ECEN 642 Digital Image Processing

Lecture 6: Image Restoration

Notes & Chapter 5

ECE Department

Texas A&M University

Image Restoration

- Image Restoration, like image enhancement, aims to improve the pictorial information in an image.
- Unlike image enhancement, it is assumed in image restoration that some information about the corrupting process is known, in the form of a degradation model.
- The degradation model allows one to mathematically obtain an operation that can best "undo" the degradation in some sense, and thus restore the original, uncorrupted image.
- Such optimal operation is called the restoration filter.

Degradation Model

- It is common to assume an additive model

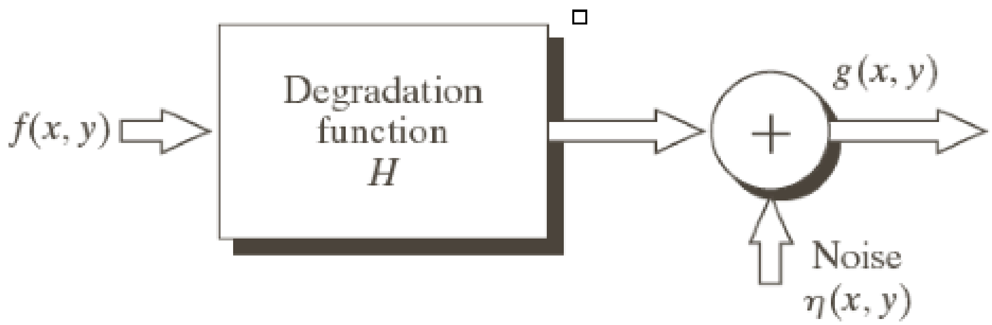
  corrupted image = degradation + noise

- The degradation is an unwanted distortion introduced by the imaging system. For example:
  - The point spread function (PSF) of an optical system (lens).
  - Motion blurring produced by a moving object.
  - Atmospheric blurring in remote sensing applications.

- The noise representes a spurious component that is added to the image. For example:
  - Statistical sensor noise (usually accentuated under low-light conditions).
  - Thermal sensor noise.
  - Etc.

Linear Degradation Model

- The linear degradation model has a long history in engineering and statistics. Here, two additional things are assumed:
  - The degradation can be represented by a shift-invariant linear operator, i.e., by convolution.
  - The noise is uncorrelated or white noise.



In the spatial domain:

