AI61201: Visual Computing With AI/ML

Module 3: Image Transformations in the Spatial Domain

Dr. Somdyuti Paul

Image Histograms

- Image histograms constitute a basic tool that facilitate point operation on images.

- The image histogram represents the distribution of pixel intensities of the image, often represented graphically.

- For a grayscale image, the histogram is thus a mapping from the gray level values to the frequency of occurrence of each gray level value in the image.

- Thus, for a $M \times N$ image having L gray levels, the histogram is given by
  $p(k) = \dfrac{n_k}{MN}$ for $0 \leq k \leq (L-1)$
  where n_k is the number of occurrences of gray level k in the image and could take values between 0 and MN.

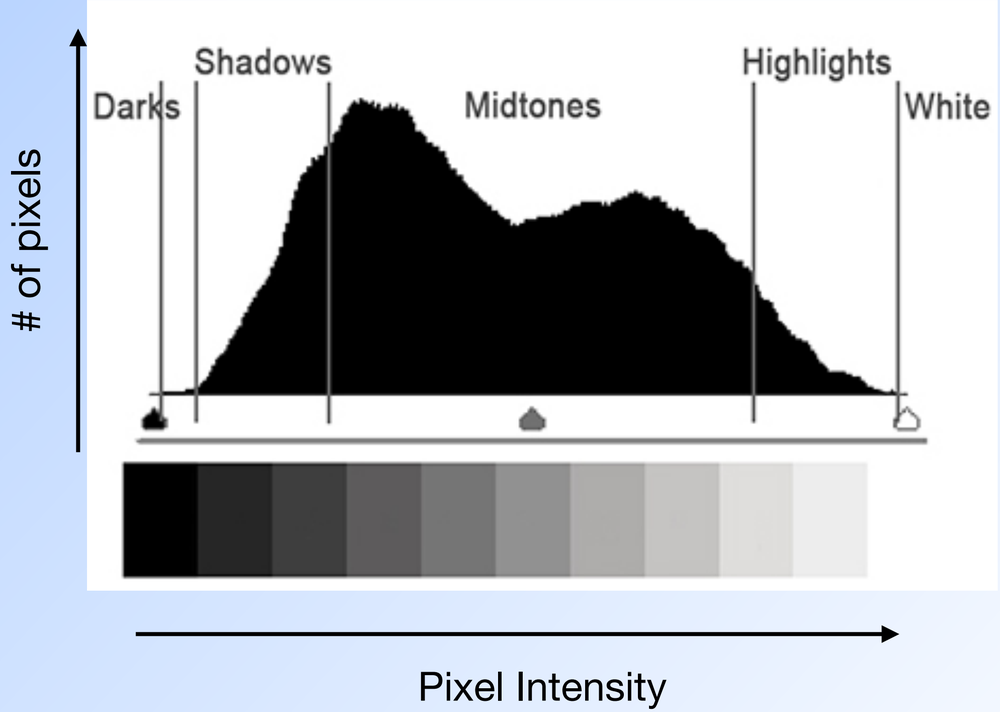
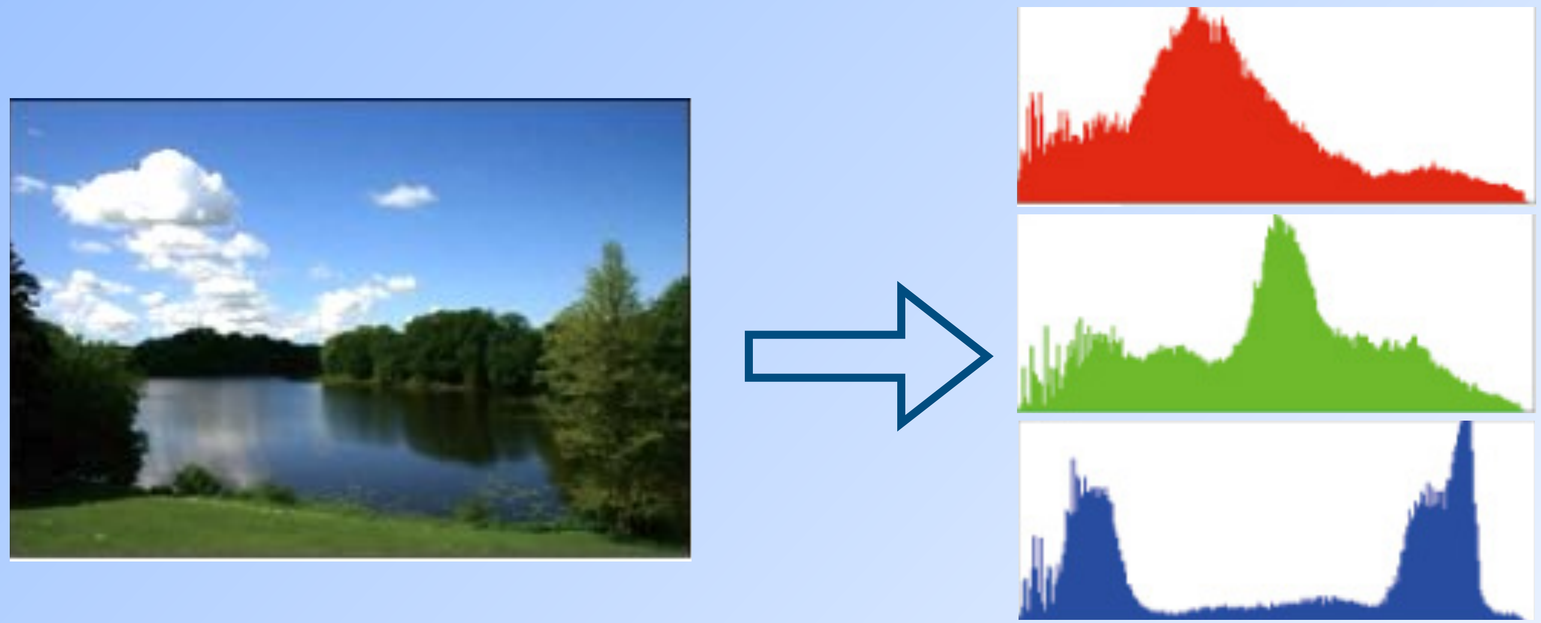


