VI Module

Digital Image Processing

1. Introduction

An image may be defined as a two dimensional function, $f(x,y)$, where x and y are spatial (plane) coordinates, and the amplitude of f at any pair of coordinates (x,y) is called the intensity or gray level of the image at that point.

Digital Image - When x, y, and the intensity values of f are all finite, discrete quantities, we call the image a digital image.

Digital Image Processing - The field of digital image processing refers to processing digital images by means of a digital computer.

Pixels - A digital image is composed of a finite number of elements, each of which has a particular location and value. These elements are called picture elements, image elements, pels, pixels.

Digitization - Digitization can be defined as the process of transforming image text, sound into digital data, that can we save, organize, retrieve and restore through electronic device.

Need for digital image processing

-> To improve the pictorial information for human interpretation

-> To process image data for storage, transmission and representation for autonomous machine perception.

Application of DIP

(i) gamma-ray imaging
(ii) X-ray imaging
(iii) imaging in the ultra-violet Band
(iv) imaging in the visible and infrared (IR) band
(v) Ultrasound imaging
(vi) Imaging in the radio band

2. Fundamental steps in image processing

There are two broad categories.

1) methods whose input and output are images
2) methods whose input may be images but whose outputs are attributes extracted from those images.

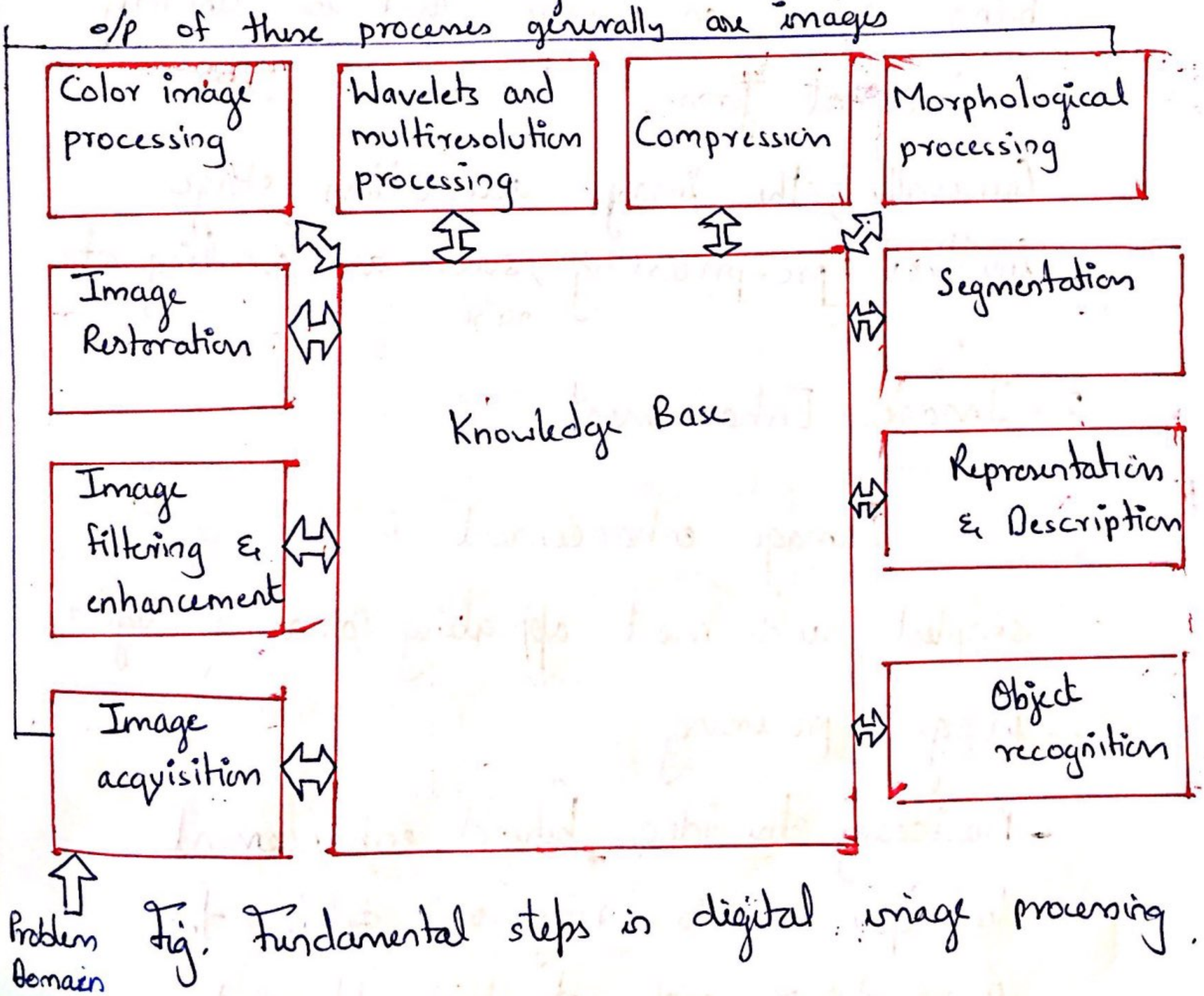


Fig. Fundamental steps in digital image processing.

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 pre-processing, such as scaling etc.
noise

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 and 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

Colour 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 modelling and processing in a digital domain etc.

5) Wavelets and Multi - Resolution 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 - 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.

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.

3: Digital Image Representation

Let an image f(x,y) is sampled and the resulting digital image has M rows and N columns. Now the values of the coordinates (x,y) become discrete

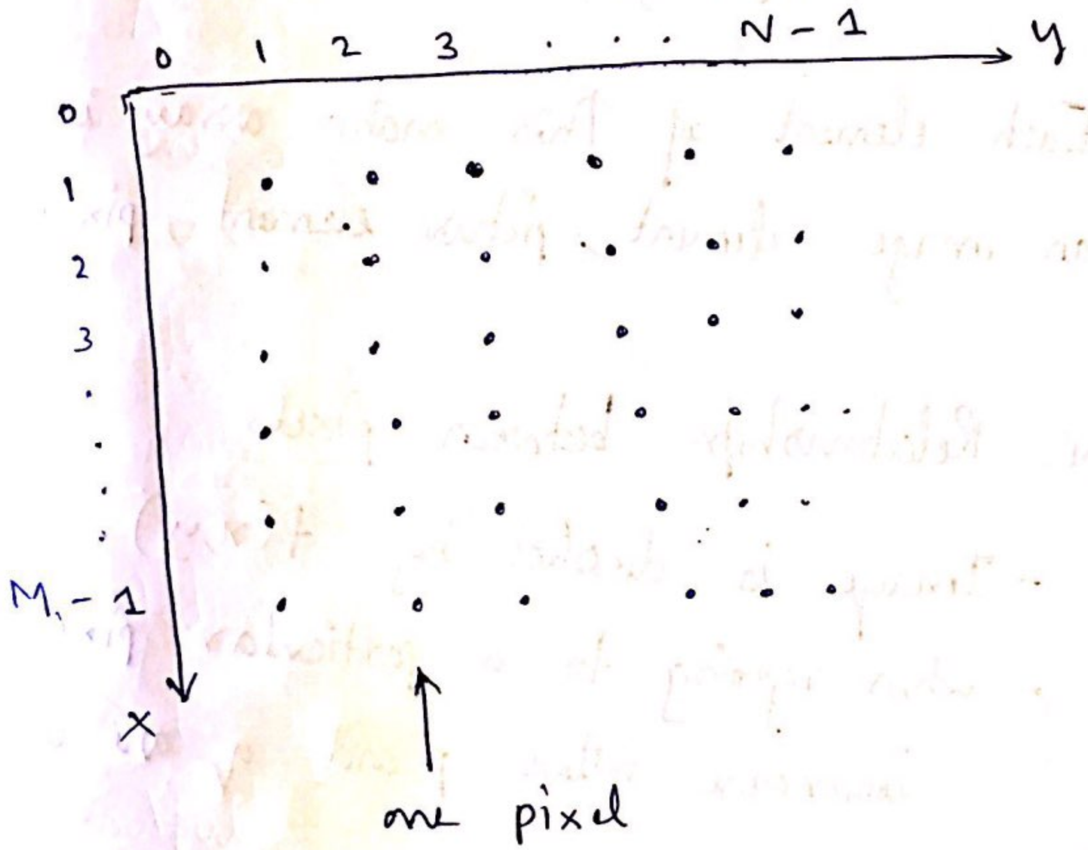


Fig. Coordinate convention to represent digital image is

The values of the coordinate at the origin (x,y) = (0,0)

The complete M×N digital image can be written in a matrix form as:

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix}$$

Each element of this matrix array is called an image element, picture element, pixel ;

4. <u>Basic Relationships between pixels</u>

- Image is denoted by $f(x,y)$.
- when referring to a particular pixel, lowercase letters p and q are used.