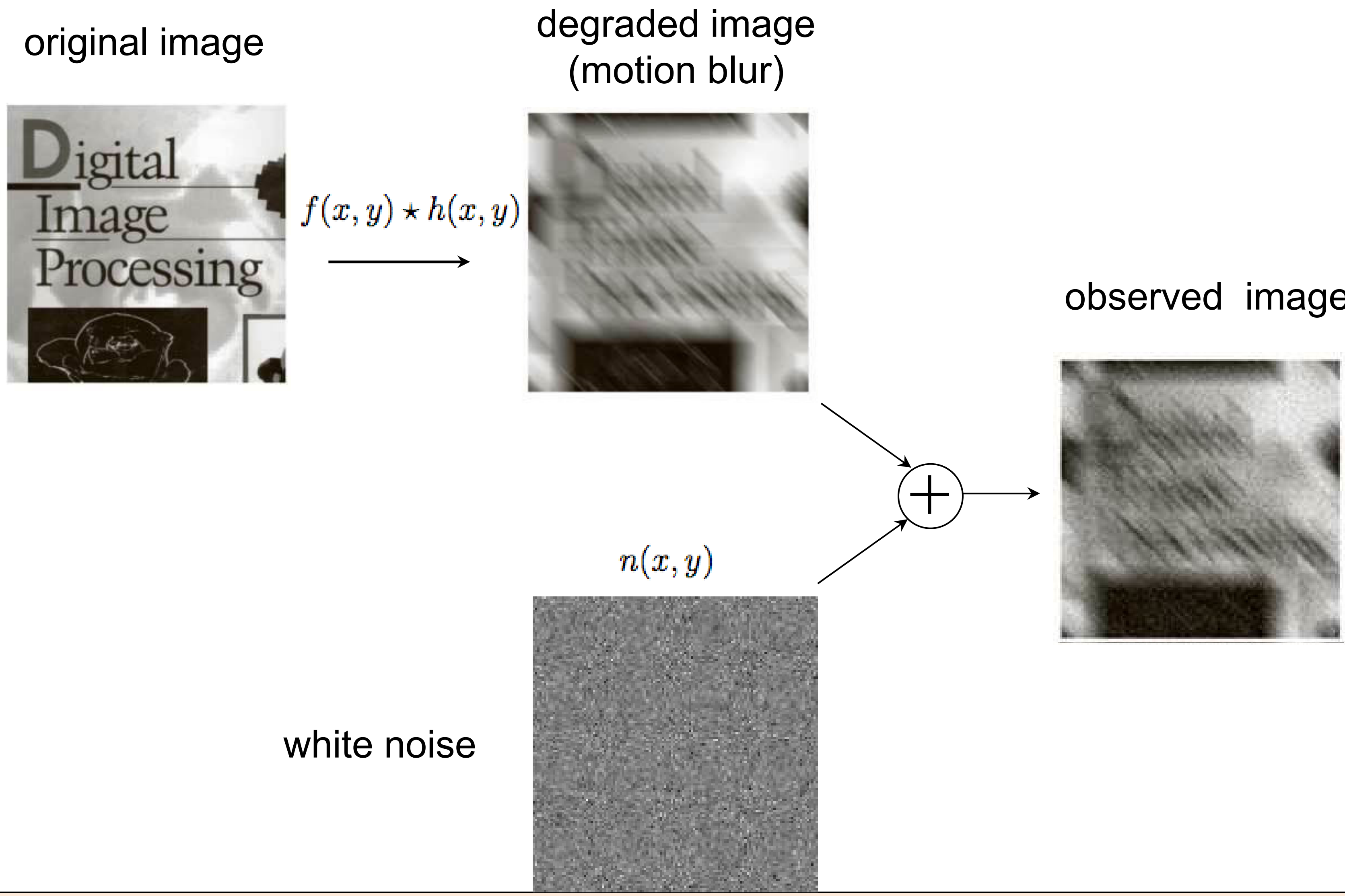
$$g(x,y) = f(x,y) * h(x,y) + n(x,y)$$

or, in the frequency domain:

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

Example –Image Model

corrupted image = degradation + noise



Statistical Properties of Noise

- We assume that the noise is a realization from an underlying random process (random image) n(x,y).

- We assume that n(x,y) is identically distributed, meaning that the noise has the same distribution for all pixels. In particular, we have

$$E[n(x,y)] = \mu \qquad \text{and} \qquad \mathrm{Var}[n(x,y)] = \sigma^2$$

  independently of the pixel (x,y).

- It is also often assumed that the noise is zero-mean, that is, $\mu = 0$.

- We also assume that n(x,y) is uncorrelated, meaning that the noise affecting one pixel is uncorrelated with the noise affecting any other distinct pixel. Assuming zero-mean noise, this is equivalent to

$$E[n(x,y)n(x+s,y+t)] = 0\,, \quad \text{for all } s,t \neq 0$$

Cross-Correlation and Auto-Correlation

- Given two random images f(x,y) and g(x,y) one defines the cross-correlation function as the image

$$R_{fg}(s,t) = E[f(x,y)g(x+s,y+t)]$$

- When f = g, this defines the auto-correlation function of f(x,y):

$$R_{ff}(s,t) = E[f(x,y)f(x+s,y+t)]$$

- This implicitly assume that the correlations do not depend on (x,y), but only on the shift (s,t).

- We talk about the wide-sense stationarity (WSS) property if:
  - Mean is constant
  - Correlation depends only on the shift.

Relationship with Correlation Operator

- Under WSS, the cross-correlation and auto-correlation functions can be approximated as

$$R_{fg}(s,t) \approx \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m,n) g(s+m, t+n) = \frac{1}{MN} [f \odot g](s,t)$$

$$R_{ff}(s,t) \approx \frac{1}{MN} [f \odot f](s,t)$$

We can compute the Fourier Transform (continuous or DFT) of the auto-correlation, which gives the power spectral density (PSD) of f:

$$S_f(u,v) = \Im \left[R_{ff}(s,t) \right]$$

The PSD receives this name because it can be shown that its integration (summation) over a range of frequencies gives the power of the signal contained in that frequency range.

Power Spectral Density - II

- Using the correlation theorem

$$f(x,y) \odot g(x,y) \stackrel{\Im}{\longleftrightarrow} F^*(u,v)G(u,v)$$

it follows that

$$S_f(u,v) = \Im\left[R_{ff}(s,t)\right] \approx \frac{1}{MN}\Im\left[[f \odot f](s,t)\right]$$

$$= \frac{1}{MN} F^*(u,v)F(u,v) = \frac{1}{MN}\left|F(u,v)\right|^2 = I(u,v)$$

- The quantity I(u,v) is called the periodogram of f(x,y). It is an approximation to the PSD.

- While the correlation scaled by 1/MN (which can be obtained as the IDFT of I(u,v)) yields a good approximation to the autocorrelation function, the periodogram is usually a poor approximation to the PSD.

