14. ASSIGNMENT TOPICS WITH MATERIALS

Unit-1

Topic 1: Image representation and its properties

We will use two principal ways to represent digital images. Assume that an image \( f(x,y) \) is sampled so that the resulting digital image has \( M \) rows and \( N \) columns. The values of the coordinates \((x, y)\) now become discrete quantities. For notational clarity and convenience, we shall use integer values for these discrete coordinates. Thus, the values of the coordinates at the origin are \((x, y) = (0, 0)\). The next coordinate values along the first row of the image are represented as \((x, y) = (0, 1)\). It is important to keep in mind that the notation \((0, 1)\) is used to signify the second sample along the first row. It does not mean that these are the actual values of physical coordinates when the image was sampled. Figure shows the coordinate convention used.
The notation introduced in the preceding paragraph allows us to write the complete M*N digital image in the following compact matrix form: The right side of this equation is by definition a digital image. Each element of this matrix array is called an image element, picture element, pixel, or pel.

The notation introduced in the preceding paragraph allows us to write the complete M*N Digital image in the following compact matrix form:

\[
\begin{bmatrix}
  f(0,0) & f(0,1) & \cdots & f(0,N-1) \\
  f(1,0) & f(1,1) & \cdots & f(1,N-1) \\
  \vdots & \vdots & \ddots & \vdots \\
  f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1)
\end{bmatrix}
\]

The right side of this equation is by definition a digital image. Each element of this matrix Array is called an image element, picture element, pixel, or pel.

**Geometric Transform**

Geometric image transformation functions use mathematical transformations to crop, pad, scale, rotate, transpose or otherwise alter an image array to produce a modified view of an image. The transformations described in this chapter are linear transformations. For a description of non-linear geometric transformations, When an image undergoes a geometric transformation, some or all of the pixels within the source image are relocated from their original spatial coordinates to a new position in the output image. When a relocated pixel does not map directly onto the centre of a pixel location, but falls somewhere in between the centres of pixel locations, the pixel's value is computed by sampling the values of the neighbouring pixels. This resampling, also known as interpolation, affects the quality of the output image.

**Cropping Images:**

Cropping an image extracts a rectangular region of interest from the original image. This focuses the viewer's attention on a specific portion of the image and discards areas of the image that contain less useful information. Using image cropping in conjunction with image magnification allows you to zoom in on a specific portion of the image. This section describes how to exactly define the portion of the image you wish to extract to create a cropped image.
Padding Images:

Image padding introduces new pixels around the edges of an image. The border provides space for annotations or acts as a boundary when using advanced filtering techniques. This exercise adds a 10-pixel border to left, right and bottom of the image and a 30-pixel border at the top allowing space for annotation. The diagonal lines in the following image represent the area that will be added to the original image. For an example of padding an image, complete the following steps.

<table>
<thead>
<tr>
<th>Transformation Name</th>
<th>Affine Matrix, T</th>
<th>Coordinate Equations</th>
<th>Example</th>
</tr>
</thead>
</table>
| Identity            | \[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
\end{bmatrix}
\] | \[x = v \] \[y = w \] | ![Identity Example] |
| Scaling             | \[
\begin{bmatrix}
c_x & 0 & 0 \\
0 & c_y & 0 \\
0 & 0 & 1 \\
\end{bmatrix}
\] | \[x = c_x v \] \[y = c_y w \] | ![Scaling Example] |
| Rotation            | \[
\begin{bmatrix}
\cos \theta & \sin \theta & 0 \\
-\sin \theta & \cos \theta & 0 \\
0 & 0 & 1 \\
\end{bmatrix}
\] | \[x = v \cos \theta - w \sin \theta \] \[y = v \cos \theta + w \sin \theta \] | ![Rotation Example] |
| Translation         | \[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
f_x & f_y & 1 \\
\end{bmatrix}
\] | \[x = v + f_x \] \[y = w + f_y \] | ![Translation Example] |
| Shear (vertical)    | \[
\begin{bmatrix}
1 & 0 & 0 \\
s_x & 1 & 0 \\
0 & 0 & 1 \\
\end{bmatrix}
\] | \[x = v + s_x w \] \[y = w \] | ![Shear (vertical) Example] |
| Shear (horizontal)  | \[
\begin{bmatrix}
1 & s_h & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
\end{bmatrix}
\] | \[x = v \] \[y = s_h v + w \] | ![Shear (horizontal) Example] |
\[ \alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} & \text{for } u = 0 \\ \sqrt{\frac{2}{N}} & \text{for } u = 1, 2, 3, \ldots, N - 1 \end{cases} \] (3.78)
Topic 2: DISCRETE COSINE TRANSFORM

The popular transform used for image compression is discrete Cosine transform (DCT) which is explained in detail in this section. The one-dimensional discrete Cosine transform is defined by equation (3.76)

\[ C(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) \cos \left( \frac{(2x + 1)u\pi}{2N} \right) \]  
(3.76)

for \( u = 0, 1, 2, \ldots, N-1 \)

Similarly, the inverse DCT is defined by equation (3.77), that is,

\[ f(x) = \sum_{u=0}^{N-1} \alpha(u) C(u) \cos \left( \frac{(2x + 1)u\pi}{2N} \right) \]  
(3.77)

for \( x = 0, 1, 2, \ldots, N-1 \)

The \( \alpha(u) \) used in equations (3.76) and (3.77) is defined in equation (3.78).

The corresponding two-dimensional DCT pair is

\[ C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos \left( \frac{(2x + 1)u\pi}{2N} \right) \cos \left( \frac{(2y + 1)v\pi}{2N} \right) \]  
(3.79)

for \( u, v = 0, 1, 2, \ldots, N-1 \)

\[ f(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v) C(u, v) \cos \left( \frac{(2x + 1)u\pi}{2N} \right) \cos \left( \frac{(2y + 1)v\pi}{2N} \right) \]  
(3.80)

for \( x, y = 0, 1, 2, \ldots, N-1 \) where \( \alpha \) is defined by equation (3.78).
3.6.1 Properties of Cosine Transform

1. The Cosine transform is real and orthogonal, that is,

\[ C = C^* \Rightarrow C^{-1} = C^T \]  \hspace{1cm} (3.81)

It is not the real part of the unitary DCT. However, the cosine transform of a sequence is related to the DFT of its symmetric extension.

2. The Cosine transform is a fast transform. The Cosine transform of a vector of \( N \) elements can be calculated in \( O(N \log_2 N) \) operations via \( N \)-point FFT. It has excellent energy compaction for correlated data.
**Topic 3: HADAMARD TRANSFORM**

The Hadamard transform is based on basis functions that are simply +1 or −1, instead of the more complex sine and cosine functions used in the Fourier transform. The one-dimensional kernel for the Hadamard transform is given in the relation

\[
H(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x)(-1)^{\sum_{i=0}^{n-1} b_i(x)b_i(u)}
\]  
(3.92)

Where the summation in the exponent is performed in modulo 2 arithmetic and \(b_i(x)\) is the \(i\)th bit in binary representation of \(x\). The one-dimensional equation for the Hadamard transform is

\[
g(x, u) = \frac{1}{N} \sum_{i=0}^{n-1} [b_i(x)b_i(u)]
\]  
(3.91)

where \(N = 2^n\) and \(u = 0, 1, 2, \ldots, N-1\).

As the Hadamard kernel forms the matrix having orthogonal rows and columns the inverse kernel exists and is given by

\[
h(x, u) = (-1)^{\sum_{i=0}^{n-1} b_i(x)b_i(u)}
\]  
(3.93)

Hence, the inverse Hadamard transform expression is as in equation (3.94).

\[
f(x) = \sum_{u=0}^{N-1} H(u)(-1)^{\sum_{i=0}^{n-1} b_i(x)b_i(u)}
\]  
(3.94)

for \(x = 0, 1, 2, \ldots, N-1\).

Similarly, the two-dimensional kernels are given by the relations

\[
g(x, y, u, v) = \frac{1}{N} (-1)^{\sum_{i=0}^{n-1} [b_i(x)b_i(u)+b_i(y)b_i(v)]}
\]  
(3.95)

and

\[
h(x, y, u, v) = \frac{1}{N} (-1)^{\sum_{i=0}^{n-1} [b_i(x)b_i(u)+b_i(y)b_i(v)]}
\]  
(3.96)

From equations (3.94) and (3.95) Hadamard transforms kernels are identical. Hence, the two-dimensional Hadamard transform pair
\[ H(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y)(-1)^{i=0} \left( \sum_{i=0}^{N-1} [b_i(x)b_i(u)+b_i(y)b_i(v)] \right) \quad (3.97) \]

and

\[ f(x, y) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} H(u, v)(-1)^{i=0} \left( \sum_{i=0}^{N-1} [b_i(x)b_i(u)+b_i(y)b_i(v)] \right) \quad (3.98) \]

The Hadamard kernel satisfies the separable and symmetric properties and they are

\[ g(x, y, u, v) = g_1(x, u)g_1(y, v) \]

\[ = h_1(x, u)h_2(y, v) \]

\[ = \left[ \frac{1}{\sqrt{N}}(-1)^{i=0} \sum_{i=0}^{n-1} [b_i(x)b_i(u)] \right] \left[ \frac{1}{\sqrt{N}}(-1)^{i=0} \sum_{i=0}^{n-1} [b_i(y)b_i(v)] \right] \quad (3.99) \]
As the two-dimensional Hadamard kernels are separable, the two-dimensional transform pair may be obtained by successive applications of the one-dimensional Hadamard transform algorithm.

The use of Walsh and Hadamard transforms is intermixed in the image processing literature. The term Walsh-Hadamard is often used to denote either transform. The reason for this is that the Hadamard transform can be obtained from the Walsh transform. For example, the Hadamard transform kernel (Table 3.2) can be obtained from the Walsh transform kernel Table(3.1) by reordering the columns.

**Table 3.2** The one-dimensional Hadamard transformation kernel for \(N=8\)

<table>
<thead>
<tr>
<th>(X)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>1</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>5</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>7</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

By interchanging columns 1 and 4 and interchanging columns 3 and 6 in the Walsh transform, the Hadamard transform can be obtained. The important feature of the Walsh transform is that it can be expressed directly in a successive doubling format in order to obtain fast Walsh transform, whereas the Hadamard transform due to its ordering does not allow the successive doubling format to obtain fast Hadamard transform.

However, a Hadamard transform leads to a simple recursive relationship for generating the transformation matrices from the lower order to higher order. For example, \(n = 2\), the Hadamard matrix can be given as
To obtain the Hadamard matrix for $N = 4$, the recursive relationship given in equation (3.100) can be used.

In general,

$$H_{2N} = \begin{pmatrix} H_N & H_N \\ H_N & -H_N \end{pmatrix}$$  \hspace{1cm} (3.102)
Unit-II

Topic 1: SPATIAL DOMAIN AND FREQUENCY DOMAIN APPROACHES

In the spatial domain method, the pixel composing of image details are considered and the various procedures are directly applied on these pixels. The image processing functions in the spatial domain may be expressed as

\[ g(x, y) = T[f(x, y)] \]  \hspace{1cm} (4.1)

Where \( f(x, y) \) is the input image, \( g(x, y) \) is the processed output image and \( T \) represents an operation on \( f \) defined over some neighborhood of \( (x, y) \). Sometimes \( T \) can also be used to operate on a set of input images. Consider an image representation shown in Figure 4.2.
In Figure 4.2, a sub image of size \((3 \times 3)\) about a point \((x, y)\) is given. The center of the sub image is moved from pixel to pixel starting at the top left corner and applying the operator at each location \((x, y)\) to give the output image 'g' at that location. The subimage considered may be circle, square, or rectangular arrays.

If we consider the simplest case, where the neighborhood is \((1 \times 1)\), the output image \(g\) depends only on the value of \(f\) at \((x, y)\) and \(T\) is called a \textit{gray level transformation function}. The gray level transformation function will then be given in the form

\[
S = T(r) \tag{4.2}
\]

Where, \(r\) and \(s\) are variables denoting the gray level of \(f(x, y)\) and \(g(x, y)\) at any point \((x, y)\).
If $T(r)$ is of the form shown in Figure 4.3(a) the effect of this transformation is to produce an image of higher contrast compared to the original image by darkening the levels below $m$ and brightening the levels above $m$ in the original image. This approach is called contrast stretching, because the values of $r$, below $m$ are compressed, and the opposite effect takes place for the values above $m$.

**Contrast Stretching:** An enhancement technique used to increase the dynamic range of gray levels in the image being processed.

