Lecture 10:

Variational Methods I: Basic Ideas

Contents

Why Variational Methods?

- image transformation that satisfies an optimality criterion
- clear description of all model assumptions without any hidden constraints
- want to find an image restoration model where
  - the filtered image u is close to the original image f
  - the filtered image u should be (at least piecewise) smooth
- treatment at the boundaries should result in a natural way from the model assumptions

Modelling

Basic Idea

(Bertero et al. 1988, Nordström 1990, Schnörr 1994, Charbonnier et al. 1994)

- variational restoration of some image f: minimiser u_α of

$$E_f(u) := \int_\Omega \Big(\underbrace{(u-f)^2}_{\text{similarity}} + \alpha \underbrace{\Psi(|\nabla u|^2)}_{\text{smoothness}} \Big)\, d\mathbf{x}$$

- first term *(data term, similarity term, fidelity term)*: rewards similarity to the original image
- second term *(smoothness term, regulariser, penaliser)*: penalises deviations from (piecewise) smoothness
- $\alpha > 0$: *regularisation parameter (smoothness weight)*
- Variational methods of this type are also called *regularisation methods*.

M I A

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

Technical Assumptions on the Penalising Function Ψ:

- Ψ is differentiable and increasing: $\Psi'(s^2) > 0$.
- $\Psi(s^2)$ is convex in s.
- There exist constants $c_1, c_2 > 0$ such that $c_1 s^2 \le \Psi(s^2) \le c_2 s^2$ for all s^2.

Examples:

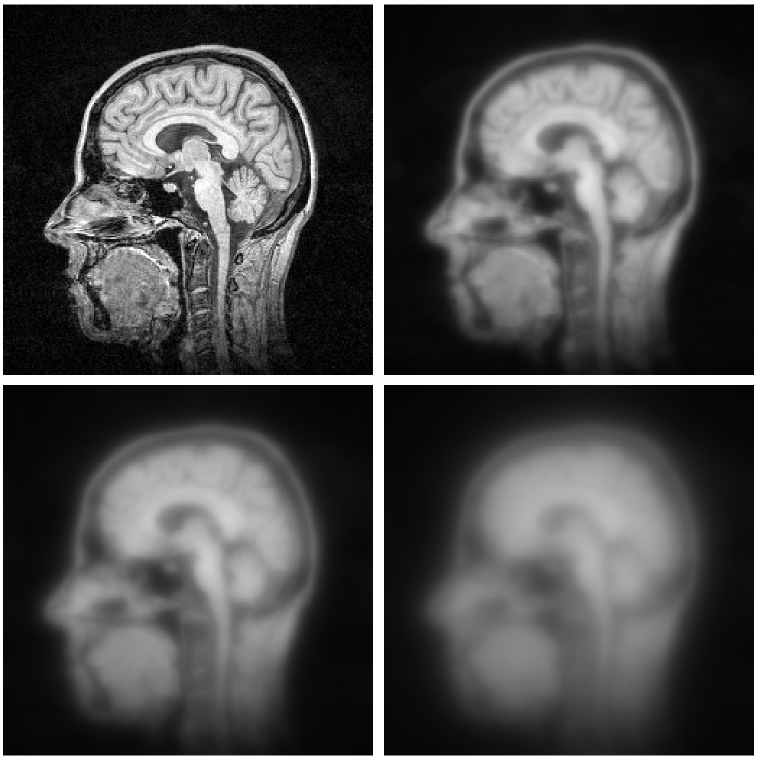
- Whittaker 1923, Tikhonov 1963:

$$\Psi(s^2) \;=\; s^2$$

- Charbonnier et al. 1994:

$$\Psi(s^2) \;=\; 2\lambda^2 \sqrt{1 + s^2/\lambda^2} - 2\lambda^2$$

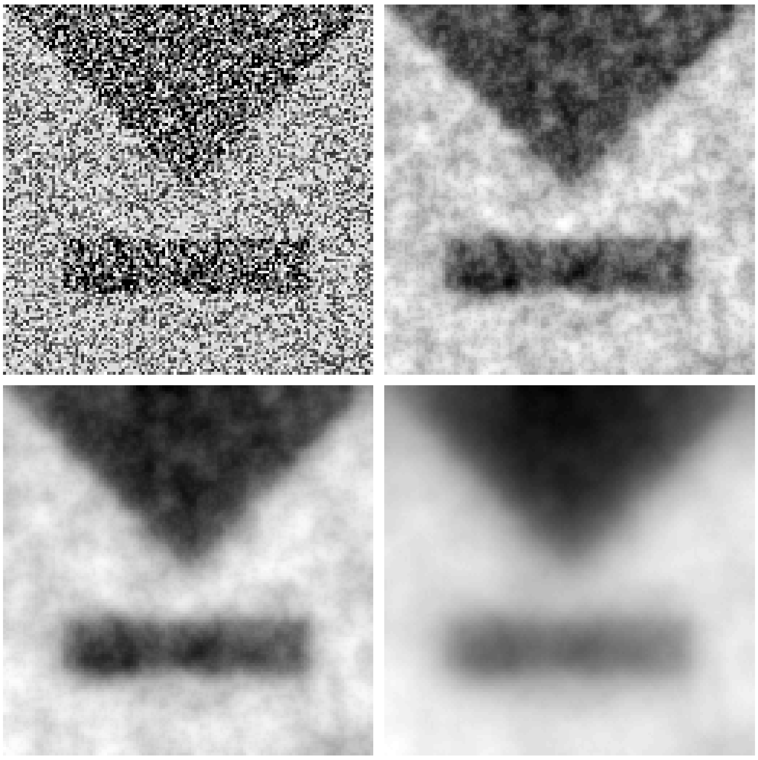
Modelling (3)



Variational image restoration with the Whittaker–Tikhonov penaliser. **Top left to bottom right:** Regularisation parameters $\alpha = 0,\ 10,\ 30,\ 100$.

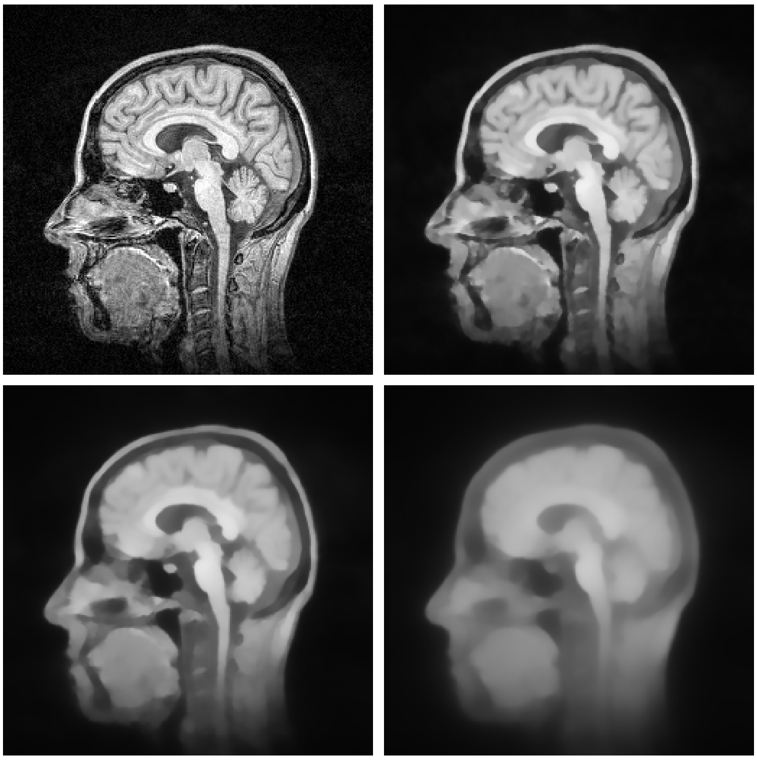
M I A 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