Uncorrelated Noise (White Noise)

- If the noise n(x,y) is identically-distributed, zero-mean, and uncorrelated (which in particular imply WSS), then it follows that

$$R_{nn}(s,t) = E[n(x,y)n(x+s,y+t)] = 0\,, \quad \text{for all } s,t \neq 0$$

and

$$R_{nn}(0,0) = E[n(x,y)n(x,y)] = \sigma^2$$

- In other words:

$$R_{nn}(s,t) = \sigma^2 \delta(s,t)$$

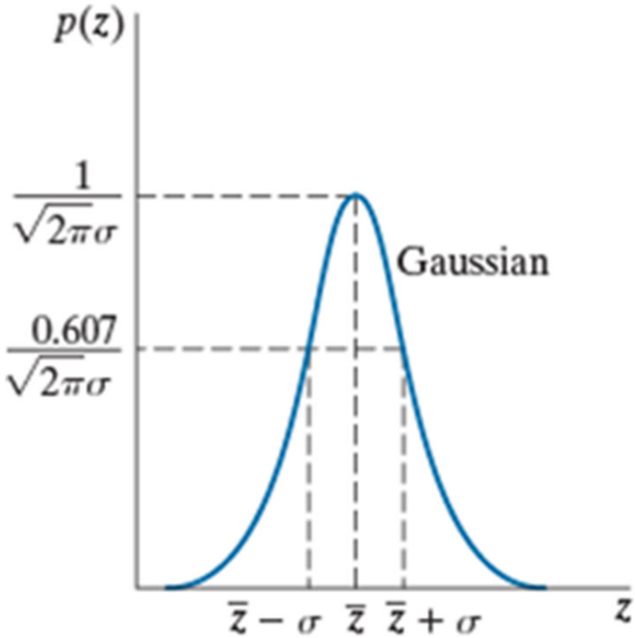
- This means that the PSD of the noise is a constant

$$S_n(u,v) = \sigma^2$$

that is, all frequency ranges contain the same power, hence the term white noise (as opposed to "colored noise" for correlated noise).

Gaussian Noise

- The most common distribution assumed for noise is the Gaussian distribution:



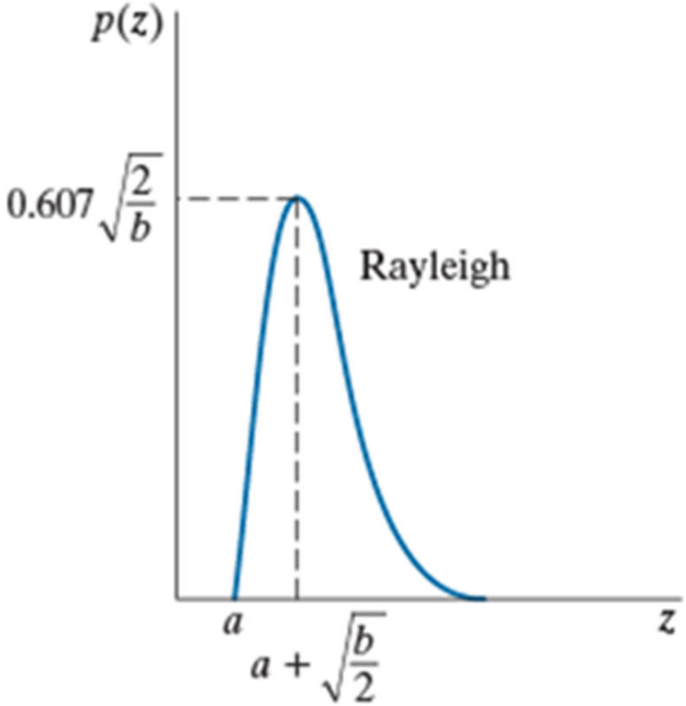
$$p(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-\bar{z})^2}{2\sigma^2}} \qquad -\infty < z < \infty$$

where z represents intensity
$\bar{z}$ is the mean
σ is the standard deviation

- For Gaussian noise, uncorrelated noise means independent noise.

Rayleigh Noise

- This distribution is useful for modeling the shape of skewed histograms:



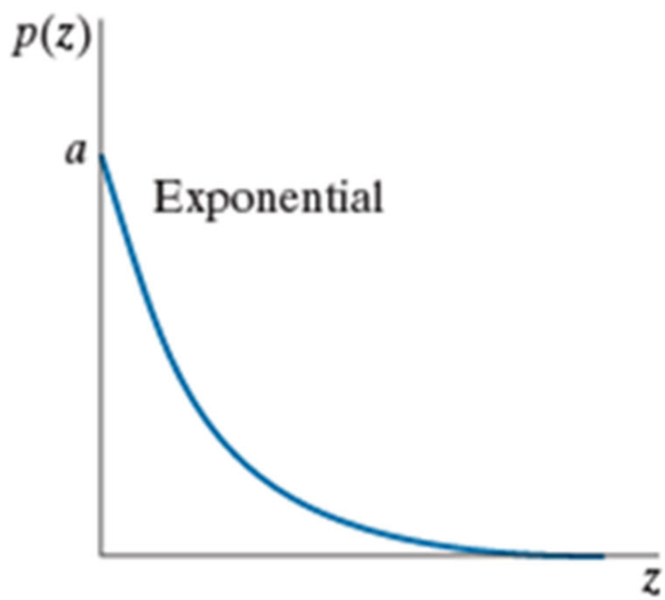
$$p(z) = \begin{cases} \frac{2}{b}(z-a)e^{-(z-a)^2/b} & z \geq a \\ 0 & z < a \end{cases}$$