4.1 <u>Neighbours of a pixel</u>

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 4 - neighbours of p, denoted by $N_4(p)$.

- Each pixel is a unit distance from (x,y) and some of the neighbour locations 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 $N_D(p)$.

- These points, together with the 4-neighbours are called the 8-neighbours of p, denoted by $N_8(p)$.

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

4.2 Adjacency, Connectivity, Regions and Boundaries

4.2.1 Adjacency

Let V be the set of intensity values used to define adjacency.

- In a binary image, $V = \{1\}$ if we are referring to adjacency of pixels with value 1.
- In the adjacency of pixels with a range of possible intensity values 0 to 255, set V could be any subset of these 256 values.

We consider three types of adjacency

a) 4-adjacency : Two pixels p and q with values from V are 4-adjacent if q is in the set $N_4(p)$.

b) 8-adjacency : Two pixels p and q with values from V are 8-adjacent if q is in set $N_8(p)$.

c) m-adjacency (mixed adjacency) – Two pixels p and q with values from V are m-adjacent if ~~q is in the set $N_8(p)$~~.

(i) q is in $N_4(p)$ or

(ii) q is in $N_D(p)$ and the set $N_4(p) \cap N_4(q)$ has no pixels whose values are from V.

- m-adjacency is a modification of 8-adjacency.
- Its introduced to eliminate the ambiguities that often arise when 8-adjacency is used.

eg: Consider the pixel arrangement shown in figure (a) for $V = \{1\}$

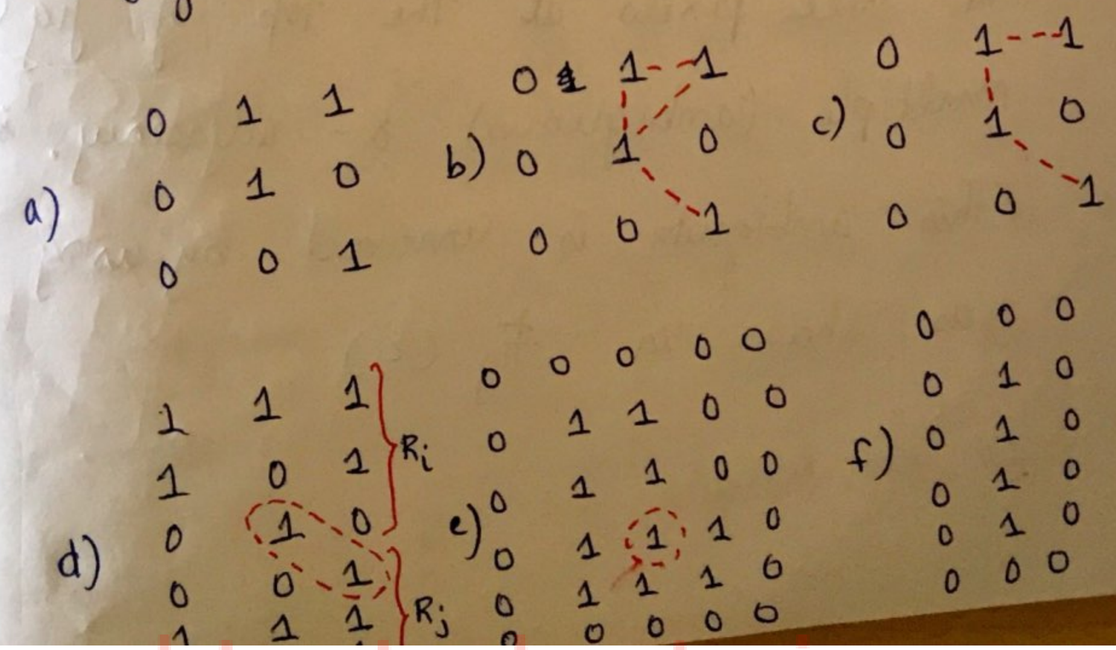


Fig. a) An arrangement of pixels
b) Pixels that are 8-adjacent (adjacency is shown by dashed lines; note the ambiguity
c) m-adjacency
d) Two regions (of 1s) that are adjacent if 8-adjacency is used.
e) The circled point is part of the boundary of the 1-valued pixels only if 8-adjacency between the region and background is used.
f) The inner boundary of the 1-valued does not form a closed path, but its outer boundary does.

