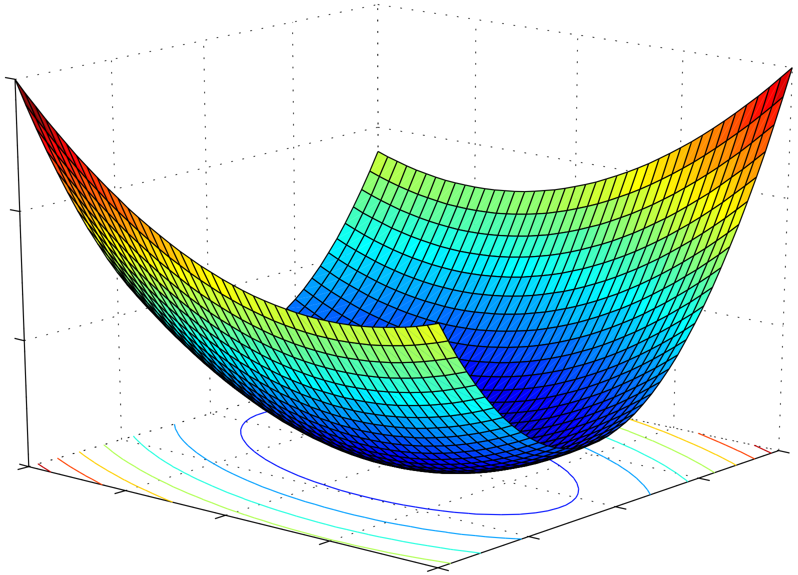
Variational Methods

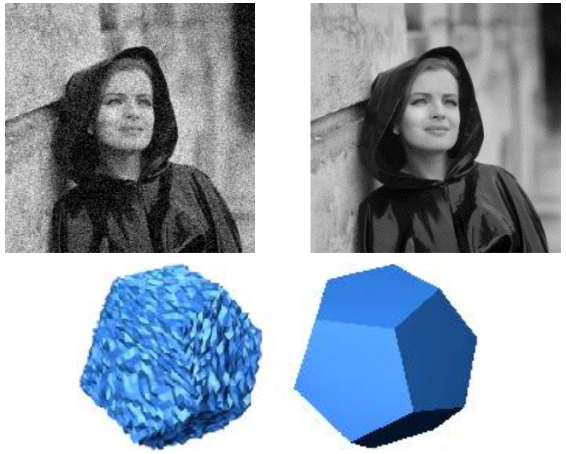


Dr. Martin Oswald

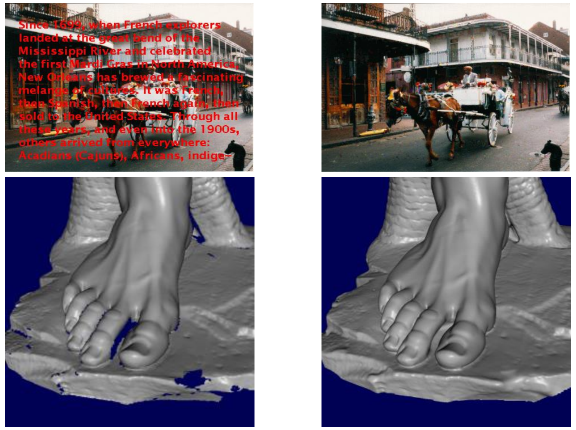
Computer Vision and Geometry Group

Overview and Applications

Denoising

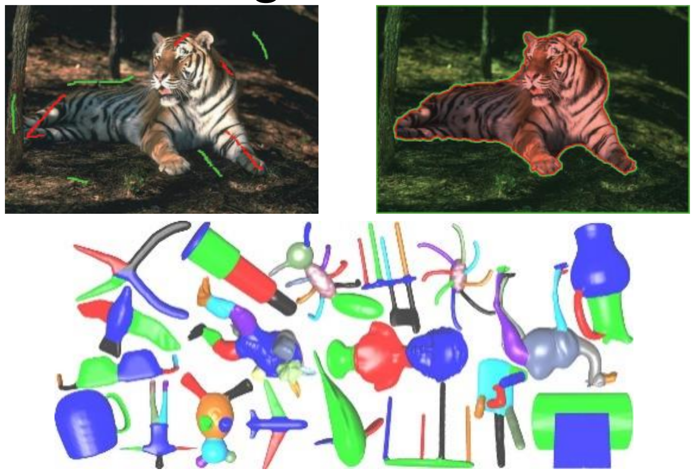


Inpainting

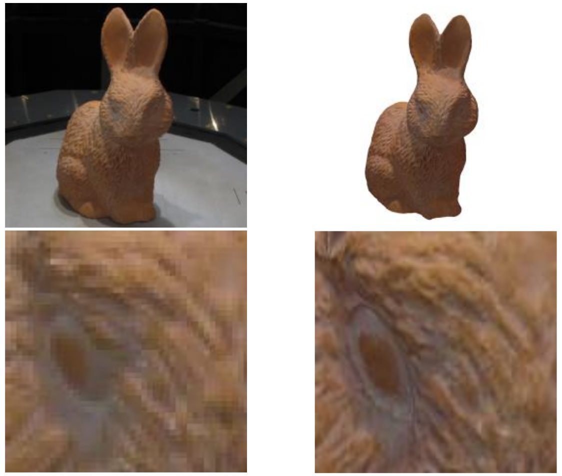


Optical Flow

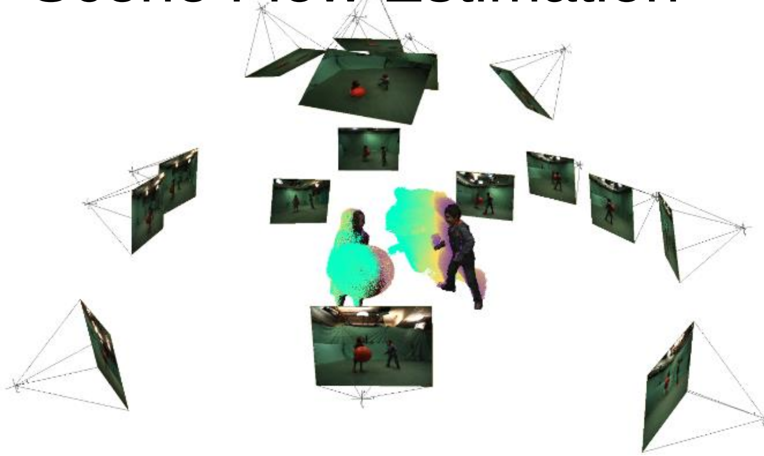
Segmentation



Super-resolution

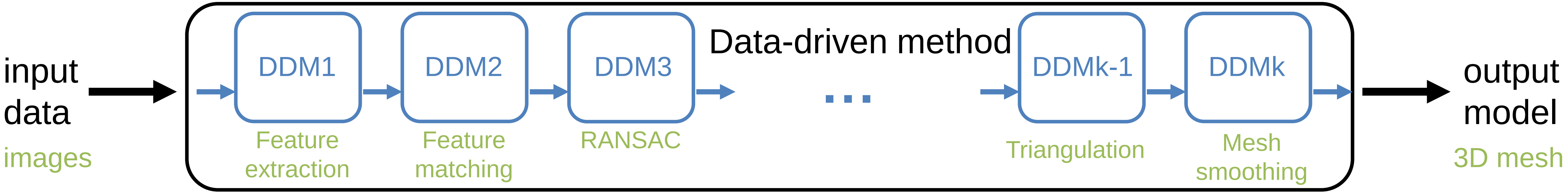


3D Reconstruction & Scene Flow Estimation



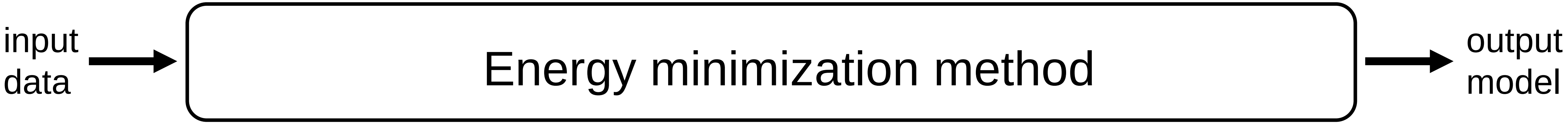
Energy Minimization Methods

Data-driven (Bottom-up) Approach (Example: 3D Reconstruction)



$$\text{some\_output} = f(\text{input})$$

Model-driven (Top-down) Approach



$$\text{best\_output} = \arg\min_{\text{model}} E(\text{model}, \text{input})$$

Energy Minimization Methods

Why energy minimization?

- A very common problem is to deal with noise and outliers. EMM's usually deal well with them and provide a transparent framework
- Mathematical analysis of the cost function allows statements regarding the existence, uniqueness and stability of solutions to a given problem.
- In traditional multistep processes the interplay of consecutive steps is often complex and intransparent. It is typically unclear how modifying or replacing one component affects the subsequent steps.
- Optimization methods are based on transparent and explicitly formulated assumptions, with no "hidden" assumptions.
- In general, optimization methods have fewer parameters. The meaning of each parameter is mostly obvious.
- Optimization methods are easily combined in a transparent manner (by adding respective cost functions).

Image Denoising

We are given a noisy image $f : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ which we assume is corrupted with additive Gaussian noise.

$$f = u + \eta \qquad\qquad \eta \sim \mathcal{G}(0, \sigma)$$

The goal is to recover the clean image $u : \Omega \rightarrow \mathbb{R}$

Original

Noisy Image

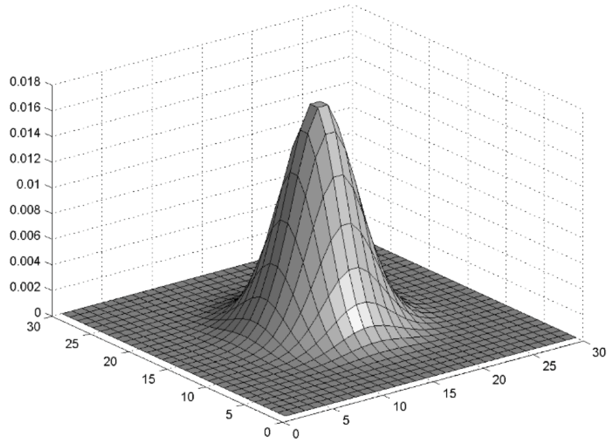