$$\bar{z} = a + \sqrt{\pi b/4}$$

$$\sigma^2 = \frac{b(4-\pi)}{4}$$

Exponential noise

- Gaussian noise is symmetric around the mean (for zero-mean noise, this implies that the noise is equally likely to be positive or negative).
- Some processes introduce noise that is asymmetric (usually strictly positive). A common case is exponential noise:



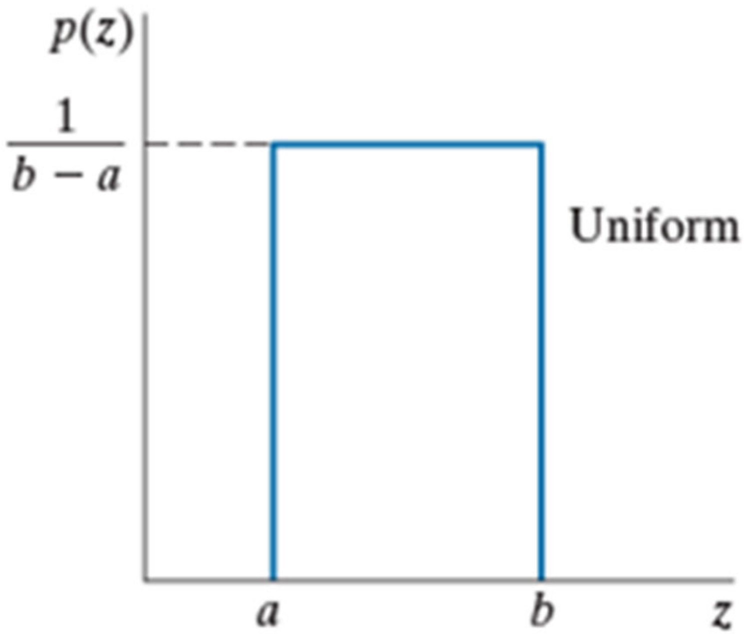
$$p(z) = \begin{cases} ae^{-az} & z \geq 0 \\ 0 & z < 0 \end{cases}$$

$$\bar{z} = \frac{1}{a}$$

$$\sigma^2 = \frac{1}{a^2}$$

Uniform Noise

- Other processes produce noise that take equally likely values over a range [a,b]. This is the case of uniform noise:



$$p(z) = \begin{cases} \dfrac{1}{b-a} & a \le z \le b \\ 0 & \text{otherwise} \end{cases}$$

