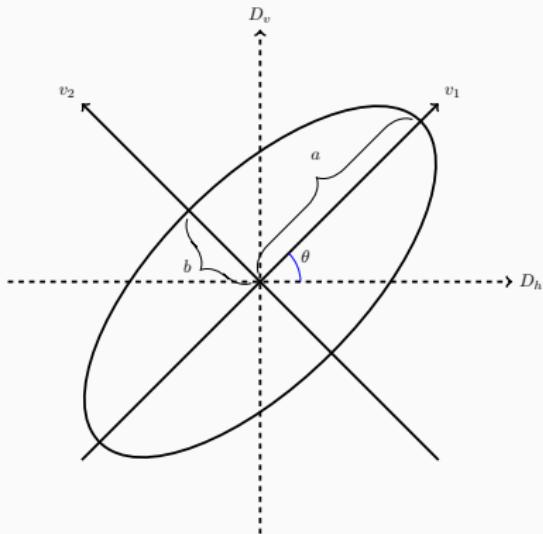
ML parameter selection results

TEST: single BGGD with fixed parameters (p, ϕ, ϱ) .

- **Unbiased estimator** + empirical variance and RMSE going to 0 as $N \rightarrow +\infty$.

From $(p_i^*, \phi_i^*, \varrho_i^*)$, get θ_i , eigenvalues $\{e_1, e_2\}_i$ & eigenvectors $\{v_1, v_2\}_i$ of Σ_i .

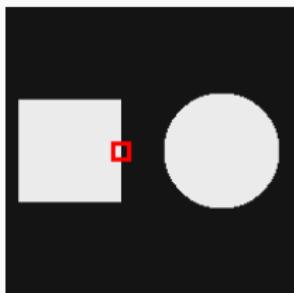
$$e_1 = 1 + \varrho =: a^2, \quad e_2 = 1 - \varrho =: b^2$$



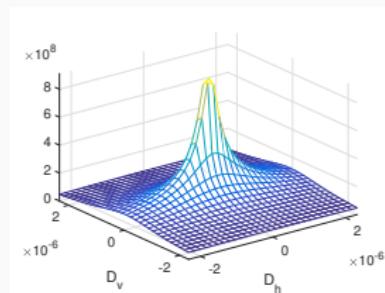
Anisotropy BGGD ellipses in the plane $D_h - D_v$.

ML parameter selection results

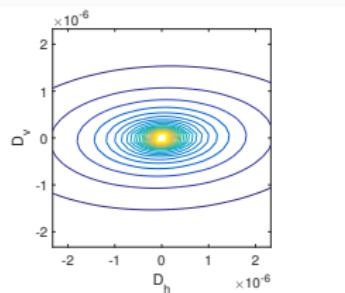
- **Unbiased estimator** + empirical variance and RMSE going to 0 as $N \rightarrow +\infty$.
- Functional form of the gradient PDF adapts to local image structures.



(a) Image



(b) Estimated PDF

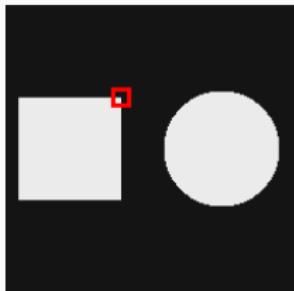


(c) Level lines

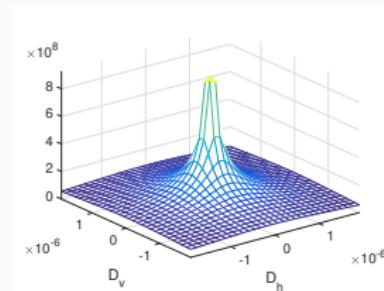
Test for **edge pixel**. Neighbourhood size 11×11 . Estimated $p^* = 0.07$, $\theta = -177.82^\circ$.

ML parameter selection results

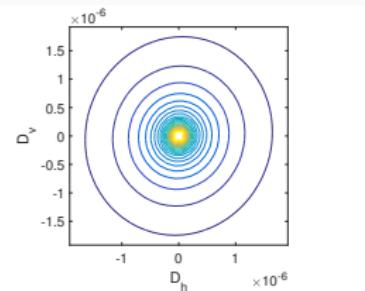
- **Unbiased estimator** + empirical variance and RMSE going to 0 as $N \rightarrow +\infty$.
- Functional form of the gradient PDF adapts to local image structures.



