$$\nu = 6 \times 10^{14} \text{ Hz}$$

i.e., the frequency of green light is 600,000,000,000,000 cycles/sec !

Hence it is more convenient to discuss electromagnetic radiation in terms of wavelengths (nm) rather than frequencies (Hz).

1.5 Units of Intensity

We know that for an object to be seen, some amount of light has to fall on it.

The unit of luminous intensity (I) is candela (cd) (SI unit) which by definition, is $(1/60)^{th}$ of a square centimetre of the surface of a black body radiator at the absolute temperature of 2045 K.

Note : A black body radiator is one that absorbs all the radiation incident upon it and has no reflecting power. The radiation from a heated black body source is called black body radiation and is always quoted in Kelvin.

The unit is called candela because initially it was defined as a standard candle burning a specific amount of wax per hour. There is another important unit apart from candela that we encounter. This unit is called lux.

1.6 The Human Visual System

It is important for designers and users of image processing to understand the characteristics of the human vision system. For efficient design of algorithms whose output is a photograph or a display viewed by a human observer, it is beneficial to have an understanding of the mechanism of human vision.

Many interesting studies have been carried out and the subject of visual perception has grown over the years. We begin with the structure of the human eye.

The Fig. 1.6.1 shows the horizontal as well as the vertical cross section of a human eye ball. The front of the eye is covered by a transparent surface called cornea. The remaining outer cover, sclera, is composed of a fibrous coat that surrounds the choroid, a layer containing blood capillaries. Sclera is the white portion of the eye.

Inside the choroid, is the retina which is composed of two types of photoreceptors. These photoreceptors are specialised to convert light rays into receptor potentials. These photoreceptors are called rods and cones (named due to their shape). Rods are long and slender while cones are shorter and thicker. Nerves connecting the retina leave the eyeball through the optic nerve bundle. Light entering the cornea is focused on the retina surface by a lens that changes shape under muscular control to perform proper focusing of near and distant objects. The iris acts as a diaphragm to control the amount of light entering the eye.

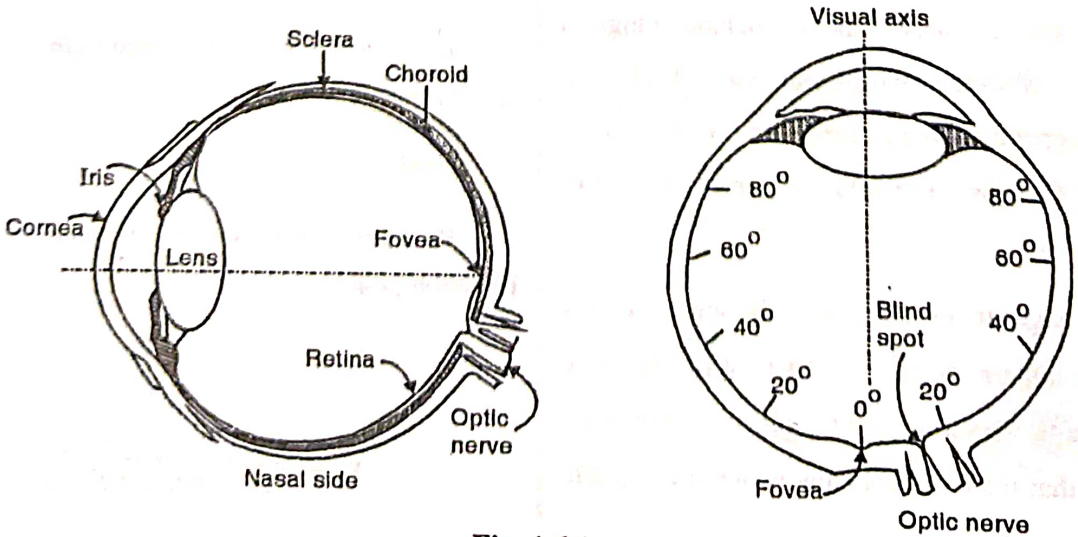


Fig. 1.6.1

Each retina has about 6 million cones and 120 million rods. Rods are most important for black and white vision in dim light. They also allow us to discriminate between different shades of dark and light and permit us to see shapes and movements. At low levels of illumination, the rods provide a visual response called Scotopic Vision. The fact that we can see in the dark is due to this scotopic vision.

Cones provide colour vision and high visual acuity (sharpness of vision) in bright light. Cones respond to higher levels of illumination, their response is called Photopic Vision. In the intermediate range of illumination, both rods and cones are active and provide Mesopic Vision.

In moon light, we cannot see colour because only the rods are functioning. The distribution of rods and cones in the retina are shown Fig. 1.6.2.

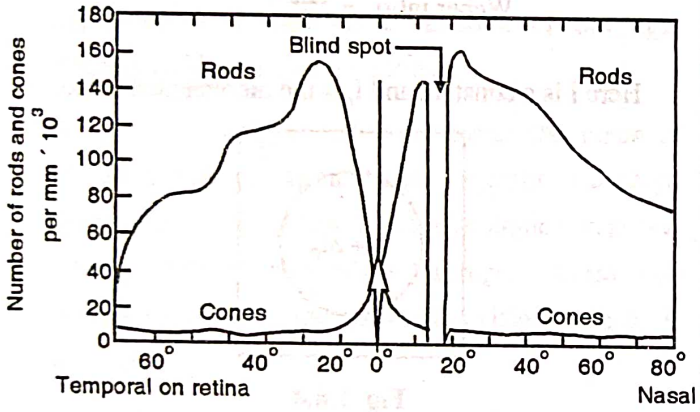


Fig. 1.6.2 : Permetric angle, degrees

At a point near the optic nerve, called fovea, the density of the cones is the greatest. Due to this, it is the region of sharpest photopic vision. The photoreceptors are connected to the nerves and these nerves, combined together, leave the eyeball as the optic nerve. (It is not as simple as stated and interested students can refer to medical journals). The optical nerve is a bundle of nerves (approx 800,000 nerve fibres) leaving the eyeball. The region from where this optic nerve leaves the eyeball has no rods and cones. This is called the blind spot. We cannot see an image that strikes the blind spot. Normally we are not aware of having a blind spot but we can easily convince ourselves of its presence.

☞ Try this experiment

Cover your left eye and gaze directly at the word 'Deep' shown below. You should also see the square drawn next to it. Now increase or decrease the distance between the book and your eye. At some point, the square will disappear. This is because the square falls on the blind spot !!

Deep ■

When a light stimulus activates a rod or a cone, a photochemical transition occurs, producing a nerve impulse. The manner in which these nerve impulses propagate through the visual system has yet to be fully understood by the medical fraternity and hence let us end our discussion on the anatomy of the human eye here. As stated earlier, this discussion is in no way exhaustive and interested students can refer medical journals. But as students of image processing this information would be sufficient.

1.6.1 Image Formation in Eye

The main difference between the lens of the eye and an ordinary optical lens is that the former is flexible. The shape of the lens is controlled by the tension of the ciliary muscles. These muscles help in focussing of objects near or far. The distance between the centre of the lens and the retina (focal length) varies from 14 mm to about 17 mm. When the eye focuses on an object far away (approx. > 3 m) the lens exhibits its lowest refractive power (Focal length = 17 mm). When the eye focuses on nearby objects, the lens of the eye is strongly refractive (Focal length = 14 mm)

Let us take an example :

Ex. 1.6.1

Consider an observer looking at a lamp-post which is at a distance of 50 m. If the height of the lamp post is 10 m, find the size of the image formed in the retina (Retinal image).

Soln. :

Since the image is far away, the focal length is approximately 17 mm. We use the similarity of triangles.

Let the retinal image be r

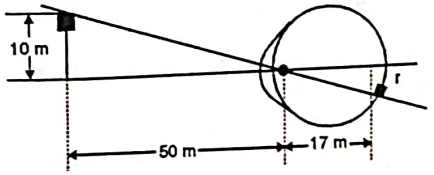


Fig. P.1.6.1

$$\frac{10}{50} = \frac{r}{17}$$

$$r = 3.4 \text{ mm}$$

∴The height of the retinal image is 3.4 mm

If the same image is observed at a distance of 100 m, we get

$$\frac{10}{100} = \frac{r}{17}$$

$$r = 1.7 \text{ mm}$$

1.6.2 Visual Phenomena

The human eye is a complex system and the images that we perceive are also equally complex. We shall now explain the visual phenomena.

What the human eye senses are in general intensity images- Intensity, Brightness and Contrast are three different phenomena. Intensity of a light source depends on the total amount of light emitted by it. Hence intensity is a physical property and can be measured. Brightness on the other hand is a psycho-visual concept and hence is actually a sensation to light intensity. Contrast may be defined as the difference in perceived brightness.

(1) Contrast Sensitivity : The response of the human eye to changes in the intensity of illumination is known to be non-linear.

Consider the simple experiment :

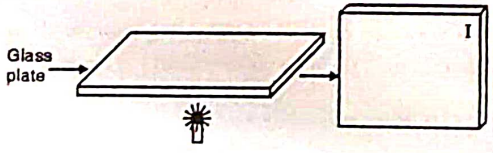


Fig. 1.6.3

We take a diffuser (opaque) glass plate and illuminate it from the bottom with constant illumination I.

The observer is asked to look at this glass plate, from the top. At the centre of this glass plate, the intensity of illumination is increased from I to I + ΔI. The observer is now asked to observe whether he or she can detect this increase.

If the observer cannot detect the change, the intensity is further increased by another increment of ΔI. This procedure is continued till the observer detects the difference. This is known as the Just Noticeable Difference (JND).

This ratio of $\frac{\Delta I_c}{I}$ is called the Weber ratio

$$\text{Weber ratio} = \frac{\Delta I_c}{I}$$

Here I is a constant and I_c is the incremented value.

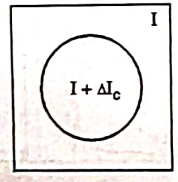


Fig. 1.6.4

A low Weber ratio implies that even a small variation (ΔI) was detected by the observer. A high Weber ratio implies that large variations (ΔI + ΔI + ...) were required for the observer to notice the change.

Hence a low Weber ratio means that the observer has good discernible vision. At proper illuminations, the Weber ratio of a group of people is more or less constant at 0.02. Weber ratio also depends on the illumination I.

As can be seen from the graph, the weber ratio is large at very low as well as very high levels of illumination. This should not come as a surprise. Our discrimination quality reduces when we are in a room that is not well lit. At the same time our discrimination quality also reduces when there is too much light.

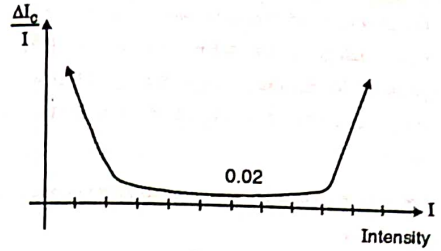


Fig. 1.6.5

(2) Brightness Adaptation : The range of light intensity levels to which the human visual system can adapt is enormous, of the order of 10^{10} from the scotopic threshold to the glare limit. It simply means we can see things in the dark and also when there is a lot of illumination. It has been shown that the intensity of light perceived by the human visual system (subjective brightness) is a logarithmic function of the light intensity incident on the human eye.

The curve in the Fig. 1.6.6 represents the range of intensities that the human visual system can adapt. When illumination is low, it is the scotopic vision that plays a dominant role while for high intensities of illumination, it is the photopic vision which is dominant.

As can be seen from the figure, the transition from scotopic vision to photopic vision is gradual and at certain levels of illuminations, both of them play a role. To cut a long story short, the dynamic range of the human eye is enormous. But there is a catch here. The eye cannot operate over the entire range simultaneously i.e., at a given point of time, the eye can only observe a small range of illuminations. This phenomena is known as Brightness Adaptation. The fact that our eyes can operate only on a small range, can be proved by a simple experiment on ourselves.

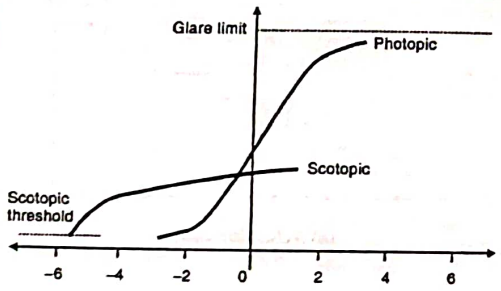


Fig. 1.6.6

Stare at the sun for a couple of seconds, the eye adapts itself to the upper range of the intensity values. Now look away from the sun and you will find that you cannot see anything for sometime. This is because the eye takes a finite time to adapt itself to this new range. A similar phenomenon is observed when the power supply of our homes is cut-off in the night. Everything seems to be pitch dark and nothing can be seen for sometime. But gradually our eyes adjust to this low level of illumination and then things start getting visible even in the dark.

(3) Acuity and Contour (Mach Bands) : Mach bands are named after the Austrian physicst Ernst Mach (1838 - 1916).

Consider a set of grey scale strips shown in Fig. 1.6.7.

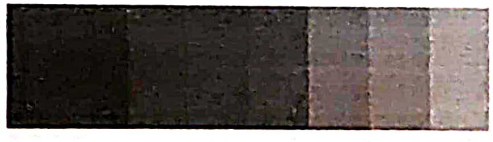
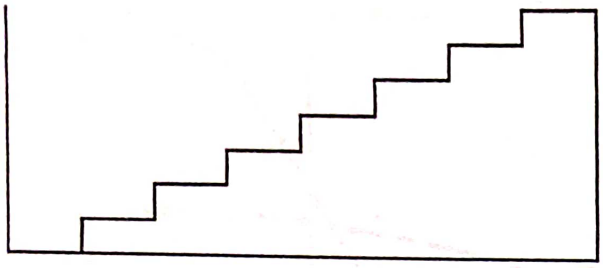


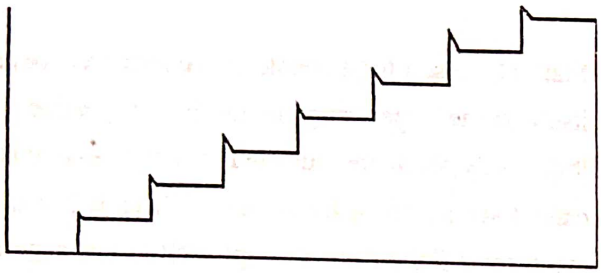
Fig. 1.6.7

Each of the eight strips have uniform intensities within the strip. This can be seen if we place our hand on all the other strips and observe only one strip at a time. But when we look at all the eight strips together, visual appearance is that each strip looks darker at its right side than its left. This is called the Mach band effect.

The actual and the perceived intensity charts are shown in Fig. 1.6.8 (a) and (b).



(a) Actual intensity



(b) Perceived intensity

Fig. 1.6.8

The intensities tend to overshoot at the right hand edges. The overshoot is a consequence of the spatial frequency response of the eye. The human eye possesses a lower sensitivity to high and low spatial frequencies than to mid range frequencies (spatial frequencies will be covered in chapter of Image Enhancement in Spatial Domain). The implication of this is that perfect edge contours in an image can be sacrificed to a certain extent as the human eye has an imperfect response to high spatial frequency brightness transitions.

(4) **Simultaneous Contrast :** Consider the image shown in Fig. 1.6.9.

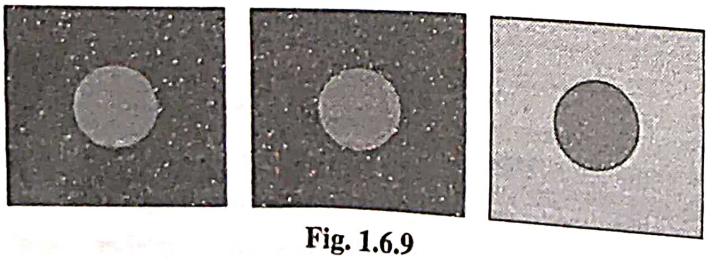


Fig. 1.6.9

Each of the small circles have the same intensity, but because the surrounding grey level of each of the circles is different, the circles do not appear equally bright. Hence the intensity that we perceive are not actually the absolute values.

(5) **Integration :** Another important property of our eye is that it integrates the scene as a whole. This is a property that facilitates programming. When we look at a picture, we do not look at each point, but the image as a whole.

Let us take an example, read the following lines as fast as you can.

"Accodring to a recheearch at an Elingsh Unvertisy, it deosn't mttaer in what oreder the ltters in a word are, the only iprmoetnt thing is that the frist and the lsat ltteer is at the rghit place. The rset can be a total mses and you can still raed it wouthit a porbelm. This is bcusease we do not raed ervey lteter by itself but the word as a wlohe."

You could read it, inspite of the spelling mistakes only because of the integration property of the eye.

Before we end, one issue needs to be sorted out. A very common question that people ask is what is the difference between image processing, computer graphics and computer vision ?

Image processing deals with manipulation of images. A input image is modified into a new image. Computer graphics deals with creation of images. In computer graphics, models (2D or 3D) are created using mathematical functions (Descriptors). Computer vision which is also known as Machine vision deals with the analysis of image content. Computer vision is used to automate a process.

Table 1.6.1

Input \ Output	Image	Description (Mathematical function)
Image	I.P.	C.V.
Description (Mathematical function)	C. G.	-

From the Table 1.6.1,

Review Questions

Q. 1 What do we mean by image processing ?

Q. 2 Why do we say that an image is a 2D-function ?

Q. 3 Calculate the frequency of oscillation of red light.

Q. 4 Explain the structure of the human eye.

Q. 5 Explain the term scotopic vision, photopic vision and mesopic vision.

Q. 6 Brightness discrimination is poor at low levels of illumination. Explain.

Q. 7 Images are processed either for human perception or for machine perception. Explain.

Q. 8 What does the Weber ratio imply ?

Q. 9 What do Mach bands imply ?

Q. 10 If an observer is looking at an object 12 m high and a distance of 2 m. Find out the size of the retinal image.

Q. 11 Distinguish between image processing and graphics.

Q. 12 Explain the term brightness adaptation.

CHAPTER 2

Image Sensing and Acquisition

2.1 Introduction

☞ Basic elements of an image processing system

Digital image processing is basically modification of images on a computer. The basic components of an image processing system are shown below.

