Coherence enhancing variational image restoration

Markus Grasmair

Department of Mathematical Sciences,
Norwegian University of Science and Technology,
Trondheim, Norway

Trondheim,
June 2019

Joint work with Peter Elbau (Vienna)

Outline

Variational Image Denoising

Given a distorted image

$$f \approx Ku_0 + \delta,$$

(standard) variational image denosing methods try to approximate the noise-free image u_0 by solving

$$\frac{1}{2}\|Ku - f\|_{L^2}^2 + \alpha\mathcal{R}(u) \to \min,$$

where:

- K ... imaging operator, e.g.
  - $K = \mathrm{Id}$ ... *denoising*.
  - $K = k*$ ... *deblurring*.
- δ ... *noise*, e.g. realisation of Gaussian i.i.d. random variable.
- $\mathcal{R}(u)$ ... regularisation term, encodes structural assumptions on the true image u_0.
- $\alpha > 0$ ... regularisation parameter balancing fidelity and regularity.

Local Regularisation Terms

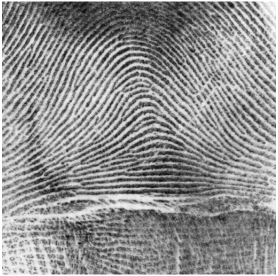
Typically, structural assumptions on u_0 are concerned with purely local regularity properties:

- Smoothness:
$$\mathcal{R}(u) := \|\nabla u\|_{L^2}^2 = \int_\Omega |\nabla u(x)|^2 \, dx.$$
- Piecewise smooth with strong edges: Total variation and variants,
$$\mathcal{R}(u) := |Du|(\Omega) \approx \int_\Omega |\nabla u(x)| \, dx.$$
- Projections of uniform objects: Mumford–Shah regulariser,
$$\mathcal{R}(u) := \int_{\Omega \setminus S_u} |\nabla u(x)|^2 \, dx + \beta \operatorname{len}(S_u)$$
  with S_u ... jump set of u.

Coherence Enhancement

True image u_0 is assumed to include several *coherent* regions, where the gradients of u_0 are almost parallel.

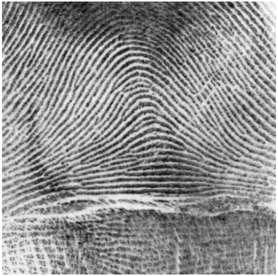
Examples:



Coherence Enhancement

True image u_0 is assumed to include several *coherent* regions, where the gradients of u_0 are almost parallel.

Examples:



Coherence Enhancement

True image u_0 is assumed to include several *coherent* regions, where the gradients of u_0 are almost parallel.

Mathematical characterisation possible via the structure tensor

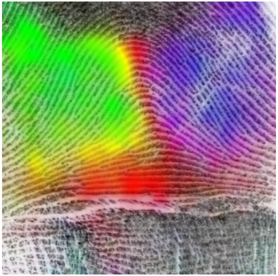
$$S_{\rho,\sigma}(u) := \eta_\rho * \big(\nabla u_\sigma \otimes \nabla u_\sigma\big).$$

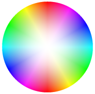
- Eigenvectors of $S_{\rho,\sigma}(u)$ indicate the main directions around a given point.
- Eigenvalues $\lambda_1 > \lambda_2 \geq 0$ of $S(u)$ indicate the strengths of the different directions.
- If $\lambda_1 \gg \lambda_2$, then u is locally highly coherent.

Coherence Enhancement

True image u_0 is assumed to include several *coherent* regions, where the gradients of u_0 are almost parallel.

Examples:





Weickert's Anisotropic Diffusion

- Original PDE setting (only for denoising):

$$\frac{\partial}{\partial t} u = \operatorname{div} g(S_{\rho,\sigma}(u)) \nabla u, \qquad u(\cdot, 0) = f,$$

where

$$g \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} \left(1 + \frac{(\lambda_1 - \lambda_2)^2}{\gamma}\right)^{-1} & 0 \\ 0 & 1 \end{pmatrix}$$

and $g(U \Lambda U^T) = U g(\Lambda) U^T$, leading to strong diffusion along edges and weak diffusion across edges.

- Related variational setting (for general image restoration):

$$\min_u \frac{1}{2} \|Ku - f\|_{L^2}^2 + \alpha \int_\Omega \langle \nabla u, g(S_{\rho,\sigma}(u)) \nabla u \rangle \, dx.$$

Outline

Alternative Characterisation of Coherence

Local parallelity of the gradients of u around a point $x \in \Omega$ can be measured by assessing the integral

$$\int_\Omega w(|x-y|)\big\langle \nabla u(x), \nabla u(y)^\perp \big\rangle^2 \, dy$$

for some decreasing weight function $w\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ defining the size of local neighbourhoods.