(a) Image



(b) Estimated PDF

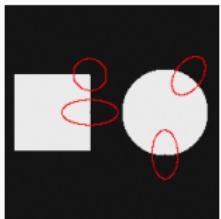


(c) Level lines

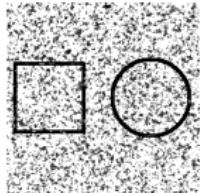
Test for **corner**. Neighbourhood size 11×11 . Estimated $p^* = 0.07$, $\theta = -72.49^\circ$.

ML parameter selection results

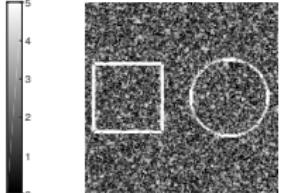
- **Unbiased estimator** + empirical variance and RMSE going to 0 as $N \rightarrow +\infty$.
- Functional form of the gradient PDF adapts to local image structures.



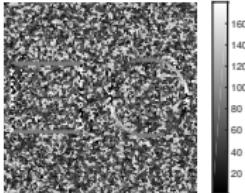
(a) Anisotropy.



(b) ρ map.



(c) e_1 map.



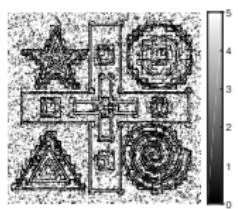
(d) θ map.

ML parameter selection results

- **Unbiased estimator** + empirical variance and RMSE going to 0 as $N \rightarrow +\infty$.
- Functional form of the gradient PDF adapts to local image structures.



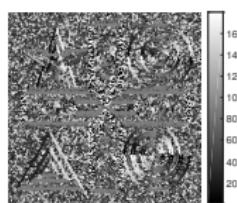
(a) Anisotropy.



(b) ρ map.



(c) e_1 map.



(d) θ map.

A flexible directional & space-variant regularisation

Image reconstruction

Well-posedness of the model

$$\min_u \underbrace{\sum_{i=1}^n \|\Lambda_i R_{\theta_i}(\nabla u)_i\|_2^{p_i}}_{\text{non-convex}} + \underbrace{\frac{\mu}{2} \|Ku - f\|_2^2}_{\text{convex}}, \quad \mu > 0, \quad p_i \in (0, 2], \quad \theta_i \in [0, 2\pi)$$

Well-posedness of the model

$$\min_u \underbrace{\sum_{i=1}^n \|\Lambda_i R_{\theta_i}(\nabla u)_i\|_2^{p_i}}_{\text{non-convex}} + \underbrace{\frac{\mu}{2} \|Ku - f\|_2^2}_{\text{convex}}, \quad \mu > 0, \quad p_i \in (0, 2], \quad \theta_i \in [0, 2\pi)$$

Proposition

The DTV_p^{sv} - L_2 functional is continuous, bounded from below by zero and coercive, hence it admits global minimisers.

Uniqueness holds if $p_i > 1$ for all i .

Based on a general result on the sum of proper, l.s.c. and coercive functions (see, e.g., [Ciak, PhD Thesis '15](#)).

Optimisation via ADMM

$$\min_u \left\{ \text{DTV}_p^{\text{sv}}(u) + \frac{\mu}{2} \|Ku - f\|_2^2 \right\}, \quad \mu > 0$$

Optimisation via ADMM

$$\min_u \quad \left\{ \sum_{i=1}^n \|\Lambda_i R_{\theta_i} t_i\|_2^{p_i} + \frac{\mu}{2} \|r\|_2^2 \right\}, \quad \mu > 0$$

with: $t := Du$, $r := Ku - f$.

Optimisation via ADMM

$$\min_u \quad \left\{ \sum_{i=1}^n \|\Lambda_i R_{\theta_i} t_i\|_2^{p_i} + \frac{\mu}{2} \|r\|_2^2 \right\}, \quad \mu > 0$$

with: $t := Du$, $r := Ku - f$.

