Lecture 11:
Image denoising using energy minimization

Let g be a noisy image corrupted by additive noise n.

Then: $g(x,y) = \underbrace{f(x,y)}_{\text{Clean image}} + \underbrace{n(x,y)}_{\text{noise}}$

Recall: Laplacian masking: $g = f - \Delta f$ (Obtain a (non-smooth) sharp image from a smooth image)

Conversely, to get a smooth image f from a non-smooth image g, we can solve the PDE for f: $\underbrace{-\Delta f + f}_{\text{unknown}} = \underbrace{g}_{\text{known}}$

We will show that solving the above equation is equivalent to minimizing something:

$$E(f) = \iint \left(f(x,y) - g(x,y)\right)^2 dx\,dy + \iint \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 dx\,dy$$

In the discrete case, the PDE can be approximated (discretized) to get:

$$f(x,y) = g(x,y) + [f(x+1,y) + f(x,y+1) + f(x-1,y) + f(x,y-1) - 4f(x,y)]$$

for all (x,y) (Linear System)

Solved by

Direct method (Big linear system)

Iterative method

Simple iterative scheme: Let g be an $N \times N$ image.

Step 1: Let $f^0(x,y) = g(x,y)$ (Initial guess of the solution)

Step 2: For $n \geq 0$ and for all $(x,y), x = 1, \cdots, M, y = 1, \cdots, N$,

$$f^{n+1}(x,y) = g(x,y) + [f^n(x+1,y) + f^n(x-1,y) + f^n(x,y+1) + f^n(x,y-1) - 4f^n(x,y)]$$

Impose boundary conditions by reflection:

$$\begin{aligned}
&f^{n+1}(0,y) = f^{n+1}(2,y); \quad f^{n+1}(M+1,y) = f^{n+1}(M-1,y) \quad \text{for } y = 1, \cdots, N \\
&f^{n+1}(x,0) = f^{n+1}(x,2); \quad f^{n+1}(x,N+1) = f^{n+1}(x,N-1) \quad \text{for } x = 1, \cdots, M \\
&f^{n+1}(0,0) = f^{n+1}(2,2); \quad f^{n+1}(0,N+1) = f^{n+1}(2,N-1) \\
&f^{n+1}(M+1,0) = f^{n+1}(M-1,2); \quad f^{n+1}(M+1,N+1) = f^{n+1}(M-1,N-1)
\end{aligned}$$

(similar for f^n)

Step 3: Continue the process until $||f^{n+1} - f^n|| \leq$ tolerance. (Convergence depends on the spectral radius of a matrix)

Consider: $$E_{\text{discrete}}(f) = \sum_{x=1}^{N}\sum_{y=1}^{N}\left(f(x,y) - g(x,y)\right)^2 + \sum_{x=1}^{N}\sum_{y=1}^{N}\left[\left(f(x+1,y) - f(x,y)\right)^2 + \left(f(x,y+1) - f(x,y)\right)^2\right]$$

Suppose f is a minimizer of E_{discrete}. Then, for each (x,y),

E_{discrete} depends on $f(x,y)$ for each (x,y)

$$\frac{\partial E_{\text{discrete}}}{\partial f(x,y)} = 0.$$

$$= 2\left(f(x,y) - g(x,y)\right) + 2\left(f(x+1,y) - f(x,y)\right)(-1) + 2\left(f(x,y+1) - f(x,y)\right)(-1) + 2\left(f(x,y) - f(x-1,y)\right) + 2\left(f(x,y) - f(x,y-1)\right)$$

By simplification, we get:

$$f(x,y) = g(x,y) + \left[f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1) - 4f(x,y)\right]$$

The continuous version of E_{discrete} can be written as:

$$E(f) = \iint \left(f(x,y) - g(x,y)\right)^2 + \iint \left[\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2\right] dx\,dy$$

$|\nabla f|^2$

Remark:

- Solving $f = g + \Delta f$ is equivalent to energy minimization
- The first term in Ediscrete is called the fidelity term.
  Aim to find f that is close to g.
- The second term is called the regularization term. Aim to enhance smoothness.
- $-\nabla f + f = g$ can also be solved in the frequency domain:

$$DFT(f) = DFT(g + \underbrace{\Delta f}_{p * f})$$

$$\therefore DFT(f)(u,v) = DFT(g)(u,v) + c\,DFT(p)(u,v)\,DFT(f)(u,v)$$

$$\Longleftrightarrow DFT(f)(u,v) = \left[\frac{1}{1 - c\,DFT(p)(u,v)}\right] DFT(g)(u,v)$$

$\downarrow$ inverse DFT

$f(x,y)$!!