The three pixels at the top of fig (b) show multiple (ambiguous) 8-adjacency. ~~is used~~.
This ambiguity is removed by using m-adjacency as shown in Fig (c).

4.2.2 Connectivity

Let S represent a subset of pixels in an image. Two pixels p and q are said to be connected in S if there exists a path between them consisting entirely of pixels in S.

- For any pixel p in S, the set of pixels that are connected to it in S is called a connected component of S.
- If it has only one connected component, then set S is called a connected set.

4.2.3 Regions

Let R be a subset of pixels in an image. We call R a region of the image if R is a connected set.

- Two regions R_i and R_j are said to be adjacent if their union forms a connected set.
- Regions that are not adjacent are said to be disjoint.

- We consider 4- and 8- adjacency when referring to regions.

For eg: Two regions (of 1s) in fig (d) are adjacent only if 8 - adjacency is used.

4.2.4 Boundary

The boundary of a region R is the set of points that are adjacent to points in the complement of R.

- Border of a region is the set of pixels in the region that have atleast one background neighbour.

- For example,

the point circled in fig (e) is not a member of the border of the 1 - valued region if 4 - connectivity is used between the region and its background.

- Inner border of the 1-valued region in Fig (f) is the region itself.
- Outer border is the corresponding border in the background.

4.3 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

a) $D(p,q) \geq 0$ $(D(p,q) = 0$ iff $p = q)$

b) $D(p,q) = D(q,p)$

c) $D(p,z) \leq D(p,q) + D(q,z)$

The Euclidean distance between p and q is defined as

$$D_e(p,q) = \left[(x-s)^2 + (y-t)^2\right]^{1/2}$$

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

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

- In this case, the pixels having a D_4 distance from (x,y) less than or equal to some value r form a diamond centered at (x,y).

- For example,

the pixels at a D_4 distance ≤ 2 from (x,y) form the following contours of constant distance:

```
        2
     2  1  2
  2  1  0  1  2
     2  1  2
        2
```

The pixels with $D_4 = 1$ are the 4-neighbours of (x,y)

- The D_8 distance (called the chessboard distance) between p and q is defined as

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

- In this case, the pixels with D_8 distance from (x,y) less than or equal to some value r

form a square centered at (x,y).

- For example,

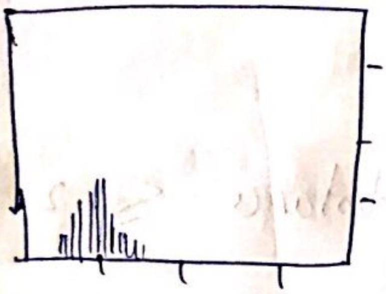
the pixels with D_8 distance $\leq$ 2. from (x,y) (the center point) form the following contours of constant distance

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

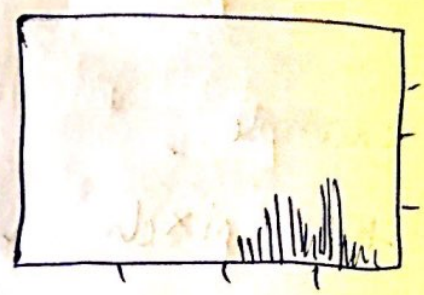
The pixels $D_8 = 1$ are the 8-neighbours of (x,y)

5. Gray Level Histogram

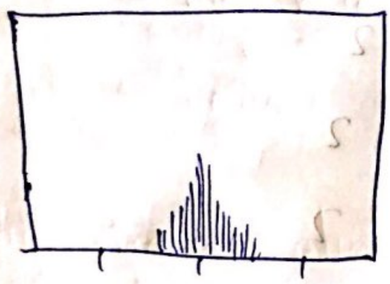
Histogram – The histogram of a digital image with intensity levels in the range [0, L−1] is a discrete function $h(r_k) = n_k$, where r_k is the k^{th} intensity value and n_k the no. of pixels in the image with intensity r_k.