Modelling (4)



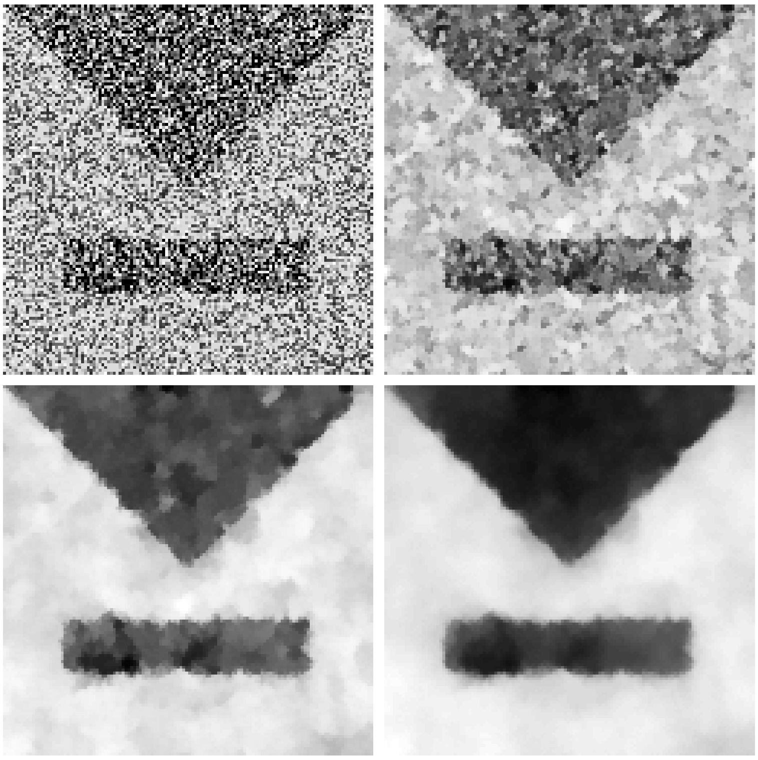
Variational image restoration with the Whittaker–Tikhonov penaliser. **Top left to bottom right:** Regularisation parameters $\alpha = 0,\ 5,\ 20,\ 100$.

Modelling (5)



Variational image restoration with the Charbonnier penaliser ($\lambda = 2$). **Top left to bottom right:** Regularisation parameters $\alpha = 0, 10, 30, 100$.

Modelling (6)



Variational image restoration with the Charbonnier penaliser ($\lambda = 1$). **Top left to bottom right:** Regularisation parameters $\alpha = 0, 20, 50, 100$.

M I A 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

Well-Posedness and Scale-Space Properties

It seems that the regularisation parameter α plays the role of a scale parameter. The following well-posedness and scale-space properties can be established:

(a) **Well-Posedness and Regularity** (Schnörr 1994)

- Existence of a unique minimizer u_α for bounded image f.
- Smoothness: $u_\alpha \in H^2(\Omega)$ ("twice differentiable")
- u_α depends continuously on f.
- No contrast enhancement for convex regularisers!

(b) **Average Grey Level Invariance**

$$\frac{1}{|\Omega|}\int_\Omega u_\alpha(\mathbf{x})\,d\mathbf{x} \;=\; \frac{1}{|\Omega|}\int_\Omega f(\mathbf{x})\,d\mathbf{x} \;=:\; \mu.$$

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

(c) **Maximum–Minimum Principle**

$$\inf f \;\le\; u_\alpha(\mathbf{x}) \;\le\; \sup f \qquad \forall\,\mathbf{x}\in\Omega,\ \forall\,\alpha>0$$

(d) **Lyapunov Property** (Scherzer/W. 2000)
Let $r \in \mathrm{C}^2$ be convex. Then

$$V(\alpha) \;:=\; \int_\Omega r(u_\alpha(\mathbf{x}))\,d\mathbf{x}$$

is bounded from below and satisfies $V(\alpha) \le V(0)$.

