Lecture 15

Method 4: Constrained least square filtering

Disadvantages of Wiener's filter:

① $|N(u,v)|^2$ and $|F(u,v)|^2$ must be known / guessed

② Constant estimation of ratio is not always suitable

Goal: Consider a least square minimization model.

Let $g = h * f + n$

degradation

noise

In matrix form, $\vec{g} = D\vec{f} + \vec{n}$

$\vec{g} = \mathcal{S}(g)$, $\vec{f} = \mathcal{S}(f)$, $\vec{n} = \mathcal{S}(n)$

$\vec{g}, \vec{f}, \vec{n} \in \mathbb{R}^{N^2}$, $D \in M_{N^2 \times N^2}$

Stacked image of g

transformation matrix of $h*f$ (or f)

Constrained least square problem:

Given $\vec{g}$, we need to find an estimation of $\vec{f}$ such that it minimizes:

$$E(f) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \left| \Delta f(x,y) \right|^2$$ subject to the constraint:

$$\| \vec{g} - D\vec{f} \|^2 = \varepsilon$$

What is Δf ?

In the discrete case, we can estimate:

$$\Delta f(x,y) \approx f(x+1,y) + f(x,y+1) + f(x-1,y) + f(x,y-1) - 4f(x,y)$$

Taylor expansion:

$$\frac{\partial^2 f}{\partial x^2}(x,y) \approx \frac{f(x+h,y) - 2f(x,y) + f(x-h,y)}{h^2}$$

$$\frac{\partial^2 f}{\partial y^2}(x,y) \approx \frac{f(x,y+h) - 2f(x,y) + f(x,y-h)}{h^2}$$

$$\xrightarrow{\text{Put } h=1} \quad \Delta f(x,y) = \left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \right)(x,y)$$

$\therefore$ Δ is the Laplacian in the discrete case

Remark:

- More generally, $\Delta f = p * f$ ← discrete convolution

  where

$$p = \begin{pmatrix} 0 & \cdots & & 0 \\ & & 1 & & \\ \vdots & 1 & -4 & 1 & \vdots \\ & & 1 & & \\ 0 & \cdots & & 0 \end{pmatrix} \; x=0$$

$$y=0$$

- Minimizing $\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} |\Delta f(x,y)|^2$ is to denoise.

- Also, $\vec{g} = D\vec{f} + \vec{n} \Leftrightarrow \vec{g} - D\vec{f} = \vec{n} \Rightarrow \|\vec{g} - D\vec{f}\|^2 = \|\vec{n}\|^2 \approx \varepsilon$

  ↑ noise level

  $\therefore$ the constraint $\|\vec{g} - D\vec{f}\|^2$ is to solve the deblurring problem.

- $\displaystyle\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} |\Delta f(x,y)|^2$ ← Denoise
- $\|\vec{g} - D\vec{f}\|^2 = \varepsilon$ ← Deblur

In the discrete case, we can estimate:

$$\Delta f(x,y) \approx f(x+1,y) + f(x,y+1) + f(x-1,y) + f(x,y-1) - 4f(x,y)$$

Taylor expansion:

$$\frac{\partial^2 f(x,y)}{\partial x^2} \approx \frac{f(x+h,y) - 2f(x,y) + f(x-h,y)}{h^2} \xrightarrow{\text{Put } h=1} \nabla^2 f(x,y) = \left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}\right)(x,y)$$

$$\frac{\partial^2 f}{\partial y^2}(x,y) \approx \frac{f(x,y+h) - 2f(x,y) + f(x,y-h)}{h^2}$$

More generally, $\Delta f = p * f$ ← discrete convolution

where

$$p = \begin{array}{cc} \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \begin{matrix} & 1 & \\ 1 & -4 & 1 \\ & 1 & \end{matrix} & \vdots \\ 0 & \cdots & 0 \end{pmatrix} & x=0 \\ y=0 & \end{array}$$

<u>Remark:</u> $\underbrace{\|\vec{g} - D\vec{f}\|^2}_{\|\vec{n}\|^2} = \varepsilon$ means we allow some fixed level of noise.

Let $\overrightarrow{p*f} = S(p*f) = L\vec{f}$

transformation matrix representing the convolution with p.

Then: $E(\vec{f}) = (L\vec{f})^T(L\vec{f})$

We will prove:

Theorem: The constrained least square problem has the optimal solution in the spatial domain that satisfies:

$$(D^TD + \gamma L^TL)\vec{f} = D^T\vec{g}$$

for some suitable parameter γ.

In the frequency domain,

$$\hat{F}(u,v) := DFT(f)(u,v) = \frac{1}{N^2}\,\frac{\overline{H(u,v)}}{|H(u,v)|^2 + \gamma\,|P(u,v)|^2}\,G(u,v)$$

($H = DFT(h)$; $G(u,v) = DFT(g)$; $P(u,v) = DFT(p)$ where

$$p = \begin{pmatrix} 0 & \cdots & & \cdots & 0 \\ \vdots & & 1 & & \vdots \\ \vdots & 1 & -4 & 1 & \vdots \\ \vdots & & 1 & & \vdots \\ 0 & \cdots & & \cdots & 0 \end{pmatrix}$$

)

<u>Remark</u>: Constrained least square filtering:

$$T(u,v) = \frac{1}{N^2} \frac{H(u,v)}{|H(u,v)|^2 + \gamma |P(u,v)|^2}$$

Let $\tilde{F}(u,v) = T(u,v)\, G(u,v)$

Compute Inverse DFT of $\tilde{F}(u,v)$.