(1) Image Acquisition.

(2) Image Storage.

(3) Image Processing.

(4) Display.

(5) Transmission (if required).

We shall discuss each one in detail.

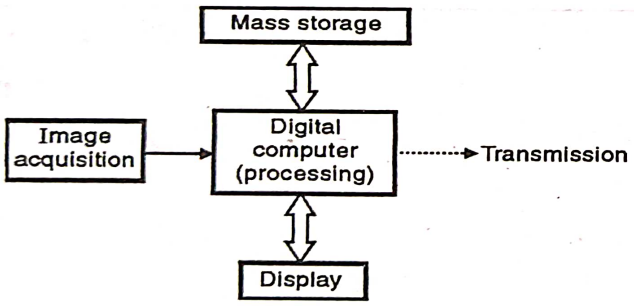


Fig. 2.1.1

(1) **Image Acquisition :** Image acquisition is the first step in any image processing system. The general aim of image acquisition is to transform an optical image (real world data) into an array of numerical data which could be later manipulated on a computer.

Image acquisition is achieved by suitable cameras. We use different cameras for different applications. If we need an X-ray image, we use a camera (film) that is sensitive to X-rays. If we want an infra-red image, we use cameras which are sensitive to infra-red radiations. For normal images (family pictures etc.) we use cameras which are sensitive to the visual spectrum. In this book we shall discuss cameras (sensors) which are sensitive only to the visual range.

Quantum detectors

Quantum detection is the most important mechanism of image sensing and acquisition. It relies upon the energy of absorbed photons being used to promote electrons from their stable state to a higher state above an energy threshold. Whenever this occurs, the properties of that material get altered in some measurable way. Planck/Einstein came up with a relationship between the wavelength of the incident photon and the energy that it carries.

$$e = \frac{h\,c}{\lambda} \qquad \ldots(2.1.1)$$

Where, e = Energy carried by a photon

h = Planck's constant, 6.626×10^{-34} Js

c = Speed of light, 3×10^{8} m/sec

λ = Wavelength of the incident radiation.

On collision, the photon transfers all or none of this quantum of energy to the electron.

The Equation (2.1.1) is very important as it tells us that the maximum wavelength to which the quantum detector will respond is determined by the energy threshold and hence by the material selection.

Of the several modes of operation of quantum detectors, the most important ones are

(a) Photoconductive (b) Photovoltaic.

(a) **Photoconductive :** The resistance of photoconductive materials drop in the presence of light due to the generation of free charge carriers. An external bias is applied across the material to measure this change. This principle of photoconductivity is used in Vidicon imaging tubes.

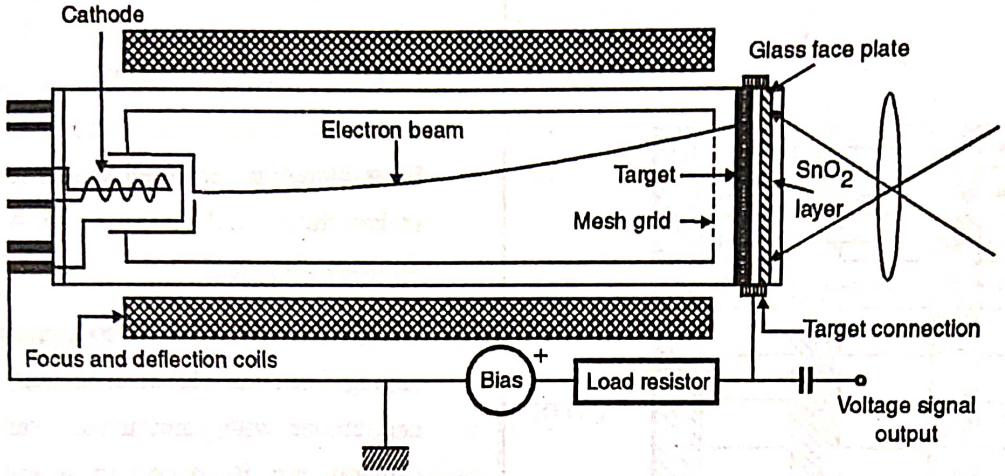


Fig. 2.1.2 : Vidicon imaging tubes

(b) **Photovoltaic :** Photovoltaic devices consist of semiconductor junctions. They are solid state arrays composed of discrete silicon imaging elements known as *Photosites*. Photovoltaic devices give a voltage output signal that is proportional to the intensity of the incident light. No external bias is required as was in the case of photoconductive devices.

The technology used in solid-state imaging sensors is based principally on charge-coupled devices, commonly known as Charged Coupled Devices CCDs. Hence the imaging sensors are called CCD sensors.

The solid state array (CCD) can be arranged in two different configurations.

(b.1) Line array CCD (b.2) Area array CCD.

(b.1) Line Arrays : The line array represents the simplest form of CCD imager and has been employed since the early 1970s. Line arrays consist of a one-dimensional array of photosites.

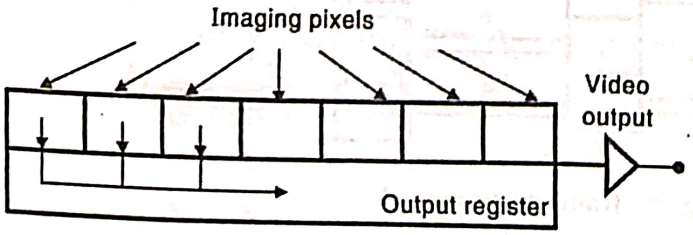


Fig. 2.1.3

A single line of CCD pixels are clocked out into the parallel output register as shown in Fig. 2.1.3. The amplifier outputs a voltage signal proportional to the contents of the row of photosites. One thing to note is that line array CCD scans only one line (hence is one-dimensional). In order to produce a two-dimensional image, the line array CCD imager has to be used as a scanning device by moving this array over the object by some mechanical activity.

This technique is used in flat bed scanners (the scanners that you come across in your laboratory or in a cyber cafe). A line array CCD can have anything from a few elements to upto 6000 or more.

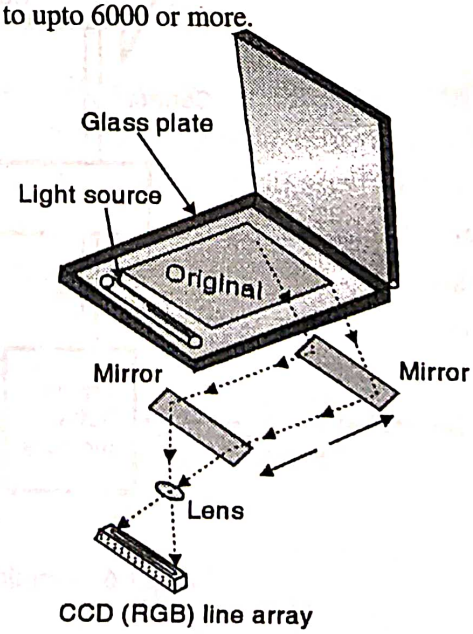


Fig. 2.1.4

(b.2) Area Arrays : The problem with line arrays is that it scans only one line. To get a two-dimensional image, we need to mechanically move the array over the entire image.

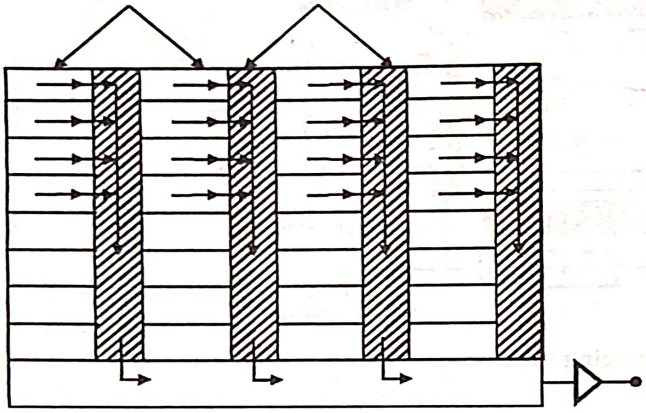


Fig. 2.1.5

Area arrays or matrix arrays consist of a two-dimensional array of photosites. They make it possible to investigate static real world scenes without any mechanical scanning. Thus much more information can be deduced from a single real time glance than would be possible with line arrays. Area arrays can be seen in all the digital cameras that we use for video imaging. The area arrays are more versatile than the line arrays, but there is a price to be paid for this. Area arrays are higher on cost and complexity.

Area sensors come in different ranges. i.e. 256 × 256, 490 × 380, 640 × 480, 780 × 575. CCD arrays are typically packaged as TV cameras. A significant advantage of solid state array sensors is that they can be shuttered at very high speeds (1/10,000 secs). This makes them ideal for applications in which freezing motion is required.

(2) Image Storage : All video signals are essentially in analog form i.e. electrical signals convey luminance and colour with continuously variable voltage. The cameras are interfaced to a computer where the processing algorithms are written. This is done by a frame grabber card. Usually a frame grabber card is a printed circuit board (PCB) fitted to the host computer with its analog entrance port matching the impedance of the incoming video signal. The A/D converter translates the video signals into digital values and a digital image is constructed. Frame grabber cards usually have a A/D card with resolution of 8 - 12-bits (256 to 4096 gray levels). Hence a frame grabber card is an interface between the camera and the computer.

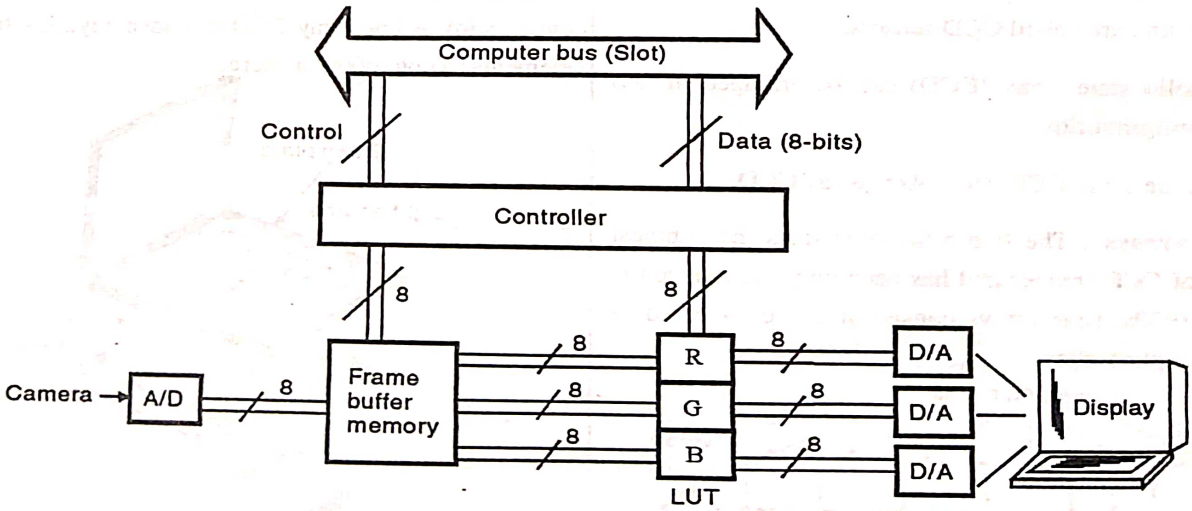


Fig. 2.1.6 : Simplified diagram of a frame grabber card

Frame grabber card has a block of memory, separate from the computers own memory, large enough to hold any image. This is known as the *frame buffer memory*. Image data, usually in the form of bytes (8-bits), is written into the frame grabber memory under the computer control, usually by DMA transfers. The contents of the memory are continuously read out at video rate (30 frames/sec), passed through the D/A converter and displayed on the monitor.

The output has a colour map or a colour Look Up Table (LUT) to permit pseudo colouring. The table consists of 8-bit values for each of the red, green and blue guns of the monitor. The commercially available frame grabber cards have additional features such as ability to zoom regions of the image and pan the images. The frame buffer memory can be upto 5 MB. Images being 2-dimensional functions, occupy a lot of space and hence the storage space on the host computer has to be large. Let us try to find out as to how much space is actually taken by images.

Each image is stored as a matrix, where every value of the matrix represents the grey level at that point. Suppose the image (matrix) is of size A × B and suppose we require C bits to represent each element of the matrix. Memory space required to store this image is approximately equal to A × B × C bits.

Suppose we have D such images then total number of bits ≈ A × B × C × D ...(2.1.2)

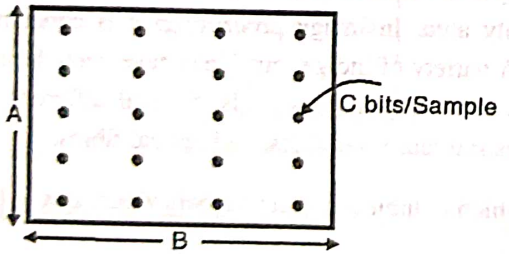


Fig. 2.1.7 (a)

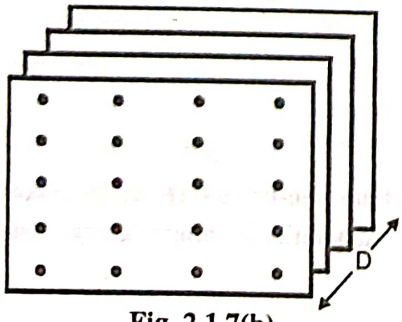


Fig. 2.1.7(b)

Thus the formula ≈ ABCD gives us a fairly good idea of the amount of space required.

	A	B	C	D	Number of bits
Low quality video phone	64	64	8	6	≈ 0.2×10^6 bits
Digital broadcast TV	720	576	25	8	≈ 83×10^6 bits
HDTV	1920	1150	50	8	≈ 883×10^6 bits

With the advent of the PCI bus, the host computers memory could now be used as the frame buffer memory. This is because the PCI bus facilitates fast communication. Hence the image could now directly be stored on the computer RAM. Due to this, the frame buffer memory has become more or less obsolete. Only the very high end frame grabber cards still have a small frame grabber memory embedded on them.

Images being two-dimensional functions, occupy a lot of space and hence it is advisable to have a computer with a sizeable hard disk and also a good RAM. Processed images could be stored on magnetic floppies or CDs. Archival data are stored on magnetic tapes.

(3) **Image Processing :** Systems ranging from microcomputers to general purpose large computers are used in image processing. Dedicated image processing systems connected to host computers are very popular now-a-days. Processing of digital images involve procedures that are usually expressed in algorithmic form due to which most image processing functions are implemented in software. The only reason for specialized image processing hardware is the need for speed in some applications or to overcome some fundamental computer limitations. The trend though is to merge general purpose small computers with image processing hardware. As stated in the earlier section, frame grabber cards play the important role of merging the image processing hardware with

the host computer. One thing that we should remember is, image processing is characterized by specific solutions. Hence there is no one way to process images. A technique that works well in one area can be totally useless in some other applications.

The image processing software that is used in this book is MATLAB.

(4) **Display :** A display device produces and shows a visual form of numerical values stored in a computer as an image array. Principal display devices are printers, TV monitors and CRTs. Any erasable raster graphics display can be used as a display unit with an image processing system. However monochrome and colour TV monitors are the principal display devices used in modern image processing systems.

These raster devices convert image data into a vi frame. One major problem is, it must refresh the scr at a rate of about 30 frames per second to avoid flicl Since some computers are unable to transfer data such high speeds, it is a good idea to buy a high frame grabber card which has frame buffer memory it.

If you want to build a image processing the systen home, you should have a Video Graphics Card (VG attached to a computer and a monitor that supports t VGA card. One could then use a simple camera t comes along with the system.

(5) **Transmission :** There are a lot of applications whe we need to transmit images. The stages in t transmission of an image over a channel or network a shown in Fig. 2.1.8.

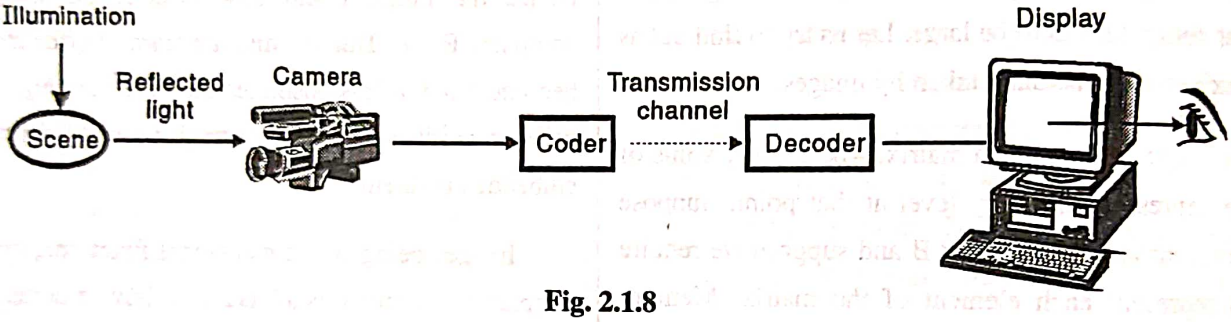


Fig. 2.1.8

The image sequence from the camera is coded into as concise a representation as possible for transmission over t channel.