Figure 4.3(b) shows a typical case of transfer function $T(r)$ that produces a binary image. In general, the methods just described use a transfer function to produce the gray level in output image at the location $(x, y)$ that depends only on the gray level of the input image at that location. These techniques are often referred to as point processing techniques. The larger neighborhoods allow a variety of processing function that go beyond just image enhancement. The neighborhood of pixels considered about the center pixel $(x, y)$ is called as mask, template, or window. A mask is a small two-dimensional array (say, $3 \times 3$) shown in Figure 4.2. The enhancement techniques based on this type are often called as mask processing or filtering.
1. FREQUENCY DOMAIN TECHNIQUES

Ans: The convolution theorem is the basis for the frequency domain approaches. Let $G(x, y)$ be an image formed by the convolution of the image $f(x, y)$ and a linear position invariant operator $H(x, y)$ and is given by,

$$G(x, y) = H(x, y) \ast f(x, y)$$  \hspace{1cm} (4.3)

Where ‘$\ast$’ represents the convolution operation.

Then, from the convolution theorem, the equation 4.3 can be written in the frequency domain as,

$$G(u, v) = H(u, v)F(u, v)$$  \hspace{1cm} (4.4)

In the linear system theory, the transform $H(u, v)$ is called the Transfer function. The various image enhancement problems can be expressed in the form of equation (4.3). In a typical image enhancement application, the image $f(x, y)$ is given and the objective after the computation of $F(u, v)$ is to select $H(u, v)$ so that the desired image can be given by the equation

$$g(x, y) = F^{-1} [H(u, v)F(u, v)]$$  \hspace{1cm} (4.5)

This equation shows some highlighted feature of the original image $f(x, y)$. For example, the edges in $f(x, y)$ can be highlighted using a function $H(u, v)$ that emphasize the high frequency component of $F(u, v)$. 

FIGURE 4.3 A typical case of transfer function
Figure 4.4 illustrates the various steps involved in the enhancement approach based on frequency domain.

![Diagram](image)

**FIGURE 4.4 Enhancement steps in frequency domain approaches**

**Topic 2: Histogram and Histogram Equalization**

**Histogram** The histogram of a digital image is the probability of occurrence associated with the gray levels in the range 0 to 255. It can be expressed using a discrete function

\[
P(r_k) = \frac{n_k}{n}
\]

where \( r_k \) is the \( k \)th gray level, \( n_k \) is the number of pixels in the image with that gray level, \( n \) is the total number of pixels in the image and \( k = 0, 1, 2, \ldots, 255 \). In general, \( P(r_k) \) gives an estimate of the probability of occurrence of gray level \( r_k \). The plot of \( P(r_k) \) for all values of \( k \) is called histogram of the image and it gives a global description of the appearance of an image. The histograms for four different types of images are shown in Figure 4.11.

**Histogram**: A plot between the probability associated with each gray level versus gray levels in the image. From this one can infer whether the given image is

- A dark image or
- Bright image or
- Low contrast image or High contrast image.

The histogram shown in Figure 4.11(a) shows that the gray levels are concentrated towards the dark end of the gray scale range. Thus this histogram corresponds to an image with overall dark characteristics. Figure 4.11(b) shows the histogram for a bright image. Figures 4.11(c) and (d) are the histograms for low-contrast and high-contrast images, respectively. Thus the shape of the histogram gives useful information about the possibility for contrast enhancement. The following discussion is for image enhancement based on histogram manipulation.
**Histogram equalization** Let $r$ be the variable representing the gray levels in the image to be enhanced. Assume that the gray levels in this image after normalization range from 0 to 1. For any value of $r$ in the original image in the interval $(0, 1)$ the transformation in the form

$$s = T(r)$$  \hspace{1cm} (4.7)

produces a gray level $s$. It is assumed that equation (4.7) satisfies the following two conditions:

1. $T(r)$ is single-valued and monotonically increasing in the interval $0 \leq r \leq 1$
2. $0 \leq T(r) \leq 1$ for $0 \leq r \leq 1$.

The first condition preserves the order from black to white in the gray scale, whereas the second condition guarantees a mapping that is consistent with the allowed range of pixel values.
An example transfer function given in Figure 4.13 satisfies these conditions. The inverse transfer function from $s$ back to $r$ is given as
Unit-III

Topic1: Model of the Image Restoration Process

The Fig. 6.3 shows, the degradation process is modeled as a degradation function that, together with an additive noise term, operates on an input image $f(x, y)$ to produce a degraded image $g(x, y)$. Given $g(x, y)$, some knowledge about the degradation function $H$, and some knowledge about the additive noise term $\varsigma(x, y)$, the objective of restoration is to obtain an estimate $\hat{f}(x, y)$ of the original image. The estimate should be as close as possible to the original input image and, in general, the more we know about $H$ and $\varsigma$, the closer $\hat{f}(x, y)$ will be to $f(x, y)$.

The degraded image is given in the spatial domain by

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

where $h(x, y)$ is the spatial representation of the degradation function and, the symbol * indicates convolution. Convolution in the spatial domain is equal to multiplication in the frequency domain, hence

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

where the terms in capital letters are the Fourier transforms of the corresponding terms in above equation.
Where \(T^{-1}(s)\) also satisfies the conditions (1) and (2) with respect to the variable \(s\). The gray levels in an image may be viewed as random quantities in the interval \((0,1)\). The original and transformed gray levels can be characterized by their probability density functions \(P_r(r)\) and \(P_s(S)\), respectively. If \(P_r(r)\) and \(T(r)\) are known and \(T^{-1}(s)\) satisfies condition (1), the probability density function of the transformed image gray levels is then given by

\[
P_s(s) = \left[ P_r(r) \frac{dr}{ds} \right]_{r = T^{-1}(s)} \quad (4.9)
\]

We now discuss the approach which is based on modifying the appearance of an image by controlling the probability density function of its gray levels using the transformation function \(T(r)\) as shown in Figure 4.12.

Consider the transformation function,

\[
S = T(r) = \int_0^r P_r(r) dr \quad 0 \leq r \leq 1 \quad (4.10)
\]

The right side of equation (4.10) is called as cumulative distribution function (CDF) of \(r\). The CDF satisfies the conditions (1) and (2) already stated. Differentiating equation (4.10) with respect to \(r\) results,

\[
\frac{ds}{dr} = P_r(r) \quad (4.11)
\]

**FIGURE 4.12** Gray level transformation function
Topic 2: Restoration Filters Used When The Image Degradation Is Due To Noise Only

If the degradation present in an image is only due to noise, then,

\[ g(x, y) = f(x, y) + \zeta(x, y) \]
\[ G(u, v) = F(u, v) + N(u, v) \]

The restoration filters used in this case are,

1. Mean filters
2. Order static filters and
3. Adaptive filters

(i) Arithmetic mean filter

This is the simplest of the mean filters. Let \( S_{xy} \) represent the set of coordinates in a rectangular subimage window of size \( m \times n \), centered at point \((x, y)\). The arithmetic mean filtering process 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 simply the arithmetic mean computed using the pixels in the region defined by \( S_{xy} \). In other words

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

This operation can be implemented using a convolution mask in which all coefficients have value \( 1/mn \)

(ii) 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 pixels in the subimage window, raised to the power 1/mn. A geometric mean filter achieves smoothing comparable to the arithmetic mean filter, but it tends to lose less image detail in the process.

(iii) Harmonic mean filter

The harmonic mean filtering operation is given by the expression

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

The harmonic mean filter works well for salt noise, but fails for pepper noise. It does well also with other types of noise like Gaussian noise.

(iv) Contra harmonic mean filter

The contra harmonic mean filtering operation yields a restored image based on the expression

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

where Q is called the order of the filter. This filter is well suited for reducing or virtually eliminating the effects of salt-and-pepper noise. For positive values of Q, the filter eliminates pepper noise. For negative values of Q it eliminates salt noise. It cannot do both simultaneously. Note that the contra harmonic filter reduces to the arithmetic mean filter if Q = 0, and to the harmonic mean filter if Q = -1.
Topic 3: Inverse Filtering

The simplest approach to restoration is direct inverse filtering, where \( F(u, v) \), the transform of the original image is computed simply by dividing the transform of the degraded image, \( G(u, v) \), by the degradation function.

\[
\hat{F}(u, v) = \frac{G(u, v)}{H(u, v)}.
\]
The divisions are between individual elements of the functions.

But \( G(u, v) \) is given by

\[
G(u, v) = F(u, v) + N(u, v)
\]

Hence

\[
\hat{F}(u, v) = F(u, v) + \frac{N(u, v)}{H(u, v)}.
\]

It tells that even if the degradation function is known the undegraded image cannot be recovered [the inverse Fourier transform of \( F(u, v) \)] exactly because \( N(u, v) \) is a random function whose Fourier transform is not known.

If the degradation has zero or very small values, then the ratio \( N(u, v)/H(u, v) \) could easily dominate the estimate \( F(u, v) \).

One approach to get around the zero or small-value problem is to limit the filter frequencies to values near the origin. \( H(0, 0) \) is equal to the average value of \( h(x, y) \) and that this is usually the highest value of \( H(u, v) \) in the frequency domain. Thus, by limiting the analysis to frequencies near the origin, the probability of encountering zero values is reduced.
UNIT-IV

Topic 1: Image Processing Is Segmentation Process

One of the most difficult task in image processing is segmentation process. It plays a vital role in any application and its success is based on the effective implementation of the segmentation technique.

The segmentation algorithms can be divided into two broad categories based on the two important properties, namely, discontinuity and similarity.

The various segmentation techniques based on (1) gray level discontinuity and (2) gray level similarity are well depicted in a graph as shown in Figure 6.1.

The forthcoming sections deal with the detection of isolated points, lines, and edges.

![Segmentation Techniques Diagram](image)

**FIGURE 6.1** Segmentation techniques
6.2 DETECTION OF ISOLATED POINTS

In order to detect the isolated points due to noise or any other interference, the general mask employed is shown in Figure 6.2(a) and the typical values of the weights 'W' are shown in Figure 6.2(b). This mask consists of coefficients ‘−1’ everywhere except at the center which is

8. The sum of these coefficients is 0. When we place the mask over an image it covers 9 pixels in the image. The average response to the mask is computed as

\[ R = \frac{1}{9} \{ W_1 Z_1 + W_2 Z_2 + \cdots + W_9 Z_9 \} = \frac{1}{9} \sum_{i=1}^{9} W_i Z_i \]  \hspace{1cm} (6.1)

where \( W_i \) is the coefficient in the mask and \( Z_i \) denotes the gray level values of the pixel in the image under the mask. Now the mask is placed at the top left corner of the image and the response to the mask is computed using equation (6.1).
FIGURE 6.2 (a) The general representation of the mask (b) The mask with coefficient values

If the mask is over a uniform intensity area, the response due to this mask is equal to 0. This means there are no isolated pixels with different gray level values. On the other hand, if the mask is placed over the area having an isolated point with different gray levels, the response to the mask will be a nonzero value. The average response will be maximum when the isolated points are just below the center of the mask. Therefore, from the mask response it is possible to locate the isolated points resulting due to noise.
Topic 2: Line Detection

The various masks used for detecting horizontal, vertical, +45°, and −45° slanting line are shown in Figure 6.3.

![Masks for detecting lines](image)

**FIGURE 6.3** Masks for detecting lines. (a) Mask for horizontal line detection (b) Mask for +45° slanting line detection (c) Mask for vertical line detection (d) Mask for −45° slanting line detection

If the first mask shown in Figure 6.3(a) is moved around an image, its response will be a large value to lines oriented horizontally. The response will be maximum when the line passes through the middle row of the mask with constant background. For example, when we move the
Mask over an image consisting of all ‘1’s as background and with a line of different gray level ‘10’s, then the response due to the first mask is computed as

\[(1 * -1) + (1 * -1) + (1 * -1) + (2 * 10) + (2 * 10) + (2 * 10) + (+1 * -1) + (+1 * -1) + (+1 * -1) = 54\]

This high response indicates that the mask is moving along a horizontal line with different gray levels compared to the background pixel gray level values. Similar experiments with second mask results in high response to vertical lines and the third mask to the lines +45° and the fourth mask to the lines in the −45° direction.

Suppose all the masks are applied to an image and the responses computed are denoted as \(R_1, R_2, R_3,\) and \(R_4\). If at a certain point in the image \(|R_i| > |R_j|\) for all \(j \neq i\), then the point is more likely to be associated with the line in the direction of mask \(i\). For example, if a point in the image where \(|R_1| > |R_j|\) for \(j = 2, 3,\) and \(4\) then that particular point is more likely to be associated with a horizontal line.

1. **SEGMENTATION USING THRESHOLD**

Thresholding is one of the most important techniques used for image segmentation. In this section we discuss the various thresholding techniques for image segmentation. The merits and demerits of various techniques are also discussed.