Denoised Result

Gaussian Filtering

Given a kernel $K : \mathbb{R}^2 \to \mathbb{R}$ and an image $u : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^d$ the convolution

$$(K * u)(x,y) := \int_\Omega K(a,b)u(x-a, y-b)\; dadb$$

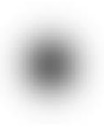
sums up the values of u around (x,y), weighted by K.



Gaussian kernel

$$K(a,b) = G_\sigma := \frac{1}{2\pi\sigma^2} \exp\left[\frac{a^2+b^2}{2\sigma^2}\right]$$

Gaussian convolution



*

=

Motion blur with non-rotational symmetric kernel

*

=

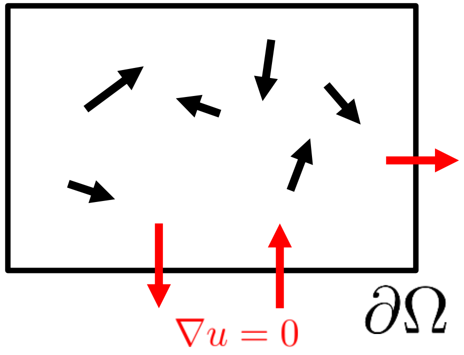
Image Denoising

Gaussian filtering:

- How to choose boundary conditions? (e.g. mirror, repeat, wrap)
- Effect of the kernel size? (e.g. $\sigma \rightarrow \infty$). Effect of kernel truncation and normalization?
- How does the local update criterion effect the image globally?
- Properties of the result image, does it get brighter or darker?
- How to smooth noise, but preserve strong edges? Adaptive kernel size?

Image Diffusion

We consider a grayscale image over time $u : \Omega \times [0, T] \rightarrow \mathbb{R}$ which gray values are diffused over time. The diffusivity $g : \Omega \rightarrow \mathbb{R}$ locally defines the amount diffusion.



Diffusion equation: $\partial_t u = \operatorname{div}(g \nabla u)$

Initial condition: $u(x, y, 0) = f(x, y) \qquad \forall (x, y) \in \Omega$

Boundary condition: $\nabla u(x, y, t) = 0 \qquad \forall t, \forall (x, y) \in \partial\Omega$

analytic solution: $u(x, y, t) = \left(\mathcal{G}_{\sqrt{2t}} * f\right)(x, y)$ for $g \equiv \mathbf{1}_\Omega$

	1	
1	-4	1
	1	

E.g. solve with iterative scheme:

$$\frac{u(x, y, t+1) - u(x, y, t)}{\tau} = \operatorname{div}(g \nabla u)$$

$$u(x,y,t{+}1) = u(x,y,t) + \tau\big(u(x{+}1,y,t) + u(x{-}1,y,t) + u(x,y{+}1,t) + u(x,y{-}1,t) - 4u(x,y,t)\big)$$

Note: The equilibrium state (or steady state) of this diffusion process is a constant image with the average gray value. To get a certain level of smoothness the diffusion process has to be stopped at the right moment.