Most transmission in broadcast television are still in analog form and analog coding methods are used to make it a efficient as possible. NTSE, PAL and SECAM are the 3 major coding systems used in various parts of the world (USA us NTSE, while India uses PAL). Digital image coding is a high activity area. In image processing; it is concerned wit efficient transmission of images over digital communication channels. A variety of indigenous ideas have been developed i recent years. Coding is influenced by the type of the channel used to carry the image signals. Several different types c transmission channels are encountered in practice including cables, terrestrial radio, satellites, and optical fibres.

Gamma : In the image transmission chain, there are many elements which exhibit a non-linear behaviour. If x is the inpu and y is the output, then

$$y = cx^{\gamma}$$

c = Constant

γ = Gamma of the device

The camera, the display device and eye all have non-unity gammas (all are non-linear). Hence to make sure that th perceived grey scale in the displayed image is correct, it is necessary to insert an additional compensating, non-linear devic called the gamma corrector.

We have discussed a lot of things so far and it is advisable at this stage to have a branch diagram of these topics.

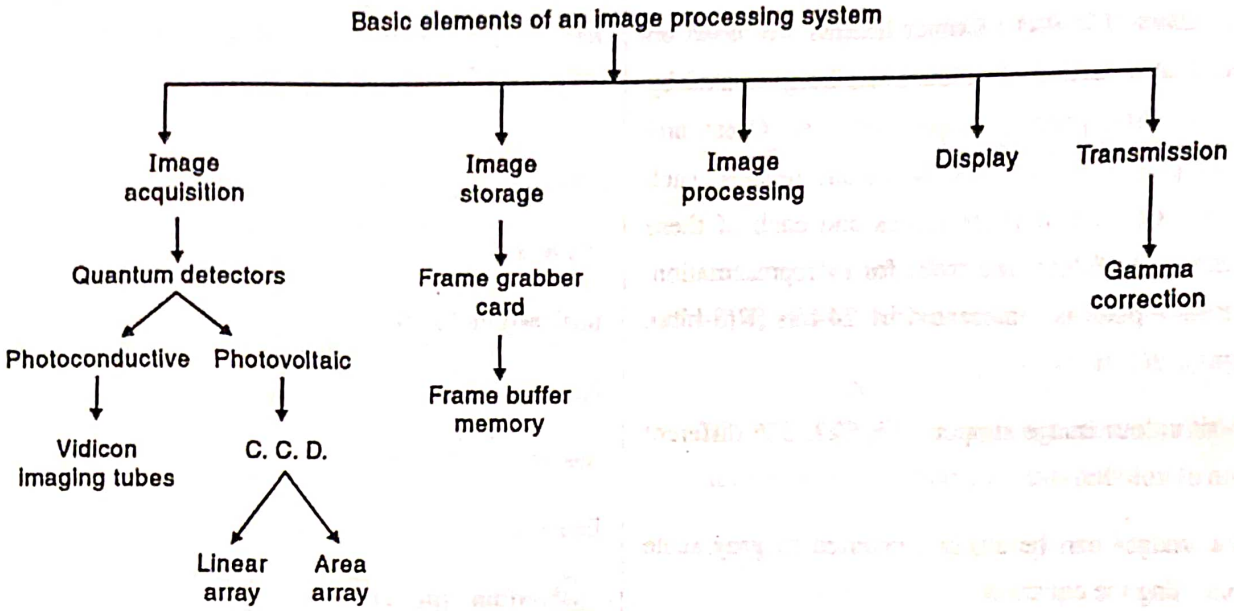


Fig. 2.1.9

2.2 Image Types

It was stated earlier that images are 2-dimensional functions. Refer Fig. 2.2.1.

Images can be classified as follows :

(1) Monochrome images (Binary images).

(2) Grey scale images.

(3) Colour (24-bit) images.

(4) Half-toned images.

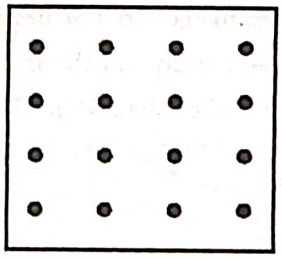


Fig. 2.2.1

(1) **Monochrome Image :** In this, each pixel is stored as a single bit (0 or 1). Here, 0 represents black while 1 represents white. It is a black and white image in the strictest sense. These images are also called bit mapped images. In such images, we have only black and white pixels and no other shades of grey. Refer Fig. 2.2.2.

Fig. 2.2.2

(2) **Grey Scale Image :** Here each pixel is usually stored as a byte (8-bits). Due to this, each pixel can have values ranging from 0 (black) to 255(white). Grey scale images, as the name suggests have black, white and various shades of grey present in the image. Refer Fig. 2.2.3.

Fig. 2.2.3

(3) Colour Image (24-bit) : Colour images are based on the fact that a variety of colours can be generated by mixing the three primary colours viz, Red, Green and Blue, in proper proportions. In colour images, each pixel is composed of RGB values and each of these colours require 8-bits (one byte) for its representation. Hence each pixel is represented by 24-bits [R(8-bits), G(8-bits), B(8-bits)].

A 24-bit colour image supports 16, 777, 216 different combination of colours.

Colour images can be easily converted to grey scale images using the equation

$$X = 0.30\,R + 0.59\,G + 0.11\,B \qquad ...(2.2.1)$$

An easier formula that could achieve similar results is

$$X = \frac{R+G+B}{3} \qquad ...(2.2.2)$$

MATLAB code for converting a colour image to a grey scale image is shown below.

```
%% Converting a colour image to a grey level image %%
clear
clc
im = imread('lily.tif');
[row col byt]=size(im);
a = im(:,:,1);        % Red plane
b = im(:,:,2);        % Green plane
c = im(:,:,3);        % Blue plane
a = double(a);
b = double(b);
c = double(c);
for x = 1:1:row
for y = 1:1:col
        new(x,y) = (a(x,y)+b(x,y)+c(x,y))/3;
new1(x,y) = 0.3*a(x,y)+0.59*b(x,y)+0.11*c(x,y);
end
end
figure(1)
imshow(uint 8(im))
figure(2)
imshow(uint 8(new))
figure(3)
imshow(uint8(new1))
```

Matlab has an inbuilt command for conversion *rgb2gray*.

(4) Half Toning : It is obvious that a grey scale image definitely looks better than the monochrome image as it utilizes more grey levels. But there is a problem in hand. Most of the printers that we use (inkjet, lasers, dot matrix) are all bi-level devices. i.e., they have only a black cartridge and can only produce two levels (black on a white back-ground). In fact, most of the printing jobs are done using bi-level devices.

You have all read newspapers at some point of time (hopefully). The images do look like grey level images. But if you look closely, all the images generated are basically using black colour. Refer Fig. 2.2.4.

Fig. 2.2.4

Even the images that you see in most of the books (including this one) are generated using black colour on a white background. In spite of this we do get an illusion of

seeing grey levels. The technique to achieve an illusion of grey levels from only black and white levels is called half-toning.

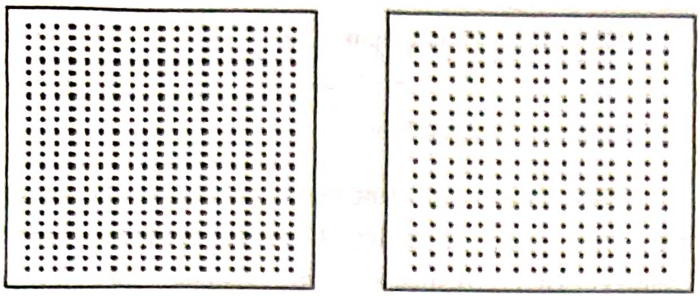


Fig. 2.2.5

The human eye integrates the scene that it sees. Consider a simple example. Consider two squares of say 0.03 × 0.03 sq. inch. One of these squares contains a lot of black dots while the other square contains fewer black dots. When we look at these squares from a distance, the two squares give us a perception of 2 different grey levels. This integration property of the eye is the basis for half toning. In this, we take a matrix of a fixed size and depending on the grey level required, we fill the matrix with black pixels.

Let us take an example.

Consider a 3 × 3 matrix. This matrix can generate an illusion of 10 different grey levels when viewed from a distance.

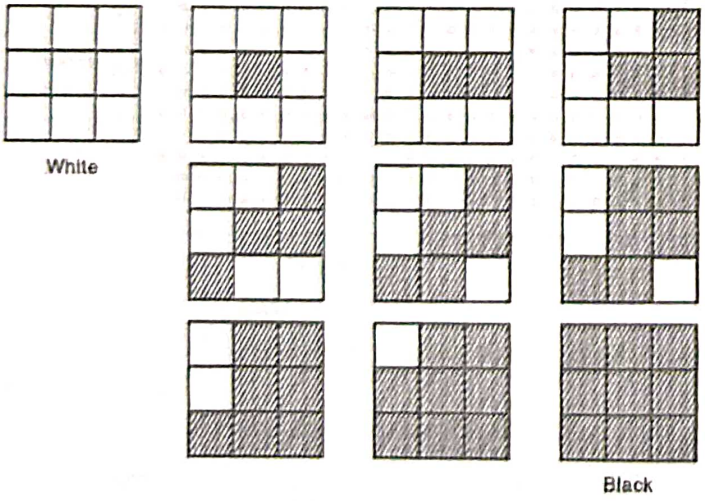


Fig. 2.2.6

Here, the first block represents white, the last block represents black and all the other blocks represent intermediate values of grey. So in this case, while printing, if we encounter grey level 0, we plot 9 pixels which are all white. If we encounter grey level 1, we plot 9 pixels of which only one pixel is black and so on.

As is evident a 3 × 3 matrix will generate 10 different grey levels.

The dots that are placed in the 3 × 3 matrix example can be in any order. But we need to follow two rules.

(1) Dots should be arranged in such a manner so that they don't form straight lines. Lines are very obvious to the viewer and hence should be avoided.

Example, suppose in an image we have 4 consecutive grey levels of value 3, and if we use a code as shown in Fig. 2.2.7, the half-toned image would look like Fig. 2.2.8.

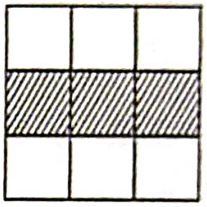


Fig. 2.2.7

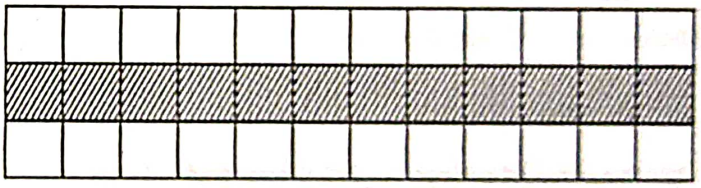


Fig. 2.2.8

It can be seen that it does not look like a grey level. It looks like a straight black line. Hence lines should be avoided while defining the matrices.

(2) If a pixel in a particular matrix is black for grey level i, it should be black for all further levels j > i. This reduces false contouring.

Half-toning gives excellent results and we do perceive a grey level image just by using black pixels on a white background.

The logic to implement a half-toned image from a grey level image is given below.

(1) Define the size of the half-toned matrices based on the number of grey levels the original image has.

(2) Generate the matrices starting from all white pixels to all black pixels.

(3) Read the original image. For every grey level value read, plot the corresponding matrix.

Remember, the physical size of the half-toned image will always be bigger than the original image as for every single grey level value, we output an entire matrix.

2.3 Image File Formats

Images obtained from the camera are stored on the host computer using different formats. A file format is a structure which defines how information is stored in the file and how that information is displayed on the monitor. There are numerous image file formats which are available. Of these only a few of them can be used universally. By universally it means, they can be understood by different operating systems.

Some of the image formats that can be used on either Macintosh or PC platforms are;

BMP (Bit Mapped Graphic Image)	JPEG (Joint Photographic Expert Group)
TIFF (Tagged Image File Format)	EPS (Encapsulated Post Script)
GIF (Graphic Interchange Format)	PICT (Macintosh Supported)

TIFF, which is one of the most well known formats was developed in 1986 and in its many versions is a standard image format for a bit-mapped graphics image. The TIFF format is also a data compression technique for monochrome as well as colour images. Images seen on the Internet sites are normally TIFF/GIF images simply because they occupy less space. GIF uses a form of Huffman coding.

BMP images developed by Microsoft can save both monochrome as well as colour images. All the wall papers that your computer has are BMP images. Similarly when we work in Paint Brush, we can save our work as a BMP image. The quality of BMP files is very good but they occupy a lot of memory space. PICT is a general purpose format supported by Macintosh and used for storing bit-mapped images. EPS is file format of the post script page description language and is device independent. This simply means that images can readily be transferred from one application to another. However EPS images can only be printed on post script compatible printers.

JPEG (Joint Photographic Expert Group) is the name of the committee that developed an image format which used compression algorithms.

JPEG images are compressed images which occupy very little space. It is based on the Discrete Cosine Transform-DCT (explained in chapter of Image Transforms). JPEG images use lossy compression algorithms which result in some loss of original data and hence the quality of JPEG images is not as good as BMP images.

Although these formats differ in technical details, they share structural similarities.

The Fig. 2.3.1 shows the typical organisation of information encoded in an image file.

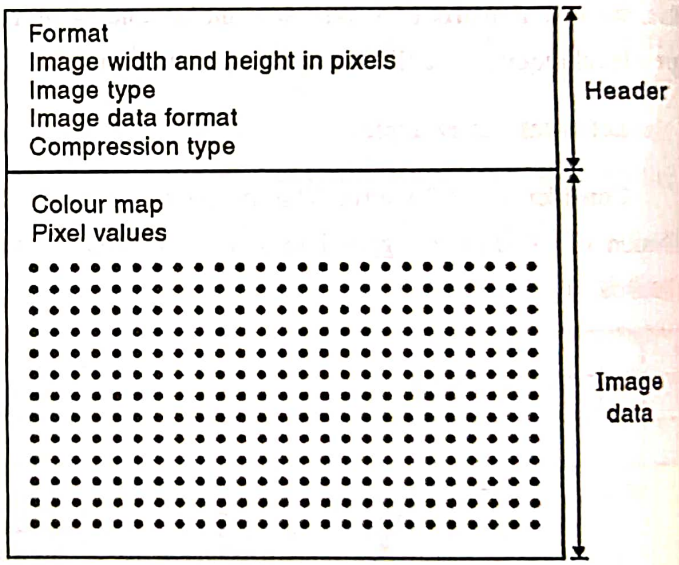


Fig. 2.3.1

The image file consists of two parts;

(1) Header (2) Image data

The header file gives us the information about the kind of image. The header file, begins with a binary code or ASCII string which identifies the format being used. The width and height of the image are given in number of pixels. Common image types include binary images, 8-bit grey scale images, 8-bit colour and 24-bit colour images. The image data format specifies the order in which pixel values

Vidicon cameras, CCD cameras and the frame grabber card are explained using simple illustrations. Importance of each is stated. Applications of line array CCD's and area array CCD's are presented. Basic concepts of image formats are also explained.

Review Questions

Q. 1 List the basic elements of an image processing system.

Q. 2 Explain working of a Vidicon camera.

Q. 3 Explain line array and area array CCD cameras.

Q. 4 Write a short note on frame grabber card.

Q. 5 Explain the concept of half toning.

Q. 6 Explain image formats.

Q. 7 Advantages of CCD over Vidicon.

Q. 8 Explain the basics of quantum detector.

Q. 9 Explain gamma of a camera.

Q. 10 How many grey levels will a half toned image have ? Explain in detail.

Chapter Ends...

CHAPTER 3

Sampling and Quantization

3.1 Introduction

We know that an image is basically a 2-dimensional representation of the 3-dimensional world. We have also studied that images can be acquired using a Vidicon or a CCD camera or using scanners. The basic requirement for image processing is that images obtained be in the digital format. For example, one cannot work with or process photographs on identity cards unless he/she scans it using a scanner.

The scanner digitizes the photograph and stores it on the hard disk of the computer. Once this is done, one can use image-processing techniques to modify the image as per requirement. In a Vidicon too, the output which is in analog form needs to be digitized in order to work with the images. To cut a long story short, to perform image processing, we need to have the images on the computer. This will only be possible when we digitize the analog pictures.

Now that we have understood the importance of digitization, let us see what this term actually means.

The process of digitization involves two steps :

(1) Sampling

(2) Quantization

In other words, Digitization = Sampling + Quantization.

3.2 Sampling

Sampling is the process of converting a continuous scene into a discrete set of numbers. A scene in the real world consists of infinite points. When we click a picture of this scene using a camera, these infinite points get mapped on to a finite set of points based on the number of pixels on the sensor of the camera. Hence if our camera is of 12 Mega pixels, then the real world scene will be stored as 12×10^6 samples in the memory. This is known as sampling. Sampling determines the spatial resolution of the digitized image and depends on the Mega pixel of the camera.

3.3 Quantization

The values obtained by sampling a continuous function usually comprise of an infinite set of real numbers ranging from a minimum to a maximum depending upon the sensors calibration. These values must be represented by a finite number of bits usually used by a computer to store or process any data. In practice, the sampled signal values are represented by a finite set of integer values. This is known as quantization. Rounding of a number is a simple example of quantization.

With these concepts of sampling and quantization, we now need to understand what these terms mean when we look at an image on the computer monitor.

Higher the spatial resolution of the image, greater is the sampling rate. Similarly, higher the grey level resolution (tonal resolution), more are the number of quantized levels.

Hence spatial resolution gives us an indication of the sampling while grey level resolution (tonal resolution) gives us an indication of the quantization.