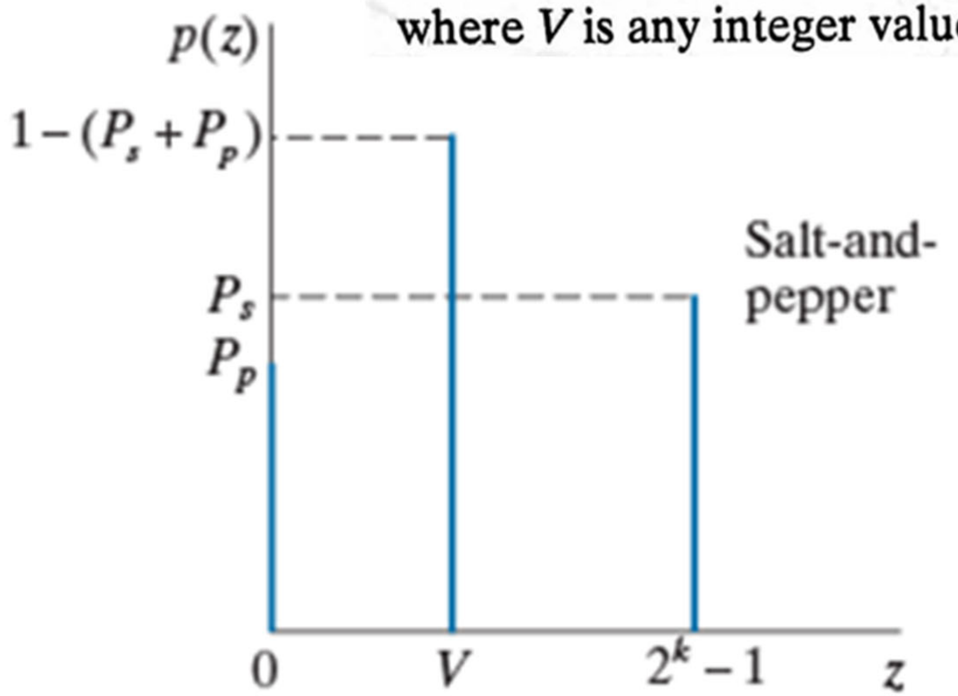
$$\bar{z} = \frac{a+b}{2}$$

$$\sigma^2 = \frac{(b-a)^2}{12}$$

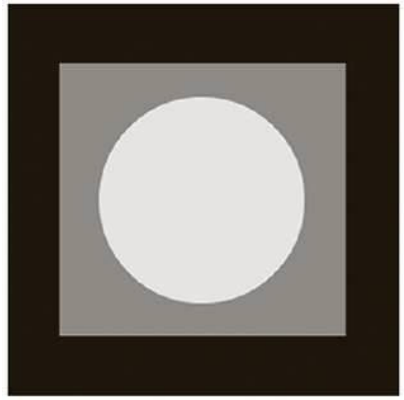
Salt and Pepper Noise

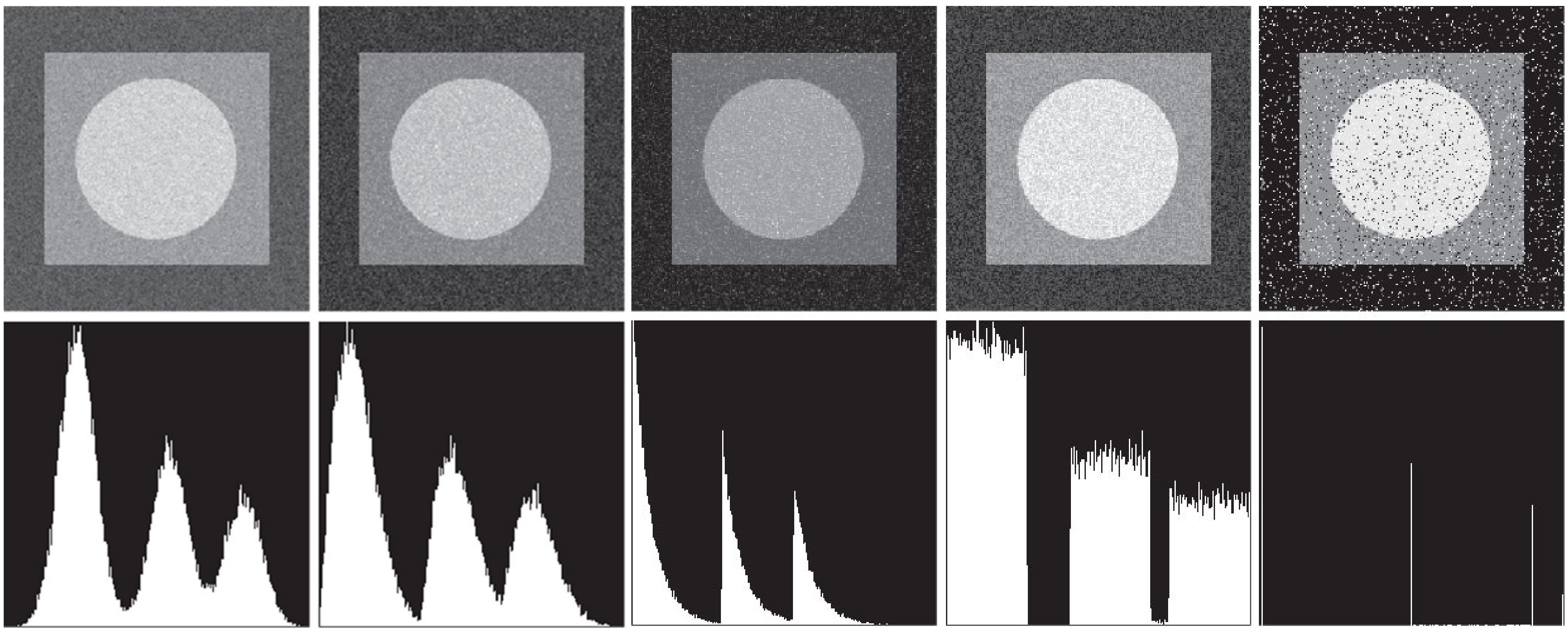
$$p(z) = \begin{cases} P_s & \text{for } z = 2^k - 1 \\ P_p & \text{for } z = 0 \\ 1 - (P_s + P_p) & \text{for } z = V \end{cases}$$

where V is any integer value in the range $0 < V < 2^k - 1$.