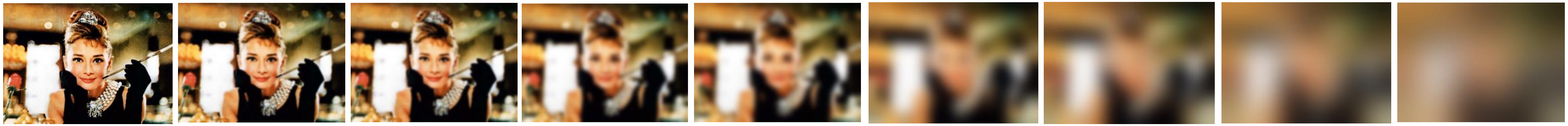


Image Denoising

Gaussian filtering:

- How to choose boundary conditions? (e.g. mirror, repeat, wrap)
- Effect of the kernel size? (e.g. $\sigma \to \infty$). Effect of kernel truncation and normalization?
- How does the local update criterion effect the image globally?
- Properties of the result image, does it get brighter or darker?
- How to smooth noise, but preserve strong edges? Adaptive kernel size?

Image diffusion:

- Small neighborhood structure, needs less memory for large σ
- Equilibrates concentration differences
- Boundary conditions are more intuitive, e.g. Neumann boundary conditions preserve total mass, i.e. avg. image brightness remains constant
- Global optimality criterion?
- When should the diffusion be stopped?
- What is a good time step size τ ? (hidden parameter), Time step size needs to be small for stability (slow convergence).

Denoising via Energy Minimization

Given the noisy image $f : \Omega \to \mathbb{R}$ and a smoothing parameter λ. A denoised image $u : \Omega \to \mathbb{R}$ can be recovered by minimizing the energy

$$E(u) = \int_\Omega \Big(\underbrace{(u-f)^2}_{\text{data similarity}} + \lambda \underbrace{|\nabla u|_2^2}_{\text{smoothness}}\Big)dx$$

Necessary condition for minimum:

$$\frac{dE}{du} = (u-f) - \lambda \operatorname{div}(\nabla u) = 0$$

→ linear system of equations: $Au = b$

→ solve with favorite linear solver

Note: The equilibrium state of this diffusion process is the image that minimizes $E(u)$, not necessarily a constant image.

Image Denoising

Gaussian filtering:

- How to choose boundary conditions? (e.g. mirror, repeat, wrap)
- Effect of the kernel size? (e.g. $\sigma \to \infty$). Effect of kernel truncation and normalization?
- How does the local update criterion effect the image globally?
- Properties of the result image, does it get brighter or darker?
- How to smooth noise, but preserve strong edges? Adaptive kernel size?

Image diffusion:

- Small neighborhood structure, needs less memory for large σ
- Equilibrates concentration differences
- Boundary conditions are more intuitive, e.g. Neumann boundary conditions preserve total mass, i.e. avg. image brightness remains constant
- Global optimality criterion?
- When should the diffusion be stopped?
- What is a good time step size τ ? (hidden parameter), Time step size needs to be small for stability (slow convergence).

Denoising via energy minimization:

- Transparent formulation with global optimality criterion
- One can show existence and uniqueness of solutions
- No boundary conditions. (no “hidden” assumptions)
- Many optimization methods available, solving a linear system with $\lambda = \tau t$ can be much faster than computing t diffusion iterations
- No time step size (no “hidden” parameters)
- Non-linear diffusivity (e.g. for edge preservation) is easy to incorporate
- Concept of separating data and smoothness cost generalizes well and is a common pattern: $E(u) = E_{data}(u) + E_{smooth}(u)$

Image Denoising

data-driven approach

Gaussian filtering:

- How to choose boundary conditions? (e.g. mirror, repeat, wrap)
- Effect of the kernel size? (e.g. $\sigma \rightarrow \infty$). Effect of kernel truncation and normalization?
- How does the local update criterion effect the image globally?
- Properties of the result image, does it get brighter or darker?
- How to smooth noise, but preserve strong edges? Adaptive kernel size?

Image diffusion:

- Small neighborhood structure, needs less memory for large σ
- Equilibrates concentration differences
- Boundary conditions are more intuitive, e.g. Neumann boundary conditions preserve total mass, i.e. avg. image brightness remains constant
- Global optimality criterion?
- When should the diffusion be stopped?
- What is a good time step size τ ? (hidden parameter), Time step size needs to be small for stability (slow convergence).

model-driven approach

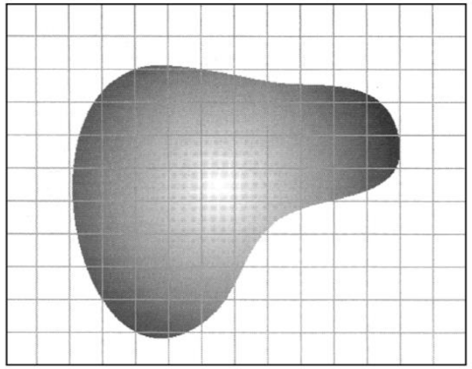
Denoising via energy minimization:

- Transparent formulation with global optimality criterion
- One can show existence and uniqueness of solutions
- No boundary conditions. (no “hidden” assumptions)
- Many optimization methods available, solving a linear system with $\lambda = \tau t$ can be much faster than computing t diffusion iterations
- No time step size (no “hidden” parameters)
- Non-linear diffusivity (e.g. for edge preservation) is easy to incorporate
- Concept of separating data and smoothness cost generalizes well and is a common pattern: $E(u) = E_{data}(u) + E_{smooth}(u)$

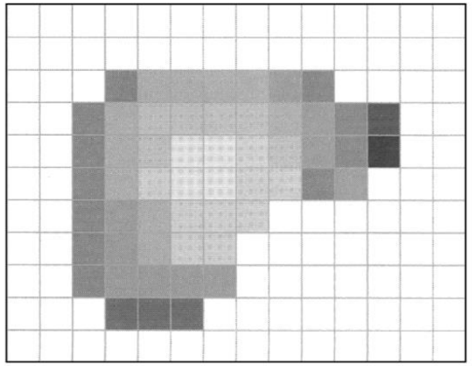
Observation: All three methods compute the same result with appropriate parameters, but with very different efficiency.
The energy minimization approach further provides several useful theoretical insights.

