

## UNIT-IV

### IMAGE RESTORATION & IMAGE SEGMENTATION

#### **IMAGE RESTORATION:**

Restoration improves image in some predefined sense. It is an objective process. Restoration attempts to reconstruct an image that has been degraded by using a priori knowledge of the degradation phenomenon. These techniques are oriented toward modeling the degradation and then applying the inverse process in order to recover the original image. Image Restoration refers to a class of methods that aim to remove or reduce the degradations that have occurred while the digital image was being obtained. All natural images when displayed have gone through some sort of degradation:

- a) During display mode
- b) Acquisition mode, or
- c) Processing mode

The degradations may be due to

- a) Sensor noise
- b) Blur due to camera mis focus
- c) Relative object-camera motion
- d) Random atmospheric turbulence
- e) Others

#### **4.1 Degradation Model:**

Degradation process operates on a degradation function that operates on an input image with an additive noise term. Input image is represented by using the notation  $f(x,y)$ , noise term can be represented as  $\eta(x,y)$ . These two terms when combined gives the result as  $g(x,y)$ . If we are given  $g(x,y)$ , some knowledge about the degradation function  $H$  or  $J$  and some knowledge about the additive noise term  $\eta(x,y)$ , the objective of restoration is to obtain an estimate  $f'(x,y)$  of the original image. We want the estimate to be as close as possible to the original image. The more we know about  $h$  and  $\eta$ , the closer  $f(x,y)$  will be to  $f'(x,y)$ . If it is a linear position invariant process, then degraded image is given in the spatial domain by

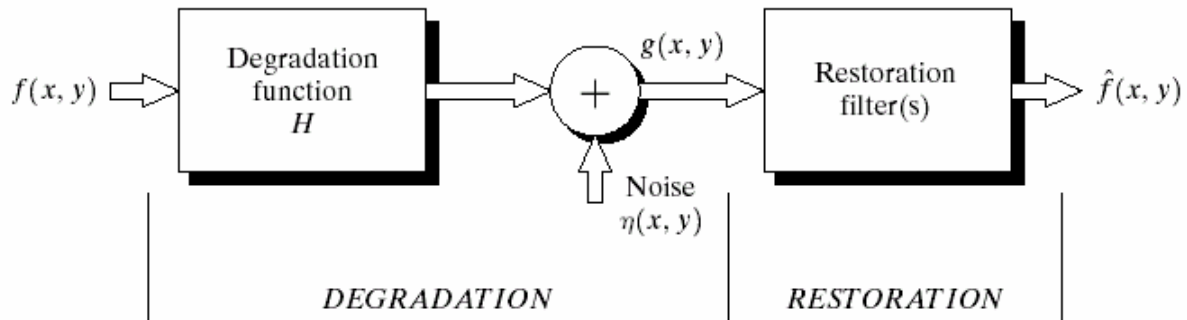
$$g(x,y)=f(x,y)*h(x,y)+\eta(x,y)$$

$h(x,y)$  is spatial representation of degradation function and symbol  $*$  represents convolution.

In frequency domain we may write this equation as

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

The terms in the capital letters are the Fourier Transform of the corresponding terms in the spatial domain.



The image restoration process can be achieved by inverting the image degradation process, i.e.,

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

where  $1/H(u,v)$  is the inverse filter, and  $G(u,v)$  is the recovered image. Although the concept is relatively simple, the actual implementation is difficult to achieve, as one requires prior knowledge or identifications of the unknown degradation function and the unknown noise source. In the following sections, common noise models and method of estimating the degradation function are presented.

#### 4.2 Noise Models:

The principal source of noise in digital images arises during image acquisition and /or transmission. The performance of imaging sensors is affected by a variety of factors, such as environmental conditions during image acquisition and by the quality of the sensing elements themselves. Images are corrupted during transmission principally due to interference in the channels used for transmission. Since main sources of noise presented in digital images are resulted from atmospheric disturbance and image sensor circuitry, following assumptions can be made:

- . The noise model is spatial invariant, i.e., independent of spatial location.
- . The noise model is uncorrelated with the object function.

##### (i) Gaussian Noise:

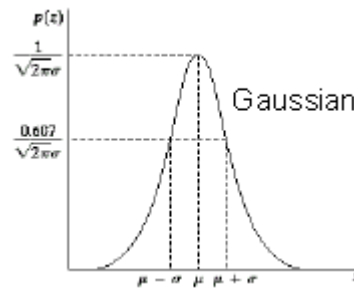
These noise models are used frequently in practices because of its tractability in both spatial and frequency domain. The PDF of Gaussian random variable,  $z$  is given by

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-\mu)^2}{2\sigma^2}}$$

$z$  = gray level

$\mu$  = mean of average value of  $z$

$\sigma$  = standard deviation



### (ii) Rayleigh Noise:

Unlike Gaussian distribution, the Rayleigh distribution is not symmetric. It is given by the formula.