6.7.1 Fundamental Concepts

The histogram of an image \(f(x, y)\) consisting of a light object on the dark background is shown in Figure 6.17(a). This histogram consists of two dominant regions, one for the object and the other for the background. For such an image it is easy to select a threshold \(T\) that separates the object and background region. Then for any point \((x, y)\), \(f(x, y) > T\) is called an object point, otherwise the point is called a background point. A general case of this approach is shown in Figure 6.17(b).
Figure 6.17 (a) Histogram of an image consisting of dark background and a light object (b) Histogram for two objects in a dark background

Figure 6.17(b) has three dominant regions that characterize the histogram of the given image. This histogram corresponds to two different light objects on a dark background. From the histogram it is possible to select two different threshold values $T_1$ and $T_2$, respectively. Then a point $(x, y)$ belongs to the first object if $T_1 < f(x, y) \leq T_2$, or it belongs to the second object if $f(x, y) > T_2$ and to the background if $f(x, y) \leq T_1$. Usually this kind of thresholding is called multilevel thresholding and is less reliable than its single threshold counterpart.
The reason for this is, that it is difficult to locate multiple thresholds in a given histogram for a real image. The thresholding technique can be put into three different types based on the function \( T \) and its associated parameters as given in equation (6.21).

\[
T = T[(x, y), p(x, y), f(x, y)] \tag{6.21}
\]

where \( f(x, y) \) is the gray level at the point \((x, y)\) and \( p(x, y) \) denotes some local property at that point (e.g. the average gray level of neighborhood center on \((x, y)\)). The threshold image \( g(x, y) \) is given in equation (6.22).

\[
G(x, y) = \begin{cases} 
1 & \text{if } f(x, y) > T \\
0 & \text{if } f(x, y) \leq T 
\end{cases} \tag{6.22}
\]

In the thresholded image the pixel labeled 1 corresponds to the object, whereas pixels labeled 0 corresponds to the background. When the threshold value \( T \) depends only on \( f(x, y) \) the threshold technique is called \textit{global}. If \( T \) depends on both \( f(x, y) \) and \( p(x, y) \), the threshold is called \textit{local}. If \( T \) depends on all the three parameters, that is, the coordinates \((x, y)\), local property \( p(x, y) \), and \( f(x, y) \) then the threshold is called \textit{dynamic}. 
Unit-V

Topic 1: IMAGE COMPRESSION MODELS

A compression system consists of two distinct structural blocks: an encoder and a decoder. An input image \( f(x, y) \) is fed into the encoder, which creates a set of symbols from the input data. After transmission over the channel, the encoded representation is fed to the decoder, where a reconstructed output image \( \hat{f}(x, y) \) is generated. In general, \( \hat{f}(x, y) \) may or may not be an exact replica of \( f(x, y) \). If it is, the system is error free or information preserving; if not, some level of distortion is present in the reconstructed image. Both the encoder and decoder shown in Fig. 3.1 consist of two relatively independent functions or subblocks. The encoder is made up of a source encoder, which removes input redundancies, and a channel encoder, which increases the noise immunity of the source encoder's output. As would be expected, the decoder includes a channel decoder followed by a source decoder. If the channel between the encoder and decoder is noise free (not prone to error), the channel encoder and decoder are omitted, and the general encoder and decoder become the source encoder and decoder, respectively.

1.

![Fig.3.1 A general compression system model](image-url)
**Topic 2: The Source Encoder and Decoder Model**

The source encoder is responsible for reducing or eliminating any coding, interpixel, or psychovisual redundancies in the input image. The specific application and associated fidelity requirements dictate the best encoding approach to use in any given situation. Normally, the approach can be modeled by a series of three independent operations. As Fig. 3.2 (a) shows, each operation is designed to reduce one of the three redundancies. Figure 3.2 (b) depicts the corresponding source decoder. In the first stage of the source encoding process, the mapper transforms the input data into a (usually nonvisual) format designed to reduce interpixel redundancies in the input image. This operation generally is reversible and may or may not reduce directly the amount of data required to represent the image.

Run-length coding is an example of a mapping that directly results in data compression in this initial stage of the overall source encoding process. The representation of an image by a set of transform coefficients is an example of the opposite case. Here, the mapper transforms the image into an array of coefficients, making its interpixel redundancies more accessible for compression in later stages of the encoding process.

The second stage, or quantizer block in Fig. 3.2 (a), reduces the accuracy of the mapper's output in accordance with some preestablished fidelity criterion. This stage reduces the psychovisual redundancies of the input image. This operation is irreversible. Thus it must be omitted when error-free compression is desired. In the third and final stage of the source encoding process, the symbol coder creates a fixed- or variable-length code to represent the quantizer output and maps
the output in accordance with the code. The term symbol coder distinguishes this coding operation from the overall source encoding process. In most cases, a variable-length code is used to represent the mapped and quantized data set. It assigns the shortest code words to the most frequently occurring output values and thus reduces coding redundancy. The operation, of course, is reversible. Upon completion of the symbol coding step, the input image has been processed to remove each of the three redundancies.

Figure 3.2(a) shows the source encoding process as three successive operations, but all three operations are not necessarily included in every compression system. Recall, for example, that the quantizer must be omitted when error-free compression is desired. In addition, some compression techniques normally are modeled by merging blocks that are physically separate in Fig. 3.2(a). In the predictive compression systems, for instance, the mapper and quantizer are often represented by a single block, which simultaneously performs both operations.

The source decoder shown in Fig. 3.2(b) contains only two components: a symbol decoder and an inverse mapper. These blocks perform, in reverse order, the inverse operations of the source encoder's symbol encoder and mapper blocks. Because quantization results in irreversible information loss, an inverse quantizer block is not included in the general source decoder model shown in Fig. 3.2(b).

**Topic 3: Lossy Predictive Coding**
In this type of coding, we add a quantizer to the lossless predictive model and examine the resulting trade-off between reconstruction accuracy and compression performance. As Fig. 9 shows, the quantizer, which absorbs the nearest integer function of the error-free encoder, is inserted between the symbol encoder and the point at which the prediction error is formed. It maps the prediction error

1. into a limited range of outputs, denoted $e_n$ which establish the amount of compression and distortion associated with lossy predictive coding.

2.

**Fig. 9** A lossy predictive coding model: (a) encoder and (b) decoder.
15. Tutorial topics and Questions: Nil

16. UNIT WISE-QUESTION BANK
UNIT-I

Two marks question answers

Q1. Define Image.

Ans:-An image may be defined as two dimensional light intensity function \( f(x, y) \) where \( x \) and \( y \) denote spatial co-ordinate and the amplitude or value of \( f \) at any point \( (x, y) \) is called intensity or grayscale or brightness of the image at that point.

Q2. What is Dynamic Range?

Ans:-The range of values spanned by the gray scale is called dynamic range of an image. Image will have high contrast, if the dynamic range is high and image will have dull washed out gray look if the dynamic range is low.

Q3. Define Brightness.

Ans:-Brightness of an object is the perceived luminance of the surround. Two objects with different surroundings would have identical luminance but different brightness.

Q4. What do you meant by Gray level?

Ans:-Gray level refers to a scalar measure of intensity that ranges from black to grays and finally to white.

Q5. Define Resolutions.

Ans:-Resolution is defined as the smallest number of discernible detail in an image. Spatial resolution is the smallest discernible detail in an image and gray level resolution refers to the smallest discernible change is gray level.

Three marks question with answers
Q1. What is meant by pixel?
Ans:- A digital image is composed of a finite number of elements each of which has a particular location or value. These elements are referred to as pixels or image elements or picture elements or pels elements.

Q2. Define sampling and quantization?
Ans:- Sampling means digitizing the co-ordinate value \((x, y)\).
Quantization means digitizing the amplitude value.

Q3. What is meant by illumination and reflectance?
Ans:- Illumination is the amount of source light incident on the scene. It is represented as \(i(x, y)\).
Reflectance is the amount of light reflected by the object in the scene. It is represented by \(r(x, y)\).

Q4. Define weber ratio.
Ans:- The ratio of increment of illumination to background of illumination is called as weber ratio \((\text{ie}) \frac{\Delta i}{i}\)
If the ratio \((\Delta i/i)\) is small, then small percentage of change in intensity is needed \((\text{ie})\) good brightness adaptation.
If the ratio \((\Delta i/i)\) is large, then large percentage of change in intensity is needed \((\text{ie})\) poor brightness adaptation.

Q5. List the categories of digital storage?
1. Short term storage for use during processing.
2. Online storage for relatively fast recall.
3. Archival storage for infrequent access.

Five marks question answers
Q1. Explain sampling and quantization.

Ans:- For computer processing, the image function \( f(x,y) \) must be digitized both spatially and in amplitude. Digitization of spatial co-ordinates is called image sampling and amplitude digitization is called grey level quantization. Sampling: Consider a digital image of size 1024*1024,256 with a display area used for the image being the same, the pixels in the lower resolution images where duplicated inorder to fulfill the entire display. The pixel replication produced a checker board effect, which is visible in the image of lower resolution. It is not possible to differentiate a 512*512 images from a 1024*1024 under this effect. But a slight increase in grainess and a small decrease in sharpness is noted. A 256*256 image shows a fine checker board pattern in the edges and more pronounced grainess there out the image. These effects is much more visible in 128*128 images and it becomes quite pronounced in 64*64 and 32*32 images.

Quantization:

It discusses the effects produced when the number of bits used to represent the level in an image is decreased. This is illustrated by reducing the grey level required to represent a 1024*1024,512 image. The 256,128, and 64 level image are visually identical for all practical purposes the 32 level images has developed a set of rigid like structure in areas of smooth grey I+I lines.
effect caused by the user insufficient number of grey levels in smooth areas of digital image is
called a false contouring. this is visible in images displayed using 16 or lesser gray level values.

Q2. What are the fundamental steps in digital image processing?

Ans:- Following are the fundamental steps in image processing.

1) Image Acquisition: This is the first step or process of the fundamental steps of digital
image processing. Image acquisition could be as simple as being given an image that is
already in digital form. Generally, the image acquisition stage involves preprocessing,
such as scaling etc.

2) Image Enhancement: Image enhancement is among the simplest and most appealing
areas of digital image processing. Basically, the idea behind enhancement techniques is to
bring out detail that is obscured, or simply to highlight certain features of interest in an
image. Such as, changing brightness & contrast etc.

3) Image Restoration: Image restoration is an area that also deals with improving the
appearance of an image. However, unlike enhancement, which is subjective, image
restoration is objective, in the sense that restoration techniques tend to be based on
mathematical or probabilistic models of image degradation.

4) Color Image Processing: Color image processing is an area that has been gaining its
importance because of the significant increase in the use of digital images over the
Internet. This may include color modeling and processing in a digital domain etc.

5) Wavelets and Multiresolution Processing: Wavelets are the foundation for representing
images in various degrees of resolution. Images subdivision successively into smaller
regions for data compression and for pyramidal representation.

6) Compression: Compression deals with techniques for reducing the storage required to
save an image or the bandwidth to transmit it. Particularly in the uses of internet it is very
much necessary to compress data.
7) Morphological Processing: Morphological processing deals with tools for extracting image components that are useful in the representation and description of shape.

8) Segmentation: Segmentation procedures partition an image into its constituent parts or objects. In general, autonomous segmentation is one of the most difficult tasks in digital image processing. A rugged segmentation procedure brings the process a long way toward successful solution of imaging problems that require objects to be identified individually.

9) Representation and Description: Representation and description almost always follow the output of a segmentation stage, which usually is raw pixel data, constituting either the boundary of a region or all the points in the region itself. Choosing a representation is only part of the solution for transforming raw data into a form suitable for subsequent computer processing. Description deals with extracting attributes that result in some quantitative information of interest or are basic for differentiating one class of objects from another.
10) Object recognition: Recognition is the process that assigns a label, such as, “vehicle” to an object based on its descriptors.

11) Knowledge Base: Knowledge may be as simple as detailing regions of an image where the information of interest is known to be located, thus limiting the search that has to be conducted in seeking that information. The knowledge base also can be quite complex, such as an interrelated list of all major possible defects in a materials inspection problem or an image database containing high-resolution satellite images of a region in connection with change-detection applications.

Q3. Describe the basic relationship between the pixels?

Ans: The basic relationship between pixels can be shown as follows:

- Neighbours of a pixel
- Adjacency, Connectivity, Regions and Boundaries
- Distance measures

Neighbours of a pixel

- A pixel $p$ at coordinates $(x, y)$ has four horizontal and vertical neighbours whose coordinates are given by

$(x+1, y), (x-1, y), (x, y+1), (x, y-1)$.