Augmented Lagrangian:

$$\begin{aligned} \mathcal{L}(u, r, t; \rho_r, \rho_t) := & \sum_{i=1}^n \|\Lambda_i R_{\theta_i} t_i\|_2^{p_i} + \frac{\mu}{2} \|r\|_2^2 - \langle \rho_t, t - Du \rangle + \frac{\beta_t}{2} \|t - Du\|_2^2 \\ & - \langle \rho_r, r - (Ku - f) \rangle + \frac{\beta_r}{2} \|r - (Ku - f)\|_2^2, \end{aligned}$$

with $\beta_r, \beta_t > 0$ and $\rho_r \in \mathbb{R}^n$, $\rho_t \in \mathbb{R}^{2n}$.

Optimisation via ADMM

$$\min_u \quad \left\{ \sum_{i=1}^n \|\Lambda_i R_{\theta_i} t_i\|_2^{p_i} + \frac{\mu}{2} \|r\|_2^2 \right\}, \quad \mu > 0$$

with: $t := Du$, $r := Ku - f$.

Solve saddle point problem:

$$\text{find } (u^*, r^*, t^*; \rho_r^*, \rho_t^*) \in (\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{2n}) \times (\mathbb{R}^n \times \mathbb{R}^{2n}) \quad \text{s.t.}$$

$$\mathcal{L}(u^*, r^*, t^*; \rho_r, \rho_t) \leq \mathcal{L}(u^*, r^*, t^*; \rho_r^*, \rho_t^*) \leq \mathcal{L}(u, r, t; \rho_r^*, \rho_t^*)$$

Optimisation via ADMM

$$\min_u \quad \left\{ \sum_{i=1}^n \|\Lambda_i R_{\theta_i} t_i\|_2^{p_i} + \frac{\mu}{2} \|r\|_2^2 \right\}, \quad \mu > 0$$

with: $t := Du$, $r := Ku - f$.

ADMM iteration for $k \geq 0$:

$$u^{(k+1)} \leftarrow \underset{u \in \mathbb{R}^n}{\operatorname{argmin}} \mathcal{L}(u, r^{(k)}, t^{(k)}; \rho_r^{(k)}, \rho_t^{(k)}) \quad (\text{linear system})$$

$$r^{(k+1)} \leftarrow \underset{r \in \mathbb{R}^n}{\operatorname{argmin}} \mathcal{L}(u^{(k+1)}, r, t^{(k)}; \rho_r^{(k)}, \rho_t^{(k)}) \quad (\text{discrepancy})$$

$$t^{(k+1)} \leftarrow \underset{t \in \mathbb{R}^{2n}}{\operatorname{argmin}} \mathcal{L}(u^{(k+1)}, r^{(k+1)}, t; \rho_r^{(k)}, \rho_t^{(k)}) \quad (*)$$

$$\rho_r^{(k+1)} \leftarrow \rho_r^{(k)} - \beta_r (r^{(k+1)} - (Ku^{(k+1)} - f))$$

$$\rho_t^{(k+1)} \leftarrow \rho_t^{(k)} - \beta_t (t^{(k+1)} - Du^{(k+1)})$$

Optimisation via ADMM

$$\min_u \left\{ \sum_{i=1}^n \|\Lambda_i R_{\theta_i} t_i\|_2^{p_i} + \frac{\mu}{2} \|r\|_2^2 \right\}, \quad \mu > 0$$

with: $t := Du$, $r := Ku - f$.

ADMM iteration for $k \geq 0$:

$$u^{(k+1)} \leftarrow \underset{u \in \mathbb{R}^n}{\operatorname{argmin}} \mathcal{L}(u, r^{(k)}, t^{(k)}; \rho_r^{(k)}, \rho_t^{(k)}) \quad (\text{linear system})$$

$$r^{(k+1)} \leftarrow \underset{r \in \mathbb{R}^n}{\operatorname{argmin}} \mathcal{L}(u^{(k+1)}, r, t^{(k)}; \rho_r^{(k)}, \rho_t^{(k)}) \quad (\text{discrepancy})$$

$$t^{(k+1)} \leftarrow \underset{t \in \mathbb{R}^{2n}}{\operatorname{argmin}} \mathcal{L}(u^{(k+1)}, r^{(k+1)}, t; \rho_r^{(k)}, \rho_t^{(k)}) \quad (*)$$

$$\rho_r^{(k+1)} \leftarrow \rho_r^{(k)} - \beta_r (r^{(k+1)} - (Ku^{(k+1)} - f))$$

$$\rho_t^{(k+1)} \leftarrow \rho_t^{(k)} - \beta_t (t^{(k+1)} - Du^{(k+1)})$$

(*) Non-convex proximal step!