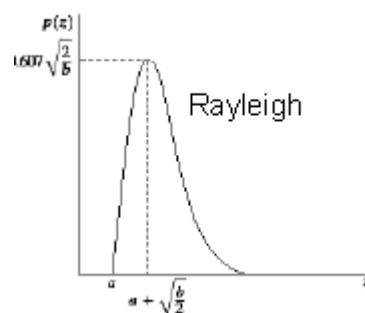
$$p(z) = \frac{2}{b} (z - a) e^{-\frac{(z-a)^2}{b}}, \quad \text{for } z \geq a$$

The mean and variance of this density is

Mean/variance

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

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



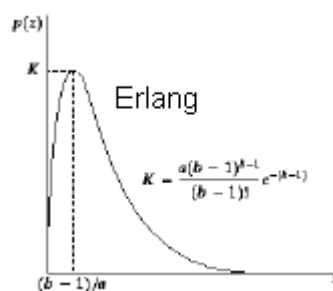
### (iii) Gamma Noise:

The PDF of Erlang noise is given by

$$p(z) = \frac{a^b z^{b-1}}{(b-a)!} e^{-az}, \quad \text{for } z \geq 0$$

The mean and variance of this noise is

Variance	
$\mu$	$\frac{b}{a}$
$\sigma^2$	$\frac{b}{a^2}$



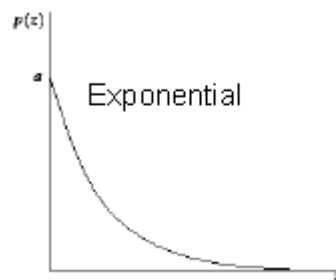
Its shape is similar to Rayleigh disruption. This equation is referred to as gamma density it is correct only when the denominator is the gamma function.

**(iv) Exponential Noise:**

Exponential distribution has an exponential shape. The PDF of exponential noise is given as

$$p(z) = ae^{-az}, \quad \text{for } z \geq 0$$

Where  $a > 0$ . It is a special case of Erlang with  $b=1$



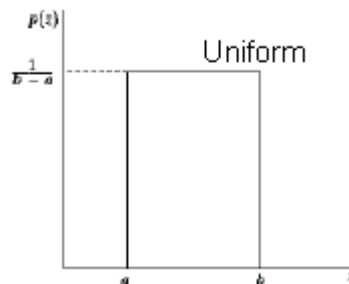
**(v) Uniform Noise:**

The PDF of uniform noise is given by

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

The mean and variance of this noise is

Mean/variance
$\mu = \frac{a+b}{2}$
$\sigma^2 = \frac{(b-a)^2}{12}$



**(vi) Impulse (salt & pepper) Noise:**

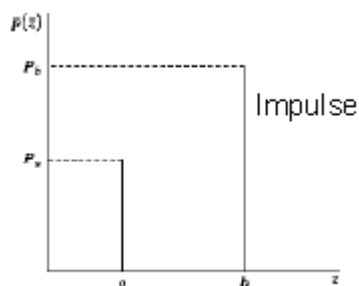
In this case, the noise is signal dependent, and is multiplied to the image.

The PDF of bipolar (impulse) noise is given by

$$p(z) = \begin{cases} P_a & \text{for } z = a \\ P_b & \text{for } z = b \\ 0 & \text{otherwise} \end{cases}$$

If  $b > a$ , gray level  $b$  will appear as a light dot in image.

Level  $a$  will appear like a dark dot.



**4.3 Restoration in the presence of Noise only- Spatial filtering:**

When the only degradation present in an image is noise, i.e.

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

or

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

The noise terms are unknown so subtracting them from  $g(x,y)$  or  $G(u,v)$  is not a realistic approach. In the case of periodic noise it is possible to estimate  $N(u,v)$  from the spectrum  $G(u,v)$ .

So  $N(u,v)$  can be subtracted from  $G(u,v)$  to obtain an estimate of original image. Spatial filtering can be done when only additive noise is present. The following techniques can be used to reduce the noise effect:

**(i) Mean Filter:**

**(a) Arithmetic Mean filter:**

It is the simplest mean filter. Let  $S_{xy}$  represents the set of coordinates in the sub image of size  $m*n$  centered at point  $(x,y)$ . The arithmetic mean filter computes the average value of the corrupted image  $g(x,y)$  in the area defined by  $S_{xy}$ . The value of the restored image  $f$  at any point  $(x,y)$  is the arithmetic mean computed using the pixels in the region defined by  $S_{xy}$ .

$$\hat{f}(x,y) = \frac{1}{mn} \sum_{(s,t) \in S_{xy}} g(s,t)$$

This operation can be using a convolution mask in which all coefficients have value  $1/mn$ . A mean filter smoothes local variations in image Noise is reduced as a result of blurring. For every pixel in the image, the pixel value is replaced by the mean value of its neighboring pixels with a weight. This will result in a smoothing effect in the image.

**(b) Geometric Mean filter:**

An image restored using a geometric mean filter is given by the expression

$$\hat{f}(x,y) = \left( \prod_{(s,t) \in S_{xy}} g(s,t) \right)^{1/mn}$$

Here, each restored pixel is given by the product of the pixel in the sub image window, raised to the power  $1/mn$ . A geometric mean filter but it loses image details in the process.

**(c) Harmonic Mean filter:**

The harmonic mean filtering operation is given by the expression

$$\hat{f}(x,y) = \frac{\sum_{(s,t) \in S_{xy}} g(s,t)^{\theta+1}}{\sum_{(s,t) \in S_{xy}} g(s,t)^{\theta}}$$

The harmonic mean filter works well for salt noise but fails for pepper noise. It does well with Gaussian noise also.

**(d) Order statistics filter:**

Order statistics filters are spatial filters whose response is based on ordering the pixel contained in the image area encompassed by the filter. The response of the filter at any point is determined by the ranking result.

