ECE 472/572 - Digital Image Processing

Lecture 8 - Image Restoration – Linear, Position-Invariant Degradations

10/10/11

Recap

- Analyze the noise
  - Type of noise
    - Spatial invariant
      - SAP, Gaussian
    - Periodic noise
  - How to identify the type of noise?
    - Test pattern
    - Histogram
  - How to evaluate noise level?
    - RMSE
    - PSNR
- Noise removal
  - Spatial domain
    - Mean filters
    - Order-statistics filters
    - Adaptive filters
  - Frequency domain
    - Band-pass/Band-reject
    - Notch filters
    - Optimal notch filter *
- Analyze the blur
  - Linear, position-invariant degradation model
    - Modeled by convolution*
    - The point spread function (PSF)
      - How to estimate?
      - Deblurring - an ill-posed problem
  - Ill-conditioning of the linear system
  - Understand why image restoration is an ill-posed problem and what it means conceptually
- Different restoration approaches
  - Frequency domain
    - Inverse filter
    - Wiener filter
  - Spatial domain
    - Unconstrained approach
    - Constrained approach
    - MAP

2

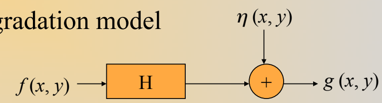
Questions

- What is PSF? How to estimate it?
- What is an ill-posed problem? What is an ill-conditioning system?
- Inverse filter and problem?
- Wiener filter and how it solved the problem?
- Unconstrained vs. Constrained approaches (572)
- What is regularization? (572)

3

Image restoration

- Degradation model



$$g(x,y) = H[f(x,y)] + \eta(x,y)$$

4

Linear vs. Non-linear

- Many types of degradation can be *approximated* by linear, space-invariant processes
- Non-linear and space-variant models are more accurate
  - Difficult to solve
  - Unsolvable

5

Linear, position-invariant degradation model

Sampling theorem →

$$f(x,y) = \int\int_{-\infty}^{\infty} f(\alpha,\beta)\delta(x-\alpha, y-\beta)d\alpha d\beta$$

$$g(x,y) = H[f(x,y)] + \eta(x,y)$$

$$= H\left[\int\int_{-\infty}^{\infty} f(\alpha,\beta)\delta(x-\alpha, y-\beta)d\alpha d\beta\right] + \eta(x,y)$$

Linearity - additivity →

$$= \int\int_{-\infty}^{\infty} H[f(\alpha,\beta)\delta(x-\alpha, y-\beta)]d\alpha d\beta + \eta(x,y)$$

Linearity - homogeneity →

$$= \int\int_{-\infty}^{\infty} f(\alpha,\beta)H[\delta(x-\alpha, y-\beta)]d\alpha d\beta + \eta(x,y)$$

$$= \int\int_{-\infty}^{\infty} f(\alpha,\beta)h(x,\alpha,y,\beta)d\alpha d\beta + \eta(x,y)$$

Space invariant
Convolution integral →

$$= \int\int_{-\infty}^{\infty} f(\alpha,\beta)h(x-\alpha, y-\beta)d\alpha d\beta + \eta(x,y)$$

6

PSF - Point Spread Function

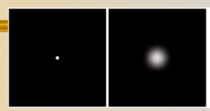
- Impulse response of system *H*

$$h(x,\alpha,y,\beta) = H[\delta(x-\alpha, y-\beta)]$$

$$g(x,y) = \int\int_{-\infty}^{\infty} f(\alpha,\beta) h(x-\alpha, y-\beta) d\alpha d\beta + \eta(x,y)$$

  - Superposition integral of the first kind
  - Convolution integral
- Point spread function (PSF)
  - Used in optics - The impulse becomes a point of light → impulse response
  - Completely characterize the linear system

Estimate the degradation



a b
FIGURE 5.24 Degradation estimation by impulse characterization. (a) An impulse of light (shown magnified). (b) Imaged (degraded) impulse.