Proposition

Problem (*) is well-posed. The computation of the non-convex proximal map can be reduced to a one-dimensional constrained optimisation problem.

Pseudo-code

Algorithm 1 ADMM scheme for $\text{DTV}_p^{\text{sv}} - L_2$

inputs: observed image $f \in \mathbb{R}^n$, noise level $\sigma > 0$
parameters: discrepancy parameter $\tau \simeq 1$, ADMM parameters $\beta_r, \beta_t > 0$
output: reconstruction $u^* \in \mathbb{R}^n$

1. **Initialisation:**
 - 2. · estimate model parameters $p_i, R_{\theta_i}, \Lambda_i, i = 1, \dots, n$, by ML approach
 - 3. · set $\delta := \tau\sigma\sqrt{n}$, $u^{(0)} = f$, $r^{(0)} = Ku^{(0)} - f$, $t^{(0)} = Du^{(0)}$, $\rho_r^{(0)} = \rho_t^{(0)} = 0$, $k = 0$
4. **while** not converging **do** **ADMM**:
 - 5. · update primal variables:
 - 6. · update dual variables:
 - 7. · $k = k + 1$
8. **end while**
9. $u^* = u^{(k+1)}$

Parameter choice:

- Regularisation parameter μ chosen by discrepancy principle $\|Ku - f\|_2 \leq \delta$
- β_r, β_t set manually.

Empirical convergence is observed, also in such non-convex regime. Proof?

Numerical results: Barbara image

TEST: Barbara image with increasing degradation.

Numerical results: Barbara image

TEST: Barbara image with increasing degradation.

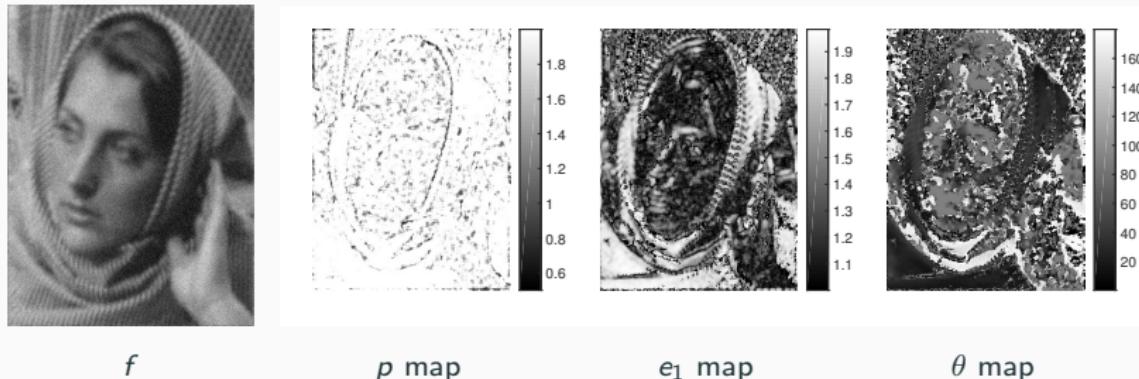
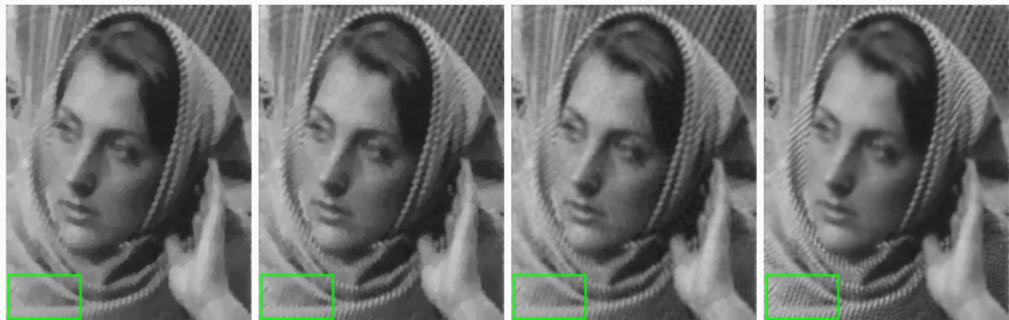