Lecture 20:

Image denoising by solving PDE (derived from energy minimisation problem)

Consider the harmonic-L2 minimization model:

$$\text{minimize} \quad E(f) = \int_{\underbrace{\Omega}_{\text{image domain}}} \big(f(x,y) - \underbrace{g(x,y)}_{\text{Observed}}\big)^2 \, dx\,dy + \int \underbrace{|\nabla f|^2}_{\text{smoothness of } f} \, dx\,dy$$

(Look for (continuous) image f)

We find f that minimizes $E(f)$.

Take any function $v(x,y)$. Consider a real-valued function $S: \mathbb{R} \to \mathbb{R}$:

$$S(\varepsilon) := E(f + \varepsilon v) = \int_\Omega \big(f(x,y) + \varepsilon v(x,y) - g(x,y)\big)^2 \, dx\,dy + \int_\Omega |\nabla f + \varepsilon \nabla v|^2 \, dx\,dy$$

$$\frac{d}{d\varepsilon} S(\varepsilon) = 2\int_\Omega \big(f(x,y) + \varepsilon v(x,y) - g(x,y)\big)^2 \, dx\,dy + 2\int_\Omega \left[\left(\frac{\partial f}{\partial x} + \varepsilon \frac{\partial v}{\partial x}\right)\frac{\partial v}{\partial x} + \left(\frac{\partial f}{\partial y} + \varepsilon \frac{\partial v}{\partial y}\right)\frac{\partial v}{\partial y}\right] dx\,dy$$

If f is a minimizer, $\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} S(\varepsilon) = 0$ for all v ($\because$ $S(0)$ is the minimum)

$$\therefore\ S'(0) = 0 = 2\int_\Omega \big(f(x,y) - g(x,y)\big)\, v(x,y)\, dx\,dy + 2\int_\Omega (f_x v_x + f_y v_y)\, dx\,dy \quad \text{for all } v.$$

Remark: If we can formulate the above equation as follows:

$$\int_\Omega T(x,y)\, v(x,y)\, dx\,dy = 0 \quad \text{for all } v(x,y)$$

then, we can conclude $T(x,y) = 0$ in Ω.

In our case, first term is okay!
second term NOT okay!

Useful Tool: (Integration by part)

$$\nabla\cdot\big(V_1(x,y), V_2(x,y)\big) = \frac{\partial V_1}{\partial x} + \frac{\partial V_2}{\partial y}$$

$$\int_\Omega \nabla f \cdot \nabla g\, dx\,dy = -\int_\Omega \big(\underbrace{\nabla\cdot}_{\text{divergence}}(\nabla f)\big)\, g\, dx\,dy + \int_{\partial\Omega} g\, (\nabla f \cdot \vec{n})\, ds$$

where $\vec{n} = (n_1, n_2)$ = outward normal on the boundary.

In our case, we get: $0 = \int_\Omega (f-g)\, v\, dx\,dy - \int_\Omega (\nabla\cdot\nabla f)\, v\, dx\,dy + \int_{\partial\Omega} (\nabla f\cdot\vec{n})\, ds$

Overall, we get: $\int_\Omega (f - g - \Delta f)\, v \, dx\, dy - \int_{\partial\Omega} (\nabla f \cdot \vec{n})\, v \, ds = 0$ for all v

We conclude: $\begin{cases} f - g - \Delta f = 0 & \text{in } \Omega \\ \nabla f \cdot \vec{n} = 0 & \text{in } \partial\Omega \end{cases}$ (PDE)

Conversely, given f that satisfies the PDE, for any h,

$$E(h) - E(f) = \int_\Omega \left[(h-g)^2 - (f-g)^2 + |\nabla h|^2 - |\nabla f|^2 \right] dx\, dy$$

$$-2(f-g)^2 + 2(f-g)(h-g)$$
$$\|$$

$$= \int_\Omega \left[(h-g) - (f-g) \right]^2 + |\nabla h - \nabla f|^2 + 2 \nabla f \cdot (\nabla h - \nabla f) + 2(f-g)(h-f)$$

$$\geq \int_\Omega 2 \nabla f \cdot (\overbrace{\nabla h - \nabla f}^{\nabla(h-f)}) + 2(f-g)(h-f)\, dx\, dy$$

$$= 2\int_\Omega \underbrace{\nabla \cdot (\nabla f)(h-f) + (f-g)(h-f)}_{(\underbrace{\nabla\cdot(\nabla f) + (f-g)}_{0})(h-f)} dx\, dy + \int_{\partial\Omega} (\underbrace{\nabla f \cdot \vec{n}}_{0})(h-f)\, dx\, dy$$

$$= 0$$

$\therefore \quad E(h) \geq E(f)$ for all h.

