Image Processing and related PDEs
Lecture 2: Image denoising

Yves van Gennip

School of Mathematical Sciences, University of Nottingham

Minicourse on Image Processing and related PDEs
University of Oxford

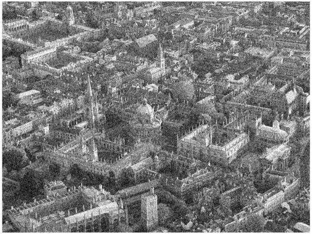
6 September, 2016

In the pipeline

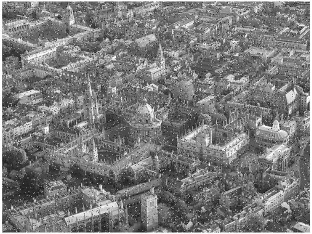
- Heat equation
- $W^{1,2}$ regularisation plus L^2 fidelity variational deblurring
- Rudin-Osher-Fatemi (ROF) variational denoising method

Examples of noise

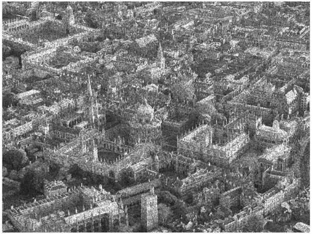
No noise



Gaussian noise



Salt and Pepper noise



Speckle noise

Modelling noise

Additive noise

$$f = u + n$$

- f is the observed image
- u is the original clean image
- n is the noise, e.g. white Gaussian noise with mean 0 and standard deviation σ^2, i.e. for all x, y, the random variables $n(x)$ and $n(y)$ are uncorrelated and all finite marginal distributions of n are Gaussian with mean 0 and standard deviation σ^2

Heat equation

Heat equation

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta u = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2}, & \text{for } t > 0 \text{ and } x \in \Omega \subset \mathbb{R}^2, \\ u(0,x) = f(x), & \text{for } x \in \Omega, \\ \text{plus boundary conditions.} \end{cases}$$

In image analysis there are two common choices for Ω:

1. $\Omega = [0, a] \times [0, b]$, i.e Ω is the image domain, or
2. $\Omega = \mathbb{R}^2$, in which case the image is first extended from $[0, a] \times [0, b]$ to $C := [-a, a] \times [-b, b]$ by mirror symmetry and then to all of $\mathbb{R}^2$ by periodicity.

In case 1 we can impose Dirichlet or Neumann boundary conditions. In general, for the heat equation on $\mathbb{R}^2$ we need an a priori bound on the growth rate of u for $|x| \to \infty$ to ensure uniqueness, but in case 2 periodicity takes care of that. ($|\cdot|$ will denote Euclidean norm.)

Solution on $\mathbb{R}^2$ [Aubert, Kornprobst]

Consider an image $f_0 \in L^1([0,a] \times [0,b])$ and extend it to $f : \mathbb{R}^2 \to \mathbb{R}$ via mirror symmetry and periodicity as described on the previous slide. Define

$$u(t,x) := (G_{\sqrt{2t}} * f)(x) := \int_{\mathbb{R}^2} G_{\sqrt{2t}}(x-y) f(y)\, dy,$$

where $G_\sigma(x) := \frac{1}{2\pi\sigma^2} e^{-|x|^2/(2\sigma^2)}$.

Then u is the unique function satisfying

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta u, & \text{for } (t,x) \in (0,\infty) \times \mathbb{R}^2, \\ \|u - f\|_{L^1(C)} \to 0, & \text{as } t \to \infty, \\ u(t,\cdot)\,|_C \in L^1(C), & \text{respecting the mirror symmetry-extension} \\ & \text{structure of } f, \text{ for } t > 0 \\ u \in C^\infty((0,T) \times \mathbb{R}^2), & \text{for } T > 0. \end{cases}$$

and

- for all $t_1 > 0$, there exists $c > 0$ such that for all $t \geq t_1$, $\sup_{x \in \mathbb{R}^2} |u(t,x)| \leq c \|f_0\|_{L^1([0,a]\times[0,b])}$, and
- if $f_0 \in L^\infty([0,a] \times [0,b])$, then $\inf_{x \in \mathbb{R}^2} f(x) \leq u(t,x) \leq \sup_{x \in \mathbb{R}^2} f(x)$, for all $t > 0$.

Convolution with a Gaussian

Using the Fourier transform

$$F[f](w) := \int_{\mathbb{R}^2} f(x) e^{-iw\cdot x}\, dx,$$

we have

$$F[u(t,\cdot)](w) = F[G_{\sqrt{2t}} * f](w) = F[G_{\sqrt{2t}}](w) F[f](w),$$

where

$$F[G_{\sqrt{2t}}](w) = e^{-|w|^2 t}.$$

Solving the heat equation/convolving with a Gaussian acts as a *low pass filter*: high frequencies (large $\|w\|$) get suppressed, with the *scale* at which this happens determined by t.

Example: Gaussian filtering on Gaussian noise

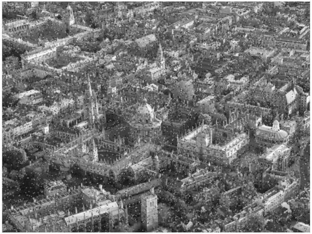
$\sigma = 0.5$

$\sigma = 0.75$

$\sigma = 1.0$

$\sigma = 1.5$

Example: Gaussian filtering on salt and pepper noise



$\sigma = 0.5$

$\sigma = 0.75$

$\sigma = 1.0$

$\sigma = 1.5$

Example: Gaussian filtering on speckle noise

$\sigma = 0.5$

$\sigma = 0.75$

$\sigma = 1.0$

$\sigma = 1.5$

Extensions and variations

- Different filters, e.g. Wiener filter (Chan, Shen, §4.3)
- Nonlinear diffusion (Perona-Malik):

$$\begin{cases} \frac{\partial u}{\partial t} = \operatorname{div}\left(c(|\nabla u|^2)\nabla u\right), & \text{for } (t,x) \in (0,T) \times \Omega, \\ \nabla u \cdot \nu = 0, & \text{for } (t,x) \in (0,T) \times \partial\Omega, \\ u(0,x) = f(x), & \text{for } x \in \Omega, \end{cases}$$

  where $\Omega \subset \mathbb{R}^2$ is bounded and open, and $c : [0,\infty) \to (0,\infty)$ will be a function designed to inhibit smoothing close to edges in the image, while allowing smoothing away from edges.
- Variational methods

The heat equation as gradient flow

Let $\Omega \subset \mathbb{R}^2$ be open, bounded, and consider the functional $F : W^{1,2}(\Omega) \to \mathbb{R}$ defined by

$$F(u) := \frac{1}{2} \int_\Omega |\nabla u|^2.$$

Let $v \in W^{1,2}(\Omega)$ and $t \in \mathbb{R}$, to compute the L^2 gradient:

$$\begin{aligned} \frac{d}{dt} F(u+tv)\Big|_{t=0} &= \frac{d}{dt}\frac{1}{2}\int_\Omega |\nabla u|^2 + 2t\nabla u \cdot \nabla v + t^2 |\nabla v|^2 \Big|_{t=0} \\ &= \int_\Omega \nabla u \cdot \nabla v = \int_\Omega v\,(-\Delta u) + \int_{\partial\Omega} v\, \nabla u \cdot \nu \\ &= \langle -\Delta u, v \rangle_{L^2(\Omega)} + \int_{\partial\Omega} v\,\nabla u \cdot \nu. \end{aligned}$$

Hence the L^2 gradient flow, given some initial u_0, is

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta u, & \text{for } (t,x) \in (0,\infty) \times \Omega, \\ \nabla u \cdot \nu = 0, & \text{for } (t,x) \in (0,\infty) \times \partial\Omega, \\ u(0,x) = u_0(x), & \text{for } x \in \Omega. \end{cases}$$

A first try at variational denoising

$W^{1,2}$ regulariser and L^2 fidelity (Aubert, Kornprobst, §3.2)

Let $\lambda > 0$, $f \in L^2(\Omega)$, and $F : W^{1,2}(\Omega) \to \mathbb{R}$, given by

$$F(u) := \int_\Omega |\nabla u|^2 + \frac{\lambda}{2} \int_\Omega (f-u)^2.$$

Minimise:

$$u^* \in \operatorname*{arg\,min}_{u \in W^{1,2}(\Omega)} F(u).$$

- $\int_\Omega (f-u)^2$ is a *fidelity term*
- $\int_\Omega |\nabla u|^2$ is a *regulariser*. In the current set-up, it prevents us from getting the unwanted solution $u^* = f$. In some future cases it will even be mathematically necessary to avoid ill-posed problems.

Gradient flow and the role of λ

Minimising F is equivalent to minimising $\frac{1}{\lambda}F$. Taking the L^2 gradient flow of the latter:

$$\begin{cases} \frac{\partial u}{\partial t} = \lambda^{-1}\Delta u - (u-f), & \text{for } (t,x) \in (0,\infty)\times\Omega, \\ \nabla u \cdot \nu, & \text{for } (t,x) \in (0,\infty)\times\partial\Omega, \\ u(0,x) = u_0(x), & \text{for } x \in \Omega. \end{cases}$$

Letting $s = t/\lambda$, we get

$$\frac{\partial u}{\partial s} = \Delta u - \lambda(u-f).$$

For $\lambda > 1$ the diffusion process described by $\frac{\partial u}{\partial s} = \Delta u$ is slower than that described by $\frac{\partial u}{\partial t} = \Delta u$. We see that increasing λ reduces smoothing (at a fixed time t).

Direct method in the calculus of variations

We let $\Omega = (0, a) \times (0, b)$.

Existence of minimisers; Evans §8.2, Dacorogna §3

Let $f \in L^2(\Omega)$ and $\lambda > 0$, then there exists a minimiser of F over $W^{1,2}(\Omega)$.

- Let $\{u_n\} \subset W^{1,2}(\Omega)$ be a minimising sequence, i.e.

$$F(u_n) \to m := \inf_{u \in W^{1,2}(\Omega)} F(u), \quad \text{as } n \to \infty.$$

- Then there is a $C > 0$ such that, for all n, $F(u_n) \leq C$.
- In particular $\int_\Omega |\nabla u_n|^2 \leq C$ and $\|u_n\|_{L^2(\Omega)} \leq \|u_n - f\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)} \leq C + \|f\|_{L^2(\Omega)}$.
- Hence $\{u_n\}$ is bounded in $W^{1,2}(\Omega)$ and so there is a $u \in W^{1,2}(\Omega)$ such that $u_n \rightharpoonup u^*$ (weakly in $W^{1,2}(\Omega)$) as $n \to \infty$.
... TBC...

Existence proof, continued

- Since $\nabla : W^{1,2}(\Omega) \to L^2(\Omega)$ is linear and bounded, $\nabla u_n \rightharpoonup \nabla u^*$ (weakly in $L^2(\Omega)$).
- Since norms are weakly lower semicontinuous with respect to their own induced natural topology (Brezis §III.3), we have

$$\int_\Omega |\nabla u^*|^2 \leq \liminf_{n\to\infty} \int_\Omega |\nabla u_n|^2.$$

- By the Rellich-Kondrachov theorem (Adams §6) $W^{1,2}(\Omega)$ is compactly embedded into $L^2(\Omega)$, thus $u_n \to u^*$ in the $L^2(\Omega)$ metric.
- Since $u \mapsto \|u - f\|^2_{L^2(\Omega)}$ is continuous w.r.t. $L^2(\Omega)$ convergence, we find

$$F(u^*) \leq \liminf_{n\to\infty} F(u_n) = m.$$

- Therefore $m = F(u^*)$ and so u^* is a minimiser of F over $W^{1,2}$.

Uniqueness of the solution

Uniqueness of the miniser

Let $f \in L^2(\Omega)$ and $\lambda > 0$, then the minimiser of F over $W^{1,2}(\Omega)$ is unique.

Uniqueness of the minimiser follows from strict convexity of F and convexity of $W^{1,2}(\Omega)$:
Assume u_1 and u_2 are two distinct minimisers of F over $W^{1,2}(\Omega)$. Let $s \in (0,1)$, then by strict convexity

$$F(su_1 + (1-s)u_2) < sF(u_1) + (1-s)F(u_2) = F(u_1),$$

since $F(u_1) = F(u_2)$. Since $W^{1,2}(\Omega)$ is convex, we have $su_1 + (1-s)u_2 \in W^{1,2}(\Omega)$, which contradicts the fact that u_1 is a minimiser of F over $W^{1,2}(\Omega)$.

Rudin-Osher-Fatemi

Total variation regulariser and L^2 fidelity (Rudin, Osher, Fatemi)

Let $\lambda > 0$, $f \in L^2(\Omega)$, and $F : BV(\Omega) \to \mathbb{R}$, given by

$$G(u) := \int_\Omega |\nabla u| + \frac{\lambda}{2} \int_\Omega (f-u)^2.$$

Minimise:

$$u^* \in \underset{u \in BV(\Omega)}{\arg\min}\, F(u).$$

- Total variation (TV):

$$\int_\Omega |\nabla u| := \sup\{ \int_\Omega u \operatorname{div} g : g \in C_c^1(\Omega, \mathbb{R}^2), \forall x \in \Omega, |g(x)| \leq 1 \}.$$

- Functions of bounded variation:

$$BV(\Omega) := \{ u \in L^1(\Omega) : \|u\|_{BV(\Omega)} := \|u\|_{L^1(\Omega)} + \int_\Omega |\nabla u| < \infty \}.$$

Total variation of an indicator function

(Giusti §1) Let $E \subset \mathbb{R}^2$ have a C^2 boundary and define

$$\chi_E(x) := \begin{cases} 1, & \text{if } x \in E, \\ 0, & \text{if } x \notin E. \end{cases}$$

Let $g \in C_c^1(\Omega, \mathbb{R}^2)$ with, for all $x \in \Omega$, $|g(x)| \leq 1$, then

$$\int_\Omega \chi_E \operatorname{div} g = \int_{E\cap\Omega} \operatorname{div} g = \int_{\partial(E\cap\Omega)} g \cdot \nu \leq \int_{\partial(E\cap\Omega)} |\nu|^2 = \mathcal{H}^1(\partial(E \cap \Omega)),$$

where we used Cauchy-Schwarz with $|g(x)| \leq 1 = |\nu(x)|$. $\mathcal{H}^1$ denotes the 1-dimensional Hausdorff measure. Hence

$$\int_\Omega |\nabla \chi_E| \leq \mathcal{H}^1(\partial(E \cap \Omega)).$$

...TBC...

Total variation of an indicator function, continued

Since E has C^2 boundary, we can extend the normal field on ∂E to a function $N \in C^1(\mathbb{R}^2, \mathbb{R}^2)$ such that for all $x \in \mathbb{R}^2$, $|N(x)| \leq 1$. Let $\eta \in C_0^\infty(\Omega)$ with, for all $x \in \Omega$, $|\eta(x)| \leq 1$, then $g := N\eta$ is admissible in the definition of TV. Hence

$$\int_\Omega |\nabla \chi_E| \geq \int_\Omega \chi_E \text{div}\,(N\eta) = \int_{\partial(\Omega \cap E)} \eta N \cdot \nu = \int_{\partial(\Omega \cap E)} \eta.$$

Taking the supremum over all such η gives

$$\int_\Omega |\nabla \chi_E| \geq \mathcal{H}^1(\partial(E \cap \Omega)).$$

We conclude

$$\int_\Omega |\nabla \chi_E| = \mathcal{H}^1(\partial(E \cap \Omega)).$$

In fact, this construction can be used to *define* the concept of perimeter for general Borel sets E. It turns out that this concept coincides with other concepts of perimeter, such as the $\mathcal{H}^1$ measures of the reduced boundary and the essential boundary (Ambrosio, Fusco, Pallara §3.5).

Lower semicontinuity and compactness

Lower semicontinuity of total variation

Let $\{u_n\} \subset BV(\Omega)$ be a sequence which converges to a function u in $L^1_{\text{loc}}(\Omega)$. Then

$$\int_\Omega |\nabla u| \leq \liminf_{n\to\infty} \int_\Omega |\nabla u_n|.$$

Proof: Let $g \in C^1_c(\Omega, \mathbb{R}^2)$ with, for all $x \in \Omega$, $|g(x)| \leq 1$, then

$$\int_\Omega u \operatorname{div} g = \lim_{n\to\infty} \int_\Omega u_n \operatorname{div} g \leq \liminf_{n\to\infty} \int_\Omega |\nabla u_n|.$$

Compactness (Evans, Gariepy §5.2)

$BV(\Omega)$ is compactly embedded in $L^1(\Omega)$.

Existence and uniqueness of solutions

Existence of a unique minimiser; Chan, Shen §4.5.4, Schönlieb §3.2

Let $f \in L^2(\Omega)$ and $\lambda > 0$, then there exists a unique minimiser of G over $BV(\Omega)$.

The existence proof follows analogously to the one for our previous functional F. The main difference is that the strong convergence of the minimising sequence now is in $L^1(\Omega)$ (by the compactness on the previous slide, using the uniform bounds on $\int_\Omega |\nabla u_n|$ and $\|u_n\|_{L^1(\Omega)} \leq C\|u_n\|_{L^2(\Omega)}$), with only weak convergence in $L^2(\Omega)$ (because of the uniform bound on $\|f - u_n\|_{L^2(\Omega)}$). Combining the $L^1(\Omega)$ lower semicontinuity of TV with the weak-$L^2(\Omega)$ lower semicontinuity of the $L^2(\Omega)$ norm then gives existence of a solution.

Since G is strictly convex and $BV(\Omega)$ is convex, uniqueness of the solution follows as before.

Formal L^2 gradient flow

Formally we compute, for $u, v \in W^{1,1}(\Omega)$, $t > 0$,

$$\frac{d}{dt}\int_\Omega |\nabla(u+tv)|\Big|_{t=0} = \int_\Omega \frac{\nabla v \cdot \nabla(u+tv)}{|\nabla(u+tv)|}\Big|_{t=0}$$
$$= -\int_\Omega v \,\text{div}\left(\frac{\nabla u}{|\nabla u|}\right) + \int_{\partial\Omega} v \frac{\nabla u}{|\nabla u|}\cdot \nu.$$

This leads to the formal gradient flow

$$\begin{cases} \frac{\partial u}{\partial t} = \text{div}\left(\frac{\nabla u}{|\nabla u|}\right) - \lambda(u-f), & \text{for } (t,x) \in (0,T)\times\Omega, \\ \nabla u \cdot \nu = 0 & \text{for } (t,x) \in (0,T)\times\partial\Omega, \\ u(0,x) = u_0(x), & \text{for } x \in \Omega. \end{cases}$$

- Convex analysis and subdifferentials needed for rigorous computations.
- Compare with Perona-Malik!
- $\text{div}\left(\frac{\nabla u}{|\nabla u|}\right)$ gives the curvature of the level sets of u!

Examples

Show example(s) on Image Processing On Line:

`http://www.ipol.im/`

Bibliography: image analysis specific references

1. G. Aubert, P. Kornprobst, *Mathematical Problems in Image Processing — Partial Differential Equations and the Calculus of Variations*, Springer, 2006
2. T.F. Chan, J. Shen, *Image Processing and Analysis*, SIAM, 2005
3. L.I. Rudin, S. Osher, E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D, 1992, 259–268
4. C.-B. Schönlieb, *Image Processing — Variational and PDE Methods*, online lecture notes, 2013/14, `http://www.damtp.cam.ac.uk/user/cbs31/Teaching_files/PDEImageLectureNotesLent2014.pdf`

Bibliography: general PDE and analysis references

1. R.A. Adams, J.J.F. Fournier, *Sobolev Spaces*, 2nd edition, Elsevier, 2008
2. L. Ambrosio, N. Fusco, D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford Science Publications, 2006
3. H. Brezis, *Analyse Fonctionnelle — Théorie et applications*, Dunod, 2005
4. B. Dacorogna, *Direct Methods in the Calculus of Variations*, Springer, 2008
5. L.C. Evans, *Partial Differential Equations*, AMS, 2002
6. L.C. Evans, R.F. Gariepy, *Measure Theory and Fine Properties of Functions*, CRC Press, 1992
7. E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*, Birkhäuser, 1984
8. F. John, *Partial Differential Equations*, 4th ed., Springer, 1982