Discrete vs. Continuous

Digital images and videos are discrete
- discrete in color or brightness space (**quantization**)
- discrete in the spatial dimensions (**space sampling**)
- discrete in time (**time sampling**)



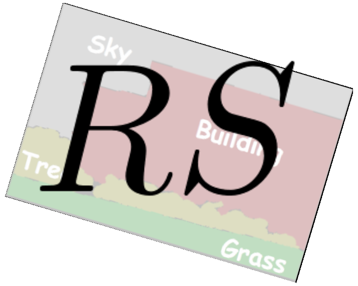
continuous



discrete

…but, the world observed by the camera sensor is continuous. Purely discrete methods usually rely on a fixed sampling, e.g. they are not invariant to resampling after image rotations or scaling.

$$S = f(I)$$



$$RS \stackrel{?}{=} f(RI)$$

The **calculus of variations** presents a theory that can accurately handle these kinds of problems. Although images are discrete they can still be analyzed in a continuous manner.

Calculus of Variations

The calculus of variations or variational methods describe a class of optimization methods that deal with minimizing functionals. The key idea is to express a problem as an optimization task by defining a suitable cost functional over a continuous solution space and by finding a problem solution as a minimizer (stationary point) of the cost functional.

Definition

A functional $E : \mathcal{V} \to \mathbb{R}$ is a mapping from a set of functions $\mathcal{V}$ to the real numbers $\mathbb{R}$.

Continuous Setting

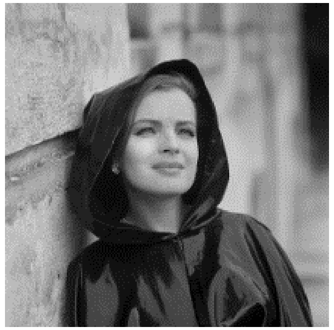
Images as functions: $u : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^d$

Domain Ω (rectangular subset of $\mathbb{R}^n$)

- $\Omega \subset \mathbb{R}^1$: signal (1D)
- $\Omega \subset \mathbb{R}^2$: image (2D)
- $\Omega \subset \mathbb{R}^3$: volume (3D)
- $\Omega \subset \mathbb{R}^4$: space-time volume (4D)

Range $\mathbb{R}^d$

- $\mathbb{R}^1$: scalar valued image (grayscale)
- $\mathbb{R}^2$: e.g. 2D-vector field
- $\mathbb{R}^3$: e.g. RGB image, HSV values, 3D-vector field
- $\mathbb{R}^4$: e.g. RGBA images, matrix valued image

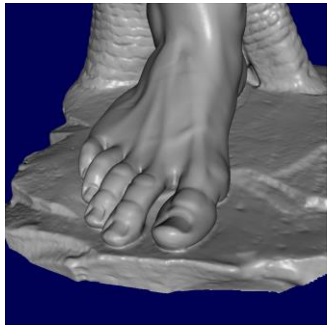


$\mathbb{R}^2 \rightarrow \mathbb{R}^1$

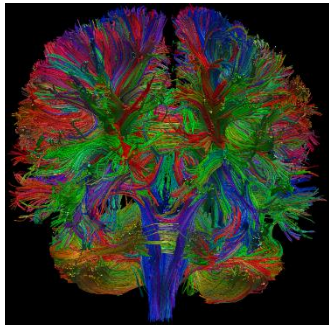
$\mathbb{R}^3 \rightarrow \mathbb{R}^1$

$\mathbb{R}^2 \rightarrow \mathbb{R}^3$

$\mathbb{R}^2 \rightarrow \mathbb{R}^2$



$\mathbb{R}^3 \rightarrow \mathbb{R}^1$



$\mathbb{R}^3 \rightarrow \mathbb{R}^3$

Discrete vs. Continuous

Energy $E : \mathcal{V} \to \mathbb{R}$ assigns each element of the space $\mathcal{V}$ a real number (energy).

	Vector space $\mathcal{V} = \mathbb{R}^n$	Function space $\mathcal{V} = \mathcal{L}^2(\Omega)$	
Elements	finitely many Elements $x_i, \quad i \in \{1, \ldots, n\}$	infinitely many Elements $u(x), \quad x \in \Omega$	
Inner Product	$\langle x, y \rangle = \sum_{i=1}^n x_i y_i$	$\langle u, v \rangle = \int_\Omega uv \ dx$	
Norm	$\lvert x \rvert_2 = \sqrt{\sum_{i=1}^n x_i^2}$	$\lVert u \rVert_2 = \left(\int_\Omega \lvert u \rvert^2 dx \right)^{\frac{1}{2}}$	
Gradient	$dE(x)/dx = \nabla E(x)$	$dE(u)/du = ?$	(Fréchet)
Directional	$\delta E(x; h) = \nabla E(x) \cdot h$	$\delta E(u; h) = ?$	(Gâteaux)
Extrema Condition	$dE(x)/dx = 0$	$dE(u)/du = ?$	

Gâteaux Derivative

Definition

Let $\mathcal{V}$ be a vector space, $E : \mathcal{V} \to \mathbb{R}$ be a functional and $u, h \in \mathcal{V}$. If the limit

$$\delta E(u; h) := \lim_{\alpha \to 0} \frac{1}{\alpha} \Big(E(u + \alpha h) - E(u) \Big)$$

exists, it is called the Gâteaux derivative of E at u with increment h.

- You can think of the Gâteaux derivative as the directional derivative (in infinite dimensions) of E at u in direction h.
- The Gâteaux differential is also called the “variation of E“ – hence the term “variational methods”, because the differential evaluates how much functional E “varies” if you go from u in direction h.
- Historical note: French mathematician René Gâteaux (1889-1914) died in WW1, his findings were published posthumously in 1919 by Paul Lévy.

Variational Principle

Theorem

If $\hat{u} \in \mathcal{V}$ is an extremum of a functional $E : \mathcal{V} \to \mathbb{R}$, then

$$\delta E(\hat{u}; h) = 0 \quad \text{for all } h \in \mathcal{V}.$$

- The variational principle is a generalization of the necessary condition for extrema for functions in $\mathbb{R}^n$ to infinite dimensional spaces.
- The necessary conditions must hold for all directions h. Similar to the finite dimensional case, the directional derivative corresponds to the projection of the functional gradient onto the direction h.
- Therefore the Gâteaux derivative can also be written as

$$\delta E(\hat{u}; h) = \left\langle \frac{dE(u)}{du}, h \right\rangle = \int \frac{dE(u)}{du}(x) h(x) \, dx$$

Euler-Lagrange Equation

Theorem

Let $\hat{u} \in \mathcal{V}$ be an extremum of the functional $E : \mathcal{C}^1(\Omega) \to \mathbb{R}$ of the form

$$E(u) = \int_\Omega L(u, \nabla u, x)dx$$

and $L : \Omega \times \Omega^n \times \Omega \to \mathbb{R}, (a, b, x) \mapsto L(a, b, x)$ be continuously differentiable.
Then $\hat{u}$ satisfies the Euler-Lagrange equation

$$\partial_a L(u, \nabla u, x) - \mathrm{div}_x(\nabla_b L(u, \nabla u, x)) = 0$$

where the divergence is computed with respect to the location variable x and

$$\partial_a L := \frac{\partial L}{\partial a} \qquad \nabla_b L := \left[\frac{\partial L}{\partial b_1} \cdots \frac{\partial L}{\partial b_n}\right]^T$$

- The Euler-Lagrange equation is a PDE which has to be satisfied for an extremal point $\hat{u}$.

Euler-Lagrange Equation

Derivation of the 1D case for functionals of the canonical form $E(u) = \int_\Omega L(u, u')dx$
the Gâteaux derivative is given by

$$\delta E(u;h) = \lim_{\alpha \to 0} \frac{1}{\alpha}\Big(E(u+\alpha h) - E(u)\Big)$$

$$= \lim_{\alpha \to 0} \frac{1}{\alpha} \int_\Omega \big(L(u+\alpha h, u' + \alpha h') - L(u, u')\big)dx$$

(apply Taylor expansion)

$$= \lim_{\alpha \to 0} \frac{1}{\alpha} \int_\Omega \Big(\big(L(u,u') + \frac{\partial L}{\partial u}\alpha h + \frac{\partial L}{\partial u'}\alpha h' + o(\alpha^2)\big) - L(u,u')\Big)dx$$

$$= \int_\Omega \left(\frac{\partial L}{\partial u}h + \frac{\partial L}{\partial u'}h'\right)dx$$

(apply integration by parts, $h = 0$ on boundary)

$$= \int_\Omega \left(\frac{\partial L}{\partial u}h - \frac{d}{dx}\frac{\partial L}{\partial u'}h\right)dx$$

$$= \int_\Omega \underbrace{\left(\frac{\partial L}{\partial u} - \frac{d}{dx}\frac{\partial L}{\partial u'}\right)}_{\frac{dE}{du}} h(x)dx$$

Euler-Lagrange Equation

Hence, the Gâteaux derivative of functional $E(u) = \int_\Omega L(u, u')dx$ in direction h is:

$$\delta E(u;h) = \int_\Omega \underbrace{\left(\frac{\partial L}{\partial u} - \frac{d}{dx}\frac{\partial L}{\partial u'}\right)}_{\frac{dE}{du}} h(x)dx$$

Combining this result with the necessary extremum condition of the variational principle, the variation of E in any direction $h(x)$ must vanish. This leads to the

Euler-Lagrange Equation (1D)

$$\frac{dE}{du} = \frac{\partial L}{\partial u} - \frac{d}{dx}\frac{\partial L}{\partial u'} = 0$$

The Euler-Lagrange equation is a differential equation expressing the necessary condition for extrema.

Edge Preserving Image Denoising

The Rudin-Osher-Fatemi (ROF) model (a.k.a. TV-L2) [Rudin Osher, Fatemi; Physica D, 1992]

Given a noisy image $f : \Omega \to \mathbb{R}$ and a smoothing parameter λ, a denoised image with preserved edges can be computed as the minimizer of the following energy

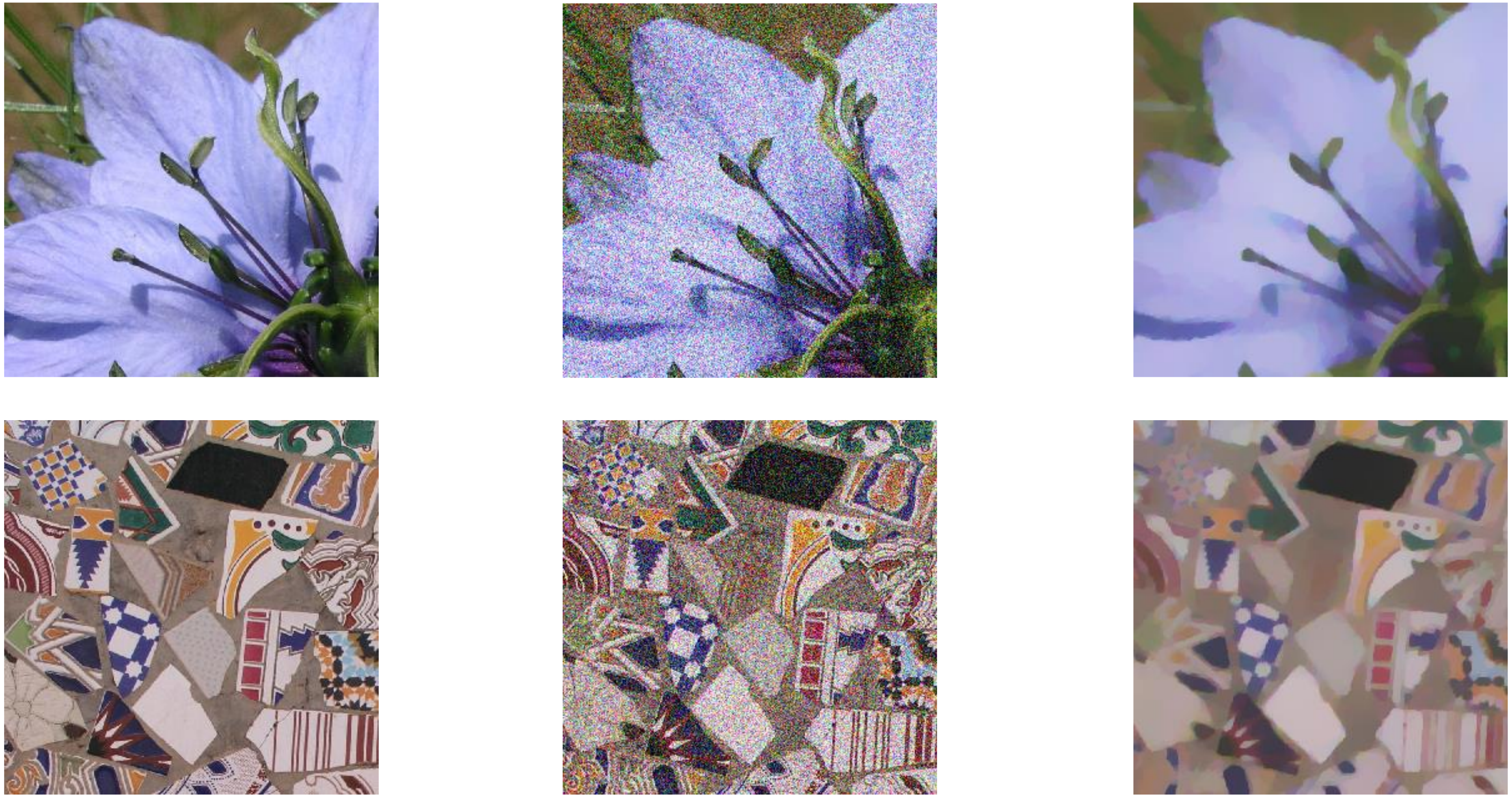
$$E(u) = \int_{\Omega} \left(|\nabla u|_2 + \frac{1}{2\lambda}(u - f)^2 \right) dx$$