$\therefore f$ is the minimizer!!

$$\frac{\partial}{\partial \sigma} f(x,y,\sigma) = \sigma \nabla \cdot \left(K(x,y) f\right)(x,y,\sigma)$$

<u>Remark</u>: • Anisotropic diffussion is related to minimizing:

$$E(f) = \int_\Omega K(x,y) \, |\nabla f(x,y)|^2 \, dx\, dy$$

• Energy minimization approach for solving imaging problem is called the Variational image processing!

$$E(f+\varepsilon v) = \int K(x,y) \, |\nabla f + \varepsilon \nabla v|^2 \, dx\, dy$$

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} E(f+\varepsilon v) = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \int K(x,y) \left(\nabla f + \varepsilon \nabla v\right) \cdot \left(\nabla f + \varepsilon \nabla v\right)$$

$$= \int_\Omega K(x,y) \left(2 \nabla f \cdot \nabla v\right)$$

$$= -\int_\Omega \left(\nabla \cdot \left(K(x,y) \nabla f\right)\right) v + \int_{\partial\Omega} \left(K(x,y) \nabla f \cdot \vec{n}\right) v$$

Total variation (TV) denoising (ROF)

Invented by: Rudin, Osher, Fatemi

Motivation: Previous model: $f = g + \Delta f$. Solve for f from noisy g.

Disadvantage: smooth out edge.

Modification: $f = g + \nabla\cdot(K\nabla f)$

K is small on edges!!

Goal: Given a noisy image $g(x,y)$, we look for $f(x,y)$ that solves:

$$f = g + \lambda\frac{\partial}{\partial x}\left(\frac{1}{|\nabla f|(x,y)}\frac{\partial f}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{1}{|\nabla f|(x,y)}\frac{\partial f}{\partial y}\right) \qquad (*)$$

Remark: Problem arises if $|\nabla f(x,y)| = 0$. Take care of it later.

We'll show that $(*)$ must be satisfied by a minimizer of:

$$J(f) = \frac{1}{2}\int_\Omega (f(x,y) - g(x,y))^2 + \underbrace{\lambda}_{\text{constant parameter } > 0.}\int_\Omega |\nabla f(x,y)|\, dx\,dy$$

Same idea: Let $S(\varepsilon) := E(f+\varepsilon v)$

$$= \frac{1}{2}\int_\Omega (f+\varepsilon v-g)^2 + \lambda \int_\Omega \underbrace{|\nabla f + \varepsilon \nabla v|}_{\sqrt{(\nabla f + \varepsilon \nabla v)\cdot(\nabla f + \varepsilon \nabla v)}}$$

$$\frac{d}{d\varepsilon} S(\varepsilon) = \left[\int_\Omega (f+\varepsilon v - g)\,v + \lambda \int_\Omega \frac{\nabla f \cdot \nabla v + 2\varepsilon \nabla v \cdot \nabla v}{\sqrt{(\nabla f + \varepsilon \nabla v)\cdot(\nabla f + \varepsilon \nabla v)}} \right]$$

If f is a minimizer, $\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} S(\varepsilon) = 0$ for all v.

$$\therefore \; S'(0) = 0 = \int_\Omega (f-g)\,v + \lambda \int_\Omega \frac{\nabla f \cdot \nabla v}{|\nabla f|}$$

$$= \int_\Omega (f-g)\,v - \lambda \int_\Omega \nabla \cdot \left(\frac{\nabla f}{|\nabla f|}\right) v + \lambda \int_{\partial\Omega} \left(\frac{\nabla f}{|\nabla f|} \cdot \vec{n}\right) v$$

$$= \int_\Omega \left[(f-g) - \lambda \nabla \cdot \left(\frac{\nabla f}{|\nabla f|}\right)\right] v + \lambda \int_{\partial\Omega} \left(\frac{\nabla f}{|\nabla f|} \cdot \vec{n}\right) v \quad \text{for all } v$$

We conclude: $(f-g) - \lambda \nabla \cdot \left(\frac{\nabla f}{|\nabla f|}\right) = 0!!$

In the discrete case,

$$J(f) = \frac{1}{2}\sum_{x=1}^{N}\sum_{y=1}^{N}(f(x,y)-g(x,y))^2+\lambda\sum_{x=1}^{N}\sum_{y=1}^{N}\sqrt{(f(x+1,y)-f(x,y))^2+(f(x,y+1)-f(x,y))^2}$$

J can be regarded as a multi-variable function depending on :
$f(1,1), f(1,2), \ldots, f(1,N), f(2,1), \ldots, f(2,N), \ldots, f(N,N)$.