Spatial resolution ⟶ Sampling

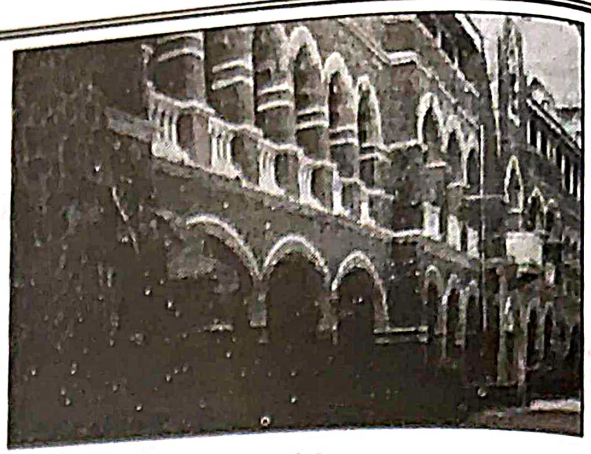
Grey level resolution ⟶ Quantization

We have already stated that an image can be considered as a 2-D array. Image f(x,y) is arranged in the form of N × M array

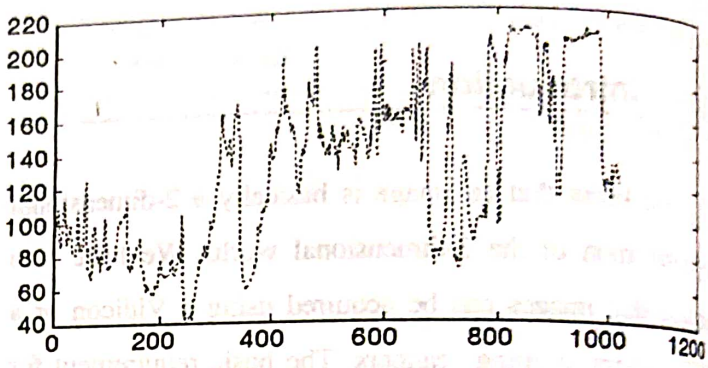
$$f(x,y) = \begin{bmatrix} f(1,1) & f(1,2) & \ldots & f(1,M) \\ f(2,1) & f(2,2) & \ldots & f(2,M) \\ f(3,1) & f(3,2) & \ldots & f(3,M) \\ \vdots & & & \vdots \\ f(N,1) & f(N,2) & \ldots & f(N,M) \end{bmatrix}_{N \times M}$$

Hence every image that is seen on the monitor, is actually this matrix. Each element of the matrix is called a *pixel*. Never forget this. Whenever we see an image on the screen of the computer it is actually a matrix which consists of N × M pixels and each pixel is considered to be a sample. Hence more the pixels, more the samples, higher the sampling rate and hence better the spatial resolution. The value of each pixel is known as the grey level.

The computer understands only ones and zeros. Hence these grey levels need to be represented in terms of zeros and ones. If we have two bits to represent the grey levels, only 4 different grey levels (2^2) can be identified viz. 00, 01 10, 11, where 00 is black, 11 is white and the other two are different shades of grey. Similarly, if we have 8 bits to represent the grey levels, we will have 256 grey levels (2^8). Hence more the bits, more are the grey levels and better is the tonal clarity (quantization). The total size of the image is N × M × m, where m is the number of bits used.



(a)



(b)

Fig. 3.3.1 : The first row of the image is plotted. The x-axis is the number of samples (pixels) in the row (sampling) while the y-axis is the grey level of each sample (quantization)

Now, comes the obvious question. As we know, more the samples and the bits, better is the image. What then should be the ideal values of sampling and quantization. This answer will vary from image to image. Given below is a table of sampling and the quantization values. As the sampling and the quantization increase, the number of bits required to store the image increases tremendously.

The clarity increases, but storage space required increases too. We hence need to get a trade-off between the two. For simplicity, we consider a square matrix of size N × N.

Another term that we need to understand is the Aspect ratio.

The ratio of the image's width to its height, measured in unit length or number of pixels is referred to as its aspect ratio. Both, a 3 × 3 inch image and a 128 × 128 image have the same aspect ratio of 1.

$$\text{Aspect ratio} = \frac{X}{Y}$$

Ex. 3.5.1

Compute the physical size of a 640 × 480 image when printed by a printer at 240 pixels per inch.

Soln. :

Since we have 240 pixels per inch, the physical size of the image is $\frac{640}{240}$ by $\frac{480}{240}$.

Ex. 3.5.2

If we want to resize a 1024 × 768 image to one that is 600 pixels wide with the same aspect ratio as the original image, what should be the height of the resized image?

Soln. :

We know

$$\text{Aspect ratio} = \frac{\text{Width}}{\text{Height}}$$

For the original image the Aspect ratio is = 1024/768 = 1.33

Now for the resized image, we want the same aspect ratio but a width of 600 pixels.

$$\text{Height} = \frac{\text{Width}}{\text{Aspect ratio}} = \frac{600}{1.33}$$

$$= 451$$

Hence the resized image will be 600 × 451.

Ex. 3.5.3

A common measure of transmission for digital data is the baud rate, defined as the number of bits transmitted per second. Transmission is accomplished in packets consisting of a start bit, a byte (8-bits) of information and a stop bit.

(a) How many minutes would it take to transmit a 1024 × image with 256 grey levels if we use a 56 k baud mode

(b) What would be the time required if we use a 750 k transmission line ?

Soln. :

(a) Since we have 256 grey levels, we need 8-bit representing each pixel.

Along with these 8-bits, we also have a start bit a stop bit.

Hence we have (8 + 2) bits per pixel.

∴ Total number of bits for transmission is

$$N = 1024 \times 1024 \times 10$$

$$N = 10485760\text{-bits}$$

These bits are transmitted at 56 k baud.

$$\therefore \text{Time taken} = \frac{N}{56 \times 10^3}$$

$$\therefore \text{Time taken} = 187.25 \text{ sec} \approx 3.1 \text{ minut}$$

(b) Now if we use a faster transmission line of 750 k rate, then

$$\text{Time taken to transmit the image} \frac{N}{750 \times 10^3} = \frac{10485760}{750 \times 10}$$

$$= 13.98 \text{ sec} \approx 14 \text{ sec}$$

Ex. 3.5.4

We have a C. T. image which we want to transmit through a grade telephone line. The CT image is of size 512 × 512 an 256 grey levels. During transmission each byte is preceded start bit and succeeded by a stop bit. How many minutes take to transmit the image ?

(Voice grade telephone lines can transmit 9600-bits/sec)

Soln. :

256 grey level required mean 8-bits.

$$\therefore \text{one packet} = [8 + 02] \text{ bits} = 10 \text{ bits}$$

$$\therefore \text{image size} = 512 \times 512 \times 10$$

Review Questions

Q. 1 Explain sampling and quantization.

Q. 2 Explain the effects of reducing sampling and quantization.

Q. 3 Explain isopreference curves.

Q. 4 Explain non-uniform sampling.

Q. 5 Compute the physical size of a 480 × 360 image when printed by a printer at 320 dpi.

CHAPTER 4

Image Enhancement in Spatial Domain

4.1 Introduction

Image enhancement is one of the first steps in image processing. As the name suggests, in this technique, the original image is processed so that the resultant image is more suitable than the original for specific applications i.e. the image is enhanced. Image enhancement is a purely subjective processing technique. By subjective we mean that the desired result varies from person to person. An image enhancement technique used to process images might be excellent for a person, but the same result might not be good enough for another. It is also important to know at the outset that image enhancement is a cosmetic procedure i.e. it does not add any extra information to the original image. It merely improves the subjective quality of the images by working with the existing data.

Image enhancement can be done in two domains :

(1) The spatial domain

(2) The frequency domain.

Let us start with explaining what is meant by the spatial domain, discuss the various methods of spatial domain enhancement and then move on to discuss the frequency domain technique in chapter of Image Enhancement in Frequency Domain.

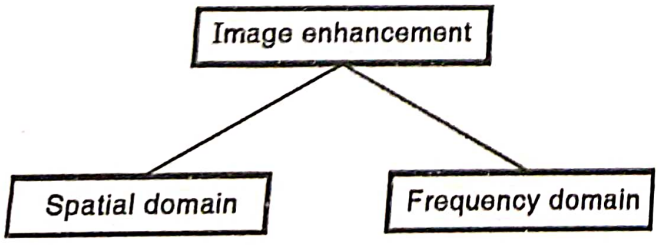


Fig. 4.1.1

4.2 Spatial Domain Methods

The term spatial domain means working in the give space, in this case, the image. It implies working with th pixel values or in other words, working directly with the ra data.

Let f(x, y) be the original image where f is the gre level value and (x, y) are the image coordinates. For a 8-b image, f can take values from 0 - 255 where 0 represen black, 255 represents white and all the intermediate valu represent shades of grey.

In a image of size 256 × 256 as shown in the Fig. 4.2.

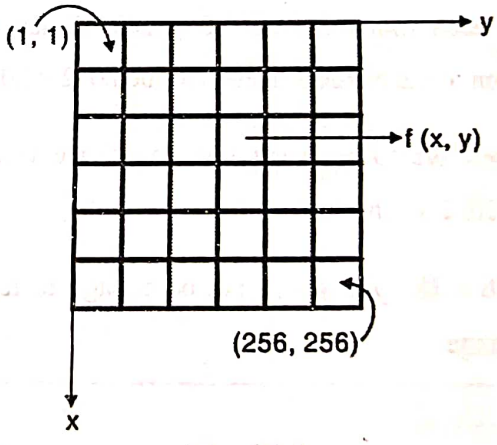


Fig. 4.2.1

Note : We always consider the axis as shown i Fig 4.2.1. The x-axis is taken in the downward vertic direction, while the y-axis is taken in the horizont direction. The reason for taking this inverted coordina system is to make sure that the image uses the standa matrix coordinates that we have been using right from o school. You will remember the first value of the matrix always at the top left corner while the last value of t matrix is at the right bottom corner.

The modified image can be expressed as

$$g(x, y) = T[f(x, y)] \quad \ldots(4.2.1)$$

Here f(x, y) is the original image and T is the transformation applied to it to get a new modified image g(x, y). For all spatial domain techniques it is simply T that changes. The general equation remains the same.

To summarize, spatial domain techniques are those, which directly work with the pixel values to get a new image based on Equation (4.2.1).

Spatial domain enhancement can be carried out in two different ways :

(1) Point processing

(2) Neighbourhood processing

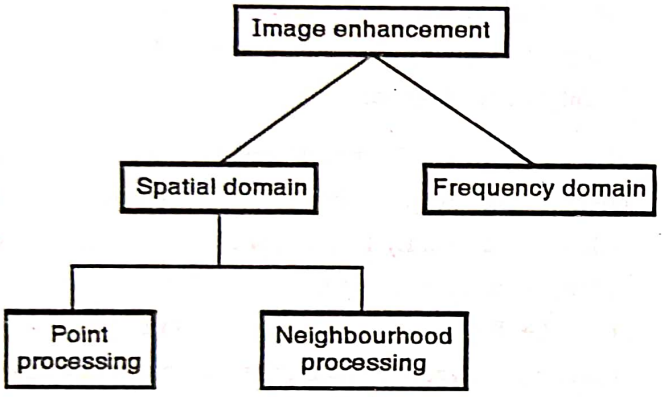


Fig. 4.2.2

4.3 Point Processing

In point processing, we work with single pixels i.e. T is 1×1 operator. It means that the new value f(x, y) depends on the operator T and the present f(x, y). This statement will be clear as we start giving some examples.

Some of the common examples of point processing are

(1) Digital negative

(2) Contrast stretching

(3) Thresholding

(4) Grey level slicing

(5) Bit plane slicing

(6) Dynamic range compression

(7) Power law transformation

Before we proceed to explain the following examples, it is important to understand the identity transformation. The identity transformation is given in the Fig. 4.3.1.

In the Fig. 4.3.1, the solid line is the transformation T. The horizontal axis represents the grey scale of the input image (r) and the vertical axis represents the grey scale of the output image (s). It is called an identity transformation because it does not modify the new image at all !!. As seen, the grey level 10 is modified to 10, 125 to 125 and finally 255 to 255. In general $s_l = r_l$. Fig. 4.3.1 will help us understand the point processing techniques better.

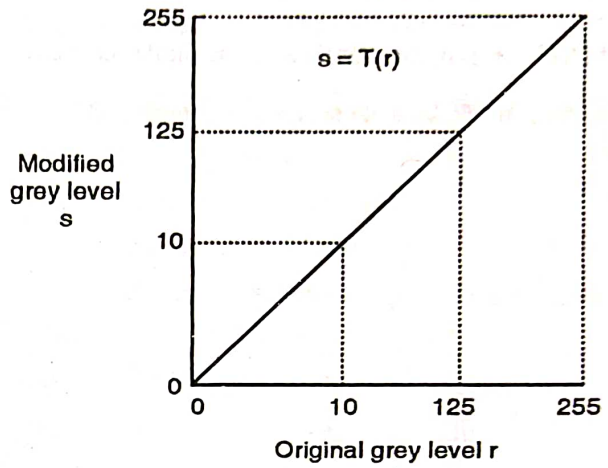


Fig. 4.3.1

(1) **Digital negative :** Digital negatives are useful in a lot of applications. A common example of digital negative is the displaying of an X-ray image. As the name suggests, negative simply means inverting the grey levels i.e. black in the original image will now look white and vice versa. The Fig. 4.3.2 is the digital negative transformation for a 8-bit image.

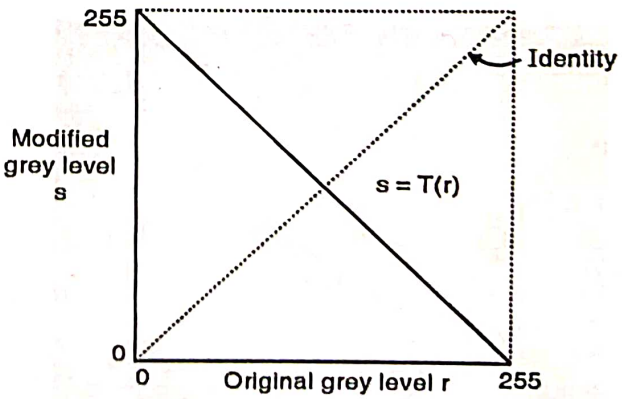


Fig. 4.3.2

The digital negative image can be obtained by using a simple transformation given by,

$$s = 255 - r \ (r_{max} = 255)$$

Hence when

$$r = 0, s = 255 \text{ and when } r = 255, s = 0.$$

In general,

$$s = (L - 1) - r \qquad \ldots(4.3.1)$$

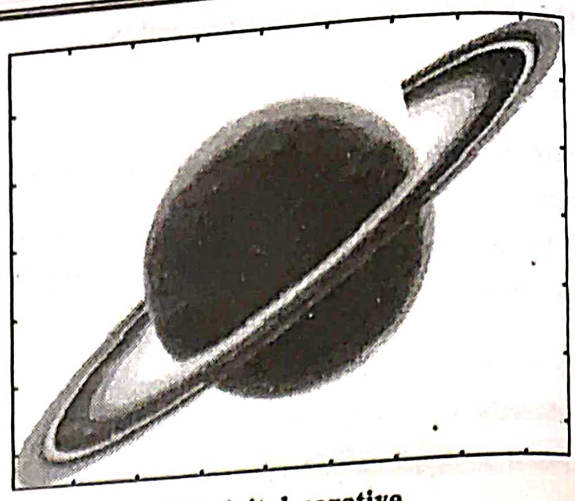
Here L is the number of grey levels. (256 in this case)

MATLAB program for finding the digital negative

```
%% MATLAB code to calculate negative %%
clear all
clc
aa=imread ('saturn.tif');
a=double (aa)
c=255; % for a 8-bit image %
b=c-a;
figure(1)
colormap(gray)
imagesc(a)
figure(2)
colormap(gray)
imagesc(b)
```

(a) Original image

Fig. 4.3.3 Continued...



(b) Digital negative

Fig. 4.3.3

(2) **Contrast stretching :** Many times we obtain low contrast images due to poor illuminations or due to wrong setting of the lens aperture. The idea behind contrast stretching is to increase the contrast of the images by making the dark portions darker and the bright portions brighter.

Fig. 4.3.4 shows the transformation used to achieve contrast stretching. In the Fig. 4.3.4, the dotted line indicates the identity transformation and the solid line is the contrast stretching transformation. As is evident from the Fig. 4.3.4, we make the dark grey levels darker by assigning a slope of less than one and make the bright grey levels brighter by assigning a slope greater than one. One can assign different slopes depending on the input image and the application. As was mentioned, image enhancement is a subjective technique and hence there is no one set of slope values that would yield the desired result.

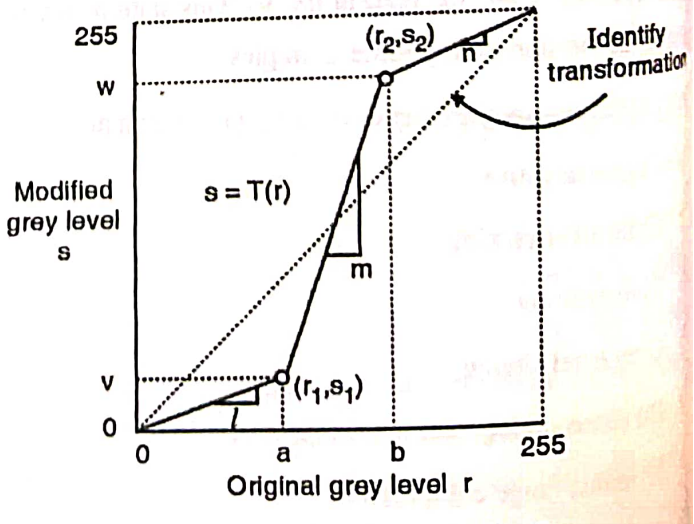


Fig. 4.3.4 : Original grey level r

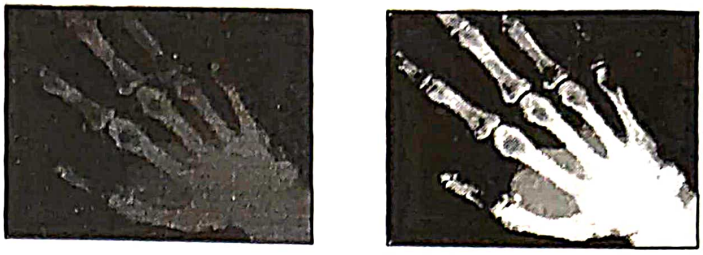
The formulation of the contrast-stretching algorithm is given below.

$$s = \begin{cases} l \cdot r & 0 \le r < a \\ m \cdot (r-a) + v & a \le r < b \\ n \cdot (r-b) + w & b \le r < L-1 \end{cases} \quad \ldots(4.3.2)$$

Where l, m and n are the slopes. It is clear from the figure that l and n are less than one while m is greater than one. The contrast stretching transformation increases the dynamic range of the modified image.

