Lecture 13:

Variational Methods IV: Functionals of Two Variables

Contents

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

Motivation

- So far, our variational approaches for image restoration obtain a processed image u as a minimiser of some energy functional $E(u)$.
- Let us now study variational approaches that minimise an energy functional $E(u, v)$ that depends on *two* functions:
  the restored image u and a corresponding edge indicator function v.
- Approaches of this type give
  - additional insights into the time-lagged diffusivity approach from Lecture 11
  - an algorithm for approximating an important variational segmentation method.
- However, first we must learn how to minimise energy functionals of two variables.

Minimisation of Functionals of Two Variables

A minimiser of the 2-D energy functional

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

satisfies necessarily *Euler–Lagrange equations for u and v*:

$$\begin{aligned} 0 &= F_u - \partial_{x_1} F_{u_{x_1}} - \partial_{x_2} F_{u_{x_2}}, \\ 0 &= F_v - \partial_{x_1} F_{v_{x_1}} - \partial_{x_2} F_{v_{x_2}} \end{aligned}$$

with boundary conditions

$$\mathbf{n}^\top \begin{pmatrix} F_{u_{x_1}} \\ F_{u_{x_2}} \end{pmatrix} = 0 \quad \text{and} \quad \mathbf{n}^\top \begin{pmatrix} F_{v_{x_1}} \\ F_{v_{x_2}} \end{pmatrix} = 0,$$

where $\mathbf{n}$ denotes a normal vector to the image boundary $\partial\Omega$.

The Euler–Lagrange equations are the steady state of the *gradient descent equations*

$$\begin{aligned} \partial_t u &= -\gamma \cdot (F_u - \partial_{x_1} F_{u_{x_1}} - \partial_{x_2} F_{u_{x_2}}), \\ \partial_t v &= -\gamma \cdot (F_v - \partial_{x_1} F_{v_{x_1}} - \partial_{x_2} F_{v_{x_2}}) \end{aligned}$$

with an arbitrary speed factor $\gamma > 0$.

This generalises results from Lectures 10 and 11 for functionals of a single variable.

Half-Quadratic Regularisation

Basic Idea (Geman / Yang 1995)

- want to avoid nonquadratic functional $E(u)$ since it creates a nonlinear problem
- replace $E(u)$ by a so-called *half-quadratic functional* $E_{HQ}(u,v)$ such that
  - $E_{HQ}(u,v)$ has the same minimiser w.r.t. u as $E(u)$,
  - the auxiliary variable v marks the location of discontinuities,
  - $E_{HQ}(u,v)$ is quadratic in u and convex in v,
  - its minimisation leads to
    - a sequence of linear problems in u,
    - simple nonlinear evaluations in v.
- This is algorithmically convenient and gives direct access to the edge information v.

How Does This Work in Practice ?

- Consider the *nonquadratic* functional from Lecture 10:

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

with the Charbonnier regulariser

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

- Its Euler-Lagrange equation is given by the *nonlinear* PDE

$$0 \;=\; (u-f) \,-\, \alpha\,\mathrm{div}\,\left(\Psi'(|\nabla u|^2)\,\nabla u\right). \qquad (1)$$

Half-Quadratic Regularisation (3)

- **Goal:** Its solution should also minimise an energy $E_{HQ}(u,v)$ that is *quadratic* in u and thus creates *only linear* problems.

- It should have the following structure:

$$E_{HQ}(u,v) \;:=\; \int_\Omega \Big((u-f)^2 \,+\, \alpha \cdot \big(v \cdot |\nabla u|^2 + \eta(v)\big) \Big)\, d\mathbf{x}$$

  with a suitable function η that we still have to find.

- $E_{HQ}(u,v)$ has two Euler-Lagrange equations:

$$\begin{aligned} 0 &= (u-f) \,-\, \alpha \,\mathrm{div}\,(v\,\nabla u) && (2) \\ 0 &= |\nabla u|^2 \,+\, \eta'(v) && (3) \end{aligned}$$

- How can we tranform the single Euler-Lagrange equation (1) into (2)–(3) ?

Half-Quadratic Regularisation (4)

- Equation (2) results by denoting the diffusivity $\Psi'(|\nabla u|^2)$ by an auxiliary variable v:

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

- To obtain the $|\nabla u|^2$ term in Equation (3), we solve v for $|\nabla u|^2$:

$$|\nabla u|^2 = -\lambda^2 \cdot \big(1 - \tfrac{1}{v^2}\big)$$

- Then we can rewrite the Euler-Lagrange equation (1) as two equations:

$$\begin{aligned} 0 &= (u-f) \,-\, \alpha \,\mathrm{div}\,(v\,\nabla u) \\ 0 &= |\nabla u|^2 \,+\, \lambda^2 \cdot \big(1 - \tfrac{1}{v^2}\big). && (4) \end{aligned}$$

- Comparing (4) with (3) shows that $\eta'(v) = \lambda^2 \cdot \big(1 - \frac{1}{v^2}\big)$ and thus $\eta(v) = \lambda^2 \cdot \big(v + \frac{1}{v}\big)$.
- This gives the following functional for u and v:

$$E_{HQ}(u,v) \;:=\; \int_\Omega \Big((u-f)^2 \,+\, \alpha \cdot \Big(v \cdot |\nabla u|^2 \,+\, \lambda^2 \cdot \big(v + \tfrac{1}{v}\big) \Big) \Big)\, d\mathbf{x}.$$

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

Minimisation by Half-Quadratic Regularisation

- Half-quadratic regularisation inspires a simple alternating minimisation algorithm for which convergence can be proven in the convex case.
- Starting with the initialisation $u^0 := f$ one proceeds iteratively.
- For $k = 0, 1, 2, \ldots$ do:
  - **Minimise $E_{HQ}(u^k, v^k)$ with respect to v^k:**
    i.e. compute the diffusivity $v^k := \Psi'(|\nabla u^k|^2)$.
  - **Minimise $E_{HQ}(u^{k+1}, v^k)$ with respect to u^{k+1}:**
    i.e. solve a linear PDE in u^{k+1}:

    $$0 \;=\; (u^{k+1} - f) \;-\; \alpha \, \mathrm{div}\,(v^k \, \nabla u^{k+1}).$$

    Can be achieved e.g. with a finite difference discretisation and solving the resulting linear system with SOR (cf. Lecture 11).
- *This is equivalent to the time lagged diffusivity method from Lecture 11.* (Weickert et al. 1999, Chan/Nikolova 2007).
- Hence, half-quadratic regularisation gives an alternative foundation for this linearisation method.

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

The Mumford–Shah Functional (1985)

Basic Structure:

- *the* prototype of energy-based image segmentation
- seeks segmentation (u, K) some image f as the minimiser of

  $$E_{MS}\,(u, K) \;:=\; \beta \int_{\Omega} (u - f)^2 \, d\mathbf{x} \;+\; \int_{\Omega \setminus K} |\nabla u|^2 \, d\mathbf{x} \;+\; \alpha \, |K|$$

  where u is smoothed version of f, and K represents its edges
- first term penalises deviations from original image f
- second term penalises variations within each segment
- third term penalises the edge length $|K|$

Theoretical Results: (e.g. Morel/Solimini 1994)

- a mathematically very difficult free boundary problem:
  unknown edge set K
- What about well-posedness?
  - existence of a minimiser with closed edge set K
  - uniqueness in general not true
  - K has piecewise C^1 arcs
- no corners or T-junctions, but triple points with $120°$ angle

Ambrosio–Tortorelli Approximation

Basic Structure

- In 1992, Ambrosio and Tortorelli suggested to approximate the Mumford–Shah functional by

$$E_{AT}(u,v) := \int_\Omega \left(\beta \cdot (u-f)^2 + v^2 \cdot |\nabla u|^2 + \alpha \cdot \left(c |\nabla v|^2 + \tfrac{(1-v)^2}{4c} \right) \right) d\mathbf{x}$$

  with an unknown smooth edge indicator function $v(\mathbf{x})$:
  $v \to 0$ at edges, and $v \to 1$ within a region
- proved convergence to the Mumford–Shah functional for $c \to 0$
  (in the sense of so-called Γ-convergence)
- advantage: smooth unknown function v is mathematically more convenient than an unknown edge set K
- Now we have to minimise an energy functional of *two* unknown functions u and v.

Ambrosio–Tortorelli Approximation (2)

- We want to minimise a functional $E(u,v) := \int_\Omega F(\mathbf{x}, u, \nabla u, v, \nabla v)\, d\mathbf{x}$ with

$$
\begin{aligned}
F &= \beta \cdot (u-f)^2 + v^2 \cdot |\nabla u|^2 + \alpha \cdot \left(c|\nabla v|^2 + \tfrac{(1-v)^2}{4c} \right), \\
F_u &= 2\beta\,(u-f), \\
F_{u_{x_1}} &= 2v^2 u_{x_1}, \\
F_{u_{x_2}} &= 2v^2 u_{x_2}, \\
F_v &= 2v|\nabla u|^2 - \tfrac{2\alpha}{4c}\,(1-v), \\
F_{v_{x_1}} &= 2\alpha c v_{x_1}, \\
F_{v_{x_2}} &= 2\alpha c v_{x_2}.
\end{aligned}
$$

- This gives the Euler-Lagrange equations

$$
\begin{aligned}
0 &= \beta(u-f) - \operatorname{div}\,(v^2\, \nabla u), \\
0 &= \tfrac{v}{\alpha}\, |\nabla u|^2 - \tfrac{1-v}{4c} - c\Delta v
\end{aligned}
$$

with homogeneous Neumann boundary conditions on $\partial\Omega$:

$$0 = \partial_{\mathbf{n}} u \quad \text{and} \quad 0 = \partial_{\mathbf{n}} v.$$

Ambrosio–Tortorelli Approximation (3)

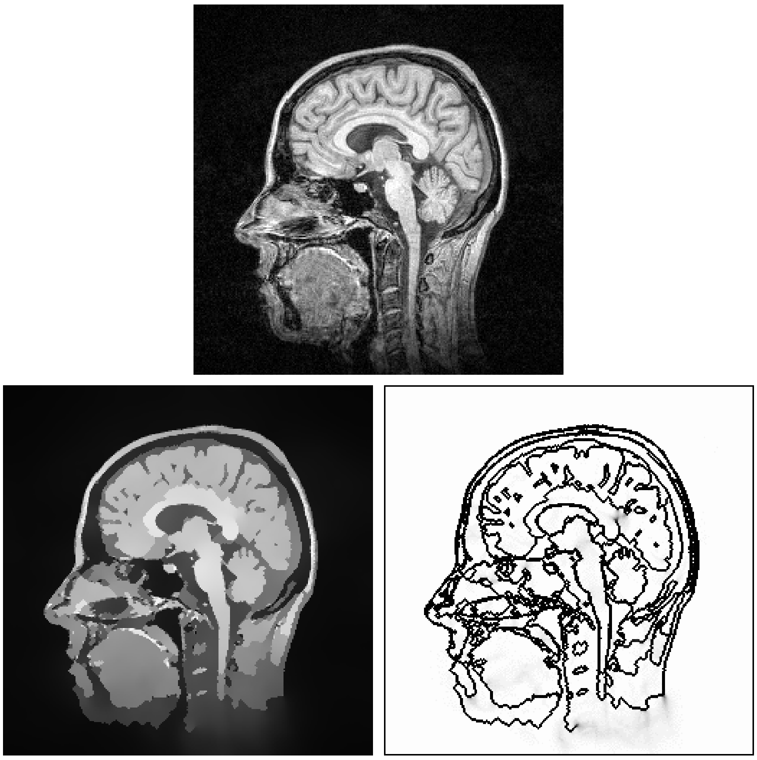
Properties

- Gradient descent interpretation:
  The Euler-Lagrange equations can be regarded as steady state $(t \to \infty)$ of the diffusion–reaction system

$$
\begin{aligned}
\partial_t u &= \operatorname{div}\,(v^2 \nabla u) + \beta \cdot (f-u), \\
\partial_t v &= c\,\Delta v - \tfrac{v}{\alpha}\,|\nabla u|^2 + \tfrac{1-v}{4c}
\end{aligned}
$$

- However, the functional $E_{AT}(u,v)$ is not jointly convex in (u,v) and may thus have multiple *local* minima.
- Consequently, the diffusion–reaction system may get stuck in a local minimum.
- no well-posedness results available yet
- simple discretisation with modified explicit scheme (cf. Lecture 11) seems to work well in practice

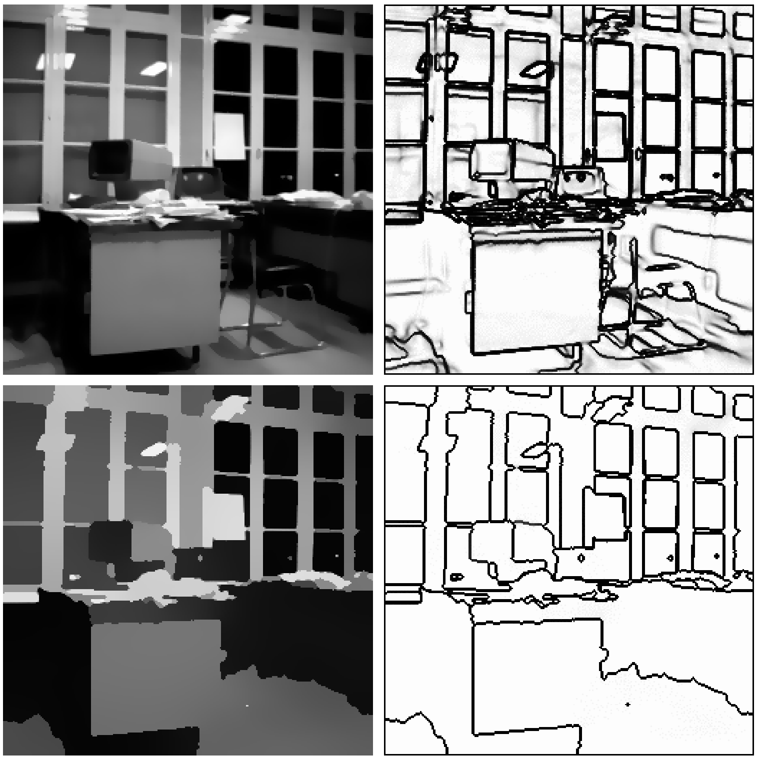
The Ambrosio–Tortorelli Approximation (4)



(a) **Top:** Test image, $\Omega = [0, 256]^2$. (b) **Bottom left:** Function u of the Ambrosio–Tortorelli segmentation with $\beta = 0.01$, $\alpha = 1$, $c = 0.01$. (c) **Bottom right:** Corresponding edge function v.

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

The Ambrosio–Tortorelli Approximation (5)



Influence of the parameter β. (a) **Top left:** Ambrosio–Tortorelli segmentation, $\beta = 0.1$, $\alpha = 1$, $c = 0.01$. (b) **Top right:** Corresponding edge function v. (c) **Bottom left:** Ambrosio–Tortorelli segmentation, $\beta = 0.001$, $\alpha = 1$, $c = 0.01$. (d) **Bottom right:** Corresponding edge function v.

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

Relations to Perona–Malik Diffusion

- Neglecting $c\,\Delta v$ in the second Euler–Lagrange equation

$$0 \;=\; \tfrac{v}{\alpha}\,|\nabla u|^2 - \tfrac{1-v}{4c} - c\Delta v$$

allows to solve directly for v (Chan/Vese 1997):

$$v \;=\; \frac{1}{1+\frac{4c}{\alpha}\,|\nabla u|^2}$$

- The remaining first Euler–Lagrange equation becomes

$$\frac{u-f}{1/\beta} \;=\; \operatorname{div}\,\big(g(|\nabla u|^2)\,\nabla u\big)$$

with the diffusivity

$$g(|\nabla u|^2) \;=\; v^2 \;=\; \frac{1}{\left(1+\frac{4c}{\alpha}\,|\nabla u|^2\right)^2}$$

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

Conclusion:

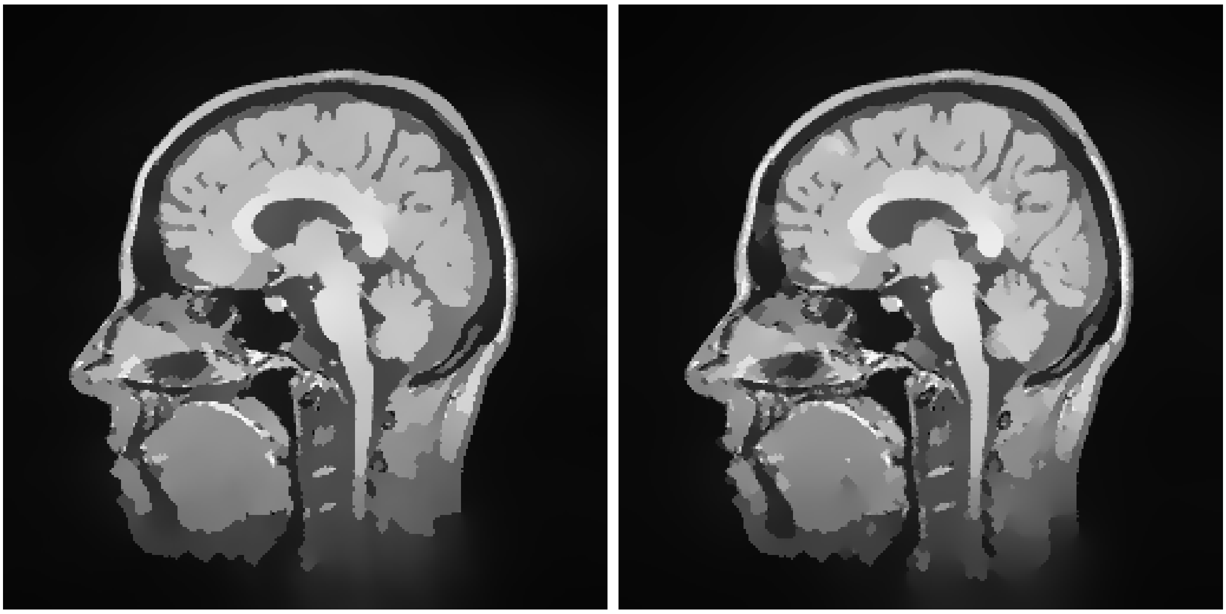
The Ambrosio–Tortorelli approximation to the Mumford–Shah functional approximates a Perona–Malik filter with diffusivity

$$g(s^2) \;=\; \frac{1}{\left(1+\frac{s^2}{3\lambda^2}\right)^2}\,,$$

contrast parameter $\lambda = \sqrt{\frac{\alpha}{12c}}$*, and diffusion time* $t = 1/\beta$*.*

(The contrast parameter λ has been chosen such that it separates forward from backward diffusion; cf. Lecture 4.)

Relations to Perona–Malik Diffusion (3)



Comparison between Ambrosio–Tortorelli and Perona–Malik. (a) **Left:** Ambrosio–Tortorelli segmentation, $\beta = 0.01$, $\alpha = 1$, $c = 0.01$. (b) **Right:** Perona–Malik filter with $g(s^2) = 1/(1+\frac{s^2}{3\lambda^2})^2$, $\lambda = 2.89$, and $t = 100$.

Summary

Summary

- Half-quadratic regularisation replaces a nonquadratic functional $E(u)$ by a functional $E_{HQ}(u, v)$ that is quadratic in u.
- Its corresponding alternating minimisation is equivalent to the time lagged diffusivity method.
- The Mumford-Shah model is the prototype of an energy-based segmentation method.
- It has well-posedness problems with respect to uniqueness.
- Its free segmentation boundaries cause mathematical problems.
- The Mumford–Shah functional can be approximated by the Ambrosio–Tortorelli functional with a fuzzy edge indicator function.
- This gives two coupled diffusion–reaction equations that can be related to Perona–Malik diffusion.
- Well-posedness questions have not been addressed for the Ambrosio–Tortorelli functional yet.

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

References

- D. Geman, C. Yang: Nonlinear image recovery with half-quadratic regularization. *IEEE Transactions on Image Processing*, Vol. 4, 932–945, 1995
  (introduced half-quadratic regularization in a stochastic setting).
- P. Charbonnier, L. Blanc–Féraud, G. Aubert, M. Barlaud: Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*, Vol. 6, 298–311, 1997
  (half-quadratic regularization from a deterministic viewpoint)
- J. Weickert, J. Heers, C. Schnörr, K.J. Zuiderveld, O. Scherzer, H.S. Stiehl: Fast parallel algorithms for a broad class of nonlinear variational diffusion approaches. *Real-Time Imaging*, Vol. 7, 31–45, 2001.
  (`http://www.mia.uni-saarland.de/weickert/publications.shtml`)
  (equivalence of half-quadratic minimisation and time-lagged diffusivity method)
- M. Nikolova, R. Chan: The equivalence of half-quadratic minimization and the gradient linearization iteration. *IEEE Transactions on Image Processing*, Vol. 16, No. 6, 1623–1627, 2007.
  (`http://www.cmla.ens-cachan.fr/fileadmin/Membres/nikolova/publisA.html`)
  (more detailed analysis of this equivalence).
- D. Mumford, J. Shah: Boundary detection by minimizing functionals, I, *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '85, San Francisco, June 19–23, 1985)*, IEEE Computer Society Press, Washington, 22–26, 1985.
  (`http://www.math.neu.edu/~shah/publications.html`)
  (introduced the Mumford–Shah functional)

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

- D. Mumford, J. Shah: Optimal approximations by piecewise smooth functions and associated variational problems. *Communications in Pure and Applied Mathematics*, Vol. 42, 577–685, 1989.
  (`http://www.math.neu.edu/~shah/publications.html`)
  (detailed theoretical study of the Mumford–Shah functional)
- J.-M. Morel, S. Solimini: *Variational Methods in Image Segmentation*, Birkhäuser, Boston, 1994.
  (monograph that covers all theoretical results for the Mumford–Shah functional until 1994)
- L. Ambrosio, V. Tortorelli: Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence. *Bolletino della Unione Matematica Italiana*, Vol. 7, 105–123, 1992.
  (introduced the Ambrosio–Tortorelli approximation)
- T. Richardson, S. Mitter: Approximation, computation, and distortion in the variational formulation. In B. M. ter Haar Romeny (Ed.): *Geometry-Driven Diffusion in Computer Vision*, Kluwer, Dordrecht, 169–190, 1994.
  (experimental results for the Ambrosio–Tortorelli approximation)
- T. Chan, L. Vese: *Variational Image Restoration and Segmentation: Models and Approximations*, CAM Report 97-47, Dept. of Mathematics, Univ. of California, Los Angeles, USA, 1997.
  (`www.math.ucla.edu/applied/cam/`)
  (studies the Perona–Malik–like simplification of the Ambrosio–Tortorelli functional)