- By observation
- By experiment
  - $g(x,y) = h(x,y)*f(x,y) + \eta(x,y)$
  - $G(u,v) = H(u,v)F(u,v) + N(u,v)$
  - $H(u,v) = G(u,v)$
- By mathematical modeling
  - Sec. 5.6.3

Image restoration – An ill-posed problem

- Degradation model

$$G(u,v) = H(u,v)F(u,v) + N(u,v)$$

$$\hat{F}(u,v) = \frac{G(u,v)}{H(u,v)} = F(u,v) + \frac{N(u,v)}{H(u,v)}$$

- *H* is ***ill-conditioned*** which makes image restoration problem an ill-posed problem
  - Solution is not stable

Ill-conditioning

$$Ax = b$$

$$A = \begin{bmatrix} 0.78 & 0.563 \\ 0.913 & 0.659 \end{bmatrix}$$

$$b = \begin{bmatrix} 0.217 \\ 0.254 \end{bmatrix}$$

$$x = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

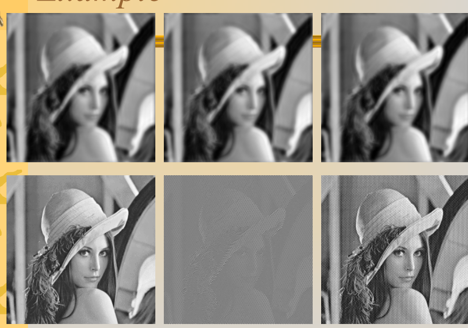
$$E = \begin{bmatrix} 0.001 & 0.001 \\ -0.002 & -0.001 \end{bmatrix}$$

$$x = \begin{bmatrix} -5 \\ 7.3085 \end{bmatrix}$$

$$cond(A) = 2.1932e+006$$

10

Example



Noise-free
Exact H | Sinusoidal noise
Exact H | Noise-free
not exact H

11

Different restoration approaches

- Frequency domain
  - Inverse filter
  - Wiener (minimum mean square error) filter
- Algebraic approaches
  - Unconstrained optimization
  - Constrained optimization
  - The regularization theory

12

The block-circulant matrix

- Stacking rows of image *f*, *g*, *n* to make *MN* x 1 column vectors **f**, **g**, and **n**. (Also called lexicographic representation of the original image). Correspondingly, **H** should be a *MN* x *MN* matrix

$$\mathbf{g} = \mathbf{Hf} + \mathbf{n}$$

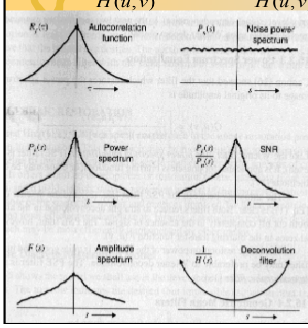
- **H** is called ***block-circulant matrix***

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_{M-1} & \mathbf{H}_{M-2} & \cdots & \mathbf{H}_1 \\ \mathbf{H}_1 & \mathbf{H}_0 & \mathbf{H}_{M-1} & \cdots & \mathbf{H}_2 \\ \mathbf{H}_2 & \mathbf{H}_1 & \mathbf{H}_0 & \cdots & \mathbf{H}_3 \\ \vdots & & & & \\ \mathbf{H}_{M-1} & \mathbf{H}_{M-2} & \mathbf{H}_{M-3} & \cdots & \mathbf{H}_0 \end{bmatrix} \quad \mathbf{H}_j = \begin{bmatrix} h_e(j,0) & h_e(j,N-1) & \cdots & h_e(j,1) \\ h_e(j,1) & h_e(j,0) & \cdots & h_e(j,2) \\ \vdots & & & \\ h_e(j,N-1) & h_e(j,N-2) & \cdots & h_e(j,0) \end{bmatrix}$$

13

Inverse filter

$$\hat{F}(u,v) = \frac{G(u,v)}{H(u,v)} = F(u,v) + \frac{N(u,v)}{H(u,v)}$$



- In most images, adjacent pixels are highly correlated, while the gray levels of widely separated pixels are only loosely correlated.
- Therefore, the autocorrelation function of typical images generally decreases away from the origin.
- Power spectrum of an image is the Fourier transform of its autocorrelation function, therefore, we can argue that the power spectrum of an image generally decreases with frequency
- Typical noise sources have either a flat power spectrum or one that decreases with frequency more slowly than typical image power spectra.
- Therefore, the expected situation is for the signal to dominate the spectrum at low frequencies while the noise dominates at high frequencies.

