COLOUR IMAGE PROCESSING

I INTRODUCTION :-

Colour :- Colour is a perceptual sensation of light in visible range incident on retina. It is a striking feature of any image & has significant bearing on scenic beauty.

Why colour image processing :-

Colour image processing is used because colour is a powerful descriptor. It simplifies object identification & extraction from a scene. Humans can discern thousands of colour shades & intensities compared to only two dozens of shades of gray.

Types of Colour Image Processing :-

There are two types of colour image processing techniques.

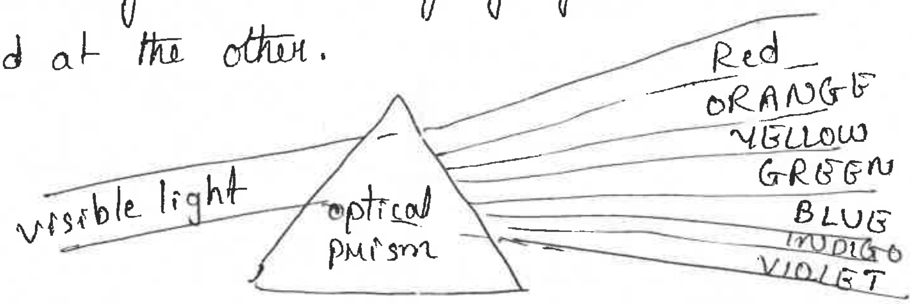
(1) Full colour Processing :- In this method, images are acquired with full colour sensor by using camera or colour scanner. This type finds applications in broad band applications such as publishing, visualization & internet.

(2) Pseudo colour Processing :- In this method, colour is assigned to monochrome intensity or range of intensities

II

COLOUR IMAGE FUNDAMENTALS :-

Light & Colour :- When a beam of sunlight passes through a prism, the emerging beam of light ~~is~~ ~~made~~ consists of colours ranging from violet at one end to Red at the other.



The frