Lecture 5

- Linear and Isotropic Nonlinear Diffusion
- Morphological Invariance and Curvature-Dependent Motion
- Variational Functionals
- Euler–Lagrange Equations and Gradient Descent Diffusion
- Processes as Gradient Descents
- Curvature Motion as Gradient Descent

Linear Diffusion, Motivation

Equilibration of quantity u

- Concentrations of substances
- Thermal energy (*heat flow*)
- **Image processing: grey value**

Linear Diffusion

- Let $u = u(x,t)$ be given in each point x of a domain $D \in \mathbb{R}^d$ and for each $t \geq 0$
- Assume that gradients of u generate a proportional mass/energy transport (**flux**)

$$j = \nabla u$$

- Flux moves mass/energy from one place to another

$$\partial_t u = \operatorname{div} j$$

- Assume that no flux over the boundary of D takes place (Neumann boundary condition)
- Initial condition: $u(x,0) = f(x)$, where f is input image
- Resulting **linear diffusion equation**

$$u_t = \operatorname{div}(\nabla u) = \Delta u$$

- Removes inhomogeneities efficiently – useful for denoising of images

Linear Diffusion Example

Figure: Original image (top left) smoothed by linear diffusion, $t = 2.5$, $t = 10$, $t = 40$.

Nonlinear Isotropic Diffusion

- Assume the flux j is no longer strictly proportional to the gradient but depends on it via a **diffusivity function** $\varphi : \mathbb{R}_0^+ \to \mathbb{R}^+$

$$j = \varphi(\|\nabla u\|)\, \nabla u$$

- Typically, φ is decreasing and positive
- Then we have nonlinear isotropic diffusion

$$u_t = \operatorname{div}(\varphi\, \nabla u)$$

Isotropic Nonlinear Diffusion Example

Figure: Original image (top left) smoothed by isotropic nonlinear diffusion, $t = 2.5$, $t = 10$, $t = 40$.

Properties of Diffusion Processes

- The mass transport $u_t = \mathrm{div}(j)$ ensures *mass conservation*
- For $t \to \infty$, u tends towards a constant (equilibrium)

$$\lim_{t\to\infty} u(x,t) = \frac{\int\limits_D f(x,0)\,\mathrm{d}x}{\int\limits_D \mathrm{d}x}$$

- Maximum–minimum principle

$$\inf_{\xi\in D} f(\xi,0) \le f(x,t) \le \sup_{\xi\in D} f(\xi,0)$$

for all $x \in D$

Morphological Invariance and Curve Evolutions

- Every grey-value image evolution can be represented as level line evolution using the equivalence

$$u_t = \beta \, \|\nabla u\| \qquad \Leftrightarrow \qquad c_t = \beta \vec{n}$$

- For example, for linear diffusion one has

$$\beta = \frac{\Delta u}{\|\nabla u\|}$$

- **For which image evolutions does morphological invariance hold?**
  - Morphological invariance means that the evolution of each level line depends only on the level line itself
  - This is the case if and only if β is a function of the curvature κ
- Consequently, one is especially interested in evolutions with $\beta = \beta(\kappa)$
- Obviously, this includes dilation and erosion ($\beta = \text{const}$) as well as curvature motion ($\beta = \kappa$)

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 28 29 30

Inverse Curvature Flow

- Consider the curve evolution

$$c_t = \pm c_{\vartheta\vartheta}$$

for a strictly convex curve c parametrised by *direction angle* ϑ

- Note that $\vartheta_s = \kappa$ where s is arc-length parameter
- This evolution is related to **inverse curvature flow**

$$c_t = \pm \kappa^{-1} \vec{n} \,, \qquad u_t = \pm \kappa^{-1} \, \|\nabla u\|$$

- In $+$ direction, inverse curvature flow is instable and generates singularities
- In $-$ direction, convex curves remain convex and are smoothed into circles

Inverse Curvature Flow, Example

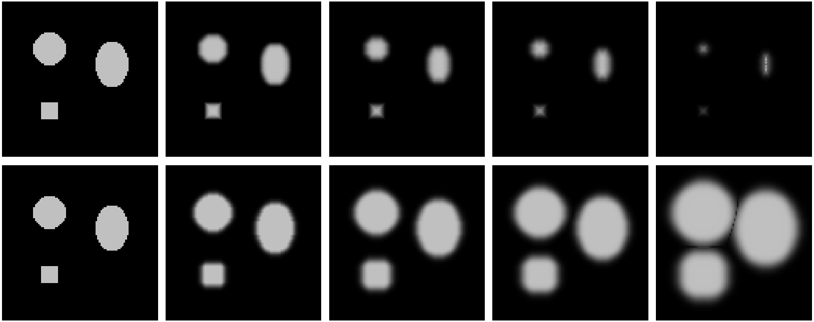

Figure: Top, left to right: A synthetic test image, after 100, 200, 300, 500 iterations of inverse Euclidean curvature flow ("+" direction, $\tau = 0.02$); **bottom:** same in "$-$" direction

Variational Problems

Formulation of image processing task as optimisation problem

- Given image $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^d$
- Search for processed image $u : D \to \mathbb{R}$ such that

$$E(u) := \int_D F(x, u, u_{x_i})\, \mathrm{d}x$$

  is minimised

  - $E(u)$ is called **functional**, often also **energy functional**
  - The integrand F depends on $x \in D$, u and the derivatives of u w.r.t. all variables x_i
  - F also depends on f, though we do not represent f as argument of F since it is considered fixed

- If f and F are sufficiently smooth, E convex, and u is restricted to a suitable space of functions, a unique minimising function u exists
- Under weaker conditions (e.g. if E is non-convex, but smoothness conditions apply), there might be multiple local minima

Variational Gradient

- We want to use **steepest descent** to find a minimum of E (global minimum if E is convex, local minimum otherwise)
- In classical analysis, the direction of steepest descent is given by the gradient
- Need therefore analog of the gradient for (infinite-dimensional) function spaces

Variational Gradient

- Functions f (images, e.g.) form some function space V (similar to a manifold but infinite-dimensional)
- A *direction vector in function space* is given by a (sufficiently smooth) function v on D. In each point $u \in V$, these functions form a vector space $\mathrm{T}_u V$ (an infinite-dimensional tangent space)
- Directional derivative of E in the direction v:

$$\delta_v E := \frac{\mathrm{d}}{\mathrm{d}\varepsilon} E(u^* + \varepsilon v)\Big|_{\varepsilon=0}$$

- Given a scalar product $\langle\,\cdot\,,\,\cdot\,\rangle$ for functions in $\mathrm{T}_u V$, there is one function $g \in \mathrm{T}_u V$ such that

$$\delta_v E = \langle g, v \rangle$$

  for all admissible functions v

- Define

$$\frac{\delta}{\delta u} E(u) := g$$

  as **variational gradient** of E w.r.t. u

Variational Gradient Computation

Compute therefore (assuming sufficient smoothness)

$$
\begin{aligned}
E(u^* + \varepsilon v) &= \int_D F(x, u^* + \varepsilon v, \partial_{x_i} u\big|_{u=u^*+\varepsilon v})\,\mathrm{d}x \\
&= \int_D \Bigg(F(x, u^*, u^*_{x_i}) + \varepsilon v \frac{\partial F}{\partial u}(x, u^*, u^*_{x_i}) \\
&\qquad + \sum_{i=1}^{d} \varepsilon v_{x_i} \frac{\partial F}{\partial u_{x_i}}(x, u^*, u^*_{x_j}) + O(\varepsilon^2) \Bigg)\,\mathrm{d}x \\
&= E(u^*) + \varepsilon \Bigg(\int_D v \frac{\partial F}{\partial u}(x, u^*, u^*_{x_i})\,\mathrm{d}x + \int_{\partial D} (\ldots)\,\mathrm{d}x \\
&\qquad - \sum_{i=1}^{d} \int_D v \frac{\mathrm{d}}{\mathrm{d}x_i} \frac{\partial F}{\partial u_{x_i}}(x, u^*, u^*_{x_j})\,\mathrm{d}x \Bigg) + O(\varepsilon^2)
\end{aligned}
$$

Variational Gradient Computation, cont.

Thus we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon} E(u^* + \varepsilon v)\Big|_{\varepsilon=0} = \int_D v \left(\frac{\partial F}{\partial u}(x, u^*, u^*_{x_i}) - \sum_i \frac{\mathrm{d}}{\mathrm{d}x_i} \frac{\partial F}{\partial u_{x_i}}(x, u^*, u^*_{x_j}) \right) \mathrm{d}x + \int_{\partial D} (\dots)\, \mathrm{d}x$$

- The integral $\int_{\partial D}$ integrates normal derivatives over the boundary of D and is responsible for the boundary conditions
- It vanishes if only functions u, v with vanishing normal derivatives on the boundary are considered (Neumann boundary conditions)
- We assume in the following that this is the case. Otherwise, the equations remain valid for inner points of D

Variational Gradient

- Then we can rewrite (if $\int_{\partial D} \ldots$ vanishes) with some $g = g(x)$

$$\delta_v E = \langle v, g \rangle$$

- g is the sought variational gradient of E w.r.t. u

Remarks

1. For the L^2 scalar product $\langle v, w \rangle := \int_D vw \, \mathrm{d}x$ one has $\delta_v E = \int_D gv \, \mathrm{d}x$
2. By Cauchy-Schwartz inequality, the unit vector

$$v^* = \frac{g}{\sqrt{\langle g, g \rangle}}$$

   maximises $\delta_v E$ among all v with $\|v\| = 1$ (i.e., vectors on the unit sphere). Thus, the gradient is the direction of largest directional derivative
3. Clearly, the scalar product $\langle \cdot \, , \, \cdot \rangle$ on $\mathrm{T}_u V$ has the role of a Riemannian metric on V

Variational Gradient Descent

- **Gradient descent.** The steepest descent for u is given by

$$u_t = -g = -\frac{\delta}{\delta u} E \; .$$

Remarks

1. Equally, $u_t = -Cg$ is a steepest descent for u provided C is arbitrary positive (may depend on t, but not on x and u). For example, $u_t = -v^*$ works, too
2. For the minimiser u^*, the variational gradient must vanish. Then one has the well-known Euler–Lagrange equations

Linear Diffusion as Gradient Descent

- Consider the functional

$$E(u) = \frac{1}{2} \int_D \|\nabla u\|^2 \, \mathrm{d}x$$

with the standard L^2 scalar product

$$\langle v, w \rangle = \int_D v(x) \cdot w(x) \, \mathrm{d}x$$

- Then we have

$$F(x, u, u_{x_i}) = \frac{1}{2} \sum_i u_{x_i}^2$$

$$g(x) = \frac{\partial F}{\partial u} - \sum_i \frac{\mathrm{d}}{\mathrm{d}x_i} \frac{\partial F}{\partial u_{x_i}} = -\operatorname{div} \nabla u = -\Delta u$$

- Gradient descent therefore

$$u_t = \Delta u$$

i.e. the linear diffusion equation

Isotropic Nonlinear Diffusion as Gradient Descent

- Consider the functional

$$E(u) = \frac{1}{2} \int_D \Phi\left(\|\nabla u\|^2 \right) \mathrm{d}x$$

  with an increasing smooth function

$$\Phi : \mathbb{R}_0^+ \to \mathbb{R}_0^+$$

  and the standard L^2 scalar product

- Then we have

$$F(x, u, u_{x_i}) = \frac{1}{2} \Phi \left(\sum_i u_{x_i}^2 \right)$$

$$g(x) = - \sum_i \frac{\mathrm{d}}{\mathrm{d}x_i} \left(\Phi' \left(\sum_j u_{x_j}^2 \right) u_{x_i} \right)$$

$$= - \operatorname{div} \left(\Phi'\left(\|\nabla u\|^2 \right) \nabla u \right)$$

Isotropic Nonlinear Diffusion as Gradient Descent, cont.

- Gradient descent therefore

$$u_t = \operatorname{div}\left(\Phi'\left(\,\|\nabla u\|^2\,\right)\nabla u\right)$$

i.e. isotropic nonlinear diffusion with diffusivity $\varphi \equiv \Phi'$

Metric for Variational Gradient

Consider in more detail the scalar product $\langle\cdot,\cdot\rangle$ (expressing the metric on the function space V):

- The function space under consideration consists of functions v on D such that $v(x) \in T_{u(x)}\mathbb{R}(\equiv \mathbb{R})$ for each $x \in D$
- The metric therefore depends on the metric of D and the metric of the range of values of u. Both can be the standard metric, but also other metrics!
- Typically,

$$\langle v, w\rangle = \int_D \omega(x, u(x))v(x)w(x)\,\mathrm{d}x$$

  with some positive-valued function ω

Remark. Changing metrics on the image range $\mathbb{R}$ is essentially a rescaling of $\mathbb{R}$. This might appear trivial at first glance but can be useful in conjunction with minimisation of energy functionals (see next slide).
Additional possibilities arise if higher-dimensional manifolds are used as image range instead of $\mathbb{R}$ (examples in later lectures).

Modified Metric on Image Range

- Consider again the functional

$$E(u) = \frac{1}{2} \int_D \Phi(\|\nabla u\|^2)\,\mathrm{d}x$$

  with Φ as before

- Assume now that the metric on the *range* of u is given by

$$\mathrm{d_h} u := \frac{1}{u}\,\mathrm{d_e} u$$

  where $\mathrm{d_e} u$ denotes the standard (Euclidean) metric

- Then we have on the function space

$$\langle v, w \rangle = \int_D \frac{1}{u(x)^2} v(x) w(x)\,\mathrm{d}x$$

Modified Metric on Image Range, cont.

- Thus

$$F(x, u, u_{x_i}) = \frac{1}{2}\Phi\left(\sum_i u_{x_i}^2\right)$$

$$g(x) = (u(x))^2 \cdot \left(\frac{\partial F}{\partial u} - \sum_i \frac{\mathrm{d}}{\mathrm{d}x_i}\frac{\partial F}{\partial u_{x_i}}\right)$$

$$= -u^2 \operatorname{div}(\Phi'(\|\nabla u\|^2)\nabla u)$$

This is the gradient in the metric given by $\mathrm{d}_h u$, i.e.

$$\frac{\delta E}{\delta_{\mathrm{h}} u} = -u^2 \operatorname{div}(\Phi'(\|\nabla u\|^2)\nabla u)$$

Modified Metric on Image Range, cont.

Transformation to standard (Euclidean) metric:

$$\frac{\delta E}{\delta_{\mathrm{e}} u} = \frac{\mathrm{d_h} u}{\mathrm{d_e} u} \cdot \frac{\delta E}{\delta_{\mathrm{h}} u} = -u\,\mathrm{div}(\Phi'(\|\nabla u\|^2)\nabla u)$$

- Gradient descent finally

$$u_t = -\frac{\delta E}{\delta_{\mathrm{e}} u} = u\,\mathrm{div}(\Phi'(\|\nabla u\|^2)\nabla u)$$

- An interesting application for this idea will be demonstrated in a later lecture

Modified Metric on Image Domain

- If the metric on the image *domain* is modified, the change also affects the functional
- Assume we have on $D \subset \mathbb{R}^2$ the metric

$$g = \begin{pmatrix} x_1^2 + x_2^2 & 0 \\ 0 & x_1^2 + x_2^2 \end{pmatrix}$$

i.e. $\mathrm{d}_{\mathrm{weighted}} x = (x_1^2 + x_2^2)\,\mathrm{d}_{\mathrm{e}} x$

- Modified metric therefore

$$\langle v, w \rangle = \int_D v(x) \cdot w(x) (x_1^2 + x_2^2)\,\mathrm{d}x$$

Modified Metric on Image Domain, cont.

- The variational functional of *linear* diffusion then becomes

$$
\begin{aligned}
E(u) &= \frac{1}{2} \int_D \|\nabla u\|^2 \, \mathrm{d}x \\
&= \frac{1}{2} \int_D \left(\left(\frac{\partial u}{\partial_{\mathrm{e}} x_1} \right)^2 + \left(\frac{\partial u}{\partial_{\mathrm{e}} x_2} \right)^2 \right) \frac{x_1^2 + x_2^2}{x_1^2 + x_2^2} \, \mathrm{d}_{\mathrm{e}} x \\
&= \frac{1}{2} \int_D \frac{1}{x_1^2 + x_2^2} \left(\left(\frac{\partial u}{\partial_{\mathrm{e}} x_1} \right)^2 + \left(\frac{\partial u}{\partial_{\mathrm{e}} x_2} \right)^2 \right) \mathrm{d}_{\mathrm{weighted}} x
\end{aligned}
$$

- Gradient:

$$
g(x) = -\frac{1}{x_1^2 + x_2^2} \Delta u \; ,
$$

gradient descent therefore

$$
u_t = \frac{1}{x_1^2 + x_2^2} \Delta u
$$

Curvature Motion as Gradient Descent

- Consider closed curve c in the plane, with variational functional

$$L(c) = \oint_c \|c_p(p)\| \,\mathrm{d}p$$

(the arc-length of c)

- Consider curve flows $c_t = v\vec{n}$ with scalar-valued functions v and metric

$$\langle v, w\rangle = \oint_c v(p)w(p)\frac{\mathrm{d}s}{\mathrm{d}p}\,\mathrm{d}p = \oint_c v(p)w(p)\,\|c_p(p)\|\,\mathrm{d}p$$

- Assuming arc-length parametrisation at $t = 0$, the variation is

$$\delta_{v\vec{n}}L = -\oint_c v\kappa\,\mathrm{d}s$$

leading to gradient descent

$$c_t = \kappa\vec{n}$$

Curvature Motion as Gradient Descent

- Curvature flow can be considered as gradient descent for the arc-length of closed level lines
- Since there are no closed geodesics in the plane, curvature flow can't stop before all closed curves have disappeared
- In other geometries, however, nontrivial steady states may be possible (see later lectures)

Different Descriptions of Curvature Motion

Image evolution

- PDE in standard notation

$$u_t = \|\nabla u\| \operatorname{div} \frac{\nabla u}{\|\nabla u\|}$$

- PDE using local gradient/level line directions:
  Curvature motion as smoothing along level lines

$$u_t = u_{\xi\xi}$$

  where in each point ξ is a unit vector parallel to the local level line (i.e. $\xi \perp \nabla u$)

well-posed, satisfies maximum–minimum principle for u

Curve evolution of level lines

- Curve evolution PDE

$$c_t = \kappa \vec{n}$$

- *Gradient descent for curve length of level lines*

References

Diffusion in image processing

- T. Iijima: Theory of pattern recognition. Electronics and Communications in Japan, 123–134, 1963 (Preceding papers in Japanese 1959, 1962)
- P. Perona, J. Malik: Scale space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Analysis and Machine Intelligence 12:629–639, 1990
- J. Weickert: A review of nonlinear diffusion filtering. In B. ter Haar Romeny, L. Florack, J. Koenderink, M. Viergever, eds., Scale Space Theory in Computer Vision, Lecture Notes in Computer Science, vol. 1252, 3–28. Springer, Berlin 1997

Variational formulations, gradient descents

- R. Kimmel: Numerical Geometry of Images. Springer, Berlin 2004
- G. Sapiro: Geometric Partial Differential Equations and Image Analysis. Cambridge University Press 2001