- If ∇u is parallel in a neighbourhood of x, this term becomes zero.
- The term becomes large, if there are many gradients $\nabla u(y)$ orthogonal to $\nabla u(x)$ in a neighbourhood of x.

Thus we can define the coherence enhancing term

$$\mathcal{R}(u) := \frac{1}{2} \int_\Omega \int_\Omega w(|x-y|)\big\langle \nabla u(x), \nabla u(y)^\perp \big\rangle^2 \, dy \, dx.$$

Reformulation with Structure Tensor

We can rewrite

$$\int_\Omega w(|x-y|)\big\langle \nabla u(x), \nabla u(y)^\perp \big\rangle^2\, dy$$
$$= \nabla u(x)^T \operatorname{adj}\Big(\int_\Omega w(|x-y|)\nabla u(y)\otimes \nabla u(y)\, dy\Big)\nabla u(x)$$

with

$$\operatorname{adj} A = (\det A)\, A^{-1}.$$

Specifically, we obtain for w ... Gaussian kernel of variance ρ^2 the expression

$$\mathcal{R}(u) = \frac{1}{2}\int_\Omega \langle \nabla u, \operatorname{adj}(S_{\rho,0}(u))\nabla u\rangle\, dx.$$

Nonlocal Integrands

We consider

$$\mathcal{R}(u) := \int_\Omega \int_\Omega s\big(x, y; \nabla u(x), \nabla u(y)\big)\, dy\, dx$$

with

$$s(x, y; \zeta, \xi) := w(|x-y|)\langle \zeta, \xi^\perp \rangle^2.$$

- The functional $\mathcal{R}$ is nonlocal, non-convex.
- The integrand s is non-convex, but separately convex.

Theorem (Boulanger et al (2011))

A nonlocal integral functional $\mathcal{R}\colon L^p(\Omega; \mathbb{R}^n) \to \mathbb{R} \cup \{+\infty\}$ *of the form*

$$\mathcal{R}(\xi) = \int_\Omega \int_\Omega s(x, y; \xi(x), \xi(y))\, dx\, dy$$

is weakly sequentially lower semi-continuous if and only if s *is separately convex.*

Existence

Note that the functional $\mathcal{R}$ is not coercive:

- For the function $u(x, y) = \sin(kx)$ we have $\mathcal{R}(u) = 0$ for all k.
- If $f(x, y) = \operatorname{sgn}(x)$, no minimiser of $\|u - f\|^2 + \mathcal{R}(u)$ exists in H^1.

Lemma

Assume that $K1 \neq 0$. Then the functional

$$\frac{1}{2}\|Ku - f\|_{L^2}^2 + \alpha\mathcal{R}(u) + \frac{\beta}{2}\|\nabla u\|_{L^2}^2$$

admits a minimiser in $H^1(\Omega)$ for all $\alpha > 0$ and $\beta > 0$.

Because of the non-convexity of $\mathcal{R}$, also this modified functional is non-convex for small β. Thus uniqueness of the minimiser is not guaranteed.

Differentiability and Gradient

Lemma

The functional $\mathcal{R}$ is Gâteaux differentiable on $W^{1,4}(\Omega)$.

Additionally, the formal L^2 gradient of $\mathcal{R}$ can be computed as

$$\nabla\mathcal{R}(u) = -2\,\mathrm{div}\big(\mathrm{adj}(S_{\rho,0}(u))\nabla u\big).$$

Thus we have the formal optimality condition

$$K^*Ku - 2\alpha\,\mathrm{div}\big(\mathrm{adj}(S_{\rho,0}(u))\nabla u\big) - \beta\Delta u = K^*f.$$

- Possible gradient flow for denoising:

$$\frac{\partial}{\partial t}u = \beta\Delta u + \mathrm{div}\big(\mathrm{adj}(S(u))\nabla u\big).$$

Outline

Numerical Approach

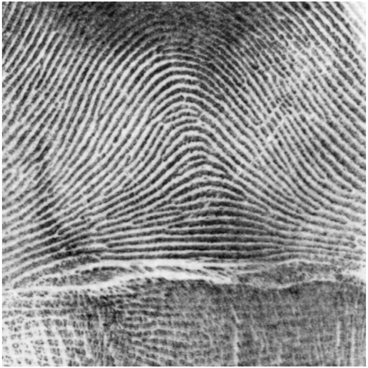
- Approximation of gradients by finite differences.
- Gradient descent method with exact line search
  - Note: We have a quartic functional.
- Convolution kernel w is Gaussian with several pixel sizes variance, that is,

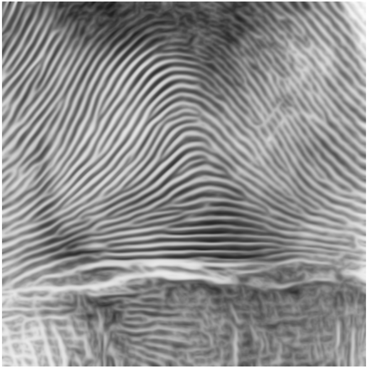
$$w(|x-y|) = C\exp(-|x-y|^2/2\rho^2).$$

  Size of ρ should correspond to size of regions where we expect u to be coherent.
- Computation of convolutions by FFT.
- Regularisation parameter $\beta \ll \alpha$ (only for some stabilisation).