MATLAB program for contrast stretching

```
%% Contrast stretching of an image %%
% Slopes taken are 0.5, 2 and 0.5 %
clear all;
clc;
a=imread('xray1.tif')
a=double(a);
[row col]=size(a);
LT=input('Enter the lower threshold value:');
UT = input('Enter the upper threshold value:');
for x=1:1:row
for y=1:1:col
 if a(x,y)<=LT
   b(x,y)=0.5*a(x,y);
 else if a(x,y)<=UT
        b(x,y)=2*(a(x,y)-LT)+0.5*LT;
 else b(x,y)=0.5*(a(x,y)-UT)+0.5*LT+2*(UT-LT);
  end
end
 end
end
subplot(2,1,1)
imshow(uint8(a))
title('Original Image');
subplot(2,1,2)
imshow(uint8(b))
title('Image after Contrast Stretching')
```



(a) Original image (b) Contrast stretched image

Fig. 4.3.5

(3) **Thresholding :** Extreme contrast stretching yields thresholding. If we observe the contrast stretching diagram closely we notice that if the first and the last slope are made zero, and the centre slope is increased, we would get a thresholding transformation i.e. if $r_1 = r_2$, $s_1 = 0$ and $s_2 = L - 1$, we get a thresholding function. It is shown in Fig. 4.3.6.

The formula for achieving thresholding is as follows.

$$\left.\begin{array}{ll} s = 0; & \text{if } r \le a \\ s = L-1; & \text{if } r > a \end{array}\right\} \quad \ldots(4.3.3)$$

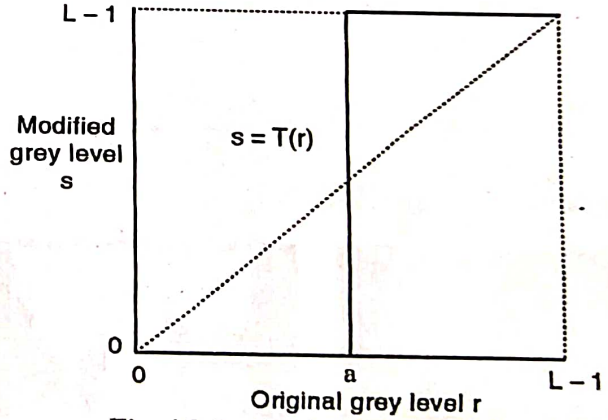


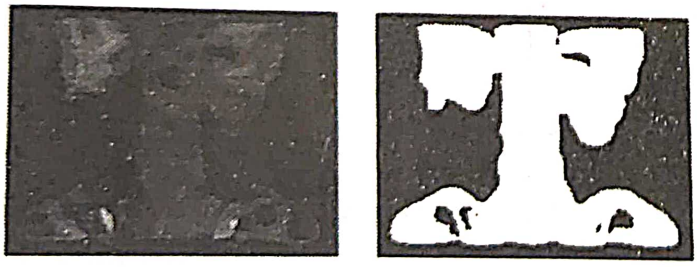
Fig. 4.3.6 : Original grey level r

Where L is the number of grey levels.

As mentioned earlier, image enhancement being a subjective phenomena, the value of a will vary from image to image and from person to person. The objective is to identify the region that he or she is interested in. An important thing to note is that a thresholded image has the maximum contrast as it has only black and white grey values.

MATLAB program for thresholding

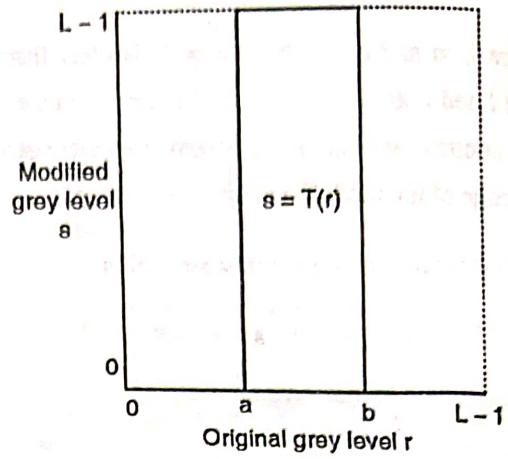
```
%%% Thresholding %%%
clear all
clc
p=imread('spine.tif');
a=p;
[row col]=size(a);
T=input('Enter value of Threshold :')
for i=1:1:row
    for j=1:1:col
        if(p(i,j)<T)
            a(i,j)=0;
        else
            a(i,j)=255;
    end
end
end
figure(1),imshow(p),
figure(2),imshow(a)
```



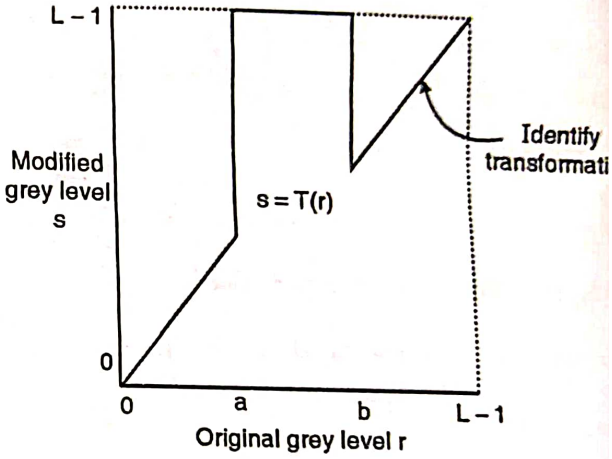
(a) Original image of spine **(b) Image obtained using threshold**

Fig. 4.3.7

(4) Grey level slicing (Intensity slicing) : What thresholding does is it splits the grey level into two parts. At times, we need to highlight a specific range of grey values like for example enhancing the flaws in an X-ray or a CT image. In such circumstances, we use a transformation known as grey level slicing. The transformation is shown in the Fig. 4.3.8(a). It looks similar to the thresholding function except that here we select a band of grey level values.



(a) Slicing without background



(b) Slicing with background

Fig. 4.3.8

This can be implemented using the formulation

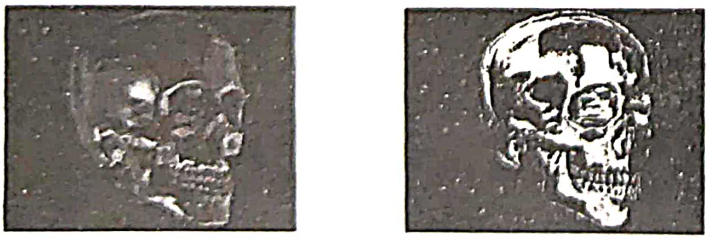
$$\left.\begin{array}{ll} s = L-1; & \text{if } a \le r \le b \\ s = 0; & \text{otherwise} \end{array}\right\} \quad ...(4.3$$

This method is known as Grey level slicing with background. This is because in this process, we have completely lost the background. In some applications, we not only need to enhance a band of grey levels but also need to retain the background. This technique of retaining the background is called as Grey-level slicing with background. The transformation is as shown in the Fig. 4.3.8(The formulation for this is

$$\left.\begin{array}{ll} s = L-1; & \text{if } a \le r \le b \\ s = r; & \text{otherwise} \end{array}\right\} \quad ...(4.3$$

MATLAB program for grey level slicing with and without background is given below,

```
%% Grey level slicing without background %%
clear all
clc
p=imread ('skull.tif');
z=double (p);
[row,col]=size(z)
for i=1:1:row
        for j=1:1:col
            if((z(i,j)>50))&&(z(i,j)<150)
            z(i,j)=255;
            else
                z(i,j)=0;
            end
        end
end
figure (1); % .....original image.
imahow (p)
figure (2);  % .....gray level slicing without background
imshow (uint8(z))
```



(a) Original image (b) Gray level slicing without background

Fig. 4.3.9

```
%% Grey level slicing with backgound %%
clear all
clc
p=imread('skull.tif');
z=double(p);
[row col]=size(p);
for i=1:1:row
    for j=1:1:col
        if(z(i,j)>50))&&(z(i,j)<150)
            z(i,j)=255;
        else
            z(i,j)=p(i,j);
        end
    end
end
figure (1); %------- original image.
imshow(p)
figure (2); %------- grey level slicing with background
imshow (uint8(z))
```

(a) Original image (b) Grey level slicing with background

Fig. 4.3.10

(5) **Bit plane slicing :** In this technique, we find out the contribution made by each bit to the final image. As mentioned earlier, an image is defined as say a 256 × 256 × 8 image. In this, 256 × 256 is the number of pixels present in the image and 8 is the number of bits required to represent each pixel. 8-bits simply means 28 or 256 grey levels.

Now each pixel will be represented by 8-bits. For example black is represented as 00000000 and white is

represented as 11111111 and between them, 254 grey levels are accommodated. In bit plane slicing, we see the importance of each bit in the final image. This can be done as follows. Consider the LSB value of each pixel and plot the image using only the LSBs. Continue doing this for each bit till we come to the MSB. Note that we will get 8 different images and all the 8 images will be binary.

Ex. 4.3.1

Given a 3 × 3 image, plot its bit planes.

1	2	0
4	3	2
7	5	2

Soln. :

Since 7 is the maximum grey level, we need only 3-bits to represent the grey levels.

Hence we will have 3-bit planes. Converting the image to binary we get,

001	010	000	1	0	0	0	1	0	0	0	0
100	011	010	0	1	0	0	1	1	1	0	0
111	101	010	1	1	0	1	0	1	1	1	0

Binary image LSB plane Middle bit plane MSB plane

```
%% MATLAB code for bit extraction %%
clear all
clc
a=imread('warne.tif');
a=double(a);
r=input('which bit image do you want to see 1=MSB 8=LSB');
[row col]=size(a);
for x=1:1:row
for y=1:1:col
c=dec2bin(a(x,y),8); % converts decimal to binary
d=c(r);
w(x,y)=double(d); %% since w is a char and cannot be plotted
if w(x,y)==49 %% since double of d will be either 49 or 48
w(x,y)=255;
else
w(x,y)=0;
end
end
end
figure(1)
imshow(uint8(a))
figure(2)
imshow(uint8(w))
```

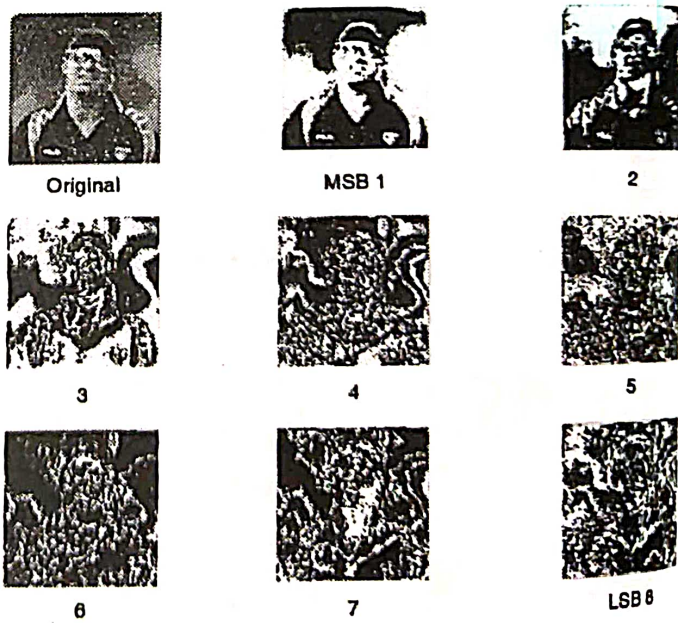


Fig. 4.3.11 : Eight images, each representing contribution of a single bit

Observing the images we come to the conclusion that the higher order bits contain majority of the visually significant data, while the lower bits contain the suitable

image compression. We can transmit only the higher order bits and remove the lower order bits. Bit plane slicing is also used in steganography.

(6) Dynamic range compression (Log transformation)

At times, the dynamic range of the image exceeds the capability of the display device. What happens is that some pixel values are so large that the other low value pixels get obscured. A simple day-to-day example of such a phenomena is that during daytime, we cannot see the stars. The reason behind this is that the intensity of the sun is so large and that of the stars is so low that the eye cannot adjust to such a large dynamic range. In image processing, a classic example of such large differences in grey levels is the Fourier spectrum (will be discussed in detail in the frequency domain enhancement technique). In the Fourier spectrum only some of the values are very large while most of the values are too small. The dynamic range of pixels is of the order of 10^6. Hence, when we plot the Fourier spectrum, we see only small dots, which represent the large values.

Something needs to be done to be able to see the small values as well. This technique of compressing the dynamic range is known as dynamic range compression. We all know that the log operator is an excellent compressing function. Hence the dynamic range compression is achieved by using a log operator. C is the normalisation constant.

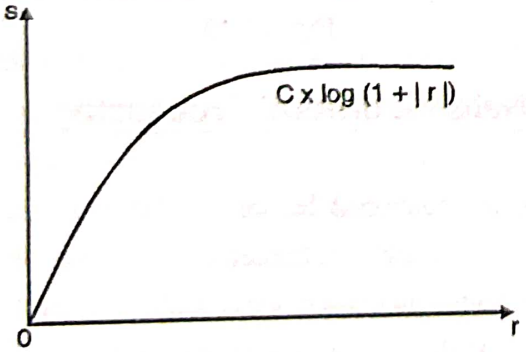
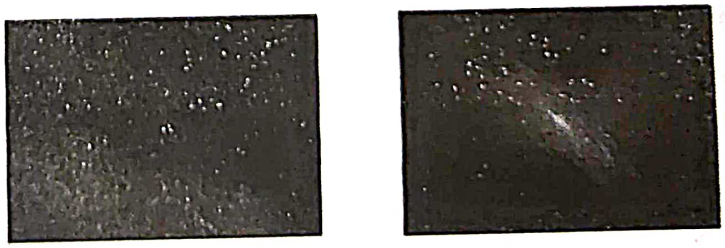


Fig. 4.3.15

MATLAB program for dynamic range compression

```
%% Dynamic range compression %%
clear all
clc
aa=imread('saturn.tif');
a=double(aa) ;
[row,col]=size(a);
for x=1:1:row
for y=1:1:col
c(x,y)=a(x,y)*((-1)^(x+y)); %% Needed to center the transform
  end
end
d=abs(fft2(c));
d_log=log(1+d);
%%% Plotting
figure(1)
colormap(gray)
imagesc(d)
figure(2)
colormap(gray)
imagesc(d_log)
```



(a) (b)

Fig. 4.3.16 : Dynamic range compression

(7) Power law transformation

The basic formula for power-law transformation is

$$g(x, y) = c \times f(x, y)^{\gamma}$$

It can also be written as $s = c\, r^{\gamma}$ ...(4.3.6)

4.4 Neighbourhood Processing

This, as mentioned before, is also a spatial domain technique in image enhancement. Unlike the point processing techniques where we consider one pixel at a time and modify it depending on our requirement, here we not only consider a pixel but also its immediate neighbours.

To cut a long story short, we change the value of the pixel f(x, y) based on the values of its 8 neighbours as shown in Fig. 4.4.1. Instead of a 3×3 neighbourhood, we could also use a 5×5 or a 7×7..... neighbourhood.

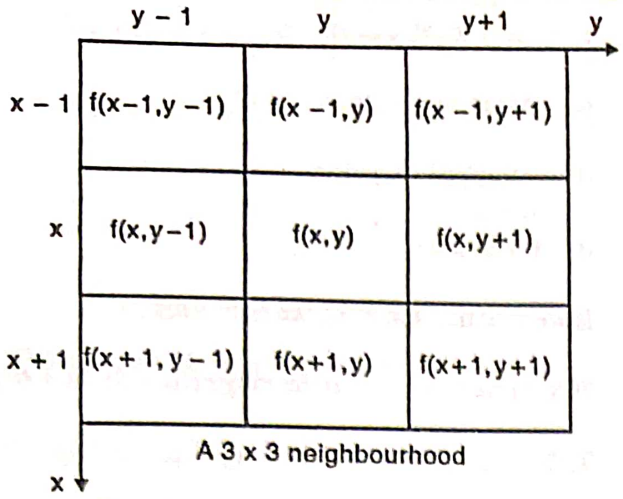


Fig. 4.4.1 : A 3 × 3 neighbourhood

w1	w2	w3
w4	w5	w6
w7	w8	w9

Fig. 4.4.2 : 3 × 3 mask

There are a lot of things that can be achieved by neighbourhood processing which are not possible with point processing. Fig. 4.4.2 shown is called a mask or a window or a template. To achieve neighbourhood processing, we place this 3 × 3 (it could also be a 5 × 5 or a 7 × 7...) mask on the image, multiply each component of the mask with the corresponding value of the image, add them up and place the value that we get, at the center.

This operation is the same as convolution. In the convolution operation, we have two signals. Of the two signals, we take one, flip it and then move it across the other signal step by step. The same thing is done here. We don't need to flip the mask as it is symmetric.