**(e)Median filter:**

It is the best order statistic filter; it replaces the value of a pixel by the median of gray levels in the Neighborhood of the pixel.

$$\hat{f}(x, y) = \text{median}_{(s,t) \in S_{xy}} \{g(s, t)\}$$

The original of the pixel is included in the computation of the median of the filter are quite possible because for certain types of random noise, the provide excellent noise reduction capabilities with considerably less blurring then smoothing filters of similar size. These are effective for bipolar and unipolar impulse noise.

**(e)Max and Min filter:**

Using the 100th percentile of ranked set of numbers is called the max filter and is given by the equation

$$\hat{f}(x, y) = \max_{(s,t) \in S_{xy}} \{g(s, t)\}$$

It is used for finding the brightest point in an image. Pepper noise in the image has very low values, it is reduced by max filter using the max selection process in the sublimated area sky. The 0th percentile filter is min filter

$$\hat{f}(x, y) = \min_{(s,t) \in S_{xy}} \{g(s, t)\}$$

This filter is useful for flinging the darkest point in image. Also, it reduces salt noise of the min operation.

**(f)Midpoint filter:**

The midpoint filter simply computes the midpoint between the maximum and minimum values in the area encompassed by

$$\hat{f}(x, y) = \left( \max_{(s,t) \in S_{xy}} \{g(s, t)\} + \min_{(s,t) \in S_{xy}} \{g(s, t)\} \right) / 2$$

It comeliness the order statistics and averaging .This filter works best for randomly distributed noise like Gaussian or uniform noise.

**4.4 Periodic Noise by Frequency domain filtering:**

These types of filters are used for this purpose-

**(i)Band Reject Filters:**

It removes a band of frequencies about the origin of the Fourier transformer.

**(ii) Ideal Band reject Filter:**

An ideal band reject filter is given by the expression

$$H(u,v) = \begin{cases} 1 & \text{if } D(u,v) < D_0 - W/2 \\ 0 & \text{if } D_0 - W/2 \leq D(u,v) \leq D_0 + W/2 \\ 1 & \text{if } D(u,v) > D_0 + W/2 \end{cases}$$

$D(u,v)$ - the distance from the origin of the centered frequency rectangle.

$W$ - the width of the band

$D_0$ - the radial center of the frequency rectangle.

**(iii) Butterworth Band reject Filter:**

$$H(u,v) = 1 / \left[ 1 + \left( \frac{D(u,v)W}{D^2(u,v) - D_0^2} \right)^{2n} \right]$$

**(iv) Gaussian Band reject Filter:**

$$H(u,v) = 1 - \exp \left[ -\frac{1}{2} \left( \frac{D^2(u,v) - D_0^2}{D(u,v)W} \right)^2 \right]$$

These filters are mostly used when the location of noise component in the frequency domain is known. Sinusoidal noise can be easily removed by using these kinds of filters because it shows two impulses that are mirror images of each other about the origin. Of the frequency transform.

**(v) Band pass Filter:**

The function of a band pass filter is opposite to that of a band reject filter. It allows a specific frequency band of the image to be passed and blocks the rest of frequencies. The transfer function of a band pass filter can be obtained from a corresponding band reject filter with transfer function  $H_{BR}(u,v)$  by using the equation

$$H_{BP}(u,v) = 1 - H_{BR}(u,v)$$

These filters cannot be applied directly on an image because it may remove too much details of an image but these are effective in isolating the effect of an image of selected frequency bands.



#### 4.5 Minimum mean Square Error (Wiener) filtering:

This filter incorporates both degradation function and statistical behavior of noise into the restoration process. The main concept behind this approach is that the images and noise are considered as random variables and the objective is to find an estimate  $\hat{f}$  of the uncorrupted image  $f$  such that the mean square error between them is minimized.

$$\hat{f}(x) = \sum_{s} h_w(x-s)g(s),$$

This error measure is given by

$$e^2 = E\{[f(x) - \hat{f}(x)]^2\} = \min$$

Where  $e(\ )$  is the expected value of the argument.

Assuming that the noise and the image are uncorrelated (means zero average value) one or other has zero mean values. The minimum error function of the above expression is given in the frequency ..... is given by the expression

$$H_w(u, v) = \frac{H^*(u, v) S_g(u, v)}{|H(u, v)|^2 S_g(u, v) + S_m(u, v)} = \frac{1}{H(u, v)} \frac{|H(u, v)|^2}{|H(u, v)|^2 + S_m(u, v) / S_g(u, v)}$$

Product of a complex quantity with its conjugate is equal to the magnitude of ..... complex quantity squared. This result is known as wiener Filter The filter was named so because of the name of its inventor N Wiener. The term in the bracket is known as minimum mean square error filter or least square error filter.

$H^*(u, v)$ -degradation function .

$H^*(u, v)$ -complex conjugate of  $H(u, v)$

$H(u, v)$   $H(u, v)$

$S_n(u, v) = I N(u, v) I^2$ - power spectrum of the noise

$S_f(u, v) = I F(u, v) I^2$ - power spectrum of the underrated image

$H(u, v)$ =Fourier transformer of the degraded function

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

The restored image in the spatial domain is given by the inverse Fourier transformed of the frequency domain estimate  $F(u, v)$ .

Mean square error in statistical form can be approved by the function

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

#### 4.6 Inverse Filtering:

It is a process of restoring an image degraded by a degradation function  $H$ . This function can be obtained by any method. The simplest approach to restoration is direct, inverse filtering. Inverse filtering provides an estimate  $F(u,v)$  of the transform of the original image simply by dividing the transform of the degraded image  $G(u,v)$  by the degradation function.