Original

Noisy Image $\sigma = 0.1$

Denoised $\lambda = 8$

Edge Preserving Image Denoising



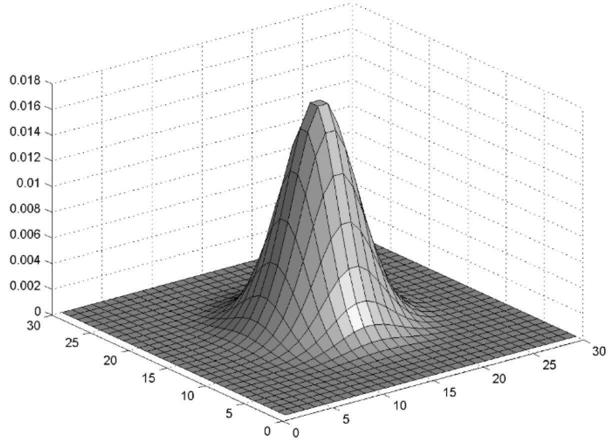
Original | Noisy Image | Denoised Result

Reminder: Image Blurring

Given a kernel $K : \mathbb{R}^2 \to \mathbb{R}$ and an image $u : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^d$ the convolution

$$(K * u)(x, y) := \int_\Omega K(a, b) u(x - a, y - b)\; dadb$$

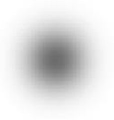
sums up the values of u around (x, y), weighted by K.



Gaussian kernel

$$K(a, b) = G_\sigma := \frac{1}{2\pi\sigma^2} \exp\left[\frac{a^2 + b^2}{2\sigma^2}\right]$$

Gaussian convolution

 *  = 

Motion blur with non-rotational symmetric kernel

 *  = 

Image Deblurring

Given an image f blurred with blur kernel b.
A deblurred image can be obtained by minimizing

$$E(u) = \int_\Omega \left(|\nabla u|_2 + \frac{1}{2\lambda}(b * u - f)^2 \right) \mathrm{d}x$$

The corresponding Euler-Lagrange equation is

$$\frac{dE}{du} = -\operatorname{div}\left(\frac{\nabla u}{|\nabla u|_2}\right) + \frac{1}{\lambda}\bar{b} * (b * u - f)$$

where $\bar{b}$ is the adjoint of kernel b, which is defined by $\bar{b}(x) = b(-x)$

Note: fine details are lost, but one can recover sharper image edges and their location.

Original blurred and noisy deblurred

Inpainting

Inpainting tries to recover missing parts of an image by using information from their surroundings.

Given a damaged image region $\Gamma \subset \Omega$ and an input image $f : \Omega \setminus \Gamma \to \mathbb{R}^d$ defined only outside the damaged region. We want to recover the full image $u : \Omega \to \mathbb{R}^d$ with $u|_{\Omega \setminus \Gamma} = f$.

Damaged Image

Inpainted Result

Inpainting

Inpainting for object removal.

Original Image

Removed region Γ

Inpainted Result

Inpainting

[Shen, Chan, SIAM 2002]

TV-Inpainting $E(u) = \int_\Gamma |\nabla u|_2 \, \mathrm{d}x$ subject to $u|_{\Omega \backslash \Gamma} = f$

Damaged Image

Inpainted Result

Note: TV-inpainting can also be easily combined with image denoising (e.g. ROF) by adding a data term and by optimizing over the full image domain, e.g. $E(u) = \int_{\Omega \backslash \Gamma} (u-f)^2 \mathrm{d}x + \int_\Omega |\nabla u|_2 \, \mathrm{d}x$

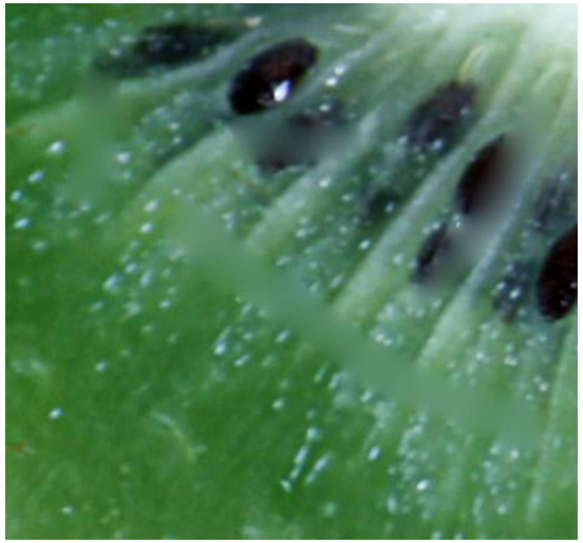
Inpainting

Original | Damaged | Inpainted Result

Inpainting

Original

Damaged



Inpainted Result

TV-inpainting is very simple fill-in technique, which works well for cartoon images, but it doesn't work well with "textured" images, because high-frequency details are not reconstructed.
Given the model's simplicity it does a pretty good job, but there are many other methods available!