- This set of pixels, called the 4-neighbours of $p$, is denoted by $N_4(p)$. Each pixel is a unit distance from $(x, y)$ and some of the neighbours of $p$ lie outside the digital image if $(x, y)$ is on the border of the image.

- The four diagonal neighbours of $p$ have coordinates

$(x+1, y+1), (x+1, y-1), (x-1, y+1), (x-1, y-1)$.
And are denoted by ND(p). These points together with the 4-neighbours are called
the 8-neighbours of p, denoted by N8(p).

Adjacency, Connectivity, Regions and Boundaries

Three types of adjacency:

- 4-adjacency. Two pixels p and q with values from V are 4-adjacent if q is in the
  set N4(p).
- 8-adjacency. Two pixels p and q with values from V are 8-adjacent if q is in the
  set N8(p).
- M-adjacency. Two pixels p and q with values from V are m-adjacent if
  q is in N4(p), or q is in ND(p) and the set N4(p) △ N4(q) has no pixels whose
  values are from V.

- A (digital) path (or curve) from pixel p with coordinates (x,y) to pixel q with
  coordinates (s,t) is a sequence of distinct pixels with coordinates
  (x0,y0), (x1,y1),.................(xn,yn)

  Where (x0,y0)= (x,y), (xn,yn)=(s,t) and pixels (xi,yi) and (xi-1,yi-1) are adjacent for
  1<=i<=n. N is the length of the path.

Distance measures

- For pixels p,q and z with coordinates (x,y), (s,t) and (v,w) respectively, D is a
distance function or metric if
  - D(p,q)>=0 (D(p,q)=0 iff p=q),
• $D(p,q) = D(q,p)$ and
• $D(p,z) \leq D(p,q) + D(q,z)$

• The Euclidean distance between $p$ and $q$ is defined as,

$D_{e}(p,q) = [(x-s)^2 + (y-t)^2]$

• The D4 distance (also called city-block distance) between $p$ and $q$ is defined as

$D_{4}(p,q) = |x-s| + |y-t|$

• The D8 distance (also called chessboard distance) between $p$ and $q$ is defined as

$D_{8}(p,q) = \max(|x-s| + |y-t|)$

**Q4. Describe the basic relationship between the pixels**

Ans: The basic relationship between pixels can be shown as follows:

• Neighbours of a pixel
• Adjacency, Connectivity, Regions and Boundaries
• Distance measures

Neighbours of a pixel

• A pixel $p$ at coordinates $(x,y)$ has four horizontal and vertical neighbours whose coordinates are given by

$(x+1,y), (x-1,y), (x,y+1), (x,y-1)$.

• This set of pixels, called the 4-neighbours of $p$, is denoted by $N_{4}(p)$. Each pixel is a unit distance from $(x,y)$ and some of the neighbours of $p$ lie outside the digital image if $(x,y)$ is on the border of the image.
• The four diagonal neighbours of \( p \) have coordinates 
\[(x+1,y+1), (x+1,y-1), (x-1,y+1), (x-1,y-1)\]

• And are denoted by \( \text{ND}(p) \). These points together with the 4-neighbours are called the 8-neighbours of \( p \), denoted by \( \text{N8}(p) \).

Adjacency, Connectivity, Regions and Boundaries

Three types of adjacency:

• 4-adjacency. Two pixels \( p \) and \( q \) with values from \( V \) are 4-adjacent if \( q \) is in the set \( \text{N4}(p) \).

• 8-adjacency. Two pixels \( p \) and \( q \) with values from \( V \) are 8-adjacent if \( q \) is in the set \( \text{N8}(p) \).

• M-adjacency. Two pixels \( p \) and \( q \) with values from \( V \) are m-adjacent if \( q \) is in \( \text{N4}(p) \), or \( q \) is in \( \text{ND}(p) \) and the set \( \text{N4}(p) \cap \text{N4}(q) \) has no pixels whose values are from \( V \).

• A (digital) path (or curve) from pixel \( p \) with coordinates \((x,y)\) to pixel \( q \) with coordinates \((s,t)\) is a sequence of distinct pixels with coordinates 
\[(x_0,y_0), (x_1,y_1), \ldots, (x_n,y_n)\]

Where \((x_0,y_0) = (x,y)\), \((x_n,y_n) = (s,t)\) and pixels \((x_i,y_i)\) and \((x_{i-1},y_{i-1})\) are adjacent for \( 1 \leq i \leq n \). \( N \) is the length of the path.

Distance measures

• For pixels \( p,q \) and \( z \) with coordinates \((x,y)\), \((s,t)\) and \((v,w)\) respectively, \( D \) is a
distance function or metric if

- $D(p,q) \geq 0$ (if $p=q$),
- $D(p,q) = D(q,p)$ and
- $D(p,z) \leq D(p,q) + D(q,z)$

• The Euclidean distance between $p$ and $q$ is defined as,

$$D_e(p,q) = \sqrt{(x-s)^2+(y-t)^2}$$

• The $D_4$ distance (also called city-block distance) between $p$ and $q$ is defined as

$$D_4(p,q) = |x-s|+|y-t|$$

• The $D_8$ distance (also called chessboard distance) between $p$ and $q$ is defined as

$$D_8(p,q) = \max(|x-s|+|y-t|)$$

**Q5. Explain the Structure of the Human eye**

*Ans: The eye is early a sphere, with an average diameter of approximately 20 mm.*

Three membranes encloses the eye,

1. Cornea

2. Sclera or Cornea:

3. Retina

*The cornea is a tough, transparent tissue that covers the anterior surface of the eye.*

*Sclera:*

*Sclera is an opaque membrane that encloses the remainder of the optical globe.*
Choroid:

- Choroid directly below the sclera. This membrane contains a network of blood vessels that serve as the major source of nutrition to the eye.

- Choroid coat is heavily pigmented and helps to reduce the amount of extraneous light entering the eye.

- The choroid is divided into the ciliary body and the iris diaphragm.

Lens:

The lens is made up of concentric layers of fibrous cells and is suspended by fibrous that attach to the ciliary body. It contains 60 to 70% of water about 60% fat and more protein than any other tissue in the eye.
Retina:

The innermost membrane of the eye is retina, which lines the inside of the wall’s entire posterior portion. There are 2 classes of receptors,

1. Cones
2. Rods

Cones:

The cones in each eye between 6and7 million. They are located primarily in the central portion of the retina called the fovea, and highly sensitive to Colour.

Rods:

The number of rods is much larger; some 75 to 150 millions are distributed over the retinal surface.

Fovea as a square sensor array of size 1.5mm*1.5mm.
Multiple Choice Questions:

1. The spatial coordinates of a digital image \((x,y)\) are proportional to:
   a) Position
   b) Brightness
   c) Contrast
   d) Noise

2. Among the following image processing techniques which is fast, precise and flexible.
   a) Optical
   b) Digital
   c) Electronic
   d) Photographic

3. An image is considered to be a function of \(a(x,y)\), where \(a\) represents:
   a) Height of image
   b) Width of image
   c) Amplitude of image
   d) Resolution of image

4. What is pixel?
   a) Pixel is the elements of a digital image
   b) Pixel is the elements of an analog image
   c) Pixel is the cluster of a digital image
   d) Pixel is the cluster of an analog image

5. The range of values spanned by the gray scale is called:
   a) Dynamic range
   b) Band range
c) Peak range
d) Resolution range

6. Which is a colour attribute that describes a pure colour?
   a) Saturation
   b) Hue
   c) Brightness
   d) Intensity

7. Which gives a measure of the degree to which a pure colour is diluted by white light?
   a) Saturation
   b) Hue
   c) Intensity
   d) Brightness

8. Which means the assigning meaning to a recognized object.
   a) Interpretation
   b) Recognition
   c) Acquisition
   d) Segmentation

9. A typical size comparable in quality to monochromatic TV image is of size.
   a) 256 X 256
   b) 512 X 512
   c) 1920 X 1080
   d) 1080 X 1080

10. A continuous image is digitised at _______ points.
    a) random
    b) vertex
c) contour
d) sampling

**Fill in the blanks**

1. ------- is the element in the image with a location and intensity value.
2. ------- is the first and foremost step in Image Processing.
3. ------- is the simplest type of connectivity.
4. ---- and ----- are required to convert a continuous sensed data into Digital form.
5. The quality of a digital image is well determined by ------- and -------.
6. An image whose gray-levels span a significant portion of gray scale have _________ dynamic range while an image with dull, washed out gray look have _________ dynamic range.
7. A ----- is required for conversion from one domain into another.
8. ----- waveform is used in slant transform.
9. The transition between continuous values of the image function and its digital equivalent is called ____________
10. Hotelling transform is also known as -------.

**Solutions**

<table>
<thead>
<tr>
<th>MCQ’s</th>
<th>Fill in the blanks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. a</td>
<td>1. Pixel</td>
</tr>
<tr>
<td>2. b</td>
<td>2. Image acquisition</td>
</tr>
<tr>
<td>3. c</td>
<td>3. 4 connectivity</td>
</tr>
<tr>
<td>4. a</td>
<td>4. Sampling , quantization</td>
</tr>
<tr>
<td>5. b</td>
<td>5. Number of samples, number of discrete gray levels</td>
</tr>
<tr>
<td>6. a</td>
<td>6. High, low</td>
</tr>
<tr>
<td>7. a</td>
<td>7. Transform</td>
</tr>
<tr>
<td>8. b</td>
<td>8. Sawtooth</td>
</tr>
<tr>
<td>9. b</td>
<td>9. Quantization</td>
</tr>
<tr>
<td>10. d</td>
<td>10. K L transform</td>
</tr>
</tbody>
</table>

**UNIT II**
Two marks question answers

Q1. Specify the objective of image enhancement technique?
Ans:- The objective of enhancement technique is to process an image so that the result is more suitable than the original image for a particular application.

Q2. List the 2 categories of image enhancement?
Ans:- Spatial domain refers to image plane itself & approaches in this category are based on direct manipulation of picture image. Frequency domain methods based on modifying the image by fourier transform.

Q3. What is the purpose of image averaging?
Ans:- An important application of image averaging is in the field of astronomy, where imaging with very low light levels is routine, causing sensor noise frequently to render single images virtually useless for analysis.

Q4. What is meant by masking?
Ans:- Mask is the small 2-D array in which the values of mask co-efficient determines the nature of process. The enhancement technique based on this type of approach is referred to as mask processing.

Q5. Define histogram?
Ans:- The histogram of a digital image with gray levels in the range [0, L-1] is a discrete function h(rk)=nk.
Where rk-kth gray level and
nk-number of pixels in the image having gray level rk.

Three marks question answers
Q1. Differentiate linear spatial filter and non-linear spatial filter?

Ans:-

<table>
<thead>
<tr>
<th>s.no.</th>
<th>Linear spatial filter</th>
<th>Non-linear spatial filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Response is a sum of products of the filter co-efficient.</td>
<td>They do not explicitly use coefficients in the sum-of-products.</td>
</tr>
</tbody>
</table>
| 2.    | $R = w(-1,-1) f(x-1,y-1) + w(-1,0) f(x-1,y) + \ldots + w(0,0) f(x,y) + \ldots + w(1,0) f(x+1,y) + w(1,1) f(x+1,y+1)$ | $R = w1z1 + w2z2 + \ldots +w9z9$ \[
\sum_{i=1}^{9} wiz_i
\]

Q2. Give the mask used for high boost filtering?

Ans:-

\[
\begin{array}{ccc}
0 & -1 & 0 \\
-1 & A+4 & -1 \\
0 & -1 & 0 \\
\end{array}
\] \[
\begin{array}{ccc}
-1 & -1 & -1 \\
-1 & A+8 & -1 \\
-1 & -1 & -1 \\
\end{array}
\]

Q3. Write the steps involved in frequency domain filtering?

Ans:-

1. Multiply the input image by (-1) to center the transform.
2. Compute $F(u,v)$, the DFT of the image from (1).
3. Multiply $F(u,v)$ by a filter function $H(u,v)$.
4. Compute the inverse DFT of the result in (3).
5. Obtain the real part of the result in (4).
6. Multiply the result in (5) by (-1)
Q4. What do you mean by Point processing?
Ans:-Image enhancement at any Point in an image depends only on the gray level at that point is often referred to as Point processing.

Q5. What are the various point operations?
Ans:-The various types of point operations are:
   a) Brightness modification
   b) Contrast stretching
   c) Histogram manipulation.
Q1. What do you mean by image enhancement. Explain its types?

Ans:-Image enhancement

Image enhancement is the process of manipulating an image so that the result is more suitable than the original for a specific application. The word specific is important here, because it establishes at the outset that enhancement techniques are problem oriented. Thus, for example, a method that is quite useful for enhancing X-ray images may not be the best approach for enhancing satellite images taken in the infrared band of the electromagnetic spectrum.

The techniques used for enhancement may be divided into two broad categories:-
   a) Frequency domain methods: based on modification of Fourier transform of an image.
   b) Spatial domain method: refers to the image plane itself and methods in this category are based on direct manipulation of the pixels an image

Q2. Explain Spatial Filtering?

Ans:- The use of spatial mask for the image processing usually called spatial filtering and spatial mask are spatial filters.

- The linear filter classified into
  a. Low pass
  b. High pass
  c. Band pass filtering

Consider 3*3 mask

<table>
<thead>
<tr>
<th>W1</th>
<th>W2</th>
<th>W3</th>
</tr>
</thead>
<tbody>
<tr>
<td>W4</td>
<td>W5</td>
<td>W6</td>
</tr>
<tr>
<td>W7</td>
<td>W8</td>
<td>W9</td>
</tr>
</tbody>
</table>

Denoting the gray level of pixels under the mask at any location by z1,z2,z3…….,z9, the response of a linear mask is

\[ R = w_1z_1 + w_2z_2 + \ldots + w_9z_9 \]

1. Smoothing Filters
a. Lowpass Spatial filtering:
The filter has to have positive coefficient.
The response would be the sum of gray levels of nine pixels which
could cause R to be out of the gray level range.
The solution is to scale the sum by dividing R by 9. The use of the
form of mask are called neighborhood averaging

b. Median filtering:
To achieve noise reduction rather than blurring.
The gray level of each pixel is replaced by the median of the gray
level in the neighbourhood of that pixel

2. Sharpening Filters
   a. Basic highpass spatial filtering:
The filter should be positive coefficient near the center and
   negative in the outer periphery.
The sum of the coefficient are 0.
This eliminate the zero-frequency term reducing significantly the
global contrast of the image

   b. High_boost filtering:

   It is used when the frequencies that need to be processed are actually below the
threshold and in cases where the value of threshold cannot be altered. In such cases,
the frequencies are multiplied by a suitable constant to raise the required low
frequency components above the threshold value. If A is the constant, then

   High_boost=(A)(Original)-Lowpass
   =(A-1) (Original)+ Original –Lowpass
   =(A-1) (Original)+Highpass

Q3. Write about image enhancement through point operation and its types?
Ans:-These processing methods are based only on the intensity of single pixels. That is why it is called point processing.

Types of point operators:

- **Contrast stretching**
  Contrast stretching (often called normalization) is a simple image enhancement technique that attempts to improve the contrast in an image by `stretching` the range of intensity values it contains to span a desired range of values, e.g. the full range of pixel values that the image type concerned allows. It differs from the more sophisticated histogram equalization in that it can only apply a linear scaling function to the image pixel values. As a result the `enhancement' is less harsh. (Most implementations accept a graylevel image as input and produce another graylevel image as output.)

- **Brightness Modification**
  The brightness of an image depends on the value associated with the pixel of the image. when changing the Brightness of an image, a constant is added or subtracted from the luminance of all sample values. The Brightness of the image can be increased by adding a constant value to each and every pixel of the image. The brightness can be decreased by subtracting a constant value from each and every pixel of the image.

**Q4. What is histogram manipulation?**
Histogram manipulation basically modifies the histogram of an input image so as to improve the visual quality of the image.

(a) Histogram The histogram of an image is a plot of the number of occurrences of gray levels in the image against the gray-level values. The histogram provides a convenient summary of the intensities in an image, but it is unable to convey any information regarding spatial relationships between pixels. The histogram provides more insight about image contrast and brightness.

1. The histogram of a dark image will be clustered towards the lower gray level.

2. The histogram of a bright image will be clustered towards higher gray level.

3. For a low contrast image, the histogram will not be spread equally, that is, the histogram will be narrow.

4. For a high-contrast image, the histogram will have an equal spread in the gray level. Image brightness may be improved by modifying the histogram of the image.

(b) Histogram Equalisation

Equalisation is a process that attempts to spread out the gray levels in an image so that they are evenly distributed across their range. Histogram equalisation reassigns the brightness values of pixels based on the image histogram.

Linear gray level transformation - A linear transformation of an image is a function that maps each pixel gray-level value into another gray-level at the same position according to a linear function.

Non linear gray level transformation – A non linear transformation maps small equal intervals into non equal intervals.

Q5. Describe local neighbourhood operations and median filtering?

Ans:- Pixels are modified based on some function of the pixels in their neighborhood.
Types of local operations are:

- **Spatial filtering**
  Spatial filtering involves passing a weighted mask or kernel over the image and replacing the original image pixel value corresponding to the center of the kernel with the sum of the original pixel values in the region corresponding to the kernel multiplied by the kernel weights.

- **Linear filtering**
  Each pixel in the input image is replaced by a linear combination of intensities of neighboring pixels.

- **Mean filter or averaging filter**
  It replaces each pixel of with the average of all the values in the local neighborhood.

**Median filter:**
Median filters are statistical non-linear filters that are often described in the spatial domain. A median smoothes the image by utilizing the median of the neighborhood. The concept of a median filter was by Tukey in 1977. Its extension to two-dimensional images was discussed by Pratt in 1978. It performs the following tasks to each pixel value in the processed image

1) All pixels in the neighborhood of the pixel in the original image which are identified by the mask sorted in the ascending (or) descending order.

2. The median of the sorted value is computed and is chosen as the pixel value for the processed image.

**Multiple Choice Questions**

1. Which of the following makes an image difficult to enhance?
a) Narrow range of intensity levels
b) Dynamic range of intensity levels
c) High noise
d) All of the mentioned

2. Which of the following is a second-order derivative operator?
   a) Histogram
   b) Laplacian
c) Gaussian
d) None of the mentioned

3. Response of the gradient to noise and fine detail ____________ the Laplacian’s
   a) is equal to
   b) is lower than
c) is greater than
d) has no relation with

4. Dark characteristics in an image are better solved using _____________
   a) Laplacian Transform
   b) Gaussian Transform
c) Histogram Specification
d) Power-law Transformation

5. What is the smallest possible value of a gradient image?
   a) e
6. Which of the following fails to work on dark intensity distributions?

a) Laplacian Transform
b) Gaussian Transform
c) Histogram Equalization
d) Power-law Transformation

7. ______________ is used to detect diseases such as bone infection and tumors.

a) MRI Scan
b) PET Scan
c) Nuclear Whole Body Scan
d) X-Ray

8. How do you bring out more of the skeletal detail from a Nuclear Whole Body Bone Scan?

a) Sharpening
b) Enhancing
c) Transformation
d) None of the mentioned

9. An alternate approach to median filtering is ______________

a) Use a mask
b) Gaussian filter

c) Sharpening

d) Laplacian filter

10. Final step of enhancement lies in _____________ of the sharpened image.

a) Increase range of contrast

b) Increase range of brightness

c) Increase dynamic range

d) None of the mentioned.

Fill in the blanks

1. In a dark image, the components of histogram are concentrated on -------- side of the grey scale.

2. Increasing the brightness of an image is a ----------- operation.

3. Multiplying the pixels with a constant “k” results in ----------- .

4. Histogram Equalisation also called as --------------.

5. To set the average value of an image zero, -------- would be set 0 in the frequency domain and the inverse transformation is done, where F(u, v) is Fourier transformed function of f(x, y)?

6. --------- filter that is used to turn the average value of a processed image zero

7. --------- attenuates high frequency while passing low frequencies of an image

8. A spatial domain filter of the corresponding filter in frequency domain can be obtained by applying ----------- operation on filter in frequency domain.

9. ----------- filtering is done in frequency domain in correspondence to lowpass filtering in spatial domain?
10. A Butterworth filter of order ---- has no ringing?

<table>
<thead>
<tr>
<th>MCQ's</th>
<th>Fill in the blanks</th>
</tr>
</thead>
<tbody>
<tr>
<td>11. D</td>
<td>1. Lower</td>
</tr>
<tr>
<td>13. B</td>
<td>3. Contrast stretching</td>
</tr>
<tr>
<td>15. C</td>
<td>5. F(0, 0)</td>
</tr>
<tr>
<td>17. C</td>
<td>7. Low pass filter</td>
</tr>
<tr>
<td>18. A</td>
<td>8. Inverse fourier transform</td>
</tr>
<tr>
<td>19. A</td>
<td>9. Gaussian</td>
</tr>
<tr>
<td>20. C</td>
<td>10. 1</td>
</tr>
</tbody>
</table>
UNIT III - IMAGE RESTORATION

Two marks question answers

Q1. What is meant by Image Restoration?
Ans:- Restoration attempts to reconstruct or recover an image that has been degraded by using a clear knowledge of the degrading phenomenon.

Q2. What are the two properties in Linear Operator?
Ans:- Additivity and homogeneity are the two properties of linear operator.

Q3. Explain additivity property in Linear Operator?
Ans:- The additive property says that if H is the linear operator, the response to a sum of two is equal to the sum of the two responses.

\[ H[f_1(x,y)+f_2(x,y)] = H[f_1(x,y)] + H[f_2(x,y)] \]

Q4. How a degradation process is modeled?
Ans:- A system operator H, which together with an additive white noise term \( n(x,y) \) a operates on an input image \( f(x,y) \) to produce a degraded image \( g(x,y) \).

Q5. What is meant by Noise probability density function?
Ans:- The spatial noise descriptor is the statistical behavior of gray level values in the noise component of the model.
Q1. Why the restoration is called as unconstrained restoration?

Ans:- In the absence of any knowledge about the noise ‘n’, a meaningful criterion function is to seek an \( f^\wedge \) such that \( H \cdot f^\wedge \) approximates of in a least square sense by assuming the noise term is as small as possible. Where \( H \) = system operator. \( f^\wedge \) = estimated input image. \( g \) = degraded image.

Q2. Which is the most frequent method to overcome the difficulty to formulate the spatial relocation of pixels?

Ans:- The point is the most frequent method, which are subsets of pixels whose location in the input (distorted) and output (corrected) imaged is known precisely.

Q3. What are the three methods of estimating the degradation function?

Ans:- The three methods of estimating the estimating the degradation function are:

a. Observation  
b. Experimentation  
c. Mathematical modeling

Q4. What is inverse filtering?

Ans:- The simplest approach to restoration is direct inverse filtering, an estimate \( F^\wedge(u,v) \) of the transform of the original image simply by dividing the transform of the degraded image \( G^\wedge(u,v) \) by the degradation function.

\[
F^\wedge(u,v) = G^\wedge(u,v)/H(u,v).
\]

Q5. Give the difference between Enhancement and Restoration?

Ans:- Enhancement technique is based primarily on the pleasing aspects it might present to the viewer. For example: Contrast Stretching, where as Removal of image blur by applying a deblurring function is considered a restoration technique.
Five marks question answers

Q1. Explain about Wiener filter used for image restoration?

Ans:- The inverse filtering approach makes no explicit provision for handling noise. This approach incorporates both the degradation function and statistical characteristics of noise into the restoration process. The method is founded on considering images and noise as random processes, and the objective is to find an estimate \( f \) of the uncorrupted image \( f \) such that the mean square error between them is minimized. This error measure is given by

\[
e^2 = E \{ (f - f)^2 \}
\]

where \( E \{ \cdot \} \) is the expected value of the argument.

It is assumed that the noise and the image are uncorrelated; that one or the other has zero mean; and that the gray levels in the estimate are a linear function of the levels in the degraded image. Based on these conditions, the minimum of the error function is given in the frequency domain by the expression,

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

where the fact that the product of a complex quantity with its conjugate is equal to the magnitude of the complex quantity squared, is used. This result is known as the Wiener filter, after N. Wiener [1942], who first proposed the concept in the year shown. The filter, which consists of the terms inside the brackets, also is commonly referred to as the minimum mean square error filter or the least square error filter. The Wiener filter does not have the same problem as the inverse filter with zeros in the degradation function, unless both \( H(u, v) \) and \( S_n(u, v) \) are zero for the same value(s) of \( u \) and \( v \).

The terms in above equation are as follows:

\( H(u, v) \) = degradation function

\( H^*(u, v) \) = complex conjugate of \( H(u, v) \)

\[
|H(u, v)|^2 = H^*(u, v)H(u, v)
\]

\( S_n(u, v) \) = power spectrum of the noise

\( S_f(u, v) \) = power spectrum of the non degraded image.