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

It shows an interesting result that even if we know the degradation function we cannot recover the undegraded image exactly because  $N(u,v)$  is not known. If the degradation value has zero or very small values then the ratio  $N(u,v)/H(u,v)$  could easily dominate the estimate  $F(u,v)$ .

#### IMAGE SEGMENTATION:

If an image has been preprocessed appropriately to remove noise and artifacts, segmentation is often the key step in interpreting the image. Image segmentation is a process in which regions or features sharing similar characteristics are identified and grouped together. Image segmentation may use statistical classification, thresholding, edge detection, region detection, or any combination of these techniques. The output of the segmentation step is usually a set of classified elements. Most segmentation techniques are either region-based or edge based.

(i) Region-based techniques rely on common patterns in intensity values within a cluster of neighboring pixels. The cluster is referred to as the region, and the goal of the segmentation algorithm is to group regions according to their anatomical or functional roles.

(ii) Edge-based techniques rely on discontinuities in image values between distinct regions, and the goal of the segmentation algorithm is to accurately demarcate the boundary separating these regions. Segmentation is a process of extracting and representing information from an image is to group pixels together into regions of similarity.

Region-based segmentation methods attempt to partition or group regions according to common image properties. These image properties consists of :

- (a) Intensity values from original images, or computed values based on an image operator
- (b) Textures or patterns that are unique to each type of region
- (c) Spectral profiles that provide multidimensional image data

Elaborate systems may use a combination of these properties to segment images, while

simpler systems may be restricted to a minimal set on properties depending of the type of data available.

#### 4.7 Categories of Image Segmentation Methods

- Clustering Methods Level Set Methods
- Histogram-Based Methods Graph partitioning methods
- Edge Detection Methods Watershed Transformation
- Region Growing Methods Neural Networks Segmentation
- Model based Segmentation/knowledge-based segmentation - involve active shape and appearance models, active contours and deformable templates.
- Semi-automatic Segmentation - Techniques like Livewire or Intelligent Scissors are used in this kind of segmentation.

##### (i) Pixel based approach:

Gray level thresholding is the simplest segmentation process. Many objects or image regions are characterized by constant reflectivity or light absorption of their surface. Thresholding is computationally inexpensive and fast. Thresholding can easily be done in real time using specialized hardware. Complete segmentation can result from thresholding in simple scenes.

$$R = \bigcup_{i=1}^S R_i \quad R_i \cap R_j = \emptyset \quad i \neq j$$

Search all the pixels  $f(i,j)$  of the image  $f$ . An image element  $g(i,j)$  of the segmented image is an object pixel if  $f(i,j) \geq T$ , and is a background pixel otherwise.

$$g(i, j) = \begin{cases} 1 & \text{for } f(i, j) \geq T \\ 0 & \text{for } f(i, j) < T \end{cases}$$

Correct threshold selection is crucial for successful threshold segmentation. Threshold selection can be interactive or can be the result of some threshold detection method.

##### (a) Multi Level Thresholding:

The resulting image is no longer binary

$$g(i, j) = \begin{cases} 1 & \text{for } f(i, j) \in D_1 \\ 2 & \text{for } f(i, j) \in D_2 \\ 3 & \text{for } f(i, j) \in D_3 \\ 4 & \text{for } f(i, j) \in D_4 \\ \dots \\ n & \text{for } f(i, j) \in D_n \\ 0 & \text{otherwise} \end{cases}$$

##### (b) Local Thresholding:

It is successful only under very unusual circumstances. Gray level variations are likely due to non-uniform lighting, non-uniform input device parameters or a number of other factors.

$$T=T(f)$$

**(c) Thresholding detection method:**

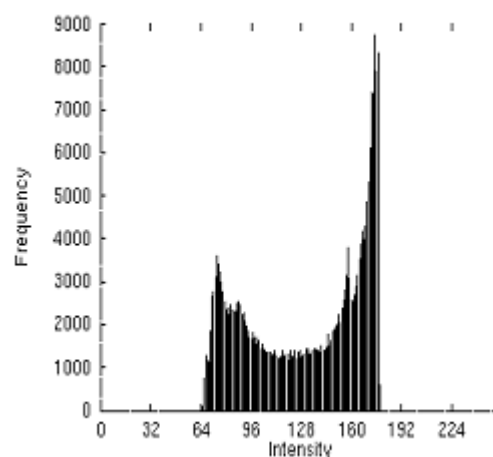
If some property of an image after segmentation is known a priori, the task of threshold selection is simplified, since the threshold is chosen to ensure this property is satisfied. A printed text sheet may be an example if we know that characters of the text cover 1/p of the sheet area.

- **P-type thresholding**

- choose a threshold T (based on the image histogram) such that 1/p of the image area has gray values less than T and the rest has gray values larger than T.
- in text segmentation, prior information about the ratio between the sheet area and character area can be used.
- if such a priori information is not available - another property, for example the average width of lines in drawings, etc. can be used - the threshold can be determined to provide the required line width in the segmented image.

- **More methods of thresholding are**

- based on histogram shape analysis
- bimodal histogram - if objects have approximately the same gray level that differs from the gray level of the background



- **Mode Method:** find the highest local maxima first and detect the threshold as a minimum between them. To avoid detection of two local maxima belonging to the same global maximum, a minimum distance in gray levels between these maxima is usually required or techniques to smooth histograms are applied.

## **(ii)Region based segmentation:**

### **(a)Region based growing segmentation:**

Homogeneity of regions is used as the main segmentation criterion in region growing.

The criteria for homogeneity:

- graylevel
- color
- texture
- shape
- model

The basic purpose of region growing is to segment an entire image  $R$  into smaller sub-images,  $R_i$ ,  $i=1,2,\dots,N$ . which satisfy the following conditions:

$$\begin{aligned} R &= \bigcup_{i=1}^N R_i; R_i \cap R_j = \Phi, i \neq j \\ H(R_i) &= True; i = 1,2,\dots, N; \\ H(R_i \cup R_j) &= False, i \neq j; \end{aligned}$$

### **(b)Region Splitting:**

The basic idea of region splitting is to break the image into a set of disjoint regions, which are coherent within themselves:

- Initially take the image as a whole to be the area of interest.
- Look at the area of interest and decide if all pixels contained in the region satisfy some *similarity constraint*.
- If TRUE then the area of interest corresponds to an entire region in the image.
- If FALSE split the area of interest (usually into four equal subareas) and consider each of the sub-areas as the area of interest in turn.
- This process continues until no further splitting occurs. In the worst case this happens when the areas are just one pixel in size.

If only a splitting schedule is used then the final segmentation would probably contain many neighboring regions that have identical or similar properties. We need to merge these regions.

### **(c)Region merging:**

The result of region merging usually depends on the order in which regions are merged. The simplest methods begin merging by starting the segmentation using regions of 2x2, 4x4 or 8x8 pixels. Region descriptions are then based on their statistical gray level properties. A region description is compared with the description of an adjacent region; if they match, they

are merged into a larger region and a new region description is computed. Otherwise regions are marked as non-matching. Merging of adjacent regions continues between all neighbors, including newly formed ones. If a region cannot be merged with any of its neighbors, it is marked 'final' and the merging process stops when all image regions are so marked. Merging Heuristics:

- Two adjacent regions are merged if a significant part of their common boundary consists of weak edges
- Two adjacent regions are also merged if a significant part of their common boundary consists of weak edges, but in this case not considering the total length of the region borders.

Of the two given heuristics, the first is more general and the second cannot be used alone because it does not consider the influence of different region sizes.

Region merging process could start by considering

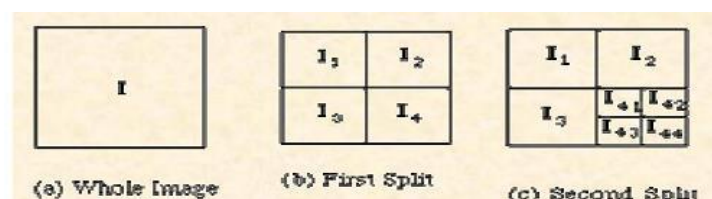
- small segments (2\*2, 8\*8) selected a priori from the image
- segments generated by thresholding
- regions generated by a region splitting module

The last case is called as "Split and Merge" method. Region merging methods generally use similar criteria of homogeneity as region splitting methods, and only differ in the direction of their application.

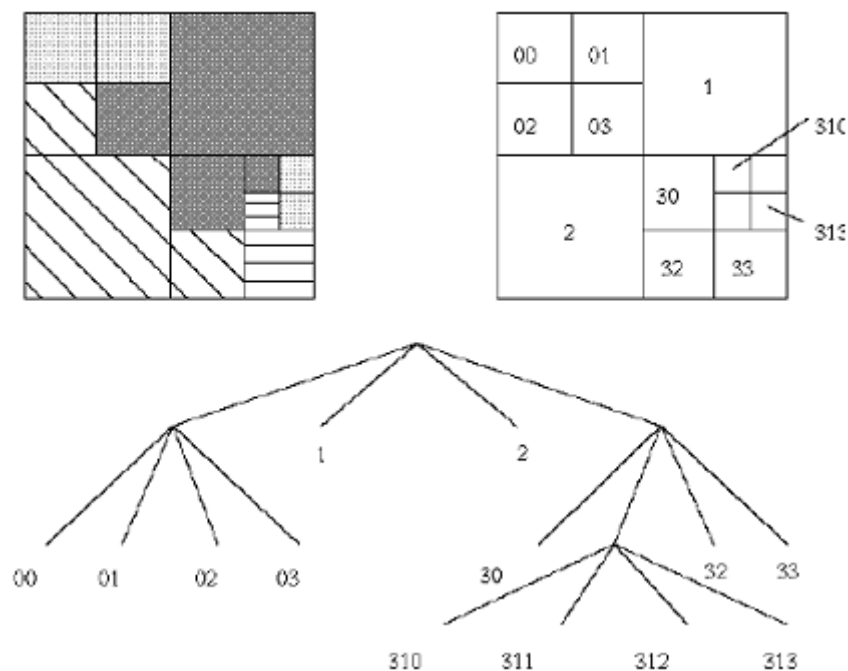
#### (d) Split & Merge:

To illustrate the basic principle of split and merge methods, let us consider an imaginary image.