Example – Test pattern for illustrating noise models





Restoration: Noise-Only Model

- Assume that the degradation operation is identitity, so that image corruption consists solely of noise:

$$g(x,y) = f(x,y) * h(x,y) + n(x,y) = f(x,y) + n(x,y)$$

Note that

$$E[g(x,y)] = E[f(x,y) + n(x,y)] = f(x,y) + E[n(x,y)] = f(x,y)$$

and

$$\mathrm{Var}[g(x,y)] = \mathrm{Var}[f(x,y) + n(x,y)] = 0 + \mathrm{Var}[n(x,y)] = \sigma^2$$

Image Averaging

- The most effective method in this case is to take multiple independent observations of the scene. If K observations $g_i(x,y)$ of the image are available, let

$$\hat{f}(x,y) = \frac{1}{K}\sum_{i=1}^{K} g_i(x,y)$$

We have that

$$E[\hat{f}(x,y)] = \frac{1}{K}\sum_{i=1}^{K} E[g_i(x,y)] = \frac{1}{K} K f(x,y) = f(x,y)$$

and

$$\mathrm{Var}[\hat{f}(x,y)] = \frac{1}{K^2}\sum_{i=1}^{K} \mathrm{Var}[g_i(x,y)] = \frac{1}{K^2} K\sigma^2 = \frac{\sigma^2}{K}$$

Image Averaging - II

- Therefore, the estimator is unbiased

$$E[\hat{f}(x,y)] = f(x,y)$$

and its variance tends to zero:

$$\mathrm{Var}[\hat{f}(x,y)] = \frac{\sigma^2}{K} \rightarrow 0 \;\text{ as }\; K \rightarrow \infty$$

We say that $\hat{f}(x,y) \rightarrow f(x,y)$ in the mean-square sense.

- Note however that the standard deviation of the noise goes to zero only as the square root of K:

$$\mathrm{std}[\hat{f}(x,y)] = \frac{\sigma}{\sqrt{K}}$$

which can be slow.

Example

1- Noise-only model:

$$g(x,y) = f(x,y) * h(x,y) + n(x,y) = f(x,y) + n(x,y)$$

The most effective method in this case is to take multiple independent observations of the scene & average!

noisy image

image averaging with K = 5

image averaging with K = 50

Restoration: Degradation-Only Model

- Assume that there is no noise, so that image distortion consists solely of the degradation functional:

$$g(x,y) = f(x,y) * h(x,y)$$

- In such a case, one attempts to find a deconvolution filter $\hat{h}(x,y)$ that will attempt to "undo" the distortion introduced by $h(x,y)$

- We have that

$$\hat{f}(x,y) = \hat{h}(x,y) * g(x,y) = \hat{h}(x,y) * [f(x,y) * h(x,y)]$$

$$= \left[\hat{h}(x,y) * h(x,y)\right] * f(x,y)$$

Therefore, we need to have

$$\hat{h}(x,y) * h(x,y) = \delta(x,y)$$

Inverse Filter

- The previous equation can be solved easily by transforming to the frequency domain:

$$\hat{H}(u,v)\,H(u,v) = 1 \Rightarrow \hat{H}(u,v) = \frac{1}{H(u,v)}$$

- In other words, one just inverts the frequency response of the degradation. For this reason, this filter is called the inverse filter.

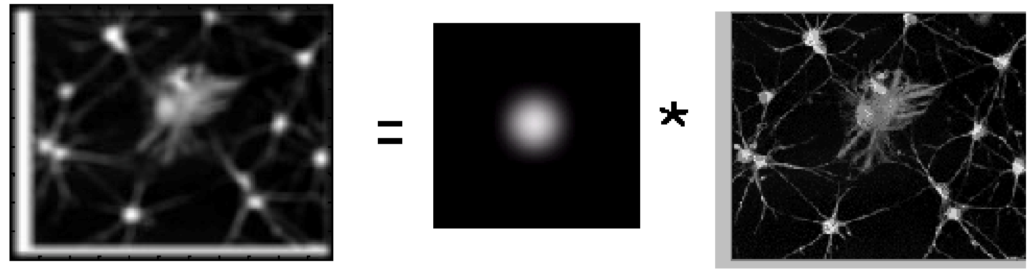
- Also for obvious reasons, this process is often called deconvolution.

- In practice, one does not know H(u,v), and it must be estimated. The restoration process is then sometimes called blind deconvolution. Estimation of H(u,v) is application-specific.