As before, \( H(u, v) \) is the transform of the degradation function and \( G(u, v) \) is the transform of the degraded image. The restored image in the spatial domain is given
by the inverse Fourier transform of the frequency-domain estimate \( F(u, v) \). Note that if the noise is zero, then the noise power spectrum vanishes and the Wiener filter reduces to the inverse filter. When we are dealing with spectrally white noise, the spectrum \( |N(u, v)|^2 \) is a constant, which simplifies things considerably. However, the power spectrum of the undergraded image seldom is known. An approach used frequently when these quantities are not known or cannot be estimated is to approximate the equation as

\[
\hat{F}(u, v) = \left[ \frac{1}{|H(u, v)|^2 + K} \right] G(u, v)
\]

where \( K \) is a specified constant.

**Q2. Explain a Model of the Image Degradation/Restoration Process?**

Ans:- The Fig. 1 shows, the degradation process is modeled as a degradation function that, together with an additive noise term, operates on an input image \( f(x, y) \) to produce a degraded image \( g(x, y) \). Given \( g(x, y) \), some knowledge about the degradation function \( H \), and some knowledge about the additive noise term \( \eta(x, y) \), the objective of restoration is to obtain an estimate \( \hat{f}(x, y) \) of the original image. the estimate should be as close as possible to the original input image and, in general, the more we know about \( H \) and \( \eta \), the closer \( \hat{f}(x, y) \) will be to \( f(x, y) \). The degraded image is given in the spatial domain by

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

where \( h(x, y) \) is the spatial representation of the degradation function and, the symbol * indicates convolution.

Convolution in the spatial domain is equal to multiplication in the frequency domain, hence

\[
G(u, v) = H(u, v) F(u, v) + N(u, v)
\]

where the terms in capital letters are the Fourier transforms of the corresponding terms in above equation.
Q3. Explain about the restoration filters used when the image degradation is due to noise only?

Ans:- If the degradation present in an image is only due to noise, then, \( g(x, y) = f(x, y) + \eta(x, y) \). The restoration filters used in this case are,

1. Mean filters
2. Order static filters and
3. Adaptive filters

Explain Mean filters.

There are four types of mean filters. They are

(i) Arithmetic mean filter: This is the simplest of the mean filters. Let \( S_{xy} \) represent the set of coordinates in a rectangular subimage window of size \( m \times n \), centered at point \( (x, y) \). The arithmetic mean filtering process computes the average value of the corrupted image \( g(x, y) \) in the area defined by \( S_{xy} \). The value of the restored image \( \hat{f}(x, y) \) at any point \( (x, y) \) is simply the arithmetic mean computed using the pixels in the region defined by \( S_{xy} \). In other words,

\[
\hat{f}(x, y) = \frac{1}{mn} \sum_{(s,t) \in S_{xy}} g(s, t).
\]
This operation can be implemented using a convolution mask in which all coefficients have value $1/mn$

(ii) Geometric mean filter: An image restored using a geometric mean filter is given by the expression.

$$f(x, y) = \left[ \prod_{(s,t) \in S_{xy}} g(s, t) \right]^\frac{1}{mn}.$$ 

Here, each restored pixel is given by the product of the pixels in the sub image window, raised to the power $1/mn$. A geometric mean filter achieves smoothing comparable to the arithmetic mean filter, but it tends to lose less image detail in the process.

(iii) Harmonic mean filter: The harmonic mean filtering operation is given by the expression

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

The harmonic mean filter works well for salt noise, but fails for pepper noise. It does well also with other types of noise like Gaussian noise.

(iv) Contra harmonic mean filter: The contra harmonic mean filtering operation yields a restored image based on the expression

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

where $Q$ is called the order of the filter. This filter is well suited for reducing or virtually eliminating the effects of salt-and-pepper noise. For positive values of $Q$, the filter eliminates pepper noise. For negative values of $Q$ it eliminates salt noise. It cannot do both simultaneously. Note that the contra harmonic filter reduces to the arithmetic mean filter if $Q = 0$, and to the harmonic mean filter if $Q = -1$. 
Q4. Write short notes on inverse filtering?

Ans:- The purpose of image restoration is to estimate or recover the scene without image degradation or distortion caused by non-ideal image system (e.g. the optics of the camera system). Inverse filtering is one of the techniques used for image restoration to obtain a recovered image $f_1(x,y)$ from the image data $g(x,y)$ so that $f_1(x,y) = f(x,y)$ in the ideal situation.

$$n(x,y) = 0 \quad \text{and} \quad h(x,y) * h'(x,y) = \delta(x,y)$$

Or

$$H(\omega_x, \omega_y)H'(\omega_x, \omega_y) = 1$$

The inverse filtering can be depicted in the block diagram format as:

![Inverse Filtering Block Diagram]

Q5. Enumerate the differences between the image enhancement and image restoration?

Ans:-

(i) Image enhancement techniques are heuristic procedures designed to manipulate an image in order to take advantage of the psychophysical aspects of the human system. Whereas image restoration techniques are basically reconstruction techniques by which a degraded image is reconstructed by using some of the prior knowledge of the degradation phenomenon.

(ii) Image enhancement can be implemented by spatial and frequency domain technique, whereas image restoration can be implement by frequency domain and algebraic techniques.

(iii) The computational complexity for image enhancement is relatively less when compared to the computational complexity for image restoration, since algebraic methods requires manipulation of large number of simultaneous equation. But, under some condition computational complexity can be reduced to the same level as that required by traditional frequency domain technique.

(iv) (Image enhancement techniques are problem oriented, whereas image restoration techniques are general and are oriented towards modeling the
degradation and applying the reverse process in order to reconstruct the original image.

(v) Masks are used in spatial domain methods for image enhancement, whereas masks are not used for image restoration techniques.

(vi) Contrast stretching is considered as image enhancement technique because it is based on the pleasing aspects of the review, whereas removal of image blur by applying a deblurring function is considered as a image restoration technique.
1. Degraded image is produced using degradation process and
   
   A. additive noise  
   B. destruction  
   C. pixels  
   D. coordinates  

2. Degraded image is given in a
   
   A. frequency domain  
   B. time domain  
   C. spatial domain  
   D. plane  

3. Purpose of restoration is to gain
   
   A. degraded image  
   B. original image  
   C. pixels  
   D. coordinates  

4. In wiener filtering it is assumed that noise and image are
   
   A. different  
   B. homogenous  
   C. correlated  
   D. uncorrelated  

5. Contraharmonic mean filter produces
   
   A. degraded image  
   B. original image  
   C. restored image  
   D. plane  

6. Mean filters reduce noise using
A. sharpening  
B. blurring  
C. restoration  
D. acquisition  

7. Approach to restoration is  
A. inverse filtering  
B. spike filtering  
C. black filtering  
D. ranking  

8. Approach that incorporates both degradation function and statistical noise in restoration is called  
A. inverse filtering  
B. spike filtering  
C. wiener filtering  
D. ranking  

9. Images usually gets corrupted during  
A. transmission  
B. degradation  
C. restoration  
D. acquisition  

10. Minimum mean square error filter is also called  
A. square error filter  
B. most square error filter  
C. least square error filter  
D. error filter  

FILL IN THE BLANKS
1) High frequency components are passed by \_________.  
2) Frequencies in predefined neighborhood are rejected by \_________.  
3) Principle sources of noise arise during image \_________.  
4) In \_________ filtering it is assumed that noise and image are uncorrelated.  
5) PSF stands for \_________.  
6) Contra-harmonic mean filter produces \_________.  
7) Restoration can not be done using \_________ projection.  
8) Degraded image is produced using \_________ function and additive noise.  
9) Spatial filtering is used in presence of \_________ random noise  
10) Salt and pepper noise also referred to term \_________.

<table>
<thead>
<tr>
<th>Solutions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MCQ’s</strong></td>
</tr>
<tr>
<td>1. A</td>
</tr>
<tr>
<td>2. C</td>
</tr>
<tr>
<td>3. B</td>
</tr>
<tr>
<td>4. D</td>
</tr>
<tr>
<td>5. C</td>
</tr>
<tr>
<td>6. B</td>
</tr>
<tr>
<td>7. A</td>
</tr>
<tr>
<td>8. C</td>
</tr>
<tr>
<td>9. A</td>
</tr>
<tr>
<td>10. C</td>
</tr>
</tbody>
</table>

**UNIT IV**
Two marks question answers

Q1. What is segmentation?
Ans: - Segmentation subdivides on image in to its constitute regions or objects. The level to which the subdivides is carried depends on the problem being solved. That is segmentation should when the objects of interest in application have been isolated.

Q2. Write the applications of segmentation.
Ans: - Detection of isolated points.
Detection of lines and edges in an image.

Q3. What are the three types of discontinuity in digital image?
Ans: - The three types of discontinuities are points, lines and edges.

Q4. How the derivatives are obtained in edge detection during formulation?
Ans: - The first derivative at any point in an image is obtained by using the magnitude of the gradient at that point. Similarly the second derivatives are obtained by using the laplacian.

Q5. Write about linking edge points?
Ans: - The approach for linking edge points is to analyze the characteristics of pixels in a small neighborhood (3x3 or 5x5) about every point (x,y) in an image that has undergone edge detection. All points that are similar are linked, forming a boundary of pixels that share some common properties.

Three marks question answers
Q1. What are the two properties used for establishing similarity of edge pixels?

Ans:- (1) The strength of the response of the gradient operator used to produce the edge pixel.
(2) The direction of the gradient.

Q2. What is edge?

Ans:- An edge is a set of connected pixels that lie on the boundary between two regions. Edges are more closely modeled as having a ramp-like profile. The slope of the ramp is inversely proportional to the degree of blurring in the edge.

Q3. Give the properties of the second derivative around an edge?

Ans:- The sign of the second derivative can be used to determine whether an edge pixel lies on the dark or light side of an edge.

It produces two values for every edge in an image.

An imaginary straightline joining the extreme positive and negative values of the second derivative would cross zero near the midpoint of the edge.

Q4. Define Gradient Operator?

Ans:- First order derivatives of a digital image are based on various approximation of the 2-D gradient. The gradient of an image f(x,y) at location (x,y) is defined as the vector

\[ \Delta f = \nabla f = [G_x, G_y] \]

\[ \| \Delta f \| = \sqrt{G_x^2 + G_y^2} \]

\[ \varphi(x,y) = \tan^{-1}(G_y/G_x) \]

\( \varphi(x,y) \) is the direction angle of vector \( \Delta f \)

Q5. What is meant by object point and background point?
Ans: To execute the objects from the background is to select a threshold $T$ that separate these modes. Then any point $(x,y)$ for which $f(x,y)>T$ is called an object point. Otherwise the point is called background point.

**Five marks question answers**

**Q1. Write short notes on image segmentation?**

Ans: Segmentation subdivides on image into its constitute regions or objects. The level to which the subdivides is carried depends on the problem being solved. Examples: In autonomous air to ground target acquisition applications identifying vehicles on a road is of interest. The first step is to segment the road from the image and then to segment the elements of the road down to objects of a range of sizes that correspond potential vehicles. In target acquisition, the system designer has control of the environment. So the usual approach is to focus on selecting the types of sensors most likely to enhance the objects of interest. Example is the use of infrared imaging to detect objects with a strong heat signature, such as tanks in motion. Segmentation algorithms for monochrome images are based on one of the two basic properties of gray level values. They are discontinuity and similarity. Based on the first category, the approach is based on abrupt changes in gray level and the areas of interest based on this category are detection of isolated points and detection of lines and edges in an image. Based on the second category the approach is based on thresholding, region growing and region splitting and merging. The concept of segmenting an image based on discontinuity or similarity of the gray level values of its pixels is applicable to both static and dynamic images.

**Q2. Write short notes on edge detection?**

Ans: Edge Detection:

Edge detection is “local” image processing methods designed to detect edge pixels. Concept that is based on a measure of intensity-level discontinuity at a point. It is possible to link edge points into edge segments, and sometimes these segments are linked in such a way that they correspond to boundaries, but this is not always the case.

**The image gradient and its properties:**
The tool of choice for finding edge strength and direction at location \((x,y)\) of an image, \(f\), is the gradient, denoted by \(\nabla f\), and defined as the vector \(\nabla f = \frac{\partial f}{\partial x}\)

\[ \nabla f = \text{grad}(f) = gy = \frac{\partial f}{\partial y} \]

The magnitude length of vector \(\nabla f\), denoted as \(M(x,y)\)

\[ M(x,y) = \text{mag}(\nabla f) = \sqrt{g_x^2 + g_y^2} \]

Is the value of the rate of change in the direction of the gradient vector.

The direction of the gradient vector is given by the angle

\[ \theta(x,y) = \tan^{-1} \frac{g_y}{g_x} \]

measured with respect to the \(x\)-axis. Follows, using these differences as our estimates of the partials, that \(\frac{\partial f}{\partial x} = -2\) and \(\frac{\partial f}{\partial y} = 2\) at the point in equation. Then

\[ g_x \frac{\partial f}{\partial x} = -2 \]

\[ \nabla f = g_y \frac{\partial f}{\partial y} = 2 \]

from which we obtain \(M(x,y) = 2\sqrt{2}\) at that point.