General Linear Inverse Problems

Proposition

Let $E(u) := \int_\Omega (Au - f)^2 \, dx$ be a general linear functional. Then the Gâteaux derivative of E is given by

$$\delta E(u; h) = \int_\Omega \big(2A^*(Au - f)\big) h \, dx$$

where A^* is the adjoint operator of A, that is, the following condition holds

$$\langle u, A^* v\rangle = \langle Au, v\rangle \quad \text{for all } u, v \in \mathcal{L}^2(\Omega)$$

Image Segmentation

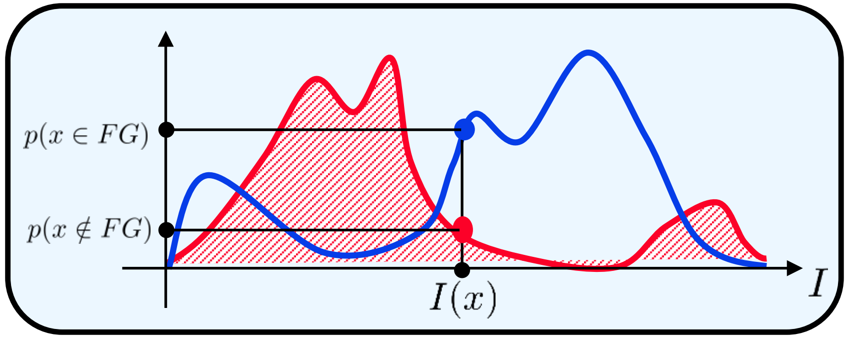
[Unger et al., BMVC 2008]

TV-Segmentation:
Given an input image and corresponding extracted pixel probabilities for being foreground and background, we can compute a binary labeling function $u : \Omega \to \{0,1\}$ which indicates for each pixel whether it belongs to the background ($u(x) = 0$) or foreground ($u(x) = 1$) by minimizing:

$$E(u) = \int_\Omega \left(|\nabla u|_2 + \lambda f u \right) \mathrm{d}x$$

Function $f : \Omega \to \mathbb{R}$ gives local preferences for pixel x being either background ($f(x) > 0$) or foreground ($f(x) < 0$).
A typical choice is:

$$f(x) = -\log \left(\frac{p(x \notin FG)}{p(x \in FG)} \right)$$



Estimated using FG / BG color models


Input Image


Segmentation Result

Multi-label Segmentation

[Chambolle et al., JIS 2012]

Given a set of labels Γ the labeling function $u : \Omega \to \{0,1\}^{|\Gamma|}$ maps to binary indicator functions defining the activation of each label. There is a separate data and smoothness cost for each label and an additional simplex constraint ensures that only one label gets activated.

$$E(u) = \sum_{\ell=1}^{|\Gamma|} \int_\Omega \left[\lambda f_\ell(x) u_\ell(x) + \frac{1}{2} |\nabla u_\ell|_2 \right] dx \qquad \text{s.t.} \sum_{\ell=1}^{|\Gamma|} u_\ell(x) = 1, \quad u(x) \geq 0 \qquad \forall x \in \Omega$$

Input Image

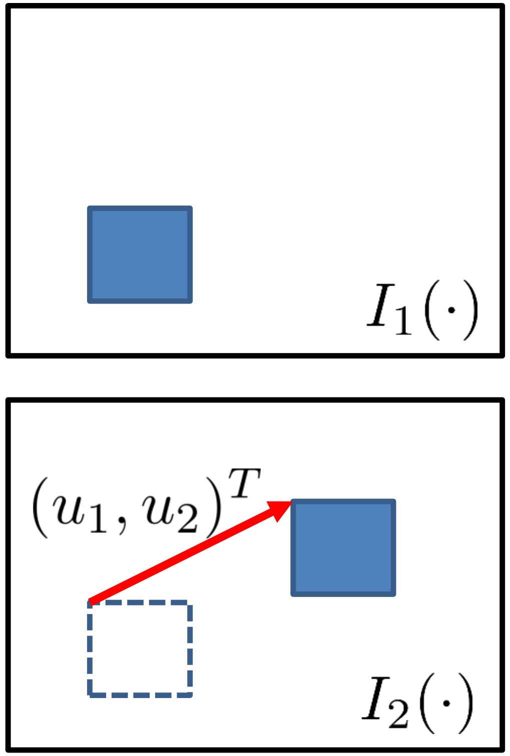
Segmentation with 10 regions

Optical Flow

[Horn, Schunck, AI, 1981]

Given two iages: $I_1, I_2 : \Omega \rightarrow \mathbb{R}$

Aim: vector field $u : \Omega \rightarrow \mathbb{R}^2$ which warps I_1 to I_2



Approach: $$E(u) = \int_\Omega (I_2(\mathbf{x} + u(\mathbf{x})) - I_1(\mathbf{x}))^2 + \lambda |\nabla u|^2 \quad d\mathbf{x}$$

Taylor-exp.: $$I_2(\mathbf{x} + u(\mathbf{x})) = I(\mathbf{x} + u(\mathbf{x}), t+1) \approx I(\mathbf{x}, t) + \nabla I^\top u + \underbrace{I_t}_{\frac{\mathrm{d}I}{\mathrm{d}t}}$$

Final energy: $$E(u) = \int_\Omega (\nabla I^\top u + I_t)^2 + \lambda |\nabla u|^2 \quad \mathrm{d}\mathbf{x}$$

As the first variational approach applied to computer vision this work influenced many other researchers.