Figure 3: Image is corrupted by AWGN and blur with BSNR = 10 dB. Zoom size: 471×361 . Gaussian blur band= 9, $\sigma = 2$.

- Parameter maps estimated on a 7×7 neighbourhood.
- In the case of large noise: pre-processing by few (5) iterations of TV.

Numerical results: Barbara image

TEST: Barbara image with increasing degradation.

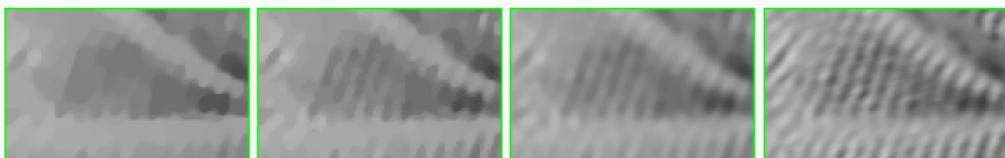


TV-L₂

TV_p-L₂, p = 0.92

TV_p^{sv}-L₂

DTV_p^{sv}-L₂



Better preservation of texture and details!

Numerical results: Barbara image

TEST: Barbara image with increasing degradation.

BSNR	TV-L ₂	TV _p -L ₂	TV _p ^{sv} -L ₂	DTV _p ^{sv} -L ₂
20	2.46	3.14	3.23	3.61
15	1.74	1.99	2.14	2.79
10	1.59	2.02	2.13	2.90

Increased SNR (ISNR) values for decreasing BSNR.

$$\text{BSNR}(u^*, u) := 10 \log_{10} \frac{\|Ku - \overline{Ku}\|_2^2}{\|u^* - Ku\|_2^2}, \quad \text{ISNR}(f, u, u^*) := 10 \log_{10} \frac{\|f - u\|_2^2}{\|u^* - u\|_2^2}$$

Numerical results: Barbara image

TEST: Barbara image with increasing degradation.

BSNR	TV-L ₂	TV _p -L ₂	TV _{p,α} ^{SV} -L ₂	DTV _p ^{SV} -L ₂
20	0.80	0.83	0.83	0.85
15	0.74	0.75	0.77	0.80
10	0.65	0.68	0.69	0.74

SSIM values for decreasing BSNR.

Numerical results: texture image

TEST: texture image with increasing degradation.

Numerical results: texture image

TEST: texture image with increasing degradation.