(e) **Convergence** (Scherzer/W. 2000)
$u_\alpha(\mathbf{x})$ converges to the average grey level μ for $\alpha \to \infty$
(in 2D: convergence in $L^p(\Omega)$, $1 \le p < \infty$).

These theoretical results resemble the ones for diffusion filtering (Lectures 4 and 8). Is there a specific reason behind this similarity?

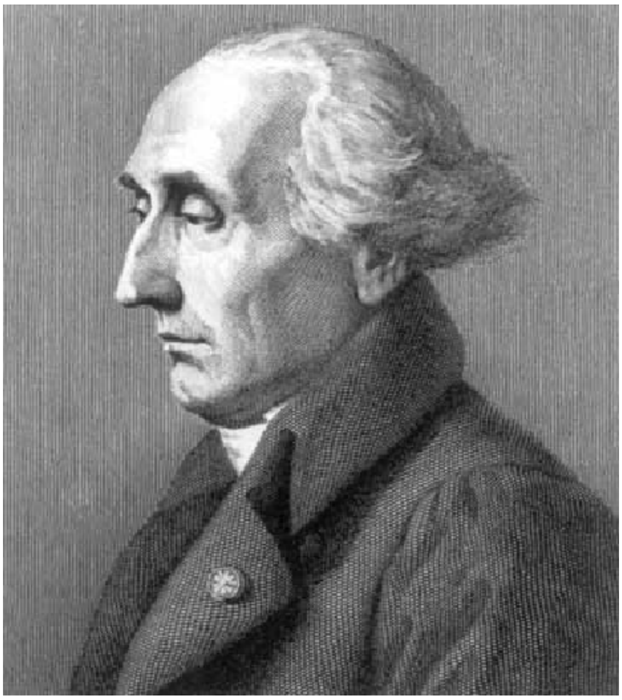
Euler–Lagrange Equations

Standard Calculus:

- considers real-valued *functions* $f(x)$, i.e. mappings from a *number* x into $\mathbb{R}$.
- If f has a minimum in ξ, then it necessarily holds $f'(\xi) = 0$.
- If f is strictly convex and $f'(\xi) = 0$, then ξ is the unique minimum of f.

Calculus of Variations:

- considers real-valued *functionals* $E(u)$, i.e. mappings from a *function* $u(x)$ into $\mathbb{R}$.
- If E is minimised by a function v, then v has to satisfy necessarily a so-called *Euler–Lagrange equation*. This is a partial differential equation in v.
- If the E is strictly convex and satisfies the Euler–Lagrange equation, then v is the unique minimiser of E.



The mathematicians Euler and Lagrange belong to the founders of the calculus of variations. **Left:** Leonhard Euler (1707–1783). **Right:** Joseph–Louis Lagrange (1736–1813).

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

Euler–Lagrange Equations (3)

Theorem (Euler-Lagrange Equations in 1-D)

A smooth function $u(x)$, $x \in [a,b]$ that minimises the 1-D energy functional

$$E(u) \;=\; \int_a^b F(x,u,u')\,dx$$

satisfies necessarily the *Euler–Lagrange equation*

$$F_u - \frac{d}{dx} F_{u'} \;=\; 0$$

and the so-called *natural boundary conditions*

$$F_{u'} = 0$$

for $x = a$ and $x = b$.

Euler–Lagrange Equations (4)

Proof

We assume that $v(x)$ is a sufficiently often differentiable minimiser of E.
We embed $v(x)$ into the family

$$u(x,\varepsilon) \;:=\; v(x) + \varepsilon h(x)$$

with some perturbation function $h(x)$.
Since $v(x)$ minimises $E(u)$, we know that the scalar-valued function

$$g(\varepsilon) \;:=\; E(u(x,\varepsilon)) \;=\; E(v+\varepsilon h)$$

has a minimum in $\varepsilon = 0$. Therefore, we have

$$0 \;=\; g'(0) \;=\; \frac{d}{d\varepsilon} E(v+\varepsilon h)\Big|_{\varepsilon=0} \;=\; \frac{d}{d\varepsilon} \int_a^b F(x, \underbrace{v+\varepsilon h}_{u(.,\varepsilon)}, \underbrace{v'+\varepsilon h'}_{u'(.,\varepsilon)})\,dx\,\Big|_{\varepsilon=0}.$$

Euler–Lagrange Equations (5)

By means of the chain rule it follows that

$$0 = \int_a^b \Big(F_u(x,u,u')\,h(x) + F_{u'}(x,u,u')\,h'(x) \Big)\,dx \,\Big|_{\varepsilon=0}$$

$$= \int_a^b \Big(F_u(x,v,v')\,h(x) + F_{u'}(x,v,v')\,h'(x) \Big)\,dx.$$

Partial integration of the second term yields

$$0 = \int_a^b \left(F_u(x,v,v')\,h(x) - \frac{d}{dx} F_{u'}(x,v,v')\,h(x) \right) dx \; + \; F_{u'}(x,v,v')\,h(x)\,\Big|_{x=a}^{x=b} \quad (1)$$

If v is a minimiser within some family of competing functions $u(\varepsilon, x) = v(x) + \varepsilon h(x)$, then v is also minimiser within the smaller class of functions where the perturbation $h(x)$ satisfies $h(a) = 0 = h(b)$. Thus,

$$0 = \int_a^b \left(F_u(x,v,v') - \frac{d}{dx} F_{u'}(x,v,v') \right) h(x)\,dx.$$

Euler–Lagrange Equations (6)

This gives the Euler–Lagrange equation

$$F_u - \frac{d}{dx} F_{u'} = 0.$$

Note that this equation does not depend on the perturbation h.

Plugging it into Equation (1) gives

$$F_{u'}(x,v,v')\,h(x)\,\Big|_{x=a}^{x=b} = 0.$$

which holds for arbitrary perturbations h (also with $h(a) \neq 0$ and $h(b) \neq 0$). Thus, one obtains the natural boundary conditions

$$F_{u'} = 0$$

for $x = a$ und $x = b$. This concludes the proof.

Theorem (Euler-Lagrange Equations in 2-D)

A minimiser of the 2-D energy functional

$$E(u) := \int_\Omega F(x_1, x_2, u, u_{x_1}, u_{x_2})\, d\mathbf{x}$$

satisfies necessarily the Euler–Lagrange equation

$$F_u - \partial_{x_1} F_{u_{x_1}} - \partial_{x_2} F_{u_{x_2}} \;=\; 0$$

with the natural boundary conditions

$$\mathbf{n}^\top \begin{pmatrix} F_{u_{x_1}} \\ F_{u_{x_2}} \end{pmatrix} \;=\; 0$$

at the image boundary $\partial\Omega$ with normal vector $\mathbf{n}$.

Euler-Lagrange Equations Applied to our Functional

For our energy functional

$$E_f(u) := \int_\Omega \Big((u - f)^2 + \; \alpha\, \Psi(|\nabla u|^2) \Big)\, d\mathbf{x}$$

we have

$$\begin{aligned} F &= (u-f)^2 + \; \alpha\, \Psi(u_{x_1}^2 + u_{x_2}^2), \\ F_u &= 2\,(u-f), \\ F_{u_{x_1}} &= 2\,\alpha\, u_{x_1}\, \Psi'(u_{x_1}^2 + u_{x_2}^2), \\ F_{u_{x_2}} &= 2\,\alpha\, u_{x_2}\, \Psi'(u_{x_1}^2 + u_{x_2}^2). \end{aligned}$$

All we have to do now is to plug this into the Euler-Langrange equation and the natural boundary conditions.

The **Euler-Lagrange equation** $F_u - \partial_{x_1} F_{u_{x_1}} - \partial_{x_2} F_{u_{x_2}} = 0$ becomes

$$2(u-f) - 2\alpha\, \partial_{x_1}\Big(\Psi'(u_{x_1}^2 + u_{x_2}^2)\, u_{x_1}\Big) - 2\alpha\, \partial_{x_2}\Big(\Psi'(u_{x_1}^2 + u_{x_2}^2)\, u_{x_2}\Big) \;=\; 0.$$

Division by 2 yields

$$u - f - \alpha\, \mathrm{div}\,(\Psi'(|\nabla u|^2)\, \nabla u) \;=\; 0.$$

The **boundary condition** $\mathbf{n}^\top \begin{pmatrix} F_{u_{x_1}} \\ F_{u_{x_2}} \end{pmatrix} = 0$ becomes

$$0 \;=\; \mathbf{n}^\top \begin{pmatrix} 2\,\alpha\, u_{x_1}\, \Psi'(u_{x_1}^2 + u_{x_2}^2) \\ 2\,\alpha\, u_{x_2}\, \Psi'(u_{x_1}^2 + u_{x_2}^2) \end{pmatrix} \;=\; 2\,\alpha\, \Psi'(|\nabla u|^2)\, \mathbf{n}^\top \nabla u.$$

Since $\alpha > 0$ and $\Psi'(|\nabla u|^2) > 0$, we obtain the Neumann boundary condition

$$0 \;=\; \mathbf{n}^\top \nabla u \;=\; \partial_{\mathbf{n}} u.$$

Relations to Diffusion Filters (Scherzer/W. 2000)

Let us rewrite the Euler–Lagrange equation as

$$\frac{u-f}{\alpha} \;=\; \mathrm{div}\,\left(\Psi'(|\nabla u|^2)\; \nabla u\right).$$

This can be regarded as fully implicit time discretisation of the diffusion filter

$$\partial_t u \;=\; \mathrm{div}\,\left(\Psi'(|\nabla u|^2)\; \nabla u\right),$$
$$u(\mathbf{x}, 0) \;=\; f(\mathbf{x})$$

with a single time step of size α.
In both cases we have Neumann boundary conditions.

A variational method with regulariser $\Psi(|\nabla u|^2)$ and regularisation parameter α approximates a diffusion filter with diffusivity $\Psi'(|\nabla u|^2)$ and stopping time α.

We may thus regard the well-posedness and scale-space theory for variational methods as a time-discrete theory for diffusion filters. This completes the continuous, space-discrete and fully discrete diffusion framework of Lecture 8.