- Let  $I$  denote the whole image shown in Fig. (a)
- Not all the pixels in Fig (a) are similar. So the region is split as in Fig. (b).
- Assume that all pixels within each of the regions  $I_1$ ,  $I_2$  and  $I_3$  are similar, but those in  $I_4$  are not.
- Therefore  $I_4$  is split next, as shown in Fig. (c).
- Now assume that all pixels within each region are similar with respect to that region, and that after comparing the split regions, regions  $I_{43}$  and  $I_{44}$  are found to be identical.
- These pair of regions is thus merged together, as in shown in Fig. (d)



A combination of splitting and merging may result in a method with the advantages of both the approaches. Split-and-merge approaches work using pyramid image representations. Regions are square-shaped and correspond to elements of the appropriate pyramid level. If any region in any pyramid level is not homogeneous (excluding the lowest level), it is split into four sub-regions -- these are elements of higher resolution at the level below. If four regions exist at any pyramid level with approximately the same value of homogeneity measure, they are merged into a single region in an upper pyramid level. We can also describe the splitting of the image using a tree structure, called a modified *quad tree*. Each non-terminal node in the tree has at most four descendants, although it may have less due to merging. Quad tree decomposition is an operation that subdivides an image into blocks that contain "similar" pixels. Usually the blocks are square, although sometimes they may be rectangular. For the purpose of this demo, pixels in a block are said to be "similar" if the range of pixel values in the block are not greater than some threshold. Quad tree decomposition is used in variety of image analysis and compression applications. An unpleasant drawback of segmentation quad trees, is the square region shape assumption. It is not possible to merge regions which are not part of the same branch of the segmentation tree. Because both split-and-merge processing options are available, the starting segmentation does not have to satisfy any of the homogeneity conditions. The segmentation process can be understood as the construction of a segmentation quad tree where each leaf node represents a homogeneous region. Splitting and merging corresponds to removing or building parts of the segmentation quad tree.

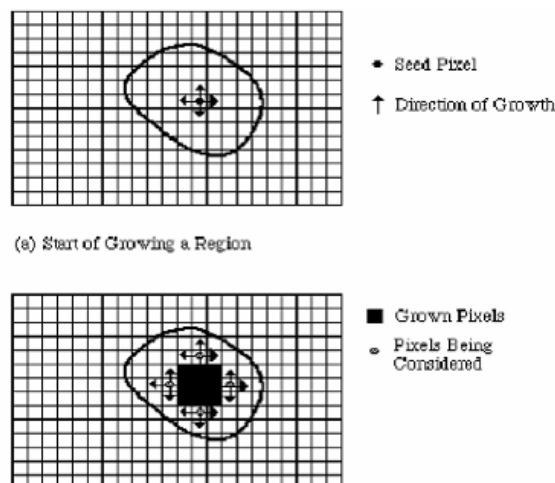


### (e)Region growing:

Region growing approach is the opposite of the split and merges approach:

- An initial set of small areas is iteratively merged according to similarity constraints.
- Start by choosing an arbitrary *seed pixel* and compare it with neighboring pixels
- Region is *grown* from the seed pixel by adding in neighboring pixels that are similar, increasing the size of the region.
- When the growth of one region stops we simply choose another seed pixel which does not yet belong to any region and start again.
- This whole process is continued until all pixels belong to some region.
- A *bottom up* method.

Region growing methods often give very good segmentations that correspond well to the observed edges.



However starting with a particular seed pixel and letting this region grow completely before trying other seeds biases the segmentation in favour of the regions which are segmented first.

This can have several undesirable effects:

- Current region dominates the growth process -- ambiguities around edges of adjacent regions may not be resolved correctly.
- Different choices of seeds may give different segmentation results.
- Problems can occur if the (arbitrarily chosen) seed point lies on an edge.

To counter the above problems, *simultaneous region growing* techniques have been developed.

- Similarities of neighboring regions are taken into account in the growing process.
- No single region is allowed to completely dominate the proceedings.



- A number of regions are allowed to grow at the same time.
- Similar regions will gradually coalesce into expanding regions.
- Control of these methods may be quite complicated but efficient methods have been developed.
- Easy and efficient to implement on parallel computers.

### **(iii)Line & Edge detection:**

#### **(a)Edge detection:**

Edges are places in the image with strong intensity contrast. Since edges often occur at image locations representing object boundaries, edge detection is extensively used in image segmentation when we want to divide the image into areas corresponding to different objects. Representing an image by its edges has the further advantage that the amount of data is reduced significantly while retaining most of the image information.

- **Canny edge detection:**

It is optimal for step edges corrupted by white noise. Optimality related to three criteria

- detection criterion ... important edges should not be missed, there should be no spurious responses
- localization criterion ... distance between the actual and located position of the edge should be minimal
- one response criterion ... minimizes multiple responses to a single edge (also partly covered by the first criterion since when there are two responses to a single edge one of them should be considered as false)

Canny's edge detector is based on several ideas:

- 1) The edge detector was expressed for a 1D signal and the first two optimality criteria. A closed form solution was found using the calculus of variations.
  - 2) If the third criterion (multiple responses) is added, the best solution may be found by numerical optimization. The resulting filter can be approximated effectively with error less than 20% by the first derivative of a Gaussian smoothing filter with standard deviation; the reason for doing this is the existence of an effective implementation.
  - 3) The detector is then generalized to two dimensions. A step edge is given by its position, orientation, and possibly magnitude (strength).
- Suppose  $G$  is a 2D Gaussian and assume we wish to convolute the image with an operator  $G_n$  which is a first derivative of  $G$  in the direction  $n$ .