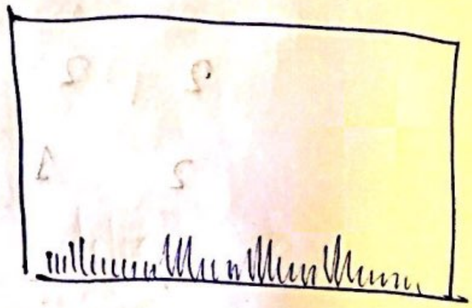
Histogram of a dark image



Histogram of light image



Histogram of low-contrast image



Histogram of high-contrast image

Fig. 4 basic histograms

Advantages of histograms

1. Histograms are the basis for numerous spatial domain processing techniques.
2. Histogram manipulation can be used for image enhancement
3. Histograms are simple to calculate in s/w
4. Popular tool for real-time image processing

Gray level transformation

There are three basic gray level transformation

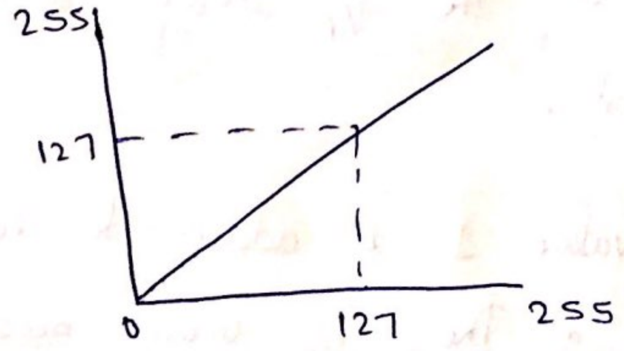
1. Linear
2. Logarithmic
3. Power - Law

1. Linear transformation

Linear transformation includes simple identity and negative transformation.

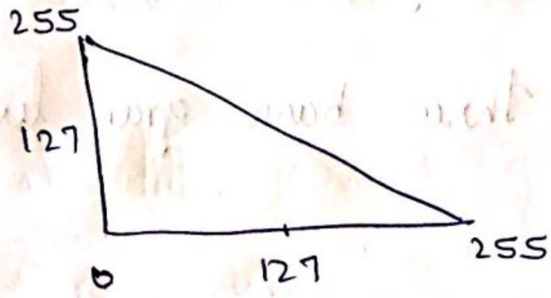
Identity transition is shown by a straight line.

In this transition, each value of the i/p image is directly mapped to each other value of o/p image.



Negative transformation is invert of identity transformation.

In negative transformation, each value of the i/p image is subtracted from the L-1 and mapped on to the o/p image.



2. Logarithmic transformations

- Log transf.
- Inverse transf.

Log transformation can be defined by this formula,

$$s = c \log (r + 1)$$

where s and r are the pixel values of the o/p and the i/p image, and c is a constant.

The value 1 is added to each of the pixel value of the i/p image because if there is a pixel intensity of 0 in the image, the log (0) is equal to infinity.

So 1 is added, to make the min. value at least 1.

Inverse log transform is opposite to log transform.

3. Power - law transformations

- n^{th} power
- n^{th} root

- These transformations can be given by the expression:

$$s = cr^{\gamma}$$

- This symbol γ is called gamma, due to which this transformation is also known as gamma transformation.

- Variation in the value of γ varies the enhancement of the images. Different display devices have their own gamma correction, that's why they display their image at different intensity.

- The gamma of different display devices is different.
- For example Gamma of CRT lies in b/w 1.8 to 2.5, that means the image displayed on CRT is dark.

6. Spatial Convolution and Correlation

Spatial Filter consists of

1) neighbourhood

2) predefined operation that is performed on the image pixels encompassed by the neighbourhood.

- Filtering creates a new pixel with coordinates equal to the coordinates of the center of the neighbourhood, and whose value is the result of filtering operation.

- A processed (filtered) image is generated as the center of the filter visits each pixel in the input image.

- If the operation performed on the image pixels is linear, then the filter is called a linear spatial filter.