**Gradient operators:**

Obtaining the gradient of an image requires computing the partial derivatives \(\frac{\partial f}{\partial x}\) and \(\frac{\partial f}{\partial y}\) at every pixel location in the image

\[ g_x = \frac{\partial f(x,y)}{\partial x} = f(x+1,y) - f(x,y) \]

\[ g_y = \frac{\partial f(x,y)}{\partial y} = f(x,y+1) - f(x,y) \]

An approach used frequently is to approximate the gradient by absolute value:

\[ \nabla f = |G_x| + |G_y| \]
The Laplacian

The laplacian of a 2-D function \( f(x,y) \) is a second order derivatives defined as

\[
\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}
\]

The first laplacian is combined with smoothing as a precursor to finding edges via zero crossings. Consider the function.

\[
\nabla^2 f = 8z_5 - (z_1 + z_2 + z_3 + z_4 + z_6 + z_7 + z_8 + z_9)
\]

0 -1 0
-1 4 -1
0 -1 0

Q3. Write short notes on edge linking by local processing?

Ans:- One of the simplest approaches for linking edge points is to analyze the characteristics of the pixels in a small neighborhood about every point in an image that has undergone edge detection. Two properties used for establishing similarity of edge pixels in the analysis are the strength of the response of the gradient operator used to produce the edge pixel, The direction of the gradient. The first property is given by the value of \( \nabla f \). Thus an edge pixel with coordinates \((x', y')\) and in the predefined neighborhood of \((x, y)\) is similar in magnitude to the pixel at \((x, y)\) if

\[
|\nabla f(x, y) - \nabla (x', y')| < T
\]

where \( T \) is a nonnegative threshold. The direction of the gradient vector is given by

\[
\alpha(x, y) = \tan^{-1} \frac{gy}{gx}
\]

Then an edge pixel at \((x', y')\) in the predefined neighborhood of \((x, y)\) has an angle similar to the pixel at \((x, y)\) if \( |\alpha(x, y) - \alpha(x', y')| < A \) where \( A \) is an angle threshold. Note that the direction of the edge at \((x, y)\) in reality is perpendicular to the direction of the gradient vector at that point. A point in the predefined neighborhood of \((x, y)\) is linked to the pixel at \((x', y')\) if both magnitude and direction criteria are satisfied. This process is repeated for every location in the image. A record must be kept of linked points as the center of the neighborhood is moved from pixel to pixel.
pixel. A simple bookkeeping procedure is to assign a different graylevel to each set of linked edge pixels.

Q4. Discuss region oriented segmentation in detail?
Ans: - The objective of segmentation is to partition an image into regions. We approached this problem by finding boundaries between based on discontinuities in gray levels, segmentation was accomplished via thresholds based on the distribution of pixels properties, such as gray level values or color.

**Basic Formulation:**

Let \( R \) represent the region of the image. We may view segmentation as a process that partitions \( R \) into \( n \) subregions, \( R_1, R_2, \ldots \), such that \( n \)

(a) \( \sum_{i=1}^{n} R_i = R \)

(b) \( R_i \) is a connected region, \( i=1,2,\ldots,n \).

(c) \( R_i \cap R_j = \emptyset \) for all \( i \) and \( j \), \( i \neq j \).

(d) \( P(R_i) = \text{TRUE} \) for \( i=1,2,\ldots,n \).

(e) \( P(R_i \cup R_j) = \text{FALSE} \) for \( i \neq j \).

Here, \( P(R_i) \) is a logical predicate defined over the points in set \( R_i \) and \( \Phi \) is the null set. Condition (a) indicates that the segmentation must be complete i.e., every pixel must be in a region.

Condition (b) requires that points in a region must be connected in some predefined sense.

Condition (c) indicates that the regions must be disjoint.

Condition (d) deals with the properties that must be satisfied by the pixels in a segmented region.

**Region Growing:**
As its name implies region growing is a procedure that groups pixel or subregions into larger regions based on predefined criteria. The basic approach is to start with a set of “seed” points and from these grow regions.

- If the result of these computation shows clusters of values, the pixels whose properties place them near the centroid of these clusters can be used as seeds.
- Descriptors alone can yield misleading results if connectivity or adjacency information is not used in the region growing process.

**Region Splitting and Merging:**

The procedure just discussed grows regions from a set of seed points. An alternative into subdivided an image initially into a set of arbitrary, disjointed regions and then merge and/or split the regions in an attempt to satisfy the conditions.

1. Split into four disjoint quadrants any region Ri for which P(Ri)=FALSE.
2. Merge any adjacent regions Rj and Rk for which P(RjURk)=TRUE.
3. Stop when no further merging or splitting is possible.

Mean and standard deviation of pixels in a region to quantify the texture of region.

**Role of thresholding:**

We introduced a simple model in which an image f(x,y) is formed as the product of a reflectance component r(x,y) and an illumination components i(x,y).
consider the computer generated reflectance function.

☐ The histogram of this function is clearly bimodal and could be portioned easily by placing a single global threshold, \( T \), in the histogram valley.

☐ Multiplying the reflectance function by the illumination function.

☐ Original valley was virtually eliminated, making segmentation by a single threshold an impossible task.

☐ Although we seldom have the reflectance function by itself to work with, this simple illustration shows that the reflective nature of objects and background can be such that they are separable.

\[
f(x,y) = i(x,y)r(x,y)
\]

Taking the natural logarithm of this equation yields a sum:

\[
z(x,y) = \ln f(x,y)
\]

\[
= \ln i(x,y) + \ln r(x,y)
\]

\[
= i(x,y) + r(x,y)
\]

☐ If \( i(x,y) \) and \( r(x,y) \) are independent random variables, the histogram of \( z(x,y) \) is given by the convolution of the histogram of \( i(x,y) \) and \( r(x,y) \).

☐ But if \( i(x,y) \) had a border histogram the convolution process would smear the histogram of \( r(x,y) \), yielding a histogram for \( z(x,y) \) whose shape could be quite different from that of the histogram of \( r(x,y) \).

☐ The degree of distortion depends on the broadness of the histogram of \( i(x,y) \),
which in turn depends on the nonuniformity of the illumination function.

☐ We have dealt with the logarithm of \( f(x,y) \), instead of dealing with the image function directly.

☐ When access to the illumination source is available, a solution frequently used in practice to compensate for nonuniformity is to project the illumination pattern onto a constant, white reflective surface.

☐ This yields an image \( g(x,y) = ki(x,y) \), where \( k \) is a constant that depends on the surface and \( i(x,y) \) is the illumination pattern.

☐ For any image \( f(x,y) = i(x,y)r(x,y) \) obtained from the same illumination function, simply dividing \( f(x,y) \) by \( g(x,y) \) yields a normalized function

\[
    h(x,y) = f(x,y)/g(x,y) = r(x,y)/k.
\]

☐ Thus, if \( r(x,y) \) can be segmented by using a single threshold \( T \), then \( h(x,y) \) can be segmented by using single threshold of value \( T/k \).

Q5. Describe the Hit-or-Miss transform?
Ans:- The Hit-or-Miss transform is used for detecting shapes. It uses two structuring elements.
1. The first one contains the foreground shape of the object which is to be detected.
2. The second structuring element contains the background shape around the object which is to be detected. It is like a window frame (a thin strip of background) to the foreground in the first structuring element. The background pixels in this mask are marked with foreground intensity and the object pixels with background intensity.
At any point on the given image,
if the foreground matches with the first structuring element AND if the complement (i.e. the background) matches with the second structuring element, then we can say that the object shape exists at that point.
The set of points where a structuring element fits can be identified by eroding the given image with the structuring element.
Figure: The set $M$ is the union of sets $A, B, C, D, E, F, G$. The object we want to detect is $D$. Hence we use a structuring element which is same as $D$. The element $W$ has been chosen slightly larger than $D$ so that $W - D$ gives a thin strip of background surrounding set $D$.

Mathematical formulation

Given a shape $A$ with 3 components shapes $C, D, E$. What is the location of $D$?

- Step 1: Detect where shape $D$ can be. We might detect larger shapes in which $D$ is contained.
- Step 2: Verify if background (fitting $D$) is present exactly around this shape.

Mathematically Hit-or-Miss transform is:

$$A \ominus B = (A \ominus D) \cap [A^c \ominus (W - D)]$$

$$A \ominus B = (A \ominus D) - \left[A \oplus (W - D)\right]$$

Several features can be extracted from the image using the five basic morphological operations of dilation, erosion, opening, closing, and hit-or-miss transform.

Multiple choice questions
1.  ----------- should stop when the objects of interest in an application have been isolated.
   a)  Segmentation
   b)  fragment
   c)  addition
   d)  none of above
2.  ----------- means dividing the image into parts for convenience of analysis.
   a)  Morphological image processing
   b)  Image Compression
   c)  Image Enhancement
   d)  Image Segmentation
3.  Combining all the edges is known as:
   a)  Edge linking
   b)  Edge joining
   c)  Edge merging
   d)  None of the above
4.  Setting up a value for separating entities based upon greater than or less than criteria is known as:
   a)  Linking
   b)  Thresholding
   c)  Digitizing
   d)  Sampling
5.  Removing the corners of the image using a structural element is:
   a)  Erosion
   b)  Dilation
   c)  Opening
   d)  Closing
6.  Increasing the corners of the image using a structural element is:
7. What does this diagram symbolize?

![Quadtrees of subregions]

a) Region merging
b) Region splitting
c) None of the above

8. What does this diagram symbolize?

![Nested quadrants]

a) Edge linking  
b) Edge joining  
c) Edge merging  
d) None of the above
9. "---------" means dividing the image for object detection.
   a) Image Compression
   b) Image Enhancement
   c) Morphological image processing
   d) Image Segmentation

10. The bottom-up region growing algorithm starts from a set of "------" pixels defined by the user and sequentially adds a pixel to a region.
   a) Edge
   b) Seed
   c) Point
   d) None of the above

Fill in the blanks:

1. "--------" provides a set of tools for extracting image features or object attributes like connected components, boundaries, skeletons, convex hull etc. which are useful for describing regions.
2. The four basic morphological operations are "-----", dilation, opening, closing.
3. "----------" deals with Form and Structure of regions.
4. Erosion of a set "--------" it while dilation "--------" it.
5. "----------" can be viewed as a morphological "reconstruction" operation
6. "--------" is performed by applying erosion followed by dilation.
7. "--------" is performed by applying dilation followed by erosion.
8. Opening "--------" contours of an image.
9. The Hit-or-Miss transform is used for detecting "--------".
10. "--------" involves grouping of the edge pixels that show evidence of belonging to the same boundary segment.

Solutions
<table>
<thead>
<tr>
<th>MCQ’s</th>
<th>Fill in the blanks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>a 1. Morphology</td>
</tr>
<tr>
<td>2.</td>
<td>d 2. Erosion, dilation, opening, closing</td>
</tr>
<tr>
<td>3.</td>
<td>a 3. morphological image processing</td>
</tr>
<tr>
<td>4.</td>
<td>b 4. shrinks, expands</td>
</tr>
<tr>
<td>5.</td>
<td>b 5. Dilation</td>
</tr>
<tr>
<td>6.</td>
<td>a 6. Opening</td>
</tr>
<tr>
<td>7.</td>
<td>a 7. Closing</td>
</tr>
<tr>
<td>8.</td>
<td>c 8. smooths</td>
</tr>
<tr>
<td>9.</td>
<td>d 9. shapes</td>
</tr>
<tr>
<td>10.</td>
<td>b 10. Edge linking</td>
</tr>
</tbody>
</table>

**UNIT V**
Two marks question answers

Q1. What is image compression?
Ans:-Image compression refers to the process of redundancy amount of data required to represent the given quantity of information for digital image. The basis of reduction process is removal of redundant data.

Q2. What is Data Compression?
Ans:-Data compression requires the identification and extraction of source redundancy. In other words, data compression seeks to reduce the number of bits used to store or transmit information.

Q3. What are two main types of Data compression?
Ans:-Lossless compression can recover the exact original data after compression. It is used mainly for compressing database records, spreadsheets or word processing files, where exact replication of the original is essential. Lossy compression will result in a certain loss of accuracy in exchange for a substantial increase in compression. Lossy compression is more effective when used to compress graphic images and digitised voice where losses outside visual or aural perception can be tolerated.

Q4. What is the need for Compression?
Ans:-In terms of storage, the capacity of a storage device can be effectively increased with methods that compress a body of data on its way to a storage device and decompresses it when it is retrieved. In terms of communications, the bandwidth of a digital communication link can be effectively increased by compressing data at the sending end and decompressing data at the receiving end.