Sketch of proof:

Recall: our problem is to minimize:

$$\vec{f}^T L^T L \vec{f} \quad \text{subject to} \quad \|\vec{g} - D\vec{f}\|^2 = \varepsilon$$

$$\|\vec{g} - D\vec{f}\|^2 = (\vec{g} - D\vec{f})^T (g - D\vec{f})$$

From calculus, the minimizer must satisfy:

$$\nabla \mathcal{L} \overset{def}{=} \nabla\left(\vec{f}^T L^T L \vec{f} + \lambda (\vec{g} - D\vec{f})^T (\vec{g} - D\vec{f})\right) = 0 \quad \text{for}$$

where $\vec{f} = (f_1, f_2, \dots, f_i, \dots, f_{N^2})^T$ and λ is the Lagrange's multiplier.

Here, $\nabla \mathcal{L} = \left(\frac{\partial \mathcal{L}}{\partial f_1}, \frac{\partial \mathcal{L}}{\partial f_2}, \dots, \frac{\partial \mathcal{L}}{\partial f_{N^2}}\right)^T$

Easy to check:

- $\nabla(\vec{f}^T \vec{a}) = \vec{a}$
- $\nabla(\vec{b}^T \vec{f}) = \vec{b}$
- $\nabla(\vec{f}^T A \vec{f}) = (A + A^T)\vec{f}$

$$\vec{f}^{\,T}\vec{a} = (f_1, f_2, \dots, f_n)\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = a_1 f_1 + a_2 f_2 + \dots + a_n f_n$$

$$\frac{\partial \vec{f}^{\,T}\vec{a}}{\partial f_j} = a_j$$

$$\therefore \nabla(\vec{f}^{\,T}\vec{a}) \overset{\text{def}}{=} \left(\frac{\partial \vec{f}^{\,T}\vec{a}}{\partial f_1}, \dots, \frac{\partial \vec{f}^{\,T}\vec{a}}{\partial f_n} \right)^T = (a_1, a_2, \dots, a_n)^T$$

etc. . ..

$$\therefore \ \mathcal{D} = 0 \Rightarrow (2L^T L)\vec{f} + \lambda(-D^T\vec{g} - D^T\vec{g} + 2D^T D\vec{f}) = 0$$

$$\Rightarrow (D^T D + \gamma L^T L)\vec{f} = D^T\vec{g}$$ where $\gamma = \frac{1}{\lambda}$ and λ is the Lagrange's multiplier.

Parameter γ can be determined by direct substitution into the equation:

$$(\vec{g} - D\vec{f})^T(\vec{g} - D\vec{f}) = \varepsilon.$$

How about in the frequency (Fourier) domain?

Two important theorems

Notation: Let $\mathcal{O}$ be a linear transformation defined by:

$$\mathcal{O}(f) = k * f \quad \text{for all } f \in M_{N \times N}(\mathbb{R}), \text{ where}$$

$k \in M_{N \times N}(\mathbb{R})$.

Let $D \in M_{N^2 \times N^2}(\mathbb{R})$ be the transformation matrix representing $\mathcal{O}$.

That is, $\mathcal{S}(\mathcal{O}(f)) = D\,\mathcal{S}(f)$. $\in \mathbb{R}^{N^2}$

Here, $\mathcal{S}$ is the stacking operator. $\mathcal{S}(I)$ is the vectorized image of I (1st col of I becomes first n entries of $\mathcal{S}(I)$, 2nd col of I becomes second n entries of $\mathcal{S}(I)$, ..., etc)

__Theorem 1:__ Let $\mathbb{K} = DFT(k)$. Then:

$$D = W \Lambda_D W^{-1} \quad \text{and} \quad D^T = W \overline{\Lambda_D} W^{-1} \quad \text{where}$$

$$\Lambda_D = \begin{pmatrix} \mathbb{K}(0,0) & & & & & & & & & \\ & \mathbb{K}(1,0) & & & & & & & & \\ & & \ddots & & & & & & & \\ & & & \mathbb{K}(N-1,0) & & & & & & \\ & & & & \mathbb{K}(0,1) & & & & & \\ & & & & & \ddots & & & & \\ & & & & & & \mathbb{K}(N-1,1) & & & \\ & & & & & & & \ddots & & \\ & & & & & & & & \mathbb{K}(0,N-1) & \\ & & & & & & & & & \ddots \\ & & & & & & & & & & \mathbb{K}(N-1,N-1) \end{pmatrix}$$

1st col of $\mathbb{K}$ becomes 1st n diagonal entries

2nd col of $\mathbb{K}$ becomes 2nd diagonal entries

for $W = W_N \otimes W_N$ where $W_N = \left(\frac{1}{\sqrt{N}} e^{\frac{2\pi j}{N} mn}\right)_{0 \le m,n \le N-1} \in M_{N \times N}(\mathbb{C})$