Image Histogram

Image Histograms

- For color images, the histogram could be formed in two different ways:
  - By individually representing the histograms of each color channel
  - As a single histogram over the range of all possible colors.



Histogram Representation of Color Images

Histogram Equalization

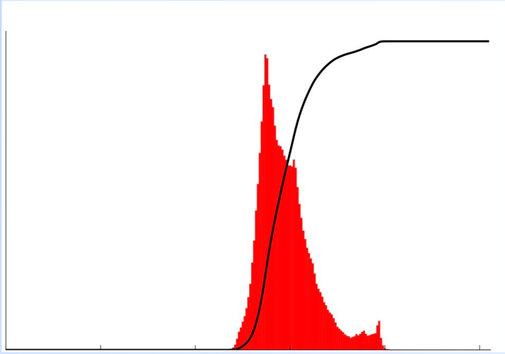
- An image suffers from low contrast when its pixel intensities vary within a small range.
- The contrast of an image can be enhanced by altering the distribution of intensity levels, in a way such that the full range of gray levels is utilized.
- It is achieved by applying a point wise transformation of the form:
  $k' = T(k)$, where $T(k)$ is a monotonically increasing function, and $0 \leq T(k) \leq (L-1)$.
- One such transformation, which gives a uniform intensity distribution is given as:

$$T(k) = (L-1)\sum_{j=0}^{k}\frac{n_j}{MN}$$

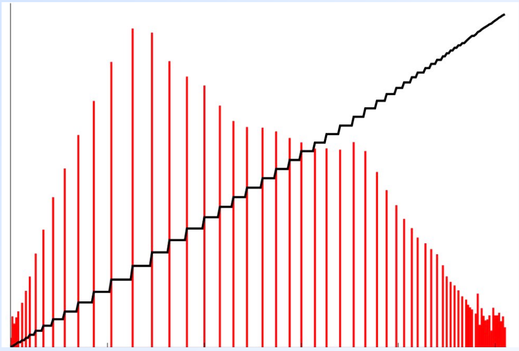
where L is the number of intensity levels.

Original Image

Histogram Equalized Image



Histogram of Original Image



Equalized Histogram

Histogram Equalization

Histogram Equalization

Advantages	Disadvantages
• Enhances visibility of details • Fully automatic • Computationally inexpensive	• Might amplify noise • Can lead to loss of local details • Processes the entire image in the same manner • Might disrupt natural color balance when applied to color images.

- Other useful operations that can be performed using image histograms include and histogram matching, adaptive thresholding, lossless compression etc.

Linear Filtering in the Spatial Domain

- A filter is said to be linear if it satisfies the principles of additivity and homogeneity, i.e.

- For an spatial signal $f(x, y)$ (such as an image), linear filtering involves performing a sum-of-products between $f(x, y)$ and a filter kernel $h(x, y)$ of size $M \times N$ as follows:

  $$g(x, y) = \sum_{i=-a}^{a} \sum_{j=-b}^{b} h(i, j) f(x - i, y - j)$$

  where $M = 2a + 1$ and $N = 2b + 1$.

- In case of spatial filtering, the filter kernel $h(x, y)$ is represented as a matrix and is also called mask, template or window.

- If the input is of size $W \times H$, and the filter kernel is of size $M \times N$, the filtered output is of size $(W - M + 1) \times (H - N + 1)$.