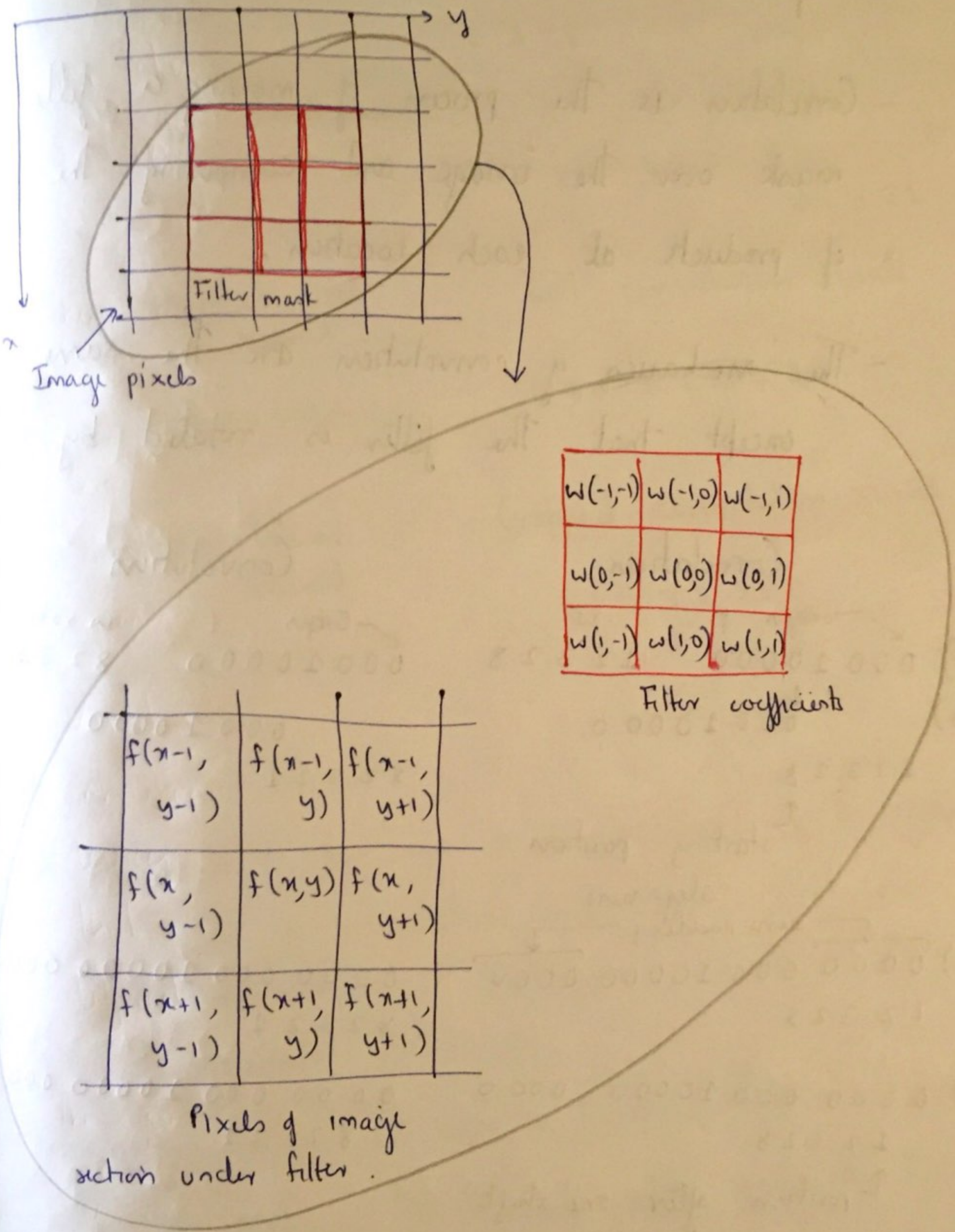


Fig. The mechanics of linear spatial filtering using 3x3 filter mask.

Spatial Correlation and Convolution

- Correlation is the process of moving a filter mask over the image and computing the sum of products at each location.

- The mechanics of convolution are the same, except that the filter is rotated by 180°.

Correlation

(origin) f   w

a) 0 0 0 1 0 0 0 0   1 2 3 2 8

b) 0 0 0 1 0 0 0 0.

1 2 3 2 8

↑ starting position alignment.

(zero padding)

c) 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0

1 2 3 2 8

d) 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0

1 2 3 2 8

↑ position after one shift

e) 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0

~~1 2 3 2 8~~

1 2 3 2 8

↑ Position after four shift

Convolution

(origin) f   w rotated 180°

0 0 0 1 0 0 0 0   8 2 3 2 1   (i)

0 0 0 1 0 0 0 0   (j)

8 2 3 2 1

0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0   (k)

8 2 3 2 1

0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0

8 2 3 2 1

0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0

8 2 3 2 1

9)

f) 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0

1 2 3 2 8

Final position ↑

0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 (n)

8 2 3 2 1

g) Full correlation result

(avoid the padding) 0 0 0 8 2 3 2 1 0 0 0 0

Full convolution result

0 0 0 1 2 3 2 8 0 0 0 0

h) Cropped correlation result

0 8 2 3 2 1 0 0

Cropped convolution result

0 1 2 3 2 8 0 0

Note

- Correlation is a function of displacement of the filter.
  - 1st value of correlation corresponds to zero displacement, 2ND corresponds to one unit displacement and so on.
- Correlating filter w with a function that contains all 0s and a single 1 yields a result that is a copy of w, but rotated by 180°.

