Diffusion Filtering

Nonlinear Filtering

- The convolution of an input image $f(x)$ with a kernel $G(x)$:

$$g(x) = (G * f)(x) = \int G(x - x') f(x') dx'$$

is a classical example of a linear filter.

- Convolutions can be efficiently implemented in frequency space because in frequency space the convolution corresponds to a simple (frequency-wise) product and because the Fast Fourier transform allows a quick conversion to and from frequency space.
- In practice, however, linear filters are often suboptimal. In smoothing/denoising, for example, the Gaussian smoothing removes both noise and signal – semantically relevant structures tend to disappear along with the noise. Instead, one would like to remove noise in an adaptive manner such that semantically important structures remain unaffected. In principle this could be done with a Gaussian smoothing where the filter width σ is adapted to the local structure (larger in noise areas, smaller at important edges).

Image Filtering by Diffusion

- Formally this would amount to the following:

$$g(x) = \int G_{\sigma(f,x)}(x')\, f(x - x')\, dx',$$

  where now the width σ of the convolution kernel G depends on the brightness values in a local neighborhood.

- It turns out that there exist other more elegant solutions to model such adaptive denoising processes by means of Diffusion filtering.
- The key observation is that image smoothing can be modeled with a diffusion process. In this process, the local brightness diffuses to neighboring pixels due to local concentration differences.
- Mathematically diffusion processes are represented by partial differential equations (PDEs).

Review: Partial Differential Equations

- A partial differential equation (PDE) is an equation containing the partial derivatives of a function of several variables.
  Example — the wave equation:

$$\frac{\partial^2 \psi(x,t)}{\partial t^2} = c^2 \Delta \psi(x,t)$$

- For functions of a single variable only one speaks of ordinary differential equations (ODEs) (dt. gewöhnliche Differentialgleichungen.
  Example — the pendulum:

$$m\frac{d^2x(t)}{dt^2} + \gamma\frac{dx(t)}{dt} + kx(t) = 0$$

- Many natural phenomena can be modeled by partial differential equations. In most cases, one can derive the respective equation from a few basic principles. A solution of a differential equation is a function for which the differential equation is true.

Analytical Solutions

- A few PDEs can be solved analytically, i.e. the solution can be written in closed form.
- Example — The wave equation (in 1D):

$$\frac{\partial^2 \psi(x,t)}{\partial t^2} = c^2 \frac{\partial^2 \psi(x,t)}{\partial x^2}$$

has the (not necessarily unique) solution: $\psi(x,t) = \sin(x - ct)$

- If solutions are not unique one can impose additional assumptions boundary conditions or initial conditions, for example $\psi(x,0) = \psi_0(x)$.
- Example — The harmonic oscillator (without friction):

$$m \frac{d^2 x(t)}{dt^2} + kx(t) = 0$$

has the (generally not unique) solution:

$$x(t) = sin(\omega t), \quad \text{with } \omega = \sqrt{k/m}.$$

The Diffusion Equation

- Diffusion is a physical process which aims at minimizing differences in the spatial concentration $u(x,t)$ of a substance.
- This process can be described by two basic equations:
  - Fick's law states that concentration differences induce a flow j of the substance in direction of the negative concentration gradient:

    $$j = -g\nabla u$$

    The diffusivity g describes the speed of the diffusion process.
  - The continuity equation

    $$\partial_t u = -\text{div}\, j$$

    where $\text{div}\, j \equiv \nabla j \equiv \partial_x j_1 + \partial_y j_2$ is called the divergence of the vector j.
- Inserting one into the other leads to the diffusion equation:

$$\boxed{\partial_t u = \text{div}\,(g \cdot \nabla u)}$$

Solution of the Linear Diffusion Equation

The one-dimensional linear diffusion equation ($g = 1$)

$$\partial_t u = \partial_x^2 u.$$

with initial condition

$$u(x, t = 0) = f(x)$$

has the unique solution:

$$u(x,t) = (G_{\sqrt{2t}} * f)(x) = \int_{-\infty}^{\infty} G_{\sqrt{2t}}(x - x') f(x') dx',$$

where

$$G_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}},$$

is a Gaussian kernel of width $\sigma = \sqrt{2t}$.

Smoothing by Diffusion

- The above result implies that smoothing of an image $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ with Gaussian kernels of increasing width σ can be realized through a diffusion process of the form

$$\begin{cases} \partial_t u(x,t) = \Delta u \\ u(x,0) = f(x) \quad \forall x \in \Omega \\ \partial_n u|_{\partial\Omega} = \langle \nabla u, n \rangle|_{\partial\Omega} = 0 \end{cases} .$$

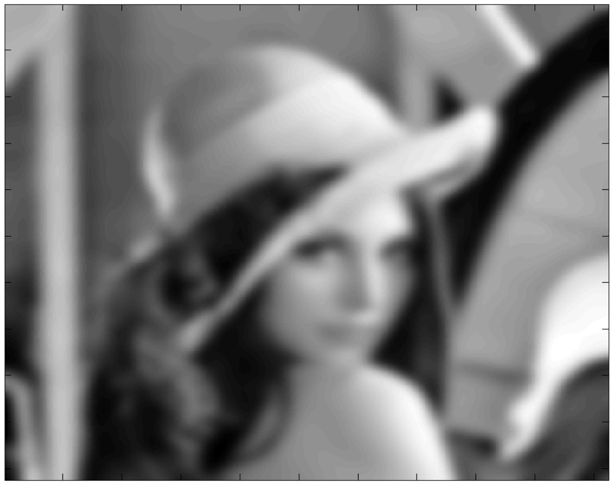
- The latter boundary condition states that the derivative of the brightness function u along the normal n at the image boundary $\partial\Omega$ must vanish. This assures that no brightness will leave or enter the image, i.e. the average brightness will be preserved.

- With increasing time t the solution $u(x,t)$ of this process will correspond to increasingly smoothed versions of the original image $f(x)$.

Image Smoothing via Diffusion

Lena original

diffusion $t = 2$



diffusion $t = 20$

diffusion $t = 100$

Nonlinear and Anisotropic Diffusion

- General diffusion equation:

$$\partial_t u = \text{div}\,(g \nabla u)$$

- For $g = 1$ (or $g = \text{const.} \in \mathbb{R}$) the diffusion process is called linear, isotropic and homogeneous.
- If the diffusivity g is space-dependent, i.e. $g = g(x)$, the process is called an inhomogeneous diffusion.
- If the diffusivity depends on u, i.e. $g = g(u)$, then it is called a nonlinear diffusion because then the equation is no longer linear in u.
- If the diffusivity g is matrix-valued then the process is called an anisotropic diffusion. A matrix-valued diffusivity leads to processes where the diffusion is different in different directions.
- Note: In the literature this terminology is not used consistently.

Edge-preserving Diffusion

- Idea: Less diffusion (smoothing) in locations of strong edge information.
- Gradient norm $|\nabla u| = \sqrt{u_x^2 + u_y^2}$ serves as edge indicator
- Diffusivity should decrease with increasing $|\nabla u|$. For example (Perona & Malik, *Scale Space and Edge Detection using Anisotropic Diffusion, PAMI 1990*):

$$g(|\nabla u|) = \frac{1}{\sqrt{1 + |\nabla u|^2/\lambda^2}}$$

- The positive parameter λ is called a contrast parameter. Areas where $|\nabla u| \gg \lambda$ will not be affected much in the diffusion process.
- The Perona-Malik model had a huge impact in image processing because it allowed a much better edge detection than classical edge detectors (such as the Canny edge detector).

Implementation with Finite Differences

- Nonlinear diffusion equation:

$$\partial_t u = \partial_x \left(g(|\nabla u|)\partial_x u\right) + \partial_y \left(g(|\nabla u|)\partial_y u\right)$$

- Discretize the operators as:

$$\partial_t u \approx \frac{u_{ij}^{t+1} - u_{ij}^t}{\tau}$$

and

$$\begin{aligned}\partial_x(g\partial_x u) &\approx \Big((g\partial_x u)_{i+1/2,j}^t - (g\partial_x u)_{i-1/2,j}^t\Big) \\ &\approx \Big(g_{i+1/2,j}^t(u_{i+1,j}^t - u_{ij}^t) - g_{i-1/2,j}^t(u_{ij}^t - u_{i-1,j}^t)\Big)\end{aligned}$$

where $g_{i+1/2,j} = \sqrt{g_{i+1,j}g_{ij}}$ assures that no diffusion takes place as soon as g is zero at one of the two pixels.

- Insert and solve for u_{ij}^{t+1}.
- Source: J. Weickert, *Anisotropic Diffusion in Image Processing*.

Nonlinear Diffusion

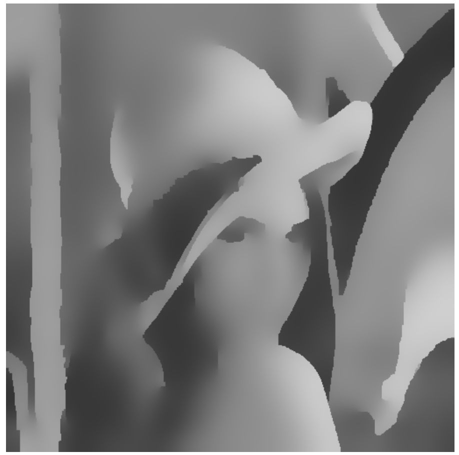
Lena original

diffusion $t = 9$

diffusion $t = 25$

diffusion $t = 100$

diffusion $t = 400$



diffusion $t = 900$

Author: D. Cremers