Q5. What are different Compression Methods?
Ans:-Run Length Encoding (RLE) Arithmetic coding Huffman coding and Transform coding

Three marks question answers
Q1. Define coding redundancy?
Ans:- If the gray level of an image is coded in a way that uses more code words than necessary to represent each gray level, then the resulting image is said to contain coding redundancy.

Q2. Define interpixel redundancy?
Ans:- The value of any given pixel can be predicted from the values of its neighbors. The information carried by is small. Therefore the visual contribution of a single pixel to an image is redundant. Otherwise called as spatial redundant geometric redundant or interpixel redundant. Eg: Run length coding

Q3. What is run length coding?
Ans:- Run-length Encoding, or RLE is a technique used to reduce the size of a repeating string of characters. This repeating string is called a run; typically RLE encodes a run of symbols into two bytes, a count and a symbol. RLE can compress any type of data regardless of its information content, but the content of data to be compressed affects the compression ratio. Compression is normally measured with the compression ratio:

Q4. Define psycho visual redundancy?
Ans:- In normal visual processing certain information has less importance than other information. So this information is said to be psycho visual redundant.

Q5. Define encoder?
Ans:- Source encoder is responsible for removing the coding and interpixel redundancy and psycho visual redundancy.

There are two components
A) Source Encoder B) Channel Encoder

Five marks question answers
Q1. Define Compression and explain data Redundancy in image compression?

Ans:- Compression: It is the process of reducing the size of the given data or an image. It will help us to reduce the storage space required to store an image or file.

Data Redundancy:

The data or words that either provide no relevant information or simply restate that which is already known. It is said to be data redundancy. Consider N1 and N2 number of information carrying units in two data sets that represent the same information

Data Redundancy Rd = 1-1/Cr

Where Cr is called the Compression Ratio.

`Cr=N1/N2.

Types of Redundancy

There are three basic Redundancy and they are classified as

1) Coding Redundancy
2) Interpixel Redundancy
3) Psychovisual Redundancy.

1. Coding Redundancy :

We developed this technique for image enhancement by histogram processing on the assumption that the grey levels of an image are random quantities. Here the grey level histogram of the image also can provide a great deal of insight in the construction of codes to reduce the amount of data used to represent it.

2. Interpixel Redundancy:

In order to reduce the interpixel redundancy in an image, the 2-D pixel array normally used for human viewing and interpretation must be transformed in to more efficient form.

3. Psychovisual Redundancy:
Certain information simply has less relative importance than other information in the normal visual processing. This information is called Psycovisual Redundant data.

Q2. Write about fidelity criteria.
Ans:- Fidelity criteria analyzes the quality of the image. The criteria for an assessment of a quality of an image are (i) Objective Fidelity Criteria and (ii) Subjective Fidelity Criteria.

i. **Objective Fidelity Criteria:**

Mean Square Error (MSE): Let \( f(x,y) \) represent an input image and let \( f^\prime(x,y) \) denote an estimate or approximate of \( f(x,y) \) for any value of \( x \) and \( y \),

\[
MSE = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x,y) - \hat{f}(x,y)]^2
\]

Signal to Noise Ratio (SNR):

\[
SNR = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x,y)]^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x,y) - \hat{f}(x,y)]^2}
\]

The mean squared error is defined as,

Signal to Noise Ratio (SNR):

ii. **Subjective Fidelity Criteria:**

- Images are viewed by human beings. Therefore measuring image quality by the subjective evaluations of a human observer is more appropriate. This can be accomplished by showing a typical decompressed image to an appropriate cross section of viewers and averaging their evaluations.
- The evaluations may be made by using an absolute rating scale or by means of side by side comparison of \( f(x,y) \) and \( f^\prime(x,y) \).
- Side by side comparisons can be done with the following scale: - \{1, 2, 3, 4, 5, 6\} to represent evaluations such as \{Excellent, Fine, Passable, Marginal, Inferior, Unusable\} respectively. - \{-3, -2, -1, 0, 1, 2, 3\} to represent subjective evaluations such as \{much worse, worse, slightly worse, the same, slightly better, better, much better\} respectively.

Q3. Define Compression and Explain the general compression system model?
**Ans:- Compression:** It is the process of reducing the size of the given data or an image. It will help us to reduce the storage space required to store an image or File.

**Image Compression Model:**

There are two Structural model and they are broadly Classified as follows

1. An Encoder
2. A Decoder.

An Input image $f(x,y)$ is fed in to encoder and create a set of symbols and after Transmission over the channel, the encoded representation is fed in to the decoder.

**A General Compression system model:**

The General system model consist of the following components, they are broadly classified as

1. Source Encoder
2. Channel Encoder
3. Channel
4. Channel Decoder
5. Source Decoder

The Source Encoder Will removes the input redundancies. The channel encoder will increase the noise immunity of the source encoder’s output. If the channel between encoder and decoder is noise free then the channel encoder and decoder can be omitted.
**MAPPER:** It transforms the input data in to a format designed to reduce the interpixel redundancy in the input image.

**QUANTIZER:**

It reduces the accuracy of the mapper’s output.

**SYMBOL ENCODER:**

It creates a fixed or variable length code to represent the quantizer’s output and maps the output in accordance with the code.

**SYMBOL DECODER:**

The inverse operation of the source encoder’s symbol will be performed and maps the blocks
Q4. Write about Lossy compression model?
Ans:- Generally most lossy compressors are three-step algorithms, each of which is in accordance with three kinds of redundancies. The block diagrams of the lossy compression model are as shown.

The first stage is a transform to eliminate the inter-pixel redundancy to pack information efficiently. Then a quantizer is applied to remove psycho-visual redundancy to represent the packed information with as few bits as possible. The quantized bits are then efficiently encoded to get more compression from the coding redundancy.

Lossy compression methods result in some loss of quality in the compressed images. It is a tradeoff between image distortion and the compression ratio. Some distortion measurements are often used to quantify the quality of the reconstructed image as well as the compression ratio (the ratio of the size of the original image to the size of the compressed image). The commonly used objective distortion measurements, which are derived from statistical terms, are the RMSE (root mean square error), the NMSE (normalized mean square error) and the PSNR (peak signal-to-noise ratio).

These measurements are defined as follows:
Where the images have \(N \times M\) pixels (8 bits per pixel), image, and \(f'(i, j)\) represents the reconstructed decompression.

Since the images are for human viewing, it leads to subjective measurements based on subjective comparisons to tell how “good” the decoded image looks to a human viewer. Sometimes, application quality can be used as a measure to classify the usefulness of the decoded image for a particular task such as clinical diagnosis in medical images and meteorological prediction in satellite images and so on.

When comparing two lossy coding methods, we may either compare the qualities of images reconstructed at a constant bit rate, or, equivalently, we may compare the bit rates used in two constructions with the same quality, if it is accomplishable.

**Q5. Describe Lossy and Lossy predictive coding.**

Ans:- For a lossy predictive encoder we have the quantizer component to quantize the error signal \(e(n)\) to produced the quantized signal \(\hat{e}(n)\). The block diagram for the mapper module of the lossy predictive encoder is shown in Fig 1.

\[
RMSE = \sqrt{\frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} [f(i, j) - f'(i, j)]^2}
\]

\[
NMSE = \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} [f(i, j) - f'(i, j)]^2}{\left[ \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} f(i, j) \right]^2}
\]

\[
PSNR = 20 \log_{10} \left( \frac{255}{RMSE} \right)
\]
A major difference between the lossy and the lossless predictive encoder is that for lossy encoder, the input to the predictor component is driven from the quantizer output $\hat{e}(n)$. The predictor output $\hat{f}(n)$ is summed up with the quantized error $\hat{e}(n)$ to give the input signal for the predictor.

Why can't we simply use the input signal $f(n)$ itself as the input to the predictor? This is because at the decoder module, the input for the predictor component will be constructed using the quantized error $\hat{e}(n)$. Hence it is better that the predictor at the encoder end also uses the same quantized error for making the predictions.

![Figure 1: Mapper module Lossy Predictive Encoder](image)
Figure 2: Inverse Mapper module for a Lossless Predictive Decoder

The mapper module of a lossy predictive decoder is shown in Fig. 12. It is exactly the same as the one for lossless decoder (Fig. 2), except that the input to the module is the quantized error $\hat{e}(n)$.

Figure 3: Inverse Mapper module for a Lossy Predictive Decoder.

1. **Describe lossless predictive coding.**

Lossless predictive coding predicts the value of each pixel by using the values of its neighboring pixels. Therefore, every pixel is encoded with a prediction error rather than its original value. Typically, the errors are much smaller compared with the original value so that fewer bits are required to store them.
DPCM (differential pulse code modulation) is a predictive coding based lossless image compression method. It is also the base for lossless JPEG compression. A variation of the lossless predictive coding is the adaptive prediction that splits the image into blocks and computes the prediction coefficients independently for each block to achieve high prediction performance. It can also be combined with other methods to get a hybrid coding algorithm with higher performance.

Predictive coding techniques try to remove the interpixel redundancy in image data. Since the neighboring pixels in an image tend to be highly correlated a predictive encoder works by encode only the new information in each pixel. This new information is usually defined as the difference (i.e. the error) between the actual and the predicted value of that pixel.

**Lossless Predictive Encoder**

![Figure 1: Mapper module for a Lossless Predictive Encoder](image)

The block diagram of a predictive encoder is shown in Fig. 1. The input to the encoder is the discrete signal \( f(n) \). The predictor component predicts the signal value to be \( \hat{f}(n) \). The output of the predictor is rounded to its nearest integer. Generally it is not possible
for the predictor to accurately predict the input signal. Some error (generally small if the prediction happens to be good) will be there. This is the prediction error $e(n)$

$$e(n) = f(n) - \hat{f}(n)$$

The mapper output can be subsequently quantized (for a lossy compression scheme) or can be directly fed to a symbol encoder. The predictor component shown in Fig 1 produces the predicted signal $\hat{f}(n)$. The input of the predictor has been tied to the input signal. Thus the predictor can use the past input samples to predict the next sample.

Multiple Choice Questions

1. Compressed image can be recovered back by

   a) image enhancement  
   b) image decompression  
   c) image contrast  
   d) image equalization

2. Image compression comprised of

   a) encoder  
   b) decoder  
   c) frames  
   d) Both A and B

3. If $P(E) = 1$, it means event

   a) does not occur  
   b) always occur  
   c) no probability  
   d) normalization

4. In image $MxN$, $N$ is

   a) rows  
   b) column  
   c) level  
   d) intensity

5. Histogram equalization refers to image
a) sampling  
b) quantization  
c) framing  
d) normalization  

6. Information lost when expressed mathematically is called  
a) markov  
b) finite memory source  
c) fidelity criteria  
d) noiseless theorem  

7. Transforming difference between adjacent pixels is called  
a) mapping  
b) image compression  
c) image watermarking  
d) image equalization  

8. Shannons theorem is also called  
a) noiseless coding theorem  
b) noisy coding theorem  
c) coding theorem  
d) noiseless theorem  

9. One that is not a type of data redundancy is  
a) coding  
b) spatial  
c) temporal  
d) facsimile  

10. Compression is done for saving  
a) storage  
b) bandwidth  
c) money  
d) Both A and B
Fill in the blanks:

1. -------- is used to reduce the data storage and bandwidth.
2. Certain information simply has less relative importance than other information in the normal visual processing. This information is called --------.
3. -------------- basically modifies the histogram of an input image so as to improve the visual quality of the image.
4. -------- reduces the accuracy of the mapper’s output.
5. -------- is a procedure that groups pixel or subregions into larger regions based on predefined criteria.
6. -------- compression methods result in some loss of quality in the compressed images.
7. -------- predicts the value of each pixel by using the values of its neighboring pixels.
8. -------- is optimal with its information packing properties.
9. -------- uses a reversible and linear transform to decorrelate the original image into a set of coefficients in transform domain.
10. -------- is an image compression standard and coding system.

<table>
<thead>
<tr>
<th>Solutions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MCQ’s</strong></td>
</tr>
<tr>
<td>11. b</td>
</tr>
<tr>
<td>12. d</td>
</tr>
<tr>
<td>13. b</td>
</tr>
<tr>
<td>14. b</td>
</tr>
<tr>
<td>15. d</td>
</tr>
<tr>
<td>16. c</td>
</tr>
<tr>
<td>17. a</td>
</tr>
<tr>
<td>18. d</td>
</tr>
<tr>
<td>19. d</td>
</tr>
<tr>
<td>20. d</td>
</tr>
</tbody>
</table>