$$G_n = \frac{\partial G}{\partial \mathbf{n}} = \mathbf{n} \cdot \nabla G$$

The direction  $\mathbf{n}$  should be oriented perpendicular to the edge

- this direction is not known in advance
- however, a robust estimate of it based on the smoothed gradient direction is available
- if  $g$  is the image, the normal to the edge is estimated as

$$\mathbf{n} = \frac{\nabla(G * g)}{|\nabla(G * g)|}$$

The edge location is then at the local maximum in the direction  $\mathbf{n}$  of the operator  $G_n$  convoluted with the image  $g$

$$\frac{\partial}{\partial \mathbf{n}} G_n * g = 0$$

- Substituting in equation for  $G_n$  from equation, we get

$$\frac{\partial^2}{\partial \mathbf{n}^2} G * g = 0$$

- This equation shows how to find local maxima in the direction perpendicular to the edge; this operation is often referred to as **non-maximum suppression**.
  - As the convolution and derivative are associative operations in equation
  - first convolute an image  $g$  with a symmetric Gaussian  $G$
  - then compute the directional second derivative using an estimate of the direction  $\mathbf{n}$
  - strength of the edge (magnitude of the gradient of the image intensity function  $g$ ) is measured as

$$|G_n * g| = |\nabla(G * g)|$$

4) Spurious responses to the single edge caused by noise usually create a so called 'streaking' problem that is very common in edge detection in general.

- Output of an edge detector is usually thresholded to decide which edges are significant.

- Streaking means breaking up of the edge contour caused by the operator fluctuating above and below the threshold.
- Streaking can be eliminated by thresholding with hysteresis
  - If any edge response is above a high threshold, those pixels constitute definite output of the edge detector for a particular scale.
  - Individual weak responses usually correspond to noise, but if these points are connected to any of the pixels with strong responses they are more likely to be actual edges in the image.
  - Such connected pixels are treated as edge pixels if their response is above a low threshold.
  - The low and high thresholds are set according to an estimated signal to noise ratio.

5) The correct scale for the operator depends on the objects contained in the image.

- The solution to this unknown is to use multiple scales and aggregate information from them.
- Different scale for the Canny detector is represented by different standard deviations of the Gaussians.
- There may be several scales of operators that give significant responses to edges (i.e., signal to noise ratio above the threshold); in this case the operator with the smallest scale is chosen as it gives the best localization of the edge.
- Feature synthesis approach.
  - All significant edges from the operator with the smallest scale are marked first.
  - Edges of a hypothetical operator with larger are synthesized from them (i.e., a prediction is made of how the large should perform on the evidence gleaned from the smaller).
  - Then the synthesized edge response is compared with the actual edge response for larger.
  - Additional edges are marked only if they have significantly stronger response than that predicted from synthetic output.
- This procedure may be repeated for a sequence of scales, a cumulative edge map is built by adding those edges that were not identified at smaller scales.

**Algorithm: Canny edge detector:**

1. Repeat steps (2) till (6) for ascending values of the standard deviation .
2. Convolve an image  $g$  with a Gaussian of scale .
3. Estimate local edge normal directions  $n$  for each pixel in the image.
4. Find the location of the edges (non-maximal suppression).
5. Compute the magnitude of the edge
6. Threshold edges in the image with hysteresis to eliminate spurious responses.
7. Aggregate the final information about edges at multiple scale using the `feature synthesis approach.

**(b)Edge Operators:**

Since edges consist of mainly high frequencies, we can, in theory, detect edges by applying a high pass frequency filter in the Fourier domain or by convolving the image with an appropriate kernel in the spatial domain. In practice, edged detection is performed in the spatial domain, because it is computationally less expensive and often yields better results.

Since edges correspond to strong illumination gradients, we can highlight them by calculating the derivatives of the image. We can see that the position of the edge can be estimated with the maximum of the 1st derivative or with the zero-crossing of the 2nd derivative. Therefore we want to find a technique to calculate the derivative of a two-dimensional image. For a discrete one dimensional function  $f(i)$ , the first derivative can be approximated by

$$\frac{df(i)}{di} = f(i + 1) - f(i)$$

Calculating this formula is equivalent to convolving the function with  $[-1 \ 1]$ . Similarly the 2<sup>nd</sup> derivative can be estimated by convolving  $f(i)$  with  $[1 \ -2 \ 1]$ . Different edge detection kernels which are based on the above formula enable us to calculate either the 1st or the 2nd derivative of a two-dimensional image. There are two common approaches to estimate the 1st derivative in a two-dimensional image, Prewitt compass edge detection and *gradient edge detection*. Prewitt compass edge detection involves convolving the image with a set of (usually 8) kernels, each of which is sensitive to a different edge orientation. The kernel producing the maximum response at a pixel location determines the edge magnitude and orientation. Different sets of kernels might be used: examples include Prewitt, Sobel, Kirsch and Robinson kernels. Gradient edge detection is the second and more widely used technique. Here, the image is

convolved with only two kernels, one estimating the gradient in the  $x$ -direction,  $G_x$ , the other the gradient in the  $y$ -direction,  $G_y$ . The absolute gradient magnitude is then given by

$$|G| = \sqrt{G_x^2 + G_y^2}$$

and is often approximated with

$$|G| = |G_x| + |G_y|$$

In many implementations, the gradient magnitude is the only output of a gradient edge detector, however the edge orientation might be calculated with The most common kernels used for the gradient edge detector are the Sobel, Roberts Cross and Prewitt operators. After having calculated the magnitude of the 1st derivative, we now have to identify those pixels corresponding to an edge. The easiest way is to threshold the gradient image, assuming that all pixels having a local gradient above the threshold must represent an edge. An alternative technique is to look for local maxima in the gradient image, thus producing one pixel wide edges. A more sophisticated technique is used by the Canny edge detector. It first applies a gradient edge detector to the image and then finds the edge pixels using *non-maximal suppression* and *hysteresis tracking*.

