Lecture 19:

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 $\begin{array}{l} x=1,2,\dots,N \\ y=1,2,\dots,N \end{array}$.

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

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

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

Note that:

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

$$= -|\nabla J(f(\cdot,\cdot;t))|^2 \leq 0.$$

$\therefore$ $J(f(\cdot,\cdot;t))$ 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)

In the continuous case, consider:

$$E(f) = \int_\Omega (f-g)^2\,dx\,dy + \lambda \int |\nabla f|\,dx\,dy$$

We want to find a sequence $f_0 = g, f_1, \ldots, f_n, \ldots$ such that:

$$E(f_0) \geq E(f_1) \geq \cdots \geq E(f_n) \geq E(f_{n+1}) \geq \cdots$$

Define $S(\varepsilon) :\overset{def}{=} E(\overbrace{f_n + \varepsilon v}^{f_{n+1}})$ for some suitable v. (Assuming that $\nabla f_n = 0$ on $\partial\Omega$)

For small $\varepsilon > 0$ by Taylor expansion,

$$\underset{E(f_{n+1})}{\underset{\shortparallel}{S(\varepsilon)}} = \underset{E(f_n)}{\underset{\shortparallel}{S(0)}} + S'(0)\underset{0}{\underset{\vee}{\varepsilon}} + \underbrace{O(\varepsilon^2)}_{\text{neglible}}$$

If $S'(0) \leq 0$, then $E(f_{n+1}) \leq E(f_n)$

Now, $$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} S(\varepsilon) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0}\left[\int_\Omega (f_n + \varepsilon v - g)^2\,dx\,dy + \lambda\int_\Omega |\nabla f_n + \varepsilon\nabla v|\,dx\,dy\right]$$

$$|\nabla f_n + \varepsilon\nabla v| = \sqrt{(\nabla f_n + \varepsilon\nabla v)\cdot(\nabla f_n + \varepsilon\nabla v)}$$

$$\therefore \quad S'(0) = \int_\Omega (f_n - g)\, v \, dx\,dy + \lambda \int_\Omega \frac{\nabla f_n \cdot \nabla v}{\underbrace{\sqrt{\nabla f_n \cdot \nabla f_n}}_{|\nabla f_n|}} \, dx\,dy$$

$$= \int_\Omega (f_n - g)\, v \, dx\,dy - \lambda \int_\Omega \nabla \cdot \left(\frac{\nabla f_n}{|\nabla f_n|} \right) v \, dx\,dy + \cancelto{0}{\lambda \int_{\partial\Omega} \frac{\nabla f_n}{|\nabla f_n|} \cdot \vec{n} \; v \, ds}$$

$$= \int_\Omega \left[(f_n - g) - \lambda \nabla \cdot \left(\frac{\nabla f_n}{|\nabla f_n|} \right) \right] v \, dx\,dy$$

Put $v = -\left[(f_n - g) - \lambda \nabla \cdot \left(\frac{\nabla f_n}{|\nabla f_n|} \right) \right]$. Then: $S'(0) \leq 0$.

$\therefore$ Gradient descent algorithm:

$$f^{n+1} = f^n + \varepsilon \left[\underbrace{- \left(2 (f_n - g) - \lambda \nabla \cdot \left(\frac{\nabla f_n}{|\nabla f_n|} \right) \right)}_{v} \right] \quad \text{for } n = 0, 1, 2, \ldots$$

This is called the gradient descent algorithm.

Image deblurring in the spatial domain

(Spatial) Deblurring / denoising model:

Consider: $\underset{\text{Observed}}{\underset{\uparrow}{g}} = \underset{\text{Degradation}}{\underset{\uparrow}{h}} * f + n$ ← noise

We aim to find f that minimizes:

$$J(f) = \frac{1}{2}\iint_D \left(h*f(x,y) - g(x,y) \right)^2 dx\,dy + \lambda \iint_D |\nabla f|\,dx\,dy$$

Goal: Use gradient descent method

Note that:

$$\iint_D h*f(x,y)\, g(x,y)\,dx\,dy = \iint_D \left[\iint_D h(\alpha,\beta) f(x-\alpha, y-\beta)\,d\alpha\,d\beta \right] g(x,y)\,dx\,dy$$

$$= \iint_D \left[\iint_D f(\underbrace{x-\alpha}_{X}, \underbrace{y-\beta}_{Y})\, g(x,y)\,dx\,dy \right] h(\alpha,\beta)\,d\alpha\,d\beta$$

Let $X = x-\alpha$, $Y = y-\beta$.

$$= \iint_D \left[\iint_D f(X,Y)\, g(X+\alpha, Y+\beta)\,dx\,dy \right] h(\alpha,\beta)\,d\alpha\,d\beta$$

$$= \iint_D f(X,Y) \underbrace{\left[\iint_D h(\alpha,\beta)\, g(\alpha+X, \beta+Y)\,d\alpha\,d\beta \right]}_{\tilde{H}(g)(X,Y)} dX\,dY$$

To minimize $J(f)$ using gradient descent, consider: $S(\varepsilon) = J(f+\varepsilon w)$. Then: $\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} S(\varepsilon) = 0$.

$$\therefore \quad S'(0) = \frac{1}{2}\iint_D \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} (h*f + \varepsilon h*w - g)^2 \, dx\,dy + \lambda \iint_D \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} |\nabla f + \varepsilon \nabla w| \, dx\,dy$$

$$= \iint_D (h*f - g)\, h*w \; dx\,dy - \lambda \iint_D \nabla \cdot \left(\frac{1}{|\nabla f|} \nabla f\right) w \, dx\,dy$$

$$= \iint_D \tilde{H}(h*f - g)(x,y)\, w(x,y)\, dx\,dy - \lambda \iint_D \nabla \cdot \left(\frac{1}{|\nabla f|} \nabla f\right)(x,y)\; w(x,y)\, dx\,dy$$

$\therefore$ Gradient descent direction:

$$w(x,y) = -\tilde{H}(h*f - g)(x,y) + \lambda\, \nabla \cdot \left(\frac{1}{|\nabla f|} \nabla f\right)(x,y)$$

Algorithm:

$$\frac{f^{n+1} - f^n}{\Delta t} = -\tilde{H}(h*f^n - g) + \lambda \underbrace{\nabla \cdot \left(\frac{1}{|\nabla f^n|} \nabla f^n\right)}_{}$$

Discrete discretization of partial derivatives.