Wiener filter (1942)

- Objective function: find an estimate $\hat{f}$ of *f* such that the mean square error between them is minimized

$$e^2 = E\{(f - \hat{f})^2\}$$

$$\hat{F}(u,v) = \frac{1}{H(u,v)} \frac{|H(u,v)|^2}{|H(u,v)|^2 + S_n(u,v)/S_f(u,v)} G(u,v)$$

K

- Potential problems:
  - Weights all errors equally regardless of their location in the image, while the eye is considerably more tolerant of errors in dark areas and high-gradient areas in the image.
  - In minimizing the mean square error, Wiener filter also smooth the image more than the eye would prefer

15

Algebraic approach – Unconstrained restoration vs. Inverse filter

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n}$$

$$\mathbf{n} = \mathbf{g} - \mathbf{H}\mathbf{f} \xrightarrow[\text{in a least squares sense}]{\text{seek } \hat{\mathbf{f}} \text{ such that } \mathbf{H}\hat{\mathbf{f}} \text{ approximates } \mathbf{g}}$$

$$J(\hat{\mathbf{f}}) = \|\mathbf{n}\|^2 = \|\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}\|^2 \xrightarrow[\text{with respect to } \hat{\mathbf{f}}]{\text{differentiate right hand side}}$$

$$\frac{\partial J(\hat{\mathbf{f}})}{\partial \hat{\mathbf{f}}} = 0 = -2\mathbf{H}^T(\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}) \Rightarrow \mathbf{H}^T\mathbf{H}\hat{\mathbf{f}} = \mathbf{H}^T\mathbf{g}$$

$$\hat{\mathbf{f}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{g} = \mathbf{H}^{-1}(\mathbf{H}^T)^{-1}\mathbf{H}^T\mathbf{g} = \mathbf{H}^{-1}\mathbf{g}$$

Compared to the inverse filter: $\hat{F}(u,v) = \frac{G(u,v)}{H(u,v)}$

16

Algebraic approach – Constrained restoration vs. Wiener filter

Minimizing $\|\mathbf{Q}\hat{\mathbf{f}}\|^2$, where $\mathbf{Q}$ is a linear operator on $\mathbf{f}$,

subject to the constraint $\|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 = \|\mathbf{n}\|^2$.

Model this problem using Lagrange optimization method

We seek $\hat{\mathbf{f}}$ that minimizes the criterion (or objective) function

$$J(\hat{\mathbf{f}}) = \|\mathbf{Q}\hat{\mathbf{f}}\|^2 + \alpha\left(\|\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}\|^2 - \|\mathbf{n}\|^2\right)$$

α is a constant, called the Lagrange multiplier.

$$\frac{\partial J(\hat{\mathbf{f}})}{\partial \hat{\mathbf{f}}} = 0 = 2\mathbf{Q}^T\mathbf{Q}\hat{\mathbf{f}} - 2\alpha\mathbf{H}^T(\mathbf{g} - \mathbf{H}\hat{\mathbf{f}})$$

$$\hat{\mathbf{f}} = (\mathbf{H}^T\mathbf{H} + \gamma\mathbf{Q}^T\mathbf{Q})^{-1}\mathbf{H}^T\mathbf{g}$$

Compared to: $\hat{F}(u,v) = \frac{1}{H(u,v)} \frac{|H(u,v)|^2}{|H(u,v)|^2 + S_n(u,v)/S_f(u,v)} G(u,v)$

17

Regularization theory

- Generally speaking, any regularization method tries to analyze a related ***well-posed*** problem whose solution approximates the original ***ill-posed*** problem.
- The well-posedness is achieved by implementing one or more of the following basic ideas
  - restriction of the data;
  - change of the space and/or topologies;
  - modification of the operator itself;
  - the concept of regularization operators; and
  - well-posed stochastic extensions of ill-posed problems.

18

Solution formulation

- For $\mathbf{g} = \mathbf{Hf} + \boldsymbol{\eta}$, the regularization method constructs the solution as

$$\min_{\mathbf{f}}\left[u(\mathbf{f},\mathbf{g}) + \beta v(\mathbf{f})\right]$$

- $u(\mathbf{f}, \mathbf{g})$ describes how the real image data is related to the degraded data. In other words, this term models the characteristic of the imaging system.
- $\beta v(\mathbf{f})$ is the regularization term with the regularization operator v operating on the original image **f**, and the regularization parameter β used to tune up the weight of the regularization term.
- By adding the regularization term, the original ill-posed problem turns into a well-posed one, that is, the insertion of the regularization operator puts some constraints on what **f** might be, which makes the solution more stable.

19

MAP (maximum a-posteriori probability)

- Formulate solution from statistical point of view: MAP approach tries to find an estimate of image **f** that maximizes the *a-posteriori* probability $p(\mathbf{f}|\mathbf{g})$ as

$$\hat{\mathbf{f}} = \arg\max_{\mathbf{f}} p(\mathbf{f} \mid \mathbf{g})$$

- According to Bayes' rule,

$$p(\mathbf{f} \mid \mathbf{g}) = \frac{p(\mathbf{g} \mid \mathbf{f})P(\mathbf{f})}{P(\mathbf{g})}$$

  - P(**f**) is the *a-priori* probability of the unknown image **f.** We call it the **prior model**
  - P(**g**) is the probability of **g** which is a constant when **g** is given
  - $p(\mathbf{g}|\mathbf{f})$ is the conditional probability density function (pdf) of **g**. We call it the **sensor model**, which is a description of the noisy or stochastic processes that relate the original unknown image **f** to the measured image **g**.