If f is the original image and g is the modified image, then

$$g(x, y) = f(x-1, y-1) \times w1 + f(x-1, y) \times w2 + f(x-1, y+1) \times w3 + f(x, y-1) \times w4 + f(x, y) \times w5 + \ldots\ldots\ldots + f(x+1, y+1) \times w9 \quad \ldots(4.4.1)$$

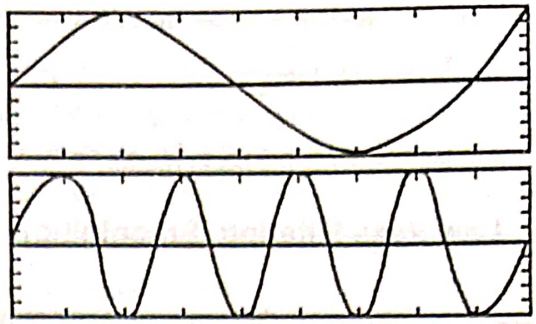


Fig. 4.4.3

Once g(x, y) is calculated, we shift the mask by one step towards the right to the next pixel. Now w5 coincides with f(x, y + 1). One of the important operations that can be achieved using neighbourhood processing is that of image filtering. We can perform low pass, high pass and band pass filtering using neighbourhood operations. Before explaining the procedure of performing filtering on images, it is imperative to understand what we mean by frequencies in an image !! We are all well versed with frequencies in 1-dimensional signals. Given two signals, we can easily distinguish between the lower frequency signal and the higher frequency signal. Refer Fig. 4.4.3.

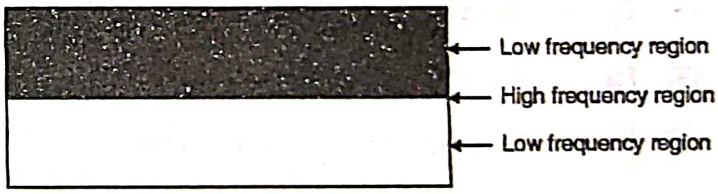


Fig. 4.4.4

We can conclude that the lower signal is of a much higher frequency, by checking the number of oscillations. That is, higher the frequency, greater are the number of oscillations. If the two signals represent voltages, then, how fast the voltages change is an indication of the frequency. The same concepts can be applied to images. In images, instead of voltages, we have grey levels. If the grey levels change slowly over a region, it is considered to have low frequency, whereas, if the grey levels change very rapidly, that region is considered to have high frequencies. Hence in images, regions where the grey levels change slowly are low frequency areas and regions which have abrupt grey level changes, are high frequency areas. Always remember this.

In most of the images, the background is considered to be a low frequency region, whereas the edges are considered to be high frequency regions. Hence low pass filtering implies removing (bluring) the edges while high pass filtering implies removing the background.

4.4.1 Low Pass Filtering (Smoothing)

Low pass filtering as the name suggests removes the high frequency content from the image. It is used to remove noise present in the image. Noise, is normally a high frequency signal and low pass filtering eliminates the noise. Before proceeding to explain the low-pass filtering technique, we shall spend some time discussing the types of noise that are fairly common in images.

4.4.2 Noise

The principal sources of noise in a digital image arise during image acquisition and during transmission. No matter how much care one takes, some amount of noise always creeps in. Based on the shapes (Probability Density Functions) of the noise, they are classified as :

(1) Gaussian noise (2) Salt and pepper noise

(3) Rayleigh noise (4) Gamma noise

(5) Exponential noise (6) Uniform noise

Of these, the first two are more common than the others. We shall explain the Gaussian and salt and pepper noise in this chapter.

(1) **Gaussian noise** : The Probability Density Function (PDF) of Gaussian noise is given by the formula,

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

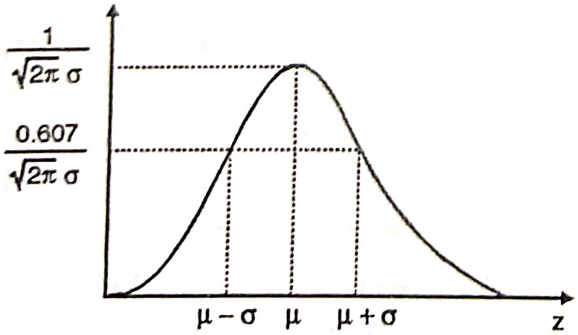


Fig. 4.4.5

$z \rightarrow$ Grey level

$\mu \rightarrow$ Mean of average value of z

$\sigma \rightarrow$ Standard deviation

$\sigma^2 \rightarrow$ Variance

If we plot this function, we notice that

70% of its value lies in the range $[(\mu - \sigma), (\mu + \sigma)]$ and

95% of its value lies in the range $[(\mu - 2\sigma), (\mu + 2\sigma)]$

Gaussian noise has a maximum value at μ and then starts falling off.

Consider the image shown in Fig. 4.4.6.

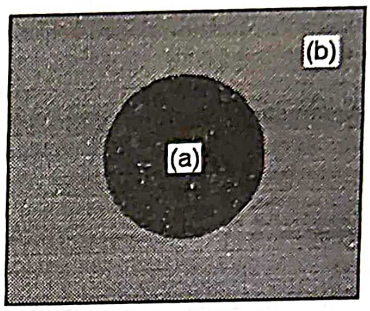


Fig. 4.4.6

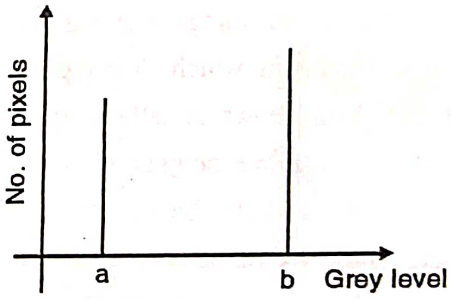


Fig. 4.4.7

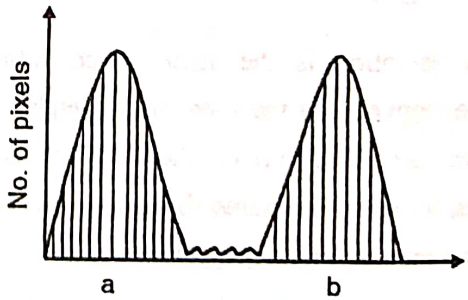


Fig. 4.4.8

There are two constant regions in the image. Hence the histogram of the image is as shown in Fig. 4.4.7. If in this image, Gaussian noise creeps in, the histogram gets modified as shown in Fig. 4.4.8. Gaussian noise arise in an image due to factors such as circuit noise, sensor noise poor illumination and high temperature.

4.4.3 Salt and Pepper Noise

Median Filter is used to remove this

The PDF of the salt and pepper noise (bipolar noise) is

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

If P_a or P_b is zero, this noise is called unipolar noise. The PDF of salt and pepper is shown in Fig. 4.4.9.

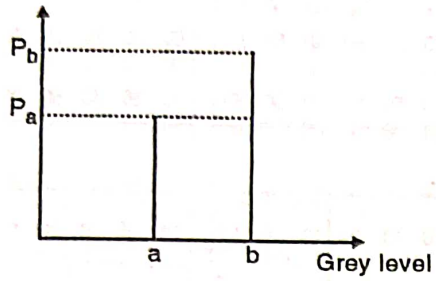


Fig. 4.4.9

Generally a and b are black and white grey levels respectively. Hence for a 8-bit image, a = 0, b = 255 because of which the noise is called salt (white) and pepper (black). Some books refer to it as speckle noise.

Take the same image as the one taken for the Gaussian example.

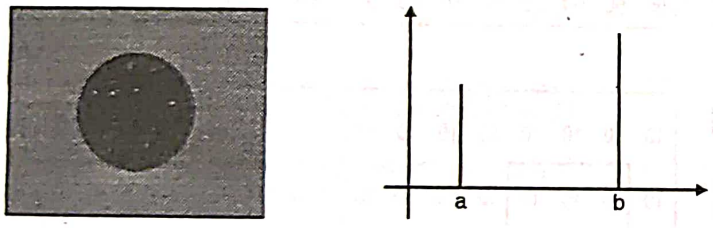


Fig. 4.4.10 (a)

When salt and pepper creeps in, the image looks like

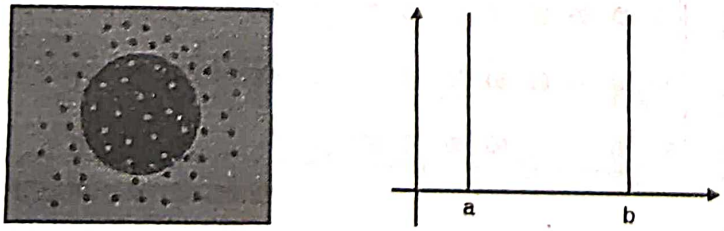


Fig. 4.4.10 (b)

Salt and pepper noise creeps into images in situations where quick transients, such as faulty switching take place.

4.4.4 Low Pass Averaging Filter (Blurring)

- Averaging / Mean Filter

If an image has Gaussian noise present in it, we use a low pass averaging filter to eliminate the noise. The frequency response of the low-pass filter along with its spatial response is shown in Fig. 4.4.11. This relationship will be proved in chapter of Image Enhancement in Frequency Domain.

*(Amp)

The edges will be removed, sharpness ↓

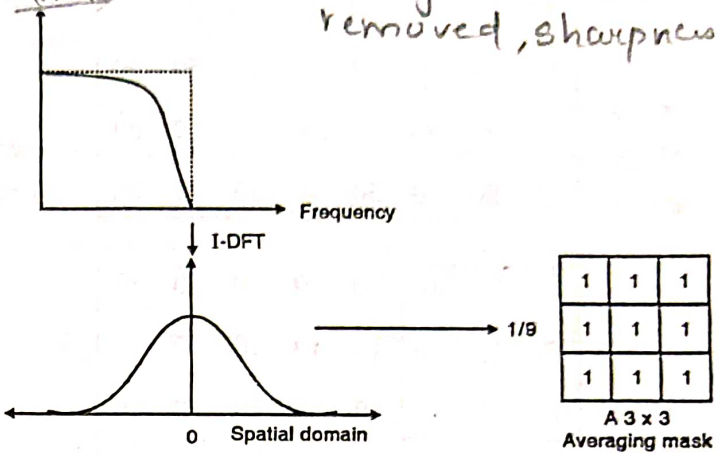


Fig. 4.4.11

From the spatial response we generate the mask that would give us the low pass filtering operation. One important thing to note from the spatial response is that all the coefficients are positive. The standard low pass averaging mask (3 × 3) is given above. As the name suggests, each element of the mask is the average value.

We could also use a 5 × 5 or a 7 × 7 mask as per our requirement.

1/25

1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1

A 5 x 5 Averaging mask

Fig. 4.4.12

Let us take an example of a 3 × 3 mask on a pseudo-image and see how it eliminates the edges. Consider a 8 × 8 size image. It is clear that the image has a single

edge between 10 and 50. To get rid of this edge (high frequency) we use a 3 × 3 averaging mask.

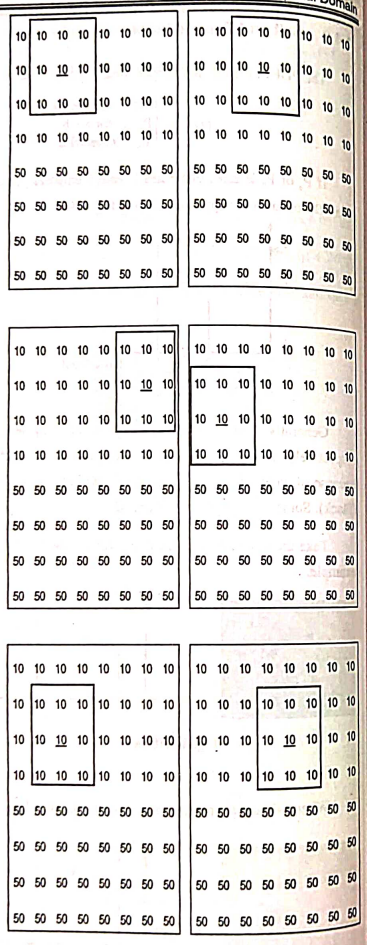
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50

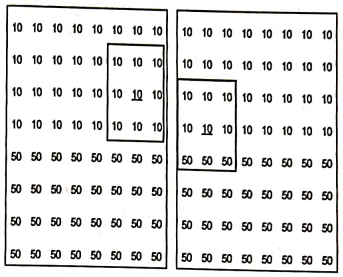
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50

We place a 3 × 3 mask on this image. We start from the left hand top corner. We cannot work with the borders and hence are normally left as they are.

We then multiply each component of the image with the corresponding value of the mask. Since all the nine values of the image are 10, the average is also ten and the centre pixel (the underlined pixel) remains ten. We now shift the mask towards the right till we reach the end of the line and then move it downwards. Some of the mask positions are shown here.

Note : The resultant should be written in a new matrix (image).





The last one in the sequence being

10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50

The result of convolving the image with the 3 × 3 averaging mask is shown.

10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
23.3	23.3	23.3	23.3	23.3	23.3	23.3	23.3
36.6	36.6	36.6	36.6	36.6	36.6	36.6	36.6
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50

As we notice, the low frequency regions have remained unchanged, but the sharp edge between 10 and 50 has become blurred. The transition has reduced. Now from 10, the grey level changes to 23.3 (23 after rounding off), from 23 it changes to 36.6 (36 after rounding off) and finally from 36 it changes to 50. Hence we can say that the sharp edge has become blurred. Check for yourself, what would happen if you took a bigger mask, say, 5 × 5. (You will have to take a bigger image to test your results).

These kinds of averaging filters are excellent when the image contains Gaussian noise. It achieves filtering by blurring the noise. Some of the other low pass averaging masks are shown.

$\frac{1}{6}\times$

0	1	0
1	2	1
0	1	0

$\frac{1}{10}\times$

1	1	1
1	2	1
1	1	1

MATLAB program for low pass filtering along with the images with Gaussian noise

```
%% Low pass filter used on an image with Gaussian noise
clear all
clc
aa=imread('xray.tif'); % size 600×800
f=double(aa);
ab=imnoise(aa,'gaussian'); % adding noise
a=double(ab);
w=[1 1 1; 1 1 1; 1 1 1]/9
[row col]=size(a);
for x=2:1:row-1
    for y=2:1:col-1
        a1(x,y)= w(1)*a(x-1,y-1)+w(2)*a(x-1,y)+w(3)*...
        a(x-1,y+1)+w(4)*a(x,y-1)+w(5)*a(x,y)+w(6)*...
        a(x,y+1)+w(7)*a(x+1,y-1)+w(8)*a(x+1,y)+w(9)*...
```

```
subplot(3,1,3)
imshow(uint8(c))
title ('Low Pass Filtered Image using MATLAB');
```

4.4.5 Low Pass Median Filtering

The averaging filter removes the noise by blurring it till it is no longer seen. But in the process, it also blurs the edges. Bigger the averaging mask more is the blurring. There are times when the image contains salt and pepper noise. If we use an averaging filter to remove the same, it will blur the noise but it would also ruin the edges. Hence when we need to eliminate salt and pepper noise, we work with a non-linear filter known as the Median filter. They are also called order-statistic filters because their response is based on the ordering or ranking of the pixels contained within the mask.

In this case, we use a mask similar to the averaging filter except that the mask has no values. So it's like working directly with the 8 neighbours of the centre pixel.

The steps to perform median filtering are as follows :

(1) Assume a 3×3 empty mask.

(2) Place the empty mask at the left hand corner.

(3) Arrange the 9 pixels in ascending or descending order.

(4) Choose the median from these nine values.

(5) Place this median at the centre.

(6) Move the mask in a similar fashion to the averaging filter.

In median filtering, the grey level of the centre pixel is replaced by the median value of the neighbourhood. Always write the resultant in a new matrix (image). We shall explain median filtering with an example. In the image shown below let 250 be the salt and pepper noise. If we use an averaging filter, the noise would spread out and the edges would also end up getting blurred. If we only want to remove the noise without disturbing the edges, we use a median filter.

10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	250	10	10	10	10	10	10
10	10	10	10	10	10	10	10
50	50	50	50	50	250	50	50
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50

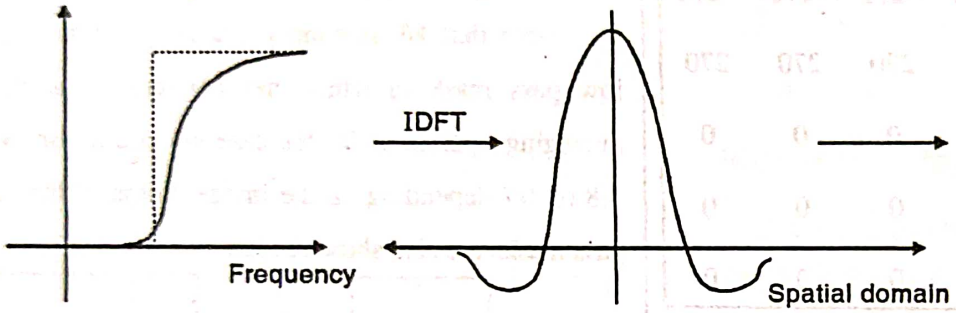
Just like the earlier case, we start by placing the mask at the left hand corner. We again ignore the borders.