Example

2- Degradation-only model:

$$g(x,y) = f(x,y) * h(x,y)$$



g(x,y) = h(x,y) * f(x,y)

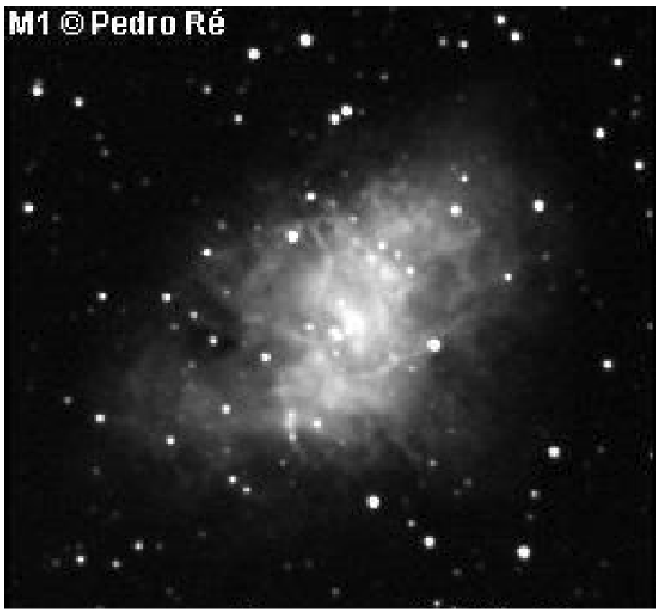
In such a case, one attempts to find a deconvolution filter that will attempt to "undo" the distortion introduced by $h(x,y)$

Estimation of the Degradation Function:

- Image Observation
- Experimentation
- Mathematical Modeling

Example of Deconvolution

- In astronomical imaging, stars are essentially points of light, that is, impulses. The degradation function, also known in this context as the point-spread function (PSF), is the system impulse response, and thus can be found from images of isolated stars.



original image of M1

deconvolved image

Restoration: Full Degradation Model

- In practice, it is unrealistic to assume a noise-free model.
- Noise is always present, and if one tries to apply the inverse filter in this case, even if one knows the exact degradation function, one gets

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

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

- The degradation frequency response H(u,v) will usually have values close to zero, which make the spurious second term large. We say then that **the noise "dominates" the restored result.**

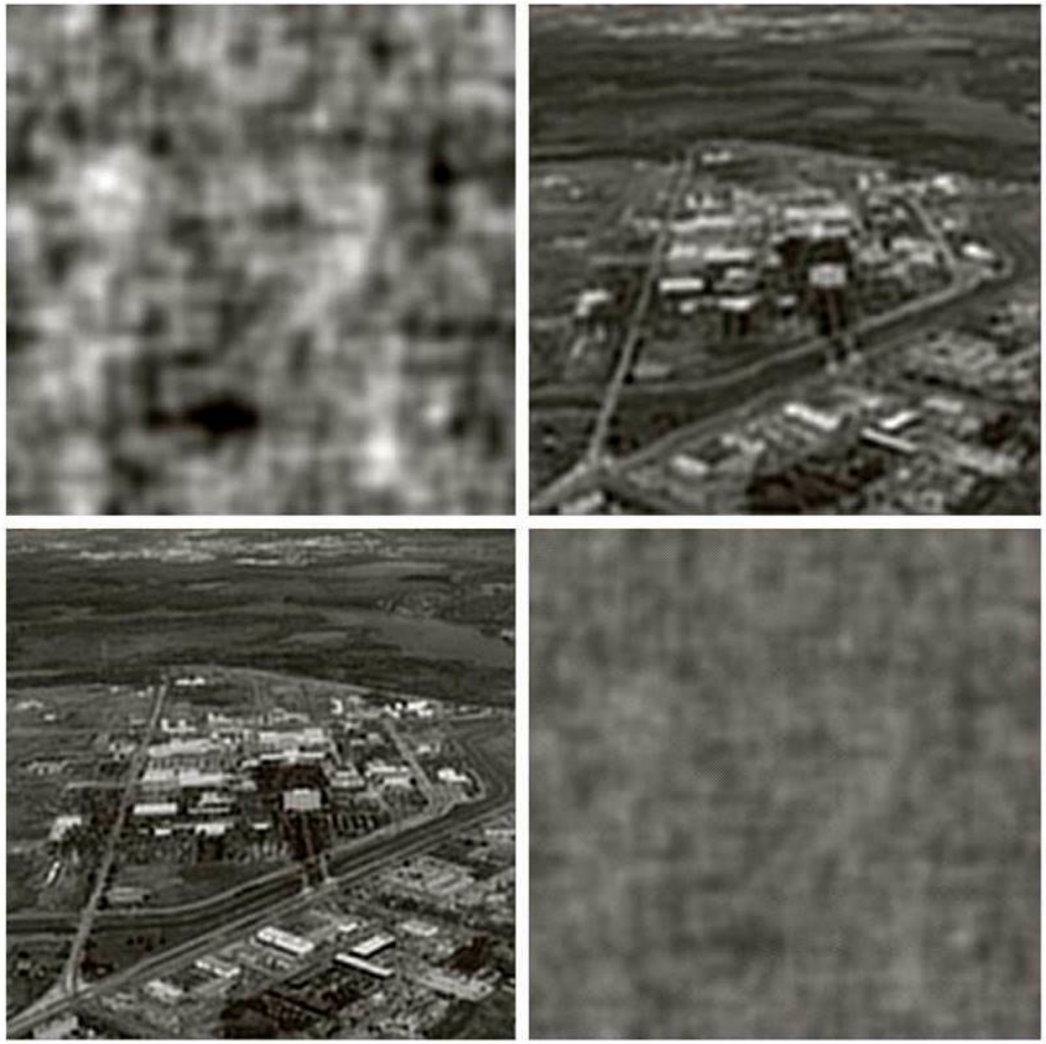
- As the troublesome spots with near-zero H(u,v) tend to occur away from the origin, the situation may be remedied by l**owpassing** the result prior to taking the IDFT.