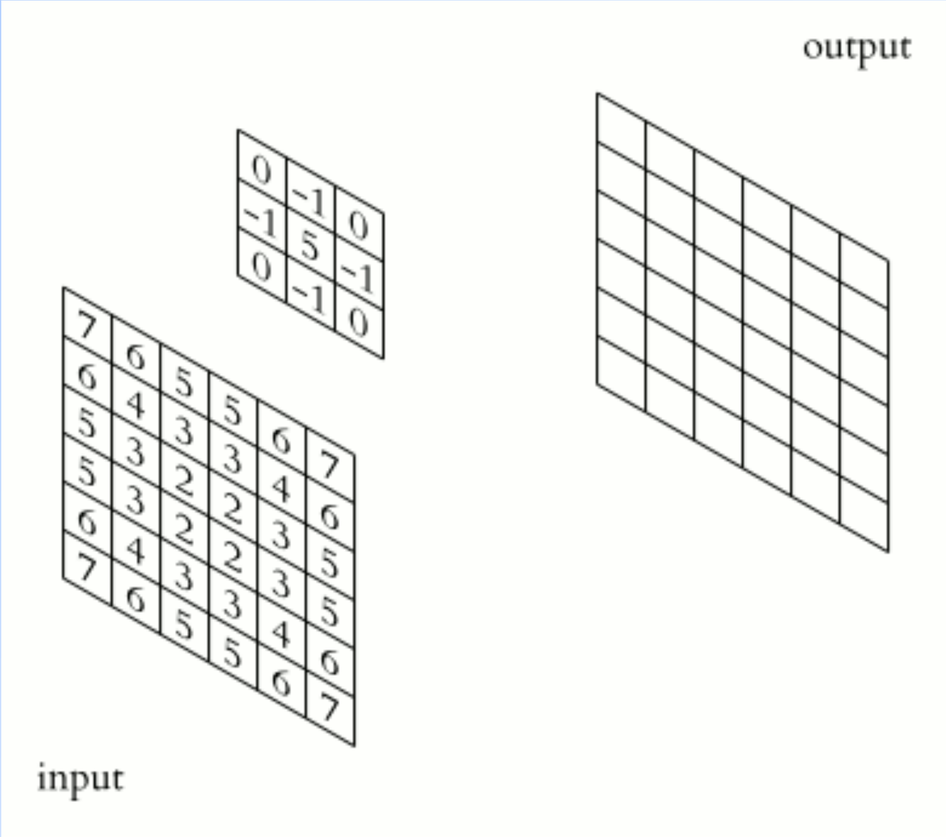
Spatial Convolution

- The process of sliding the filter kernel over the image and performing a weighted sum of pixel values in the neighborhood defined by the kernel is called spatial convolution:

$$g(x, y) = f(x, y) * h(x, y) = \sum_{i=-a}^{a} \sum_{j=-b}^{b} f(i, j) h(x - i, y - j).$$

- An associated operation is spatial correlation, which is identical to the convolution operation, except that the kernel is not flipped horizontally and vertically.

- Sometimes, a padding $(M - 1)/2$ and $(N - 1)/2$ pixels along the horizontal and vertical dimensions respectively are applied to make the output size the same as the input size.



Spatial Convolution

Spatial Convolution

- Properties of the convolution operation:
  - Commutative: $h * f = f * h$
  - Associative: $h_1 * (h_2 * f) = (h_1 * h_2) * f$
  - Distributive:$(f_1 + f_2) * h = f_1 * h + g_1 * h$
- Convolution is shift invariant and spatially localized.
- This simple operation is at the heart of convolutional layers, which are the building blocks of powerful convolutional neural network models.

Linear Smoothing Filters

- Smoothing filters are applied to reduce the effects of sharp transitions in intensity in an image.
- It is used to enhance the visual appearance of an image with the following effects
  - Blurring
  - Noise reduction
  - Removal of irrelevant details
- A smoothing filter can be implemented by simply averaging the pixels with a box kernel.

Input Image

$$*\ \frac{1}{25}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} =$$

Box Filter Kernel

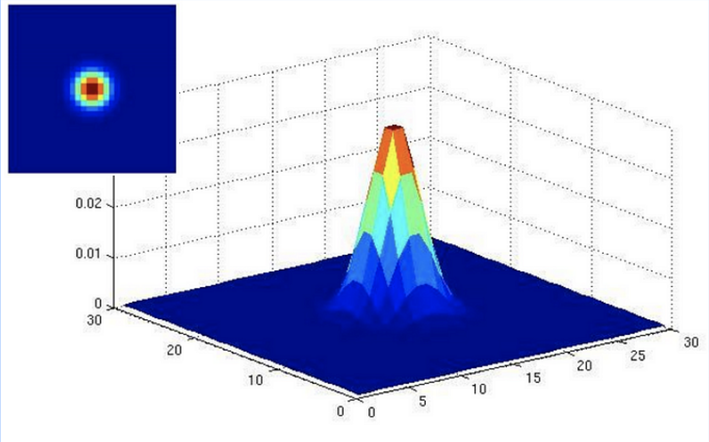
Filtered Output

Linear Smoothing Filters

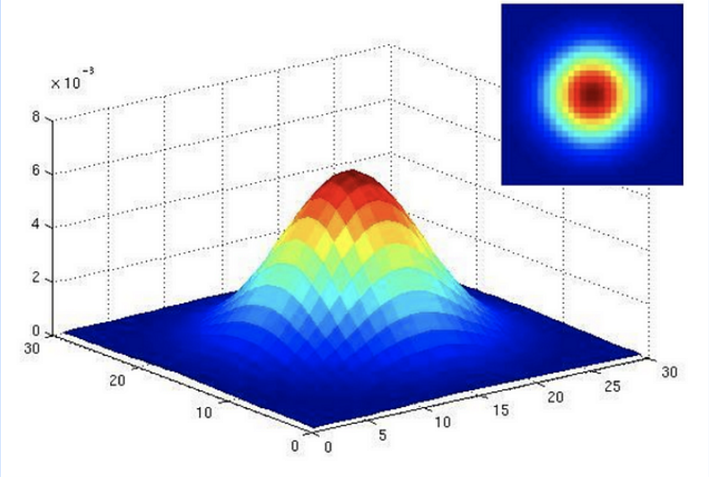
- Box filters equally weigh all pixels in the local neighborhood, which may not be ideal for many applications.
- Another commonly used smoothing filter is the Gaussian kernel given by:

$$g(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

- The Gaussian kernel is circularly symmetric
- The parameter σ controls the scale of smoothing.



$\sigma = 2$



$\sigma = 5$

Gaussian filter kernels for different values of σ

Linear Smoothing Filters

- The Gaussian kernel used for filtering is obtained by sampling the 2D Gaussian function at discrete locations, followed by normalizing such that the filter coefficients sum to 1.
- The size of the Gaussian kernel could be limited to $\lceil 6\sigma \rceil \times \lceil 6\sigma \rceil$ to achieve efficient computation without any significant change in the filtered result.

$$\frac{1}{4.8976} \times \begin{bmatrix} 0.3679 & 0.6065 & 0.3679 \\ 0.6065 & 1.0 & 0.6065 \\ 0.3679 & 0.6065 & 0.3679 \end{bmatrix}$$

A 3×3 Gaussian kernel with $\sigma = 1$

Original Image

$\sigma = 3$, with 21×21 kernel

$\sigma = 9$, with 21×21 kernel

Separable Kernels

- A 2D filter kernel is said to be separable if it can be expressed as an outer product of two 1D kernels, i.e. $g(x, y) = g_1(x)g_2(y)^T$, where $g(x, y)$ is of size $M \times N$ and $g_1(x)$ and
- It is desirable to have separable filter kernels whenever possible due to the computational advantages they offer.
- Considering an image of size $W \times H$ convolved with a kernel of size $M \times N$
  - # of multiplications and additions are of the order $W \times H \times M \times N$
  - # of multiplications and additions when the kernel is separable are of the order $W \times H \times (M + N)$
- Both the box kernel and the Gaussian kernel discussed so far are separable kernels:
  - Box kernel ($M = N = 3$): $\frac{1}{9}\begin{bmatrix}1\\1\\1\end{bmatrix}\begin{bmatrix}1 & 1 & 1\end{bmatrix}$
  - Gaussian kernel ($\sigma = 1, M = N = 3$): $\frac{1}{4.8976}\begin{bmatrix}0.6065\\1\\0.6065\end{bmatrix}\begin{bmatrix}0.6065 & 1 & 0.6065\end{bmatrix}$

Linear Edge Detection Filters

- Linear filters are also used to perform edge detection by computing image gradients.

$$\nabla f(x,y) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial x} \end{bmatrix}$$

- The image gradients along the horizontal and vertical directions can be approximated as follows:

$$\frac{\partial f}{\partial x} = \big(f(x-1,y+1) + 2f(x,y+1) + f(x+1,y+1)\big) - \big(f(x-1,y-1) + 2f(x,y-1) + f(x+1,y-1)\big)$$

$$\frac{\partial f}{\partial y} = \big(f(x+1,y-1) + 2f(x+1,y) + f(x+1,y+1)\big) - \big(f(x-1,y-1) + 2f(x-1,y) + f(x-1,y+1)\big)$$

$$g(x,y) = \sqrt{\left(\frac{\partial f(x,y)}{\partial x}\right)^2 + \left(\frac{\partial f(x,y)}{\partial y}\right)^2}$$ (gradient magnitude)

- The above expressions for the image gradients are used to implement the Sobel operator.

Linear Edge Detection Filters

-1	0	1
-2	0	2
-1	0	1

(a)

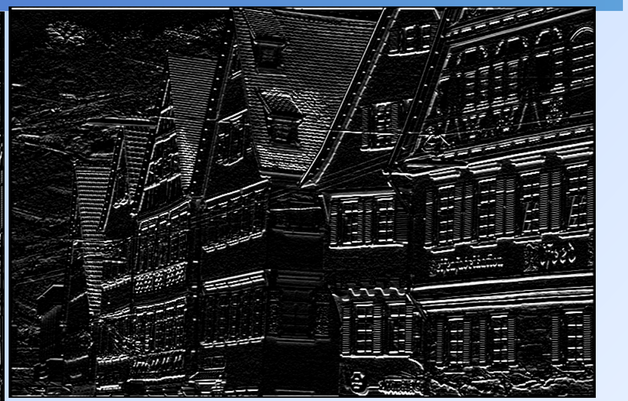
-1	-2	-1
0	0	0
1	2	1

(b)