10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	250	10	10	10	10	10	10
10	10	10	10	10	10	10	10
50	50	50	50	50	250	50	50
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50

10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	250	10	10	10	10	10	10
10	10	10	10	10	10	10	10
50	50	50	50	50	250	50	50
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50
50	50	50	50	50	50	50	50

4.5 Highpass Filtering → Edges enhanced

Highpass filtering eliminates the low frequency regions while retaining or enhancing the high frequency components. An image, which is high-passed, would have no background (as background are low frequency regions) and would have enhanced edges. Hence high pass filters are used to sharpen blurred images. Once we have understood as to how a mask moves over the entire image for the averaging filter, the same applies to the high pass filter as well. All that needs to be changed are the mask coefficients. The frequency response and the spatial response of a high pass filter are shown in Fig. 4.5.1. This will be proved in chapter of Image Enhancement in Frequency Domain.



−1	−1	−1
−1	8	−1
−1	−1	−1

A 3 x 3 high pass mask

Fig. 4.5.1

As seen from the spatial response, it is clear that the mask coefficients have to be such that they have a positive value at the centre and negative values at the periphery.

One of the high pass masks is shown in Fig. 4.5.1.

The important thing to note is that the sum of the coefficients of the high pass mask has to be equal to zero. This is because, when we place this mask over the low frequency regions, the result should be zero.

Let us take an example of a pseudo-image.

10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
100	100	100	100	100	100	100	100
100	100	100	100	100	100	100	100
100	100	100	100	100	100	100	100
100	100	100	100	100	100	100	100

By now we know what to expect when we do a high pass operation. The background, which is low frequency, gets eliminated while the edges get enhanced.

We move the mask in a similar fashion as in the low pass filter example. The output that we get is shown below. Note that the background has been eliminated and we get all zero values. This is because the sum of the mask coefficients is zero.

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
–270	–270	–270	–270	–270	–270	–270	–270
270	270	270	270	270	270	270	270
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

There are two issues to be dealt with, when working with high pass masks.

(1) As we can see, there are negative values in the output image. Pixel values *cannot be negative*. They always start with zero. Hence we need to get rid of the negative values. If we consider the mod operation, do you think it would solve the problem ? NO. The reason for that is –270 is a value that is lower than zero, that is, it is supposed to be darker than the darkest value i.e. zero. If we now take the mod value i.e. MOD [–270], we get +270, which is definitely not darker than the zero value. What would happen is that all large negative values would be shown as bright spots. Hence taking the mod of negative values will distort the image grey levels. A simple way to get rid of this problem is to let all negative values be zero. Hence, –270 would be written as 0.

(2) The values of the original image at the edge are very large and there is a tendency of them going out of range due to the centre weight of +8. We always encounter this problem in practice. To eliminate this problem, we use a mask with a scaling function.

Hence for practical purposes we use a mask shown below.

$$\frac{1}{9}\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

3×3

High pass mask

Note that 1/9 is simply a scaling function unlike in low pass mask in which the 1/9 term is a part of the averaging operation. In this case we could also work with 1/8 or 1/7 depending on the image. Some of the other high pass masks are also shown below.

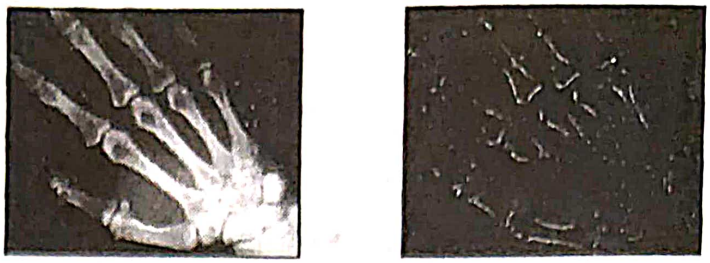
$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \qquad \begin{bmatrix} -1 & -2 & -1 \\ -2 & 12 & -2 \\ -1 & -2 & -1 \end{bmatrix}$$

The result using the first mask and putting negative values as zeros is

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
30	30	30	30	30	30	30	30
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

MATLAB program for high pass filtering

```
%% High pass filtering %%
clear all
clc
aa=imread ('xray.tif');
a=double (aa);
[row col]=size(a);
w=[-1 -1 -1; -1 8 -1, -1 -1 -1]
for x=2:1:row-1
    for y=2:1:col-1
        al(x,y)= w(1)*a(x-1,y-1)+w(2)*a(x-1,y)+w(3)* ...
        a(x-1,y+1)+w(4)*a(x,y-1)+w(5)*a(x,y)+w(6)* ...
    a(x,y+1)+w(7)*a(x+1,y-1)+w(8)*a(x+1,y)+w(9)* ...
        a(x+1,y+1);
    end
end
figure(1)
imshow(uint8(a))
figure(2)
imshow(uint8(al))
```



(a) Original image (b) High pass filtered image

Fig. 4.5.2

4.6 High-Boost Filtering

As can be seen, the high pass filter gives great results. But there is one problem. It gets rid of the complete background. There are times, when we need to enhance the edges but also retain some of the background. To do that, we use a modified version of the high pass filter known as High-Boost filtering.

In High-Boost filtering, we pass some of the background along with the high frequency content.

We know that High pass = Original – Low pass

To pass some of the background, we multiply the original image with a multiplicative factor A. This gives us high boost filtering.

Hence,

High Boost = (A) original – Low pass

= (A –1) Original + Original – Low pass

High Boost = (A –1) Original + High pass

If A = 1, then

High Boost = High pass

If A > 1, then some of the original signal is added back to the high pass result. This process restores some of the background into the high passed image. This technique is also known as unsharp masking. The mask coefficients for high boost filtering are shown below.

Here X = 9A – 1

Hence if A = 1, X = 8 which is a high pass mask.

$\frac{1}{9}$

-1	-1	-1
-1	X	-1
-1	-1	-1

3 × 3 High boost mask

If A = 1.1, X = 8.9.

We have a mask which is as shown

$\frac{1}{9}$

-1	-1	-1
-1	8.9	-1
-1	-1	-1

3 × 3 High boost mask

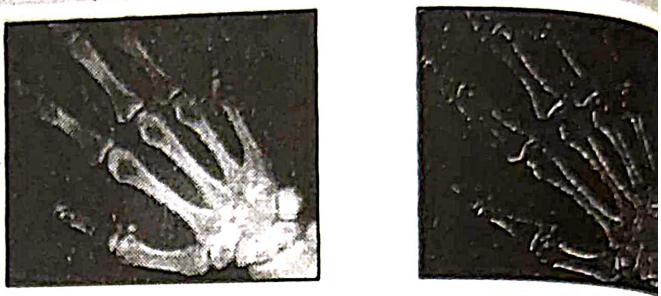
We can select different values of A and see the difference.

Use the same pseudo image as the one used in the high pass example to see the results. What you will notice is that places which had a zero in the high pass example will now have some positive values. Hence it does not eliminate the background completely. This technique is one of the basic tools that is used in the printing industry.

There are other neighbourhood techniques like Robert's filtering, Sobel's filtering and Prewitt's filtering, which are similar to the methods discussed above. They form a separate chapter called Segmentation and hence would be discussed later.

MATLAB program for high-boost filtering

```
%% High boost filtering
clear all
clc
aa=imread ('xray.tif');
a=double(aa);
[row col]=size (a);
w=[-1 -1 -1; -1 8.9 -1; -1 -1 -1]; %% High boost mask
for x=2:1:row-1
    for y=:1:col-1
a1(x,y)= w(1)*a(x-1,y-1)+w(2)*a(x-1,y)+w(3)*
a(x-1,y+1)+w(4)*a(x,y-1)+w(5)*a(x,y)+w(6)*
a(x,y+1)+w(7)*a(x+1,y-1)+w(8)*a(x+1,y)+w(9)*
a(x+1,y+1);
    end
end
figure(1)
imshow(uint8(a))
figure(2)
imshow(uint8(a1))
```



(a) Original (b) High boost

Fig. 4.6.1

The generalised MATLAB program for high filtering using any mask size is given next.

```
%% High pass filtering of an image %%
clear all;
clc;
a=imread('mri2.tif');
a=double(a);
[row col]=size(a);
m = input('Enter the mask size:');
for i=1:1:m
    for j=1:1:m
        w(i,j)=-1
    end
end
s=(m+1)/2;
w(s,s)=m^2-1;
```

```
for x=1:1:row
    for y=s:1:col
        b(x,y)=a(x,y);
    end
end
for x=s:1:row-s
    for y=s:1:col-s
        b(x,y)=0;
    end
end
for x=s:1:row-s
    for y=s:1:col-s
        for i=1:1:m
            for j=1:1:m
                b(x,y)=a(x-s+i,y-s+j)*w(i,j)+b(x,y);
            end
        end
        if b(x,y) < 0
            b(x,y)=0;
        end
    end
end
n=0
for x=s:1:row-s
    for y=s:1:col-s
        if b(x,y) > n
            n=b(x,y);
        end
    end
end
c=255/n;
for x=s:1:row-s
    for y=s:1:col-s
        b(x,y)=c*b(x,y);
    end
end
subplot(2,1,1)
imshow(uint8(a))
title('Original Image');
subplot(2,1,2)
imshow(uint8(b))
title('High Pass Filtered Image')
```

4.7 Zooming

Another important application in image enhancement where the spatial domain neighbourhood operation is used image zooming. Zooming of images is not something that is new to you. If you have used Microsoft paint or Photoshop editor, you would be aware that there is an option which allows us to zoom an image by 25%, 50%, and 100% .....Have you ever imagined as to how an image that looked small can suddenly look so big ? The technique is quite simple and after reading the next couple of pages, you would be able to do it yourself.

Zooming can be carried out using two different methods,

(1) Replication (2) Interpolation

We will discuss both these methods in detail. We shall first start with zooming using replication.

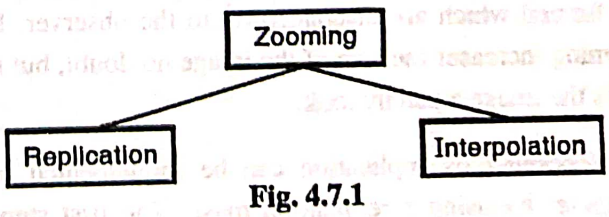


Fig. 4.7.1

4.7.1 Replication

In replication, we simply replicate each pixel and then replicate each row. Consider the pseudo-image shown below.

1	2	3	4
5	6	7	8
9	8	6	7
0	1	2	3

As stated earlier, we start from the first row. We replicate each pixel and then replicate each row. The first row now looks like

1 1 2 2 3 3 4 4

We now replicate this row to get

1 1 2 2 3 3 4 4

1 1 2 2 3 3 4 4

Performing this operation on the entire image we get

1	1	2	2	3	3	4	4
1	1	2	2	3	3	4	4
5	5	6	6	7	7	8	8
5	5	6	6	7	7	8	8
9	9	8	8	6	6	7	7
9	9	8	8	6	6	7	7
0	0	1	1	2	2	3	3
0	0	1	1	2	2	3	3

Hence a 4 × 4 image is zoomed to a 8 × 8 image. This method can be repeated to get bigger images. Remember, this is an image enhancement technique and hence no new data is added. In the zoomed pseudo-image, we observe that as we increase the size of the image, clusters of grey levels are formed which are disconcerting to the observer. Hence zooming increases the size of the image no doubt, but it also gives the image a patchy look.

Zooming by replication can be implemented on the computer by using a replication mask. The first step is to interlace the original image with zeros. This is known as zero interlacing. In this we add zeros after every pixel and then add a complete row of zeros. Adding zeros to every other pixel of the first row we get,

1 0 2 0 3 0 4 0

This is also known as zero interlacing the colum we add zeros at every other column. Now inserting full of zeros gives us. This is known as zero interlaci rows as we add zeros at every other row.

1 0 2 0 3 0 4 0

0 0 0 0 0 0 0 0

1	0	2	0	3	0	4	0
0	0	0	0	0	0	0	0
5	0	6	0	7	0	8	0
0	0	0	0	0	0	0	0
9	0	8	0	6	0	7	0
0	0	0	0	0	0	0	0
0	0	1	0	2	0	3	0
0	0	0	0	0	0	0	0

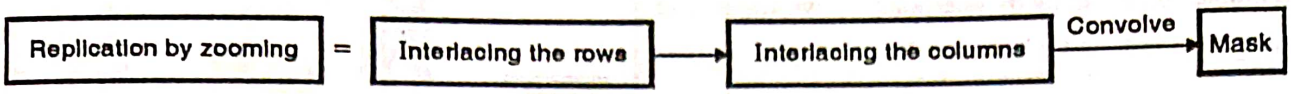
This image is known as the zero interlaced imag this image, we run a replication mask given below to g zoomed image.

1	1
1	1

The final zoomed image is shown below,

1	1	2	2	3	3	4	4
1	1	2	2	3	3	4	4
5	5	6	6	7	7	8	8
5	5	6	6	7	7	8	8
9	9	8	8	6	6	7	7
9	9	8	8	6	6	7	7
0	0	1	1	2	2	3	3
0	0	1	1	2	2	3	3

Note : To get the above result, we need to add a r zeroes at the top and a column of zeroes a beginning of the zero interlaced image ma 9 × 9.



As mentioned earlier, zooming by replication gives the final image a patchy look since clusters of grey levels are formed. This can be substantially reduced by using a better method of zooming known as interpolation.

4.7.2 Linear Interpolation

In this method, instead of replicating each pixel, average of the two adjacent pixels along the rows is taken and placed between the two pixels. The same operation is then performed along the columns. This operation is done on the interlaced image.

1	0	2	0	3	0	4	0
0	0	0	0	0	0	0	0
5	0	6	0	7	0	8	0
0	0	0	0	0	0	0	0
9	0	8	0	6	0	7	0
0	0	0	0	0	0	0	0
0	0	1	0	2	0	3	0
0	0	0	0	0	0	0	0

Interpolaiton along rows is given by,

$$v_1(m, 2n) = u(m, n); \quad 0 \le m \le M-1, 0 \le n \le N-1$$

$$v_1(m, 2n+1) = \frac{1}{2}[u(m, n) + u(m, n+1)];$$

$$0 \le m \le M-1; \quad 0 \le n \le N-1$$

Interpolation of the column is given by,

$$v(2m, n) = v_1(m, n);$$

$$v(2m+1, n) = \frac{1}{2}[v_1(m, n) + v_1(m+1, n)];$$

$$0 \le m \le M-1; 0 \le n \le N-1$$

Hence the first row now becomes the modified image is as shown

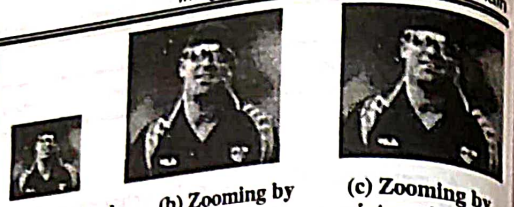
1	1.5	2	2.5	3	3.5	4	2

1	1.5	2	2.5	3	3.5	4	2
0	0	0	0	0	0	0	0
5	5.5	6	6.5	7	7.5	8	4
0	0	0	0	0	0	0	0
9	8.5	8	7	6	6.5	7	3.5
0	0	0	0	0	0	0	0
0	0.5	1	1.5	2	2.5	3	1.5
0	0	0	0	0	0	0	0

This is the first step. We now proceed to find the average values along the columns. Hence the final image is as shown. In practice, we need to round of the pixel values.

1	1.5	2	2.5	3	3.5	4	2
3	3.5	4	4.5	5	5.5	6	3
5	5.5	6	6.5	7	7.5	8	4
7	7	7	6.25	6.5	7	7.5	3.25
9	8.5	8	7	6	6.5	7	3.5
4.5	4.5	4.5	4.25	4	4.5	5	2.5
0	0.5	1	1.5	2	2.5	3	1.5
0	0.25	0.5	0.75	1	1.25	1.5	0.75

If we compare the results of replication and interpolation, it is clear that the patchiness that was present in the replicated image is much less in the interpolated image. Hence we can conclude that zooming by interpolation is more effective than zooming by replication.



(a) Original (b) Zooming by replication (c) Zooming by interpolation

Fig. 4.7.2

4.8 Solved Examples

Ex. 4.8.1

Obtain the digital negative of the following 8 Bits Per Pixel - BPP image.

121	205	217	156	151
139	127	157	117	125
252	117	236	138	142
227	182	178	197	242
201	106	119	251	240

Soln. :

It is known that it is a 8-bit image. Hence the number of grey levels that this image can hold is $2^8 = 256$

$\therefore L = 256.$

Hence the minimum grey level is 0

while the maximum grey level is 255

$$s(x,y) = (L-1) - r(x,y)$$

$$= (256-1) - r(x,y)$$

$$s(x,y) = 255 - r(x,y)$$

We get the digital negative using the above equation.

134	50	38	99	104
116	128	98	138	130
3	138	19	117	113
28	73	77	58	13
54	149	136	4	15

Ex. 4.8.2

For a given image find -

(i) Digital negative of an image.

(ii) Bit plane slicing.

4	3	2	1
3	1	2	4
5	1	6	2
2	3	5	6

Soln. :

We assume the image to be 3-bit

(i) Digital negative

$$s = (L-1) - r$$

$$\text{or } g(x,y) = (L-1) - f(x,y)$$

$$\text{Here } L = 2^3 = 8$$

$$\therefore s = 7 - r$$

$$\text{i.e., } g(x,y) = 7 - f(x,y)$$

$\therefore$ The digital negative of the image is

3	4	5	6
4	6	5	3
2	6	1	5
5	4	2	1

(ii) Bit plane slicing

In this, we convert the given image into binary and then separate the planes.

We assume the image to be 3 bit.

100	011	010	001
011	001	010	100
101	001	110	010
010	011	101	110

Separating the bit planes we get

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

MSB plane

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

Center bit plane

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

LSB plane

Ex. 4.8.3

For following image find :

(i) Contrast stretching $r_2 = 5, r_1 = 3, s_2 = 6, s_1 = 2$.