- The correlation of a filter $w(x,y)$ of size $m \times n$ with an image $f(x,y)$ denoted as $w(x,y) \star f(x,y)$, is given by the equation,

$$w(x,y) \star f(x,y) = \sum_{x=-a}^{a} \sum_{t=-b}^{b} w(s,t)\, f(x+s, y+t)$$

This equation is evaluated for all values of the displacement variables x and y so that all elements of w visit every pixel in f.

where, $a = \frac{m-1}{2}$ $b = \frac{n-1}{2}$

m and n are odd integers

- The convolution of $w(x,y)$ and $f(x,y)$ denoted by,

$$w(x,y) \bigstar f(x,y) = \sum_{x=-a}^{a} \sum_{t=-b}^{b} w(s,t)\, f(x-s, y-t)$$

where the minus signs on the right flip (rotate it by 180°)

7. Edge Detection - Robert, Prewitt, Sobel

- Edge pixels are pixels at which the intensity of an image function changes abruptly.
- Edges are sets of connected edge pixels.
- Edge detectors are local image processing methods designed to detect edge pixels.

Three fundamental steps performed in edge detection

1. Image smoothing for noise reduction
2. Detection of edge points - operation that extracts from an image all points that are potential candidates to become edge points
3. Edge Localization - The objective of this step is to select from the candidate edge points only the points that are true members of the set of points comprising an edge.

- Some of the techniques for achieving these objectives are Robert, Prewitt and sobel operators

- The tool of choice for finding edge strength and direction at location (x,y) of an image f, is the gradient, denoted by ∇f, and defined as the vector.

$$\nabla f = \text{grad}(f) = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

- The magnitude (length) of vector ∇f, denoted as $M(x,y)$ where

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

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,

$$\alpha(x,y) = \tan^{-1}\left[\frac{g_y}{g_x}\right]$$

measured w.r.t x axis.

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

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

When diagonal edge direction is of interest, we need a 2-D mask.

{mask is small matrix used for blurring, sharpenning, embossing, edge detection etc}

Roberts cross-gradient operators are one of the earliest attempts to use 2-D masks with a diagonal preference.

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

(a)

-1	0
0	1

(b)

0	-1
1	0

(c)

Roberts

$$g_x = \frac{\partial f}{\partial x} = (z_9 - z_5)$$

$$g_y = \frac{\partial f}{\partial y} = (z_8 - z_6)$$

- Consider the 3×3 regions in (a). The robert operators are based on implementing the diagonal differences.

- These derivatives can be implemented by filtering an image with masks in b and c.

- 2×2 masks are not as useful for computing edge direction.

- 3×3 masks takes into account the nature of data on opposite sides of the center point and thus carry more information regarding the direction of an edge.

- The simplest digital approximations to the partial derivatives using masks of size 3×3 are given by

$$g_x = \frac{\partial f}{\partial x} = (z_7 + z_8 + z_9) - (z_1 + z_2 + z_3)$$

$$g_y = \frac{\partial f}{\partial y} = (z_3 + z_6 + z_9) - (z_1 + z_4 + z_7)$$

These can be implemented over an entire image by filtering f with two masks given in figure d and e

−1	−1	−1
0	0	0
1	1	1

(d)

Prewitt

−1	0	1
−1	0	1
−1	0	1

(e)

- These masks are called the Prewitt operators.

Sobel operators

- A slight variation of the preceding two equations uses a weight of 2 in the center coefficient:

$$g_x = \frac{\partial f}{\partial x} = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)$$

$$g_y = \frac{\partial f}{\partial y} = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)$$

- Using a 2 in the center location provides image smoothing.

−1	−2	−1
0	0	0
1	2	1

(f)

Sobel

(g)

−1	0	1
−2	0	2
−1	0	1

- These masks are called sobel operators.
- Prewitt masks are simpler to implement than the sobel masks
- Sobel masks have better noise suppression (smoothing) characteristics makes them preferrable.