If f is a minimizer, then $\frac{\partial J}{\partial f(x,y)} = 0$ for all (x,y).

$$\begin{aligned}\frac{\partial J}{\partial f(x,y)} &= (f(x,y)-g(x,y)) + \lambda\frac{2(f(x+1,y)-f(x,y))(-1)+2(f(x,y+1)-f(x,y))(-1)}{2\sqrt{(f(x+1,y)-f(x,y))^2+(f(x,y+1)-f(x,y))^2}}\\
&+\lambda\frac{2(f(x,y)-f(x-1,y))}{2\sqrt{(f(x,y)-f(x-1,y))^2+(f(x-1,y+1)-f(x-1,y))^2}}\\
&+\lambda\frac{2(f(x,y)-f(x,y-1))}{2\sqrt{(f(x+1,y-1)-f(x,y-1))^2+(f(x,y)-f(x,y-1))^2}} = 0\end{aligned}$$

By simplification:

$$f(x,y)-g(x,y)=\lambda\left\{\frac{f(x+1,y)-f(x,y)}{\sqrt{(f(x+1,y)-f(x,y))^2+(f(x,y+1)-f(x,y))^2}}\right.$$

$$\left.-\frac{f(x,y)-f(x-1,y)}{\sqrt{(f(x,y)-f(x-1,y))^2+(f(x-1,y+1)-f(x-1,y))^2}}\right\}$$

$$+\lambda\left\{\frac{f(x,y+1)-f(x,y)}{\sqrt{(f(x+1,y)-f(x,y))^2+(f(x,y+1)-f(x,y))^2}}\right.$$

$$\left.-\frac{f(x,y)-f(x,y-1)}{\sqrt{(f(x+1,y-1)-f(x,y-1))^2+(f(x,y)-f(x,y-1))^2}}\right\}$$

$$= \frac{\left.\frac{\partial f}{\partial x}\right|_{(x,y)}}{|\nabla f|_{(x,y)}}$$

$$= \frac{\left.\frac{\partial f}{\partial x}\right|_{(x-1,y)}}{|\nabla f|_{(x-1,y)}}$$

$$= \frac{\left.\frac{\partial f}{\partial y}\right|_{(x,y)}}{|\nabla f|_{(x,y)}}$$

$$= \frac{\left.\frac{\partial f}{\partial y}\right|_{(x,y-1)}}{|\nabla f|_{(x,y-1)}}$$

Discretization of $f - g = \lambda\, \nabla \cdot \left(\frac{\nabla f}{|\nabla f|} \right)$

How to minimise $J(f)$

We consider the problem of finding f that minimizes $J(f)$.

In the discrete case, J depends on $f(x,y)$ for $\begin{matrix} x=1,2,\dots,N \\ y=1,2,\dots,N \end{matrix}$.

Consider a time-dependent image $f(x,y;\underbrace{t}_{\text{time}})$. Assuming that $f(x,y;t)$ satisfies:

$$\frac{df(\cdot,\cdot;t)}{dt} = -\nabla J(f(\cdot,\cdot;t)) \qquad (\ast\ast)$$

We can show that $J(f(\ \ ;t))$ decreases as t increases.

Note that:

$$\frac{d}{dt}J(f(\cdot,\cdot;t)) = \nabla J(f(\cdot,\cdot;t))\cdot\frac{d}{dt}f(\cdot,\cdot;t) = -\nabla J(f(\cdot,\cdot;t))\cdot\nabla J(f(\cdot,\cdot;t))$$
$$= -\left|\nabla J(f(\cdot,\cdot;t))\right|^2 \le 0.$$

$\therefore$ $J(f(\cdot,\cdot;t))$ is decreasing as t increases!!

In the discrete case,

$$\frac{f^{n+1} - f^n}{\Delta t} = -\nabla J(f^n)$$

(Gradient descent algorithm)

For the ROF model:

$$
\begin{aligned}
&\frac{f^{n+1}(x,y) - f^n(x,y)}{\Delta t} \\
&= -(f^n(x,y) - g(x,y)) + \lambda \frac{f^n(x+1,y) - f^n(x,y)}{\sqrt{(f^n(x+1,y) - f^n(x,y))^2 + (f^n(x,y+1) - f^n(x,y))^2}} \\
&- \lambda \frac{f^n(x,y) - f^n(x-1,y)}{\sqrt{(f^n(x,y) - f^n(x-1,y))^2 + (f^n(x-1,y+1) - f^n(x-1,y))^2}} \\
&+ \lambda \frac{f^n(x,y+1) - f^n(x,y)}{\sqrt{(f^n(x+1,y) - f^n(x,y))^2 + (f^n(x,y+1) - f^n(x,y))^2}} \\
&- \lambda \frac{f^n(x,y) - f^n(x,y-1)}{\sqrt{(f^n(x+1,y-1) - f^n(x,y-1))^2 + (f^n(x,y) - f^n(x,y-1))^2}}
\end{aligned}
$$

Discretization of ∇J

(Gradient descent algorithm for ROF)