$f(x,y) =$

4	3	2	1
3	1	2	4
5	1	6	2
2	3	5	6

Soln. :

(i) Contrast stretching $r_2 = 5, r_1 = 3, s_2 = 6, s_1 = 2$:

$f(x,y) =$

4	3	2	1
3	1	2	4
5	1	6	2
2	3	5	6

The contrast stretching transformation is shown in Fig. P. 4.8.3.

From the values of r_1 and s_1 we can find the slope of α. In a similar manner, we can find the value of β using the values of r_1, r_2 and s_1, s_2.

From the values of r_2 and s_2, we can find the value of γ.

$$\alpha = \frac{s_1}{r_1} = \frac{2}{3} = 0.66$$

$$\beta = \frac{y_2 - y_1}{x_2 - x_1} = \frac{6-2}{5-3} = 2$$

Finally, $$\gamma = \frac{y_2 - y_1}{x_2 - x_1} = \frac{7-6}{7-5} = 0.5$$

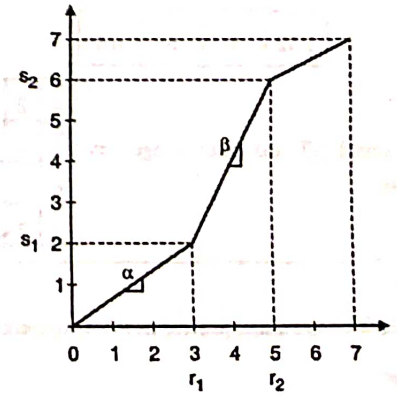


Fig. P. 4.8.3

Hence we have the required slopes to compute contrast stretching. The contrast stretching formula is given below.

$$s = \begin{cases} \alpha r & 0 \le r < 3 \\ \beta \cdot (r - r_1) + s_1 & 3 \le r < 5 \\ \gamma (r - r_2) + s_2 & 5 \le r < 7 \end{cases}$$

Here $r_1 = 3$, $r_2 = 5$, $s_1 = 2$, $s_2 = 6$

$\alpha = 0.66$, $\beta = 2$, $\gamma = 0.5$

We make a table of values for r and s using the contrast stretching formula.

r	s
0	$s = \alpha r = 0.66 \times 0 = 0$
1	$s = \alpha r = 0.66 \times 1 = 0.66$
2	$s = \alpha \cdot r = 0.66 \times 2 = 1.32$
3	$s = \beta (r - r_1) + s_1 = 2(3-3) + 2 = 2$
4	$s = \beta (r - r_1) + s_1 = 2(4-3) + 2 = 4$
5	$s = \gamma (r - r_2) + s_2 = 0.5(5-5) + 6 = 6$
6	$s = \gamma (r - r_2) + s_2 = 0.5(6-5) + 6 = 6.5$
7	$s = \gamma (r - r_2) + s_2 = 0.5(7-5) + 6 = 7$

Hence the contrast stretched image is

4	2	1.32	0.66
2	0.66	1.32	4
6	0.66	6.5	1.32
1.32	2	6	6.5

If we round off, the final image would be ⇒

4	2	1	1
2	1	1	4
6	1	7	1
1	2	6	7

Ex. 4.8.4

For the 3-bit 4 × 4 size image, perform the following operations :

(i) Negation

(ii) Thresholding with T = 4

(iii) Intensity level slicing with background $r_1 = 2$ and $r_2 = 5$

(iv) Bit plane slicing for MSB and LSB planes

(v) Clipping with $r_1 = 2$ and $r_2 = 5$

1	2	3	0
2	4	6	7
5	2	4	3
3	2	6	1

Soln. :

Let the given image be f(x, y)

(I) Negation

$$s = (L-1) - r$$

OR

$$g(x, y) = (L-1) - f(x, y)$$

Since the image is 3-bit, $L = 2^3 = 8$

$\therefore L - 1 = 7$

$\therefore g(x, y) = 7 - f(x, y)$

Hence the digital negative of the image is

$$g(x, y) = \begin{array}{|c|c|c|c|} \hline 6 & 5 & 4 & 7 \\ \hline 5 & 3 & 1 & 0 \\ \hline 2 & 5 & 3 & 4 \\ \hline 4 & 5 & 1 & 6 \\ \hline \end{array}$$

(II) Thresholding with T = 4

The thresholding function is shown below along with thresholding formula.

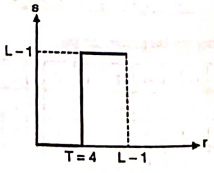


Fig. P. 4.8.4

$$g(x, y) = \begin{cases} 0 & \text{if } f(x, y) \le 4 \\ L-1 & \text{if } f(x, y) \ge 4 \end{cases}$$

Here

$L - 1 = 7$

$\therefore$ The thresholded image is shown below

7	7	7	7
7	0	0	0
0	7	0	7
7	7	0	7

(III) Bit plane slicing

We convert the given image into binary and then separate the plane. Since the given image is 3-bit, the binary image would be as shown.

001	010	011	000
010	100	110	111
101	010	100	011
011	010	110	001

Separating the bit planes we obtain.

0 0 0 0	0 1 1 0	1 0 1 0
0 1 1 1	1 0 1 1	0 0 0 1
1 0 1 0	0 1 0 1	1 0 0 1
0 0 1 0	1 1 1 1	1 0 0 1
MSB plane	Center bit plane	LSB plane

(IV) Clipping with $r_1 = 2$, $r_2 = 5$

The clipping transformation is shown in Fig. P. 4.8.4(a).

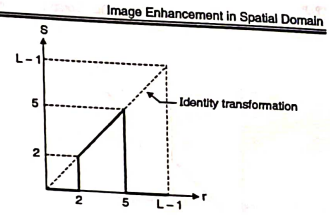


Fig. P. 4.8.4(a)

$$s = \begin{cases} r & 2 \le r \le 5 \\ 0 & \text{otherwise} \end{cases}$$

Hence the final clipped image is shown below

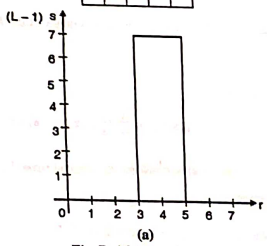
0	2	3	0
2	4	0	0
5	2	4	3
3	2	0	0

Ex. 4.8.5

Perform intensity level (grey level) slicing on the 3 BPP image. Let $r_1 = 3$ and $r_2 = 5$. Draw the modified image using with background and without background transformations.

Soln. : Let us draw the grey level transformation.

2	1	2	2	1
2	3	4	5	2
6	2	7	6	0
2	6	6	5	1
0	3	2	2	1



(a)

Fig. P. 4.8.5 Continued....

We convolve this image with a standard 3 × 3 Low pass and a High pass mask.

(i) Low Pass

$$\text{Low pass mask } h_1(x,y) = \frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$g_1(x,y) = f(x,y) * h_1(x,y) = \begin{bmatrix} 23.77 & 30.66 & 23.66 \\ 30.88 & 41.22 & 30.66 \\ 24.11 & 30.88 & 23.667 \end{bmatrix}$$

(ii) High Pass

$$\text{High pass mask } h_2(x,y) = \frac{1}{9}\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

$$g_2(x,y) = f(x,y) * h_2(x,y) = \begin{bmatrix} 6.22 & 0.33 & 8.33 \\ 2.11 & 78.77 & -0.667 \\ 7.889 & 1.11 & 7.33 \end{bmatrix}$$

We now show that High pass = Original – Low pass

$$\text{Original} - \text{Lowpass} = \begin{bmatrix} 30 & 31 & 32 \\ 33 & 120 & 30 \\ 32 & 32 & 31 \end{bmatrix} - \begin{bmatrix} 23.77 & 30.66 & 23.66 \\ 30.88 & 41.22 & 30.66 \\ 24.11 & 30.88 & 23.667 \end{bmatrix}$$

$$= \begin{bmatrix} 6.22 & 0.33 & 8.33 \\ 2.11 & 78.77 & -0.667 \\ 7.889 & 1.11 & 7.33 \end{bmatrix}$$

This is the same as the High pass result. Therefore,

High pass = Original – Low pass

Soln. :

(i) Thresholding (T = 4)

Since the given image is 3-bit, $L = 2^3 = 8$

The transformation for thresholding is shown in Fig. P. 4.8.14.

Here,

$s = L-1 = 7 \quad r \geq 4$

$s = 0 \quad r < 4$

∴ The final image would be,

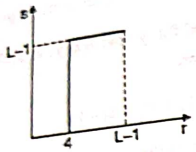


Fig. P. 4.8.14

7	0	0	0
0	0	7	7
7	0	0	0
0	7	7	7

(ii) Intensity level slicing with background for $r_1 = 2, r_2 = 5$

The transformation for intensity level slicing is shown in Fig. P. 4.8.14(a).

Here,

$s = L-1 = 7 \quad 2 \leq r \leq 5$

$s = r \quad$ otherwise

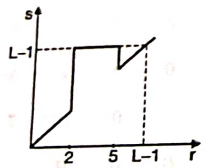


Fig. P. 4.8.14(a)

Hence all values between 2 and 5 are made equal to 7 and the remaining values remain unchanged. Hence the final result is,

7	7	7	0
1	7	7	7
7	7	7	1
7	7	6	7

(iii) Bit plane slicing

We convert the image to binary. Since there are 8 grey levels (0 to 7), we require 3 bits to represent each pixel.

4	2	3	0
1	3	5	7
5	3	2	1
2	4	6	7

Binary ⟶

100	010	011	000
001	011	101	111
101	011	010	001
010	100	110	111

The LSB and MSB planes are shown below.

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

LSB plane

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

MSB plane

(iv) Negation

The transformation for negation is shown in Fig. P. 4.8.14(b).

Here,

$s = (L-1) - r$

$s = 7 - r$

Here the grey levels get inverted.

∴ The final image is as shown below.

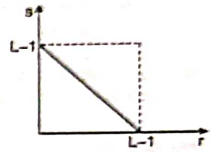
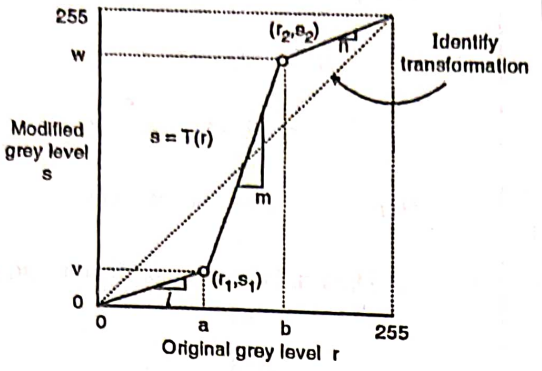
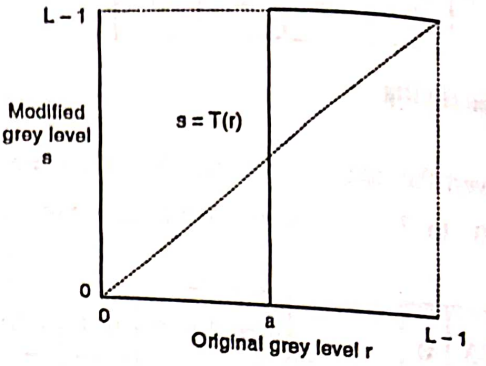


Fig. P. 4.8.14(b)

3	5	4	7
6	4	2	0
2	4	5	6
5	3	1	0

4.9 Comparison of Contrast Stretching and Thresholding

Sr. No.	Contrast Stretching	Thresholding
1.	Contrast stretching increases the dynamic range by making the dark pixels darker and bright pixels brighter.	Thresholding also increase that dynamic range by setting a threshold. All values less than the threshold are made black while values greater than the threshold are made equal to white.
2.	The formula of the contrast-stretching algorithm is given below : $s = \begin{cases} l \cdot r & 0 \leq r < a \\ m \cdot (r-a) + v & a \leq r < b \\ n \cdot (r-b) + w & b \leq r < L-1 \end{cases}$	The formula for achieving thresholding is as follows : $s = 0$; if $r \leq a$ $s = L-1$; if $r > a$

Sr. No.	Contrast Stretching	Thresholding
3.	The transformation for contrast stretching is shown below : 	The transformation for thresholding is shown below : 
4.	We can adjust the contrast in an image by changing the values of the slope.	Thresholding is extreme contrast stretching. In other words, a thresholded image has maximum contrast.
5.	The final result of contrast stretching gives us a grey scale image. The output image has different shades of grey.	The final result of thresholding gives us a binary image. Which means that the thresholded image has only black and white values.

Summary

In this chapter, we introduce the term spatial domain and image enhancement in the spatial domain. Different image enhancement techniques that improve the subjective quality of the image are explained. The methods described here can be classified into two broad groups : point processing and neighborhood processing. The concept of frequencies in an image is explained using simple illustrations. Basic filtering operations such low pass filtering and high pass filtering are presented. Advantages of each of the operations are stated using MATLAB codes. Any particular technique that may be applied depends on the subjective quality of the image as well as the purpose behind the processing. It must be noted that more than one technique can be used to get the desired result.

Review Questions

Q. 1 Distinguish between point operations and neighbourhood operations.

Q. 2 Compare between enhancement and restoration.

Q. 3 Explain spatial domain processing.

Q. 4 Compare and contrast average filtering and median filtering.

Q. 5 Explain the difference between operations involving 3 × 3 mask for median filtering and average filtering.

Q. 6 Develop a procedure for computing the m n × n neighbourhood.

Q. 7 State and explain the features of median Compute the output of the median filte following cases.

A. Y(m) = {2, 4, 8, 3, 2}

w = {–1, 0, 1, 2}

B. Y(m) = {8, 2, 4, 3, 4}

w = {–1, 0, 1}

where, Y(m) is an array and w the window.

Q. 8 What do we mean by Gaussian noise and w averaging filter used to eliminate it ?

Q. 9 What is salt and pepper noise (binary no how does a median filter remove it ?

Q. 10 Explain the output and application of the zero memory enhancement techniques.

(1) Contrast stretching

(2) Thresholding

(3) Range compression

(4) Bit extraction.

Q. 11 Compare and contrast the low pass and hi spatial filters.

Q. 12 Show that a high pass filter can be obtained as

HP = ORIGINAL – LP (assume a 3 × 3 mask)

Q. 13 Explain why a median filter is called a salt and pepper filter.

Q. 14 Develop the procedure for computing low pass filtering using a 3 × 3 mask.

Q. 15 Compare and contrast the smoothing and sharpening filters.

Q. 16 Enhancement does not add any information. Explain.

Q. 17 Low pass filter is a smoothing filter. Explain.

Q. 18 Give a brief account of enhancement filters in the spatial domain.

Q. 19 Explain why median filter is not suitable for gaussian noise. What is the limitation of median filter ?

Q. 20 What do we mean by frequencies in an image ?

Q. 21 What is meant by unsharp masking and crisping ? Explain with relevant figures.

Q. 22 Given a image, what is the output using a 3 × 3 averaging filter and a medain filter ? Explain the result.

3	3	3	3	3	3
3	3	3	3	3	3
3	3	10	10	3	3
3	10	3	3	3	3
3	3	3	3	8	3
3	3	3	3	3	3
3	3	3	3	3	3
3	3	3	3	3	3

Q. 23 How is image subtraction carried out ? Give an example of image subtraction in the field of medicine.

Q. 24 If in a sample of blood (say) the RBC's are supposed to be counted. What kind of operation would you apply ? The RBC's are visible but are not very clear. Also note that apart from the RBC's, there are no other cells present in the sample. Justify your answer.

Q. 25 What is the effect of using a larger mask in case of averaging filters e.g. 5 × 5 mask ? Explain using an example.

Q. 26 Explain the major difference between high pass and high boost filtering with an example.

Q. 27 Discuss the limiting effect of repeatedly applying a 3 × 3 low pass spatial filter to a digital image. You may ignore the border effects.

Q. 28 The figure below shows the cross sections of the transfer function of a high pass filter and its corresponding impulse response function. What would the transfer function and the impulse response look like for a high boost filter ?

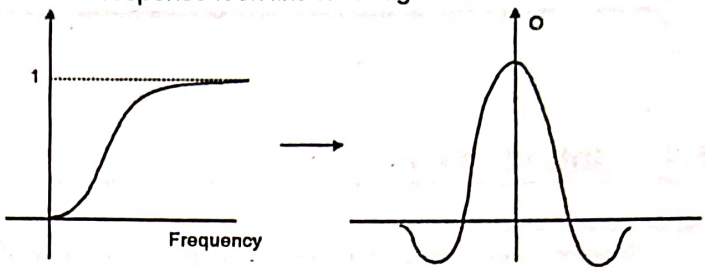


Fig. Q. 28

Q. 29 Explain the procedure of zooming an image using replication and interpolation.

Q. 30 The grey levels in an image range from 10 to 50 and we wish to display the image on a device that has a grey level range of 0 to 255. What transformation would we use ?

Q. 31 Following figure is an image with 8 grey levels. Transform it to an image with 4 grey levels without changing the brightness and the contrast.

0	1	1	1	1	4
1	1	2	2	2	3
1	1	2	2	3	3
1	2	6	6	3	3
1	2	2	4	4	4
1	2	2	3	7	5

Q. 32 Why should the sum of any high pass mask be zero ?

Q. 33 Perform discrete convolution on the following image arrays.

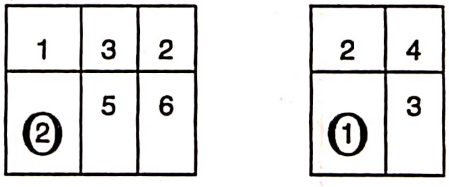


Fig. Q. 33

Q. 34 Draw the probability density functions of :

(a) Rayleigh noise

(b) Exponential noise.

Q. 35 Explain the power-law transformation.

Chapter Ends...

□□□