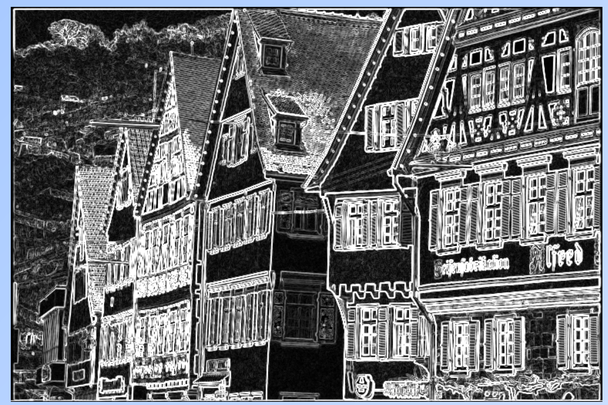
Sobel operators along (a) horizontal and (b) vertical directions

Input

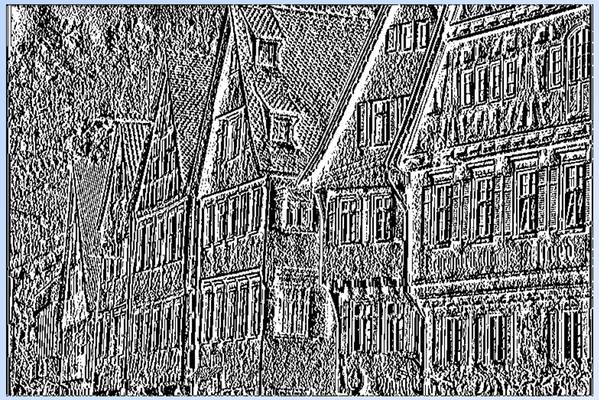
Horizontal gradient



Vertical gradient



Gradient magnitude



Gradient orientation

Linear Edge Detection Filters

- Other similar kernels used for edge detection are as follows:
  - Prewitt: simpler, but lacks the noise smoothing characteristics of Sobel
  - Scharr: more sensitive to edges and less sensitive to noise as compared to both Prewitt and Sobel.

-1	0	1
-1	0	1
-1	0	1

-1	-1	-1
0	0	0
1	1	1

Prewitt kernels

-3	0	3
-10	0	10
-3	0	3

-3	-10	-3
0	0	0
3	10	3

Scharr kernels

Linear Sharpening Filters

- The purpose of a sharpening filter is to highlight transitions in intensity within an image.
- While image smoothing involves spatial integration (averaging), sharpening involves spatial differentiation.
- Image differentiation enhances sharp discontinuities in intensity and suppresses constant intensity regions or regions having smooth intensity transitions.
- A simple isotropic derivative operator is the Laplacian, given by:

  $\nabla^2 f(x,y) = \dfrac{\partial^2 f(x,y)}{\partial x^2} + \dfrac{\partial^2 f(x,y)}{\partial y^2}$, where

  $\dfrac{\partial^2 f(x,y)}{\partial x^2} = f(x-1,y) + f(x+1,y) - 2f(x,y)$, $\dfrac{\partial^2 f(x,y)}{\partial y^2} = f(x,y-1) + f(x,y+1) - 2f(x,y)$

  Thus, $\nabla^2 f(x,y) = f(x-1,y) + f(x+1,y) + f(x,y-1) + f(x,y+1) - 4f(x,y)$
- The discrete approximation of the Laplacian can also be extended to include the diagonal elements.

Linear Sharpening Filters

- To achieve sharpening, a scaled version of the Laplacian is added back to the original image.
  $g(x, y) = f(x, y) + c\,\nabla^2 f(x, y).$

  The sign of c is the same as the sign of the central pixel of the Laplacian kernel used.

0	1	0
1	-4	1
0	1	0

1	1	1
1	-8	1
1	1	1

Laplacian kernels

Input Image

Laplacian

Sharpened Image

Linear Sharpening Filters

- Smoothing filters can also be used to perform image sharpening.
- Sharpening using image smoothing involves the following steps:
  - Blurring the original image using a smoothing filter:
    $I_{blur}(x, y) = I(x, y) * h(x, y)$
  - Subtracting the blurred image from the original image to get the mask:
    $I_{mask}(x, y) = I(x, y) - I_{blur}(x, y)$
  - Adding the mask back to the original image to get the sharpened image
    $I_{sharp}(x, y) = I(x, y) + kI_{mask}(x, y)$
- The above process is referred to as unsharp masking ($k \leq 1$) or high boost filtering ($k > 1$).