Regularisers and Their Corresponding Diffusivities

- The **Tikhonov regulariser**

$$\Psi(|\nabla u|^2) \;=\; |\nabla u|^2$$

corresponds to the linear diffusivity

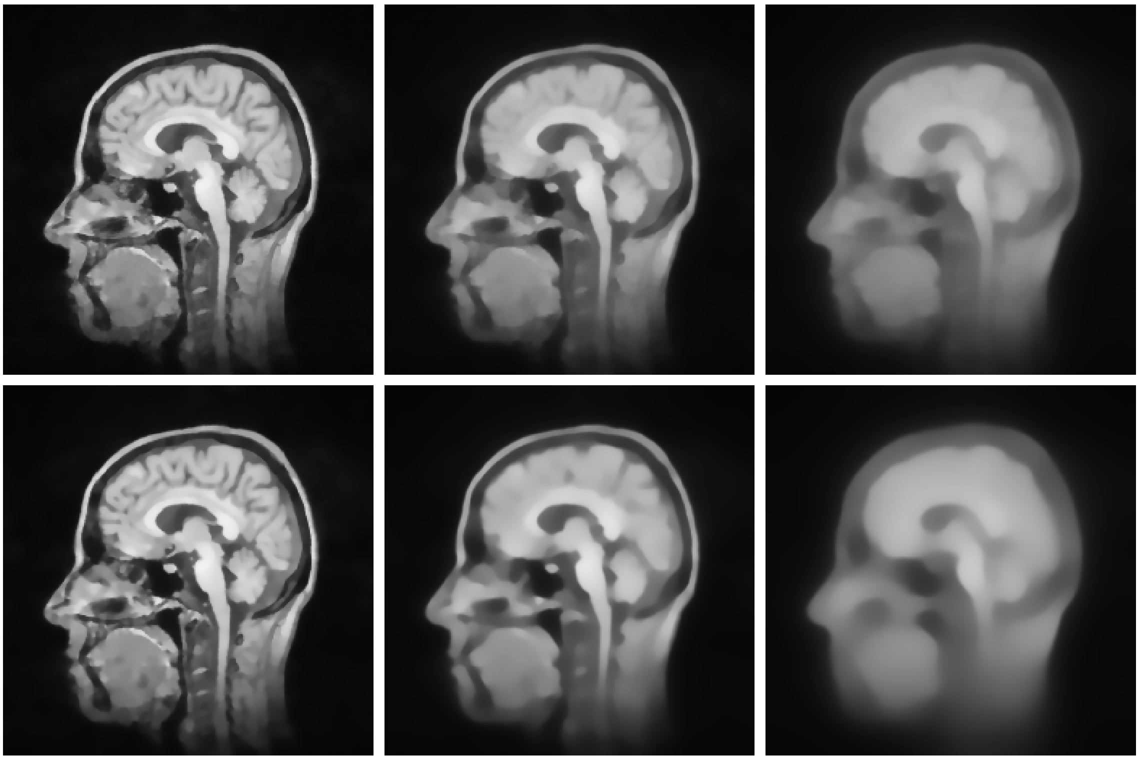
$$\Psi'(|\nabla u|^2) \;=\; 1.$$

- The **Charbonnier regulariser**

$$\Psi(|\nabla u|^2) \;=\; 2\lambda^2\sqrt{1+|\nabla u|^2/\lambda^2} - 2\lambda^2$$

leads to the nonlinear Charbonnier diffusivity

$$\Psi'(|\nabla u|^2) \;=\; \frac{1}{\sqrt{1+|\nabla u|^2/\lambda^2}}.$$



Comparison between variational image restoration and nonlinear diffusion filtering (Charbonnier diffusivity, $\lambda = 2$). (b) **Top:** Variational restoration with regularisation parameters $\alpha = 10$, 30, 100. (a) **Bottom:** Temporal evolution of the diffusion filter, $t = 10$, 30, 100.

Summary

- Variational image restoration methods minimise an energy functional with a similarity term and a smoothness term.
- They lead to a similar well-posedness and scale-space theory as nonlinear diffusion filters.
- The Euler-Lagrange equation provides a necessary condition for a minimiser of the energy functional. It is a PDE.
- The Euler-Lagrange equation for variational image restoration can be regarded as a fully implicit time discretisation of a nonlinear diffusion filter.
- The regularisation parameter α approximates the stopping time of the diffusion filter.
- The regulariser $\Psi(|\nabla u|^2)$ corresponds to the diffusivity $\Psi'(|\nabla u|^2)$.