$$|\nabla u| = \sqrt{{u_1}_x^2 + {u_1}_y^2 + {u_2}_x^2 + {u_2}_y^2}$$


Input


Result


Ground truth

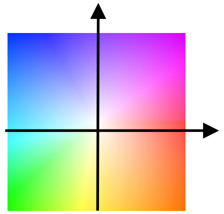

Color code

Image Super-Resolution

[Mitzel et al., GCPR 2009]

high-res low-res

$$E(I_H) = \sum_{k=1}^{N} \int_{\Omega} \left(\left| DBW_k I_H - I_L^{(k)} \right|_2 + \lambda \left| \nabla I_H \right|_2 \right) \mathrm{d}x$$

down-sampling blur warping

Input video

Low-res Images $I_L^{(k)}$ + warping W_k

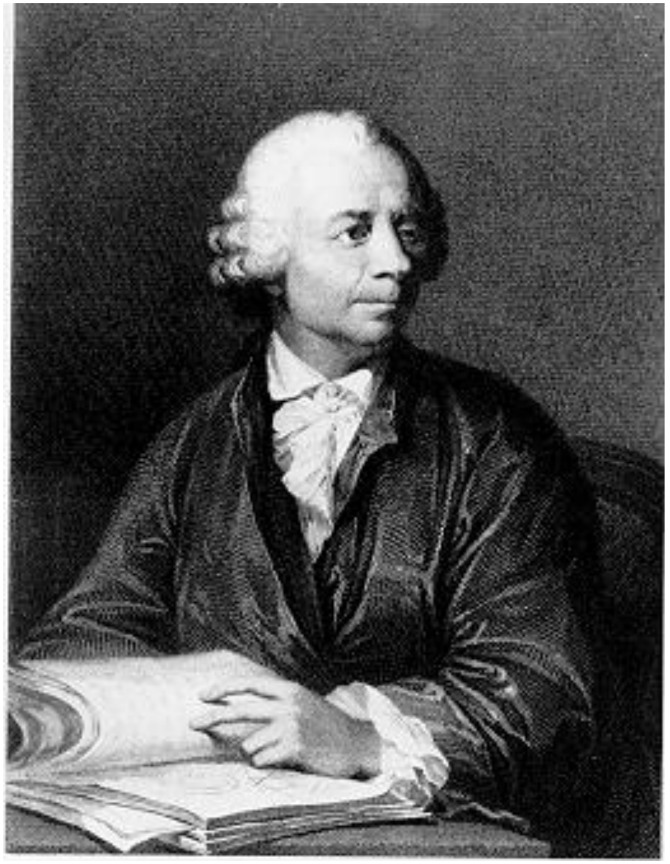
High-res Image I_H

bicubic

Mitzel et al. GCPR '09

Unger et al. GCPR '10

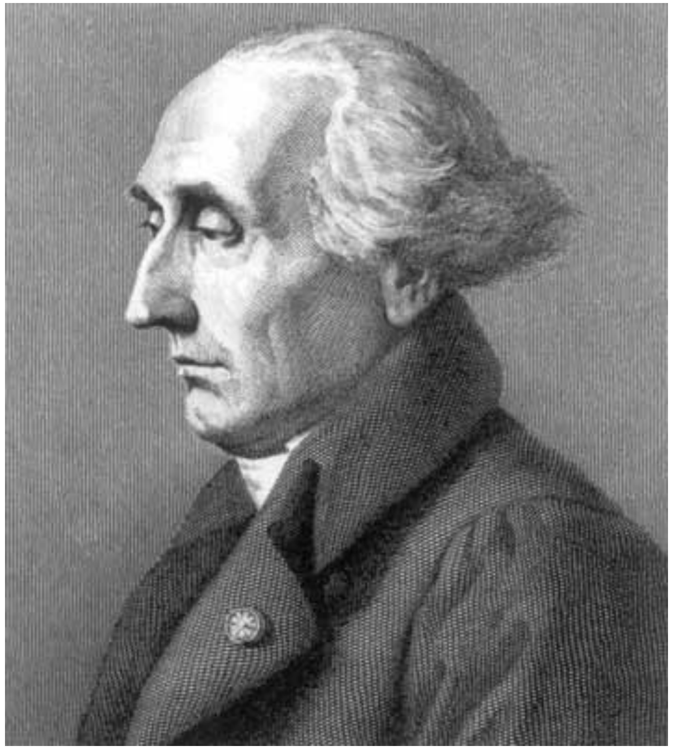
Leonhard Euler



Leonhard Euler (1707 – 1783)

- Published 866 papers and books, most of them in the last 20 years of his life. Considered the greatest mathematician of the 18th century.
- Major contributions: Euler number e, Euler angles, Euler formula, Euler theorem, Euler equations (for fluid flows), Euler-Lagrange equations,...
- 13 children

Joseph-Louis Lagrange



Joseph-Louis Lagrange
(1736 – 1813)

- born Giuseppe Lodovico Lagrangia (in Turin). self-taught.
- With 19 years: Professor for mathematics in Turin.
- Later in Berlin (1766-1787) and Paris (1787-1813).
- 1788: La Méchanique Analytique.
- 1800: Leçons sur le calcul des fonctions.

References

- Gelfand and Fomin, “*Calculus of Variations*”, Prentince-Hall, 1963
- Luenberger, “*Optimization by Vector Space Methods*”, Whiley, 1969
- B. Dacorogna, “*Introduction to the Calculus of Variations*”, Imperial College Press, 2009
- R. Tyrrell Rockafellar and Roger J-B Wets, “*Variational Analysis*”, volume 317 of Grundlehren der Mathematischen Wissenschaften. Springer, 1998
- G. Aubert and P. Kornprobst, “*Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*”, volume 147 of Applied Mathematical Sciences. Springer-Verlag, 2006
- Joachim Weickert, “*Anisotropic Diffusion in Image Processing*”, Teubner-Verlag, Stuttgart, Germany, 1998
- R. Gâteaux, “*Sur la notion d’intégrale dans le domaine fonctionnel et sur la théorie du potentiel*”, Bulletin de la S. M. F., tome 47, p. 47-70, 1919 http://archive.numdam.org/ARCHIVE/BSMF/BSMF_1919__47_/BSMF_1919__47__47_1/BSMF_1919__47__47_1.pdf
- L.I. Rudin, S. Osher, E. Fatemi, “*Nonlinear total variation based noise removal algorithms*”, Physica D 60(1-4):259–268, 1992
- Shen, J., Chan, T.F. “*Mathematical models for local nontexture inpaintings.*”, SIAM Journal on Applied Mathematics 62, 2002
- Markus Unger, Thomas Pock, Daniel Cremers, and Horst Bischof. “*TVSeg - interactive total variation based image segmentation.*”, British Machine and Vision Conference (BMVC), 2004.
- C. Zach, D. Gallup, J.-M. Frahm, and M. Niethammer, “*Fast global labeling for real-time stereo using multiple plane sweeps.*”, VMV 2008
- A. Chambolle, D. Cremers, and T. Pock, “*A convex approach to minimal partitions.*”, SIAM Journal on Imaging Sciences, 5(4):1113–1158, 2012
- Horn, B. and Schunck, B., “*Determining optical flow*”, "Artificial Intelligence, 17:185-203, 1981

Slide credits: Many slides are direct copies / are inspired by, or use material from the lecture slides from Prof. Bastian Goldlücke and Prof. Daniel Cremers.