Input Image

Gaussian blurred image (kernel size =21, $\sigma = 5$

Mask image

Sharpened image ($k = 3$)

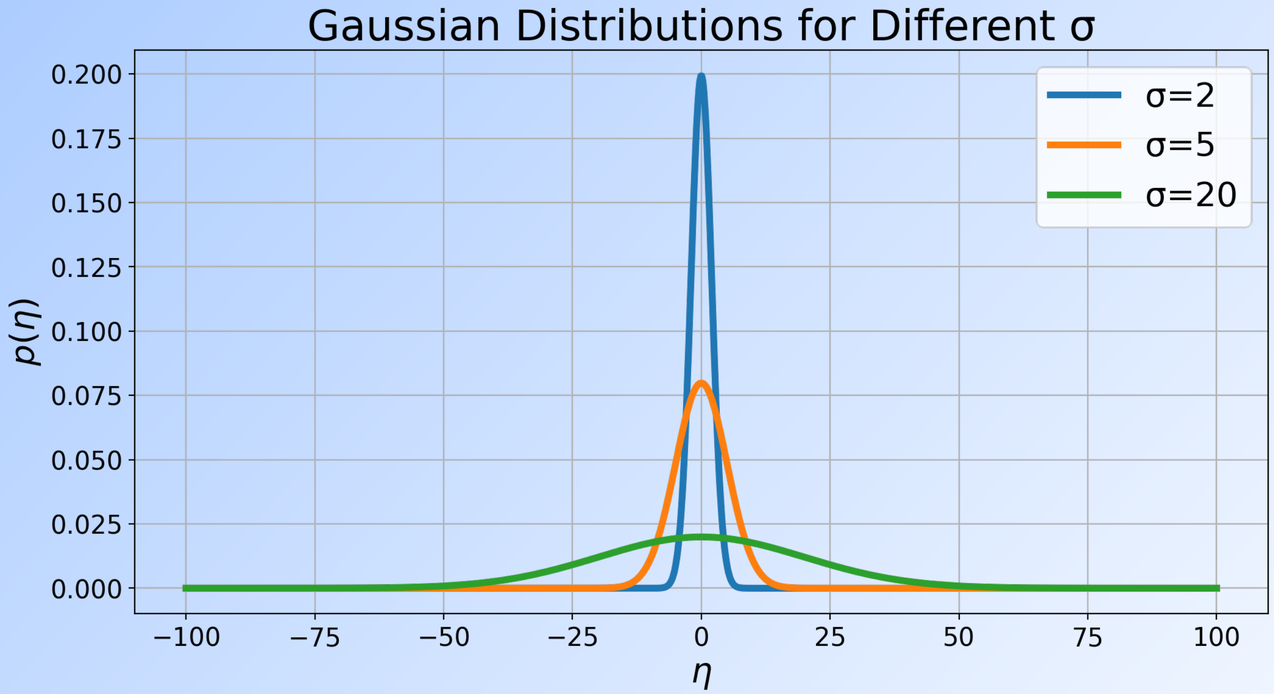
Image Noise Models

- Images are degraded by different types of noise during acquiring, processing and transmission.
- Several of these noise types could be modeled mathematically, which facilitate tasks such as noise reduction, design of robust algorithms, as well as robust compression and communication protocols.
- The following models are commonly used to describe noise in digital images
  - Gaussian Noise
  - Salt and Pepper Noise
  - Poisson Noise
  - Quantization Noise
  - Speckle Noise

Image Noise Models

Gaussian Noise:

- Gaussian noise is widely used to model many real-world noises such as thermal noise, sensor noise etc.
- An image corrupted with zero-mean additive white Gaussian noise (AWGN) is modeled as: $I_n(x, y) = I(x, y) + \eta$, where $\eta \sim \mathcal{N}(0,\sigma^2)$



$$p(\eta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\eta-\mu)^2}{2\sigma^2}\right)$$

Image Noise Models

Salt and Pepper Noise:

- Salt and pepper noise, which is also known as impulse noise randomly changes some pixels of an image to black and white pixels, while leaving other pixels unchanged.
- This type of noise can occur due to transmission over noisy digital links.

$$I_n(x,y) = \begin{cases} I(x,y) & \text{with probability } 1-(p_s+p_p) \\ I_{\text{min}} & \text{with probability } p_p \\ I_{\text{max}} & \text{with probability } p_s \end{cases}$$

Where I_{max} and I_{min} are the maximum and minimum possible intensity levels.

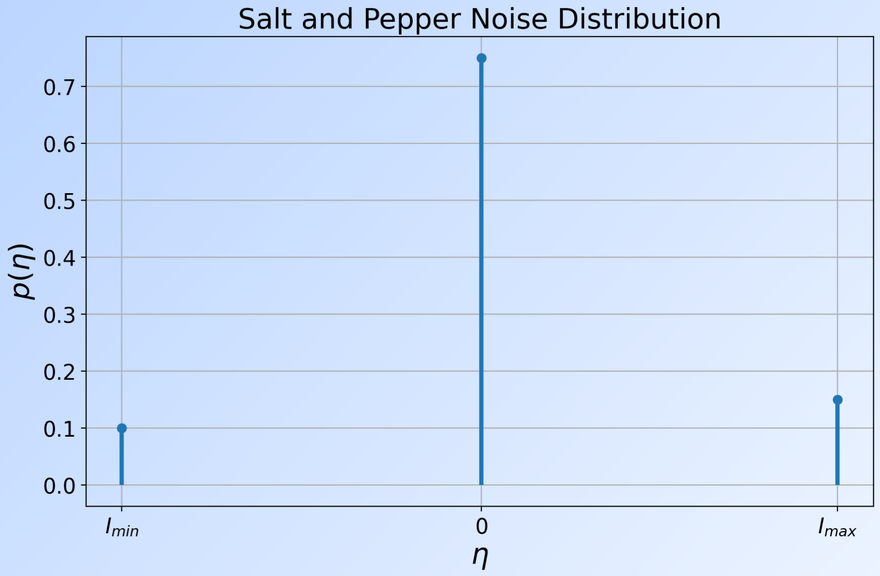


Image Noise Models

Shot Noise:

- Shot noise or photon counting noise, arises due to the discrete nature of photon detection in imaging systems.
- This type of noise is prevalent in low-light imaging conditions, and is more pronounced at faster shutter speeds.
- An image corrupted with shot noise undergoes intensity fluctuations that follow a Poisson distribution

$$I_n(x, y) = I(x, y) + \eta(x, y), \text{ where } \eta(x, y) \sim \frac{e^{-\lambda}\lambda^k}{k!}$$

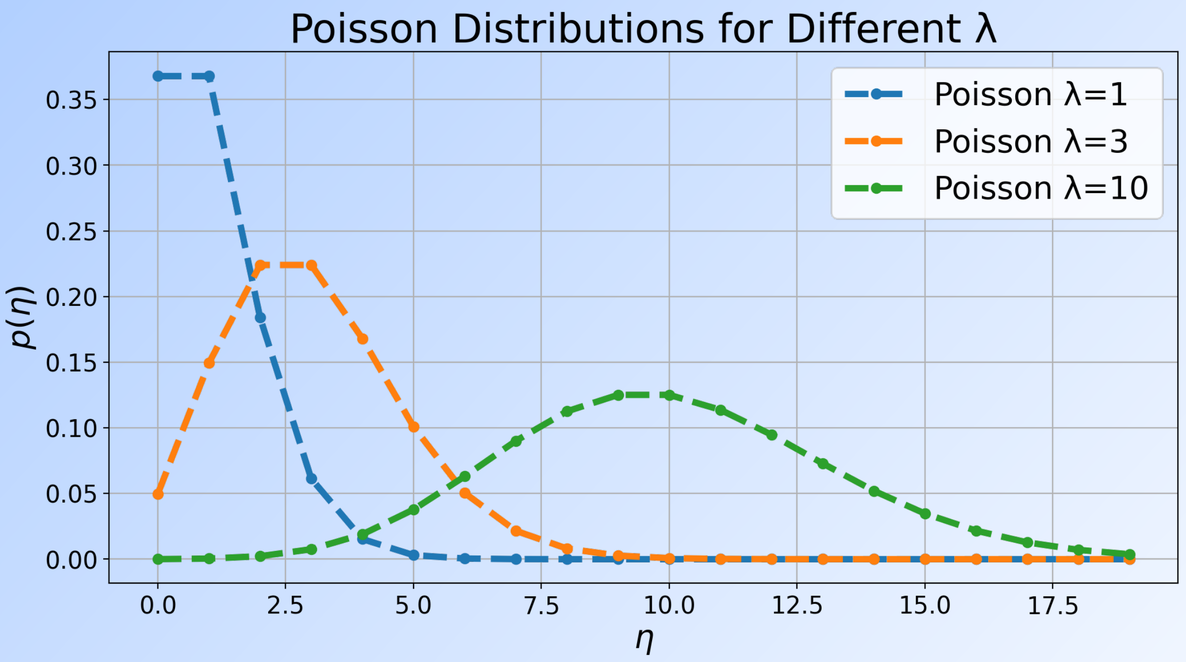
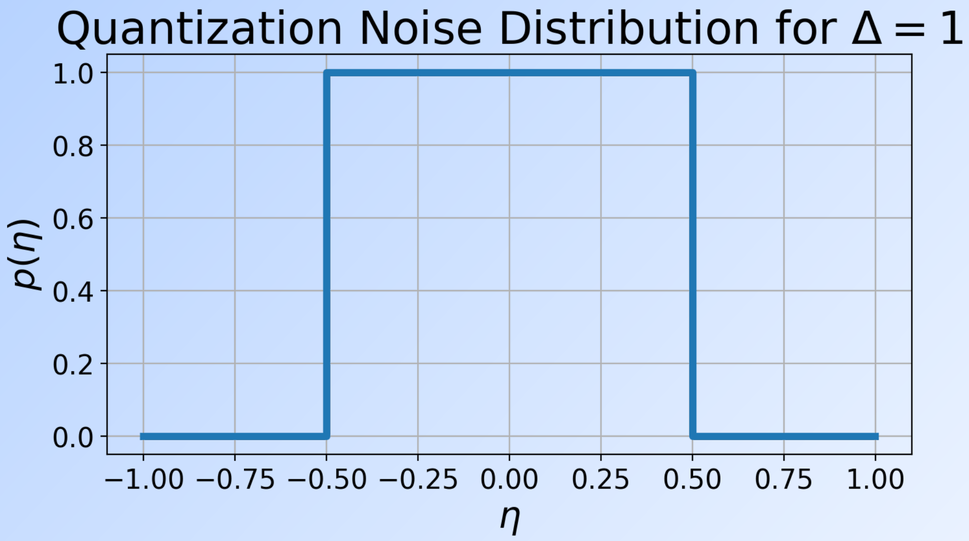


Image Noise Models

Quantization Noise:

- Quantization noise arises due to the quantization of continuous intensity levels during the process of acquisition of digital images
- Quantization noise could be modeled as: $I_n(x, y) = I(x, y) + \eta$, where $\eta \sim U(-\Delta/2, \Delta/2)$, Δ being the quantization interval.