Example of Deconvolution in Noise

Original

Degraded

Example of Deconvolution in Noise



direct
inverse filter

inverse filter w/
lowpass D_0=40

inverse filter w/
lowpass D_0=70

inverse filter w/
lowpass D_0=85

Wiener Filter

- If deconvolution does not work, or performs poorly, one should use the Wiener filter.
- The origins of the Wiener filter lie in the pioneering theoretical work of Norbert Wiener, the founder of modern Signal Processing (and Cybernetics), and of Eberhard Hopf.
- The Wiener filter is the optimal linear MSE filter. In our setting, this means that, given the original image f(x,y), the degraded image g(x,y), and the restored image

$$\hat{f}(x,y) = \hat{h}(x,y) * g(x,y)$$

the mean-square error (MSE)

$$\mathrm{MSE}(\hat{f}(x,y)) = E\left[(\hat{f}(x,y) - f(x,y))^2\right]$$

is minimized when $\hat{h}(x,y)$ is chosen to be the Wiener filter.

Wiener Filter for Linear Degradation Model

$$\hat{f}(x,y) = \hat{h}(x,y) * g(x,y)$$

$$\mathrm{MSE}(\hat{f}(x,y)) = E\left[(\hat{f}(x,y) - f(x,y))^2\right]$$

- In the frequency domain:

$$\hat{F}(u,v) = \left[\frac{H^*(u,v)S_f(u,v)}{S_f(u,v)|H(u,v)|^2 + S_\eta(u,v)}\right]G(u,v)$$

Wiener Filter for Linear Degradation Model

- Therefore:

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

$\hat{F}(u,v)$ = FT of the estimated undegraded image

$G(u,v)$ = FT of the degraded image

$H(u,v)$ = degradation transfer function (FT of the spatial degradation);

$H^*(u,v)$ = complex conjugate

$|H(u,v)|^2 = H^*(u,v)H(u,v)$

$S_\eta(u,v)$ = $|N(u,v)|^2$ = power spectrum of the noise

$S_f(u,v)$ = $|F(u,v)|^2$ = power spectrum of the undegraded image

Signal-to-Noise Ratio

- The ratio

$$\mathrm{SNR}(u,v) = \frac{S_f(u,v)}{S_n(u,v)}$$

is called the signal-to-noise ratio (SNR) at frequency (u,v).

- It is easy to see that

$$\mathrm{SNR}(u,v) \approx 0 \;\Rightarrow\; \hat{H}(u,v) \approx 0$$

that is, if the SNR is very low at a particular frequency, the Wiener filter suppresses that frequency from the output. On the other hand,

$$\mathrm{SNR}(u,v) \approx \infty \;\Rightarrow\; \hat{H}(u,v) \approx \frac{1}{H(u,v)}$$

that is, if the SNR is very high at a particular frequency, the Wiener filter behaves as an ordinary deconvolution filter at that frequency.

Signal-to-Noise Ratio - II

- In fact, if there is no noise (i.e., we have a degradation-only model), then $S_n(u,v) = 0$ for all (u,v), and the Winer filter reduces to the inverse filter, showing that the latter is indeed the optimal linear MSE filter in this case.

- In practice, estimation of SNR(u,v) may not be trivial. It is common to assume in such cases a constant SNR, in which case the frequency response of the restoration filter becomes

$$\hat{H}(u,v) = \frac{1}{H(u,v)} \frac{|H(u,v)|^2}{|H(u,v)|^2 + K}$$

where K = 1/SNR.

Wiener Filter Example

degraded image | inverse filter | Wiener filter

high noise

medium noise

low noise