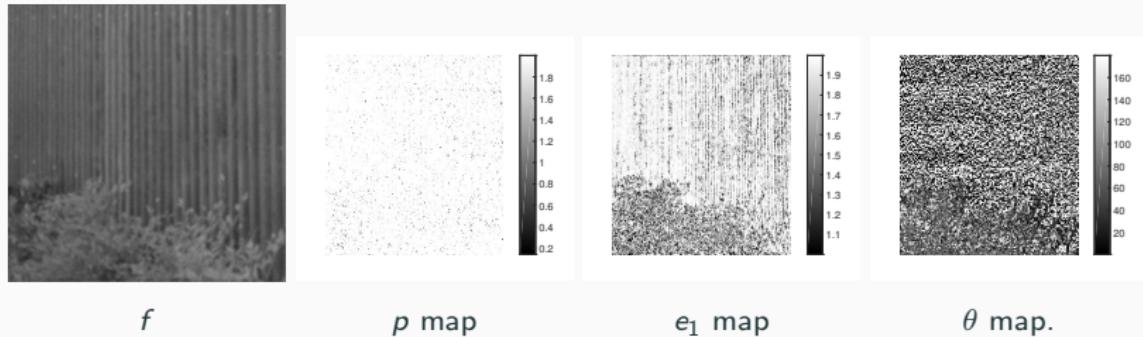
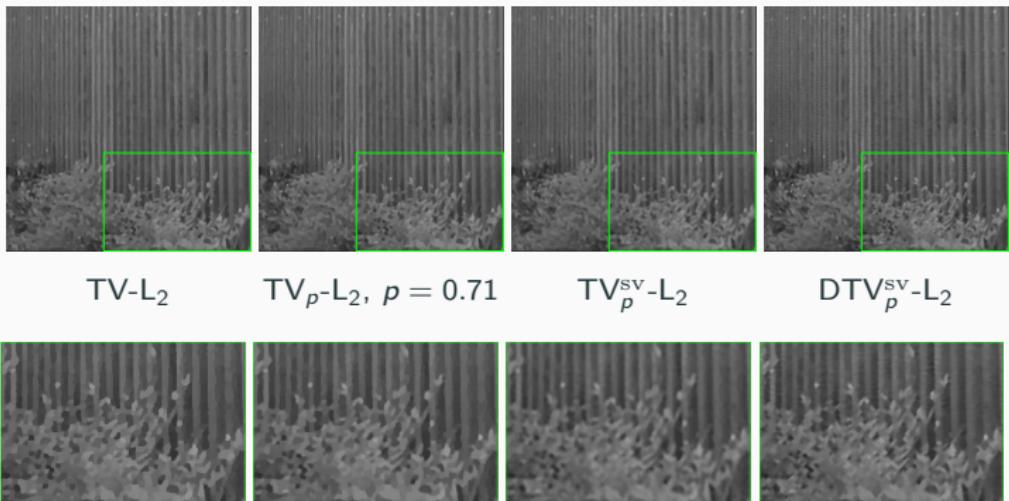


Figure 3: Image is corrupted by AWGN and blur with BSNR = 10 dB. Zoom size: 500×500 . Gaussian blur band= 9, $\sigma = 2$.

- Parameter maps estimated on a 3×3 neighbourhood.
- In the case of large noise: pre-processing by few (5) iterations of TV.

Numerical results: texture image

TEST: texture image with increasing degradation.



Numerical results: texture image

TEST: texture image with increasing degradation.

BSNR	TV-L ₂	TV _p -L ₂	TV _p ^{sv} -L ₂	DTV _p ^{sv} -L ₂
20	2.07	2.43	2.53	2.78
15	1.83	2.06	2.26	2.56
10	0.94	1.55	1.86	2.45

Increased SNR (ISNR) values for decreasing BSNR.

$$\text{BSNR}(u^*, u) := 10 \log_{10} \frac{\|Ku - \overline{Ku}\|_2^2}{\|u^* - Ku\|_2^2}, \quad \text{ISNR}(f, u, u^*) := 10 \log_{10} \frac{\|f - u\|_2^2}{\|u^* - u\|_2^2}$$

Numerical results: texture image

TEST: texture image with increasing degradation.

BSNR	TV-L ₂	TV _p -L ₂	TV _{p,α} ^{SV} -L ₂	DTV _p ^{SV} -L ₂
20	0.78	0.79	0.80	0.81
15	0.76	0.77	0.78	0.79
10	0.70	0.72	0.74	0.76

SSIM values for decreasing BSNR.

Conclusions

Conclusions

Take-home messages:

- BGGD for flexible description of **natural image statistics**.
- Variational **space-dependent, directional regularisation** adapting to local image structures (upon ML parameter estimation).
- Efficient **ADMM optimisation** (non-convex proximal step). Results show improved **texture reconstruction**.

Outlook:

- **Applications** to inverse problems with different measurement/image space (e.g. tomography → Rob's talk)?
- **Optimisation**: theoretical guarantees for non-convex ADMM? Other algorithms?



L. Calatroni, A. Lanza, M. Pragliola, F. Sgallari, *Space-variant anisotropic regularisation and automated parameter selection for image restoration problems*, SIAM Journal of Imaging Sciences, in press, 2019.

Thank you for your attention!

Questions?

luca.calatroni@polytechnique.edu