Fingerprints

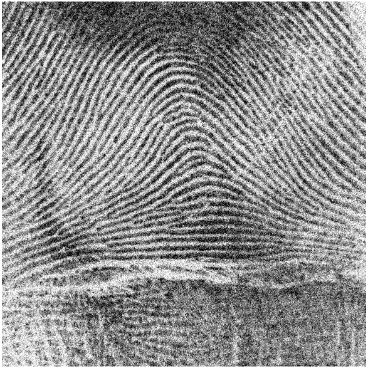
Smoothing of fingerprint image with varying levels of noise and constant regularisation parameters.

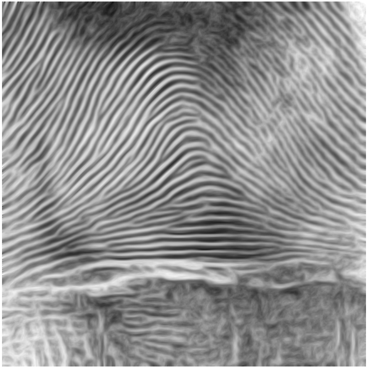




Fingerprints

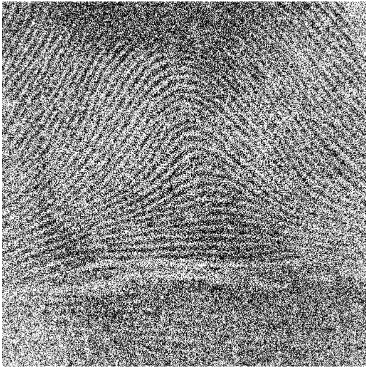
Smoothing of fingerprint image with varying levels of noise and constant regularisation parameters.

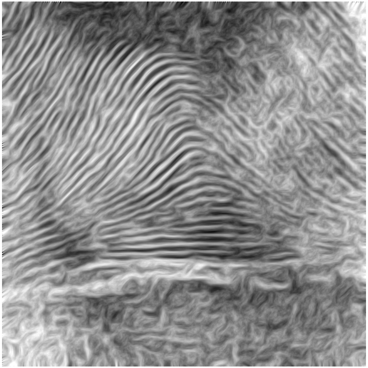




Fingerprints

Smoothing of fingerprint image with varying levels of noise and constant regularisation parameters.



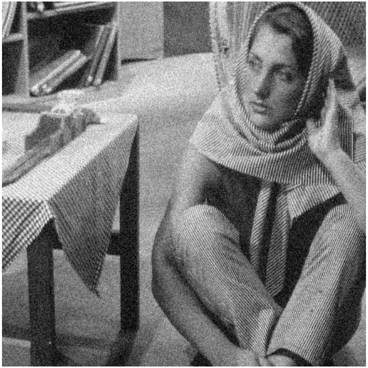


Barbara

Smoothing of barbara image with varying levels of noise and constant regularisation parameters.

Barbara

Smoothing of barbara image with varying levels of noise and constant regularisation parameters.



Barbara

Smoothing of barbara image with varying levels of noise and constant regularisation parameters.

Barbara — detail

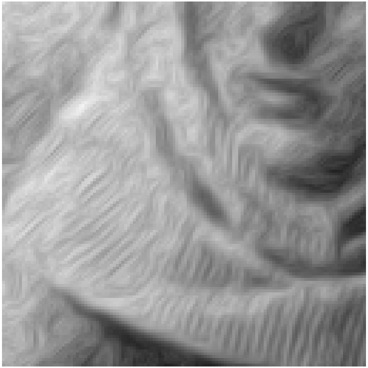
Smoothing of barbara image with varying levels of noise and constant regularisation parameters.

Barbara — detail

Smoothing of barbara image with varying levels of noise and constant regularisation parameters.

Barbara — detail

Smoothing of barbara image with varying levels of noise and constant regularisation parameters.



Conclusion and Outlook

Summary:

- Coherence enhancing image denoising with a nonlocal functional.
- Existence result for the proposed nonlocal functional.
- Numerical implementation by gradient descent with exact line search.

Outlook:

- Generalisation of the idea to total variation based regularisation term of the form

$$\mathcal{R}(u) = \int_\Omega \int_\Omega |\langle \nabla u(x), \nabla u(y)^\perp \rangle| \, dy \, dx.$$

- Efficient numerical solution of more general linear inverse problems with coherent solutions.
- Adaptive choice of purely local/coherent enhancing regularisation term in order to avoid artifacts.