An operator based on the 2nd derivative of an image is the Marr edge detector, also known as *zero crossing detector*. Here, the 2nd derivative is calculated using a Laplacian of Gaussian (LoG) filter. The Laplacian has the advantage that it is an isotropic measure of the 2<sup>nd</sup> derivative of an image, *i.e.* the edge magnitude is obtained independently from the edge orientation by convolving the image with only one kernel. The edge positions are then given by the zero-crossings in the LoG image. The scale of the edges which are to be detected can be controlled by changing the variance of the Gaussian. A general problem for edge detection is its sensitivity to noise, the reason being that calculating the derivative in the spatial domain corresponds to accentuating high frequencies and hence magnifying noise. This problem is addressed in the Canny and Marr operators by convolving the image with a smoothing operator (Gaussian) before calculating the derivative.

### **(c) Line detection:**

While edges (I. E. boundaries between regions with relatively distinct graylevels) are by far the most common type of discontinuity in an image, instances of thin lines in an image occur frequently enough that it is useful to have a separate mechanism for detecting them. A convolution based technique can be used which produces an image description of the thin

lines in an input image. Note that the Hough transform can be used to detect lines; however, in that case, the output is a P A R A M E T R I C description of the lines in an image. The line detection operator consists of a convolution kernel tuned to detect the presence of lines of a particular width  $n$ , at a particular orientation  $\theta$ . Figure below shows a collection of four such kernels, which each respond to lines of single pixel width at the particular orientation shown.

a)	<table border="1" style="display: inline-table;"><tr><td>-1</td><td>-1</td><td>-1</td></tr><tr><td>2</td><td>2</td><td>2</td></tr><tr><td>-1</td><td>-1</td><td>-1</td></tr></table>	-1	-1	-1	2	2	2	-1	-1	-1
-1	-1	-1								
2	2	2								
-1	-1	-1								

b)	<table border="1" style="display: inline-table;"><tr><td>-1</td><td>2</td><td>-1</td></tr><tr><td>-1</td><td>2</td><td>-1</td></tr><tr><td>-1</td><td>2</td><td>-1</td></tr></table>	-1	2	-1	-1	2	-1	-1	2	-1
-1	2	-1								
-1	2	-1								
-1	2	-1								

c)	<table border="1" style="display: inline-table;"><tr><td>-1</td><td>-1</td><td>2</td></tr><tr><td>-1</td><td>2</td><td>-1</td></tr><tr><td>2</td><td>-1</td><td>-1</td></tr></table>	-1	-1	2	-1	2	-1	2	-1	-1
-1	-1	2								
-1	2	-1								
2	-1	-1								

d)	<table border="1" style="display: inline-table;"><tr><td>2</td><td>-1</td><td>-1</td></tr><tr><td>-1</td><td>2</td><td>-1</td></tr><tr><td>-1</td><td>-1</td><td>2</td></tr></table>	2	-1	-1	-1	2	-1	-1	-1	2
2	-1	-1								
-1	2	-1								
-1	-1	2								

Four line detection kernels which respond maximally to horizontal, vertical and oblique (+45 and -45 degree) single pixel wide lines. These masks above are tuned for light lines against a dark background, and would give a big negative response to dark lines against a light background. If we are only interested in detecting dark lines against a light background, then we should negate the mask values. Alternatively, we might be interested in either kind of line, in which case, we could take the absolute value of the convolution output.

If  $R_i$  denotes the response of kernel  $I$ , we can apply each of these kernels across an image, and for any particular point, if for all that point is more likely to contain a line whose orientation (and width) corresponds to that of kernel  $I$ . One usually thresholds to eliminate weak lines corresponding to edges and other features with intensity gradients which have a different scale than the desired line width. In order to find complete lines, one must join together line fragments, ex: with an edge tracking operator.

**(d)Corner detection:**

Input to the corner detector is the gray-level image. Output is an image in which values are proportional to the likelihood that the pixel is a corner. The Moravec detector is maximal in pixels with high contrast. These points are on corners and sharp edges.

$$MO(i, j) = \frac{1}{8} \sum_{k=i-1}^{i+1} \sum_{l=j-1}^{j+1} |g(k, l) - g(i, j)|$$

Using ZH Operator,

The image function  $f$  is approximated in the neighborhood of the pixel  $(i, j)$  by a cubic polynomial with coefficients  $c_k$ . This is a cubic facet model. The ZH operator estimates the corner strength based on the coefficients of the cubic facet model.

$$g(i, j) = c_1 + c_2x + c_3y + c_4x^2 + c_5xy + c_6y^2 + c_7x^3 + c_8x^2y + c_9xy^2 + c_{10}y^3$$

$$ZH(i, j) = \frac{-2(c_2^2c_6 - c_2c_3c_5 - c_3^2c_4)}{(c_2^2 + c_3^2)^{\frac{3}{2}}}$$