20

MAP - Derivation

- Bayes interpretation of regularization theory

$$\hat{\mathbf{f}} = \arg\max_{\mathbf{f}} p(\mathbf{f} \mid \mathbf{g}) = \arg\max_{\mathbf{f}}\left[p(\mathbf{g} \mid \mathbf{f})P(\mathbf{f})\right]$$

$$\text{Let } \Omega = -\ln\left[p(\mathbf{g} \mid \mathbf{f})P(\mathbf{f})\right] = -\ln\left[p(\mathbf{g} \mid \mathbf{f})\right] - \ln\left[P(\mathbf{f})\right]$$

$$\Omega_n = -\ln\left[p(\mathbf{g} \mid \mathbf{f})\right] \quad \text{Noise term}$$

$$\Omega_p = -\ln\left[P(\mathbf{f})\right] \quad \text{Prior term}$$

21

The noise term

- Assume Gaussian noise of zero mean, σ the standard deviation

$$\Omega_n = -\ln\left[p(\mathbf{g} \mid \mathbf{f})\right] = -\ln\left[u\left(\|\mathbf{Hf} - \mathbf{g}\|\right)\right]$$

$$= -\ln\left[\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(\mathbf{Hf}-\mathbf{g})^2}{2\sigma^2}\right)\right]$$

$$= \frac{1}{2\sigma^2}\sum_{i=0}^{MN-1}\left(f \otimes h - g\right)_i^2$$

22

The prior model

- The *a-priori* probability of an image by a Gibbs distribution is defined as

$$P(f) = \frac{\exp\left(-\frac{U(f)}{T}\right)}{Z}$$

  - U(f) is the energy function
  - T is the temperature of the model
  - Z is a normalization constant

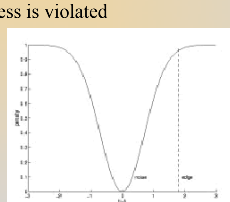
$$\Omega_p = -\ln\left[P(\mathbf{f})\right] = -\ln\left[\frac{\exp\left(-U(\mathbf{f})/T\right)}{Z}\right] = \frac{U(\mathbf{f})}{T}$$

23

The prior model (cont')

- U(f), the prior energy function, is usually formulated based on the smoothness property of the original image. Therefore, U(f) should measure the extent to which the smoothness is violated

punishment



Difference between neighborhood pixels

24

The prior model (cont')

$$\frac{U(\mathbf{f})}{T} = \left[-\frac{\beta}{\sqrt{2\pi}\tau} \exp\left(-\frac{(\nabla^k \mathbf{f})^2}{2\tau^2} \right) \right] \Big/ T$$

$$= \sum_{i=0}^{MN-1} \left[-\frac{\beta}{\sqrt{2\pi}\tau} \exp\left(-\frac{(f \otimes r)_i^2}{2\tau^2} \right) \right] \Big/ T$$

- β is the parameter that adjusts how smooth the image goes
- The k-th derivative models the difference between neighbor pixels. It can also be approximated by convolution with the right kernel

25

The prior model – Kernel r

- Laplacian kernel

$$\frac{\partial f^2}{\partial x^2} + \frac{\partial f^2}{\partial y^2}$$

$$r = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

26

The objective function

$$\Omega = \Omega_n + \Omega_p$$

$$= \frac{1}{2\sigma^2} \sum_{i=0}^{MN-1} (f \otimes h - g)_i^2 - \frac{\beta}{\sqrt{2\pi}\tau T} \sum_{i=0}^{MN-1} \exp\left(-\frac{(f \otimes r)_i^2}{2\tau^2} \right)$$

- Use gradient descent to solve f

$$f^{k+1} = f^k - \alpha \frac{\partial \Omega}{\partial f}$$

$$\frac{\partial \Omega}{\partial f} = \frac{1}{\sigma^2} \sum_{i=0}^{MN-1} \left[(f \otimes h - g) \otimes h_{rev} \right] + \sum_{i=0}^{MN-1} \left\{ \left[\frac{\beta (f \otimes r)}{\sqrt{2\pi}\tau^3 T} \exp\left(-\frac{(f \otimes r)^2}{2\tau^2} \right) \right] \otimes r_{rev} \right\}_i$$

27