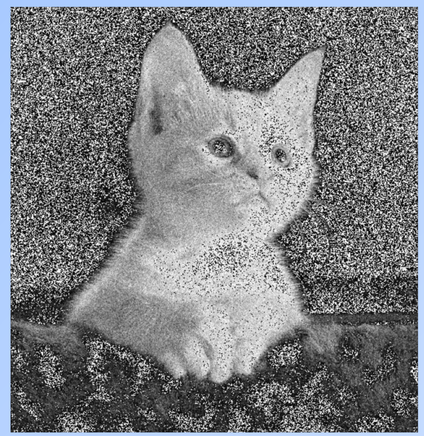
Speckle Noise:

- Speckle noise is a type of multiplicative noise that appears as granular noise over the image.
  $I(x, y) = I(x, y)(1 + \eta(x, y))$
- It results in the interference of coherent waves, such as in radar or ultrasound imaging.

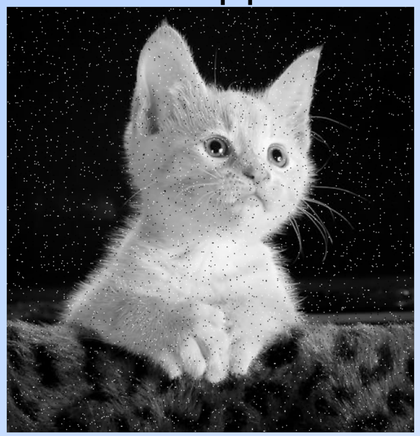
Image Noise Models

Original

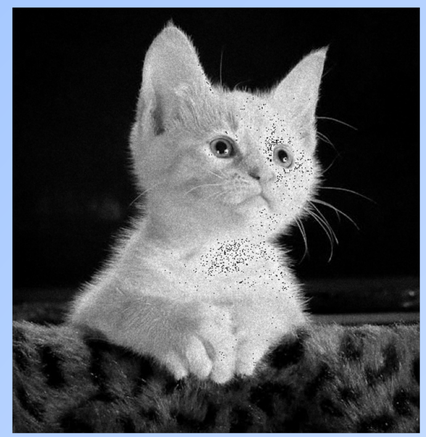
Gaussian Noise



Salt-and-Pepper Noise

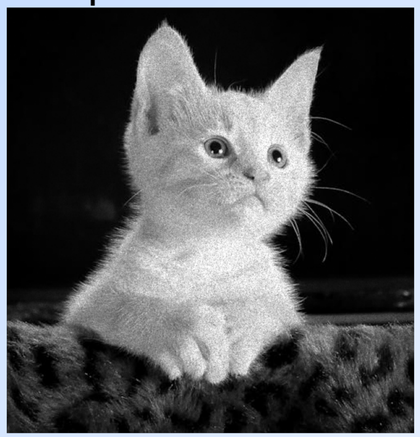


Poisson Noise



Quantization Noise

Speckle Noise



Effect of different noise types on an image

Non-Linear Filtering in the Spatial Domain

- Non-linear operations are often more effective in noise removal or image enhancement tasks.

Median filter

- The median filter is an order statistic filter, i.e. the filter's response is based on ordering or ranking the pixel's within the filter window:
  $g(x, y) = \text{median}(\{f(x+i, y+j) \mid -k \leq i, j \leq k\})$
- Unlike linear smoothing filters, median filtering is better at preserving edges and details while removing impulse (salt and pepper noise).

Original Image

Noisy Image with Impulse Noise

Gaussian Filter Applied to the Noisy Image

Median Filter Applied to the Noisy Image

Non-Linear Filtering in the Spatial Domain

- Bilateral filter is another non-linear image filtering technique that is effective at preserving edges while smoothing the effect of noise.
- While linear smoothing filters weigh pixel intensities in the neighborhood of a certain pixel by their spatial distances from that pixel, bilateral filtering, bilateral filtering combines spatial distance weighting and intensity difference weighting to smooth the image.

  $g(x,y) = \frac{1}{W(x,y)} \sum_{i,j} f(i,j) \cdot \exp\left(-\frac{(x-i)^2 + (y-j)^2}{2\sigma_d^2}\right) \cdot \exp\left(-\frac{(f(x,y) - f(i,j))^2}{2\sigma_r^2}\right)$, where σ_d and σ_r control the weight factors for spatial distance and intensity similarity, and $W(x,y)$ is the normalization factor, which is equal to the sum of the filter weights.
- Although more effective in preserving edges, bilateral filters are computationally more complex and require careful selection of the parameters σ_d and σ_r to achieve desirable results.

Non-Linear Filtering in the Spatial Domain

Original Image

Noisy Image

Using Bilateral Filter

Using Gaussian Filter

Using Box Filter