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

References

- M. Bertero, T. A. Poggio, V. Torre: Ill-posed problems in early vision. *Proc. IEEE*, Vol. 76, 869–889, 1988.
  (studies a number of quadratic variational problems)
- N. Nordström: Biased anisotropic diffusion - a unified regularization and diffusion approach to edge detection. *Image and Vision Computing*, Vol. 8, 318–327, 1990.
  (early reference on nonquadratic variational methods)
- C. Schnörr: Unique reconstruction of piecewise smooth images by minimizing strictly convex non-quadratic functionals. *Journal of Mathematical Imaging and Vision*, Vol. 4, 189–198, 1994.
  (well-posedness properties of convex nonquadratic variational methods)
- E. T. Whittaker: A new method of graduation. *Proceedings of the Edinburgh Mathematical Society*, Vol. 41, pp. 65–75, 1923. *(early paper on (discrete) variational methods with a quadratic regulariser)*
- A. N. Tikhonov: Solution of incorrectly formulated problems and the regularization method. *Soviet Mathematics Doklady*, Vol. 4, 1035–1038, 1963.
  (another classical paper on variational methods with a quadratic regulariser)
- P. Charbonnier, L. Blanc–Féraud, G. Aubert, M. Barlaud: Two deterministic half-quadratic regularization algorithms for computed imaging. *Proc. IEEE International Conference on Image Processing* (ICIP–94, Austin, Nov. 13–16, 1994), Vol. 2, 168–172, 1994.
  (variational methods with a nonquadratic regulariser)

References (2)

- O. Scherzer, J. Weickert: Relations between regularization and diffusion filtering, *Journal of Mathematical Imaging and Vision*, Vol. 12, 43–63, 2000.
  (scale-space properties of regularisation methods)
- R. Courant, D. Hilbert: *Methods of Mathematical Physics*. Vol. 1, Interscience, New York, 1953.
  (a classical book in which the calculus of variations is discussed)
- I. M. Gelfand and S. V. Fomin: *Calculus of Variations*, Dover, New York, 2000.
  (well-written and inexpensive monograph)

Assignment T3 (1)

Assignment T3 – Theoretical Home Work

Problem 1 (Discretisation of Anisotropic Nonlinear Diffusion) (4 points)

For the discretisation of anisotropic nonlinear diffusion, one has to discretise the mixed derivative expression $\partial_x(b\partial_y u) + \partial_y(b\partial_x u)$.

Discretise the mixed derivative $\partial_y(b\partial_x u)$ in two different ways:

(a)
- using forward differences for ∂_y and ∂_x
- using backward differences for ∂_y and ∂_x
- averaging both results

(b)
- using forward differences for ∂_y and backward differences for ∂_x
- using backward differences for ∂_y and forward differences for ∂_x
- averaging both results

Which approximations for the entire expression $\partial_x(b\partial_y u) + \partial_y(b\partial_x u)$ are obtained with discretisations (a) and (b)? What is the result when these two approximations to $\partial_x(b\partial_y u) + \partial_y(b\partial_x u)$ are averaged?

Problem 2 (Discretisation of Anisotropic Nonlinear Diffusion) (2 points)

Find a suitable positive definite diffusion tensor $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ where the "nonnegativity" semidiscretisation from Lecture 8 (slide 12) with $h_1 = h_2$ has a negative stencil weight outside the centre. Can this also happen if the positive definite diffusion tensor is diagonally dominant?

Problem 3 (Stopping Time Selection: Decorrelation Criterion) (4 points)

Smooth the 1-D signal (3, 5, 3, 5, -5, -3, -5, -3) by discretised linear diffusion. Use the stencil (0.25, 0.5, 0.25) and reflecting boundary conditions. Apply the decorrelation criterion to determine when to stop the iterative process.

Problem 4 (Convex Functionals and Forward Diffusion) (2 points)

Consider the 1-D energy functional

$$E(u) \;=\; \int_a^b \Big((u-f)^2 + \alpha \Psi(u_x^2) \Big)\, dx$$

and its corresponding diffusion process $\partial_t u \;=\; \partial_x(\Psi'(u_x^2)u_x)$. Assume that the regulariser $s \mapsto \Psi(s^2)$ is convex in s such that $\frac{d^2\Psi}{ds^2} \geq 0$. Show that the corresponding diffusion process always performs forward diffusion (and can therefore not act contrast-enhancing).

Deadline for submission: Friday, May 23, 10 am (before the lecture).