Lecture 22:

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:

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

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:

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

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

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

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

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

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

Gradient descent algorithm

Let $f: \mathbb{R}^n \to \mathbb{R}$. We want to find a sequence $\vec{x}_0 \in \mathbb{R}^n, \vec{x}_1 \in \mathbb{R}^n, \dots, \vec{x}_n \in \mathbb{R}^n, \dots$
such that $f(\vec{x}_0) \geq f(\vec{x}_1) \geq \dots \geq f(\vec{x}_n) \geq f(\vec{x}_{n+1}) \geq \dots$

So, $\vec{x}_0, \vec{x}_1, \dots, \vec{x}_n, \dots$ iteratively minimizes $f(\vec{x})$.

Given $\vec{x}_0$, we want to find $\vec{x}_1 = \vec{x}_0 + t\vec{v}$ $(t>0, \vec{v} \in \mathbb{R}^n)$ such that

$$f(\vec{x}_1) \leq f(\vec{x}_0).$$

Note that: $f(\vec{x}_1) = f(\vec{x}_0 + t\vec{v}) \approx f(\vec{x}_0) + t\, \nabla f(\vec{x}_0) \cdot \vec{v} + \underbrace{\frac{t^2}{2!} \vec{v}^T f''(\vec{x}_0)\vec{v} + \dots}_{\text{(neglible)}}$

Choose $\vec{v} = -\nabla f(\vec{x}_0)$. Then:

$$f(\vec{x}_1) \approx f(\vec{x}_0) - t\,|\nabla f(\vec{x}_0)|^2 \leq f(\vec{x}_0)$$

Similarly, given $\vec{x}_n$, choose $\vec{v} = -\nabla f(\vec{x}_n)$. Let $\vec{x}_{n+1} = \vec{x}_n + t\vec{v} = \vec{x}_n + t\, \nabla f(\vec{x}_n)$.
Then: for small $t>0$, we have

$$f(\vec{x}_{n+1}) \approx f(\vec{x}_n) - t\,|\nabla f(\vec{x}_n)|^2 \leq f(\vec{x}_n).$$

Therefore, we have an iterative scheme:

$$\vec{x}_{n+1} = \vec{x}_n + t\,\vec{v}_n \,, \text{ where } \vec{v}_n = -\nabla f(\vec{x}_n)$$

$t > 0$ is small, called the time step.

$\vec{v}_n \in \mathbb{R}^n$ is called the descent direction at n^{th} iteration.

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 $x = 1, 2, \dots, N$, $y = 1, 2, \dots, N$.

Our goal is to find a sequence of images:

$f^0, f^1, f^2, \dots, f^n, f^{n+1}, \dots$ such that $J(f_0) \geq J(f_1) \geq \dots \geq J(f_n) \geq J(f_{n+1}) \geq \dots$

Define: $\nabla J(f^n) = \begin{pmatrix} \frac{\partial J}{\partial f^n(1,1)} \\ \frac{\partial J}{\partial f^n(2,1)} \\ \vdots \\ \frac{\partial J}{\partial f^n(N,1)} \\ \vdots \\ \frac{\partial J}{\partial f^n(1,N)} \\ \vdots \\ \frac{\partial J}{\partial f^n(N,N)} \end{pmatrix}$

Given $\overrightarrow{f^n}$, define $\overrightarrow{f^{n+1}} = \overrightarrow{f^n} + \Delta t \vec{v}_n$

where $\vec{v}_n = -\nabla J(f^n)$.

Here, $\overrightarrow{f^n}$ is the vectorized image of f^n.

Then:

$$J(\overrightarrow{f^{n+1}}) = J(\overrightarrow{f^n} + \Delta t \vec{v}_n) \approx J(\overrightarrow{f^n}) + \Delta t \nabla J(f^n) \cdot \vec{v}_n = J(\overrightarrow{f^n}) - \Delta t |\nabla J(f^n)|^2 \leq J(\overrightarrow{f^n}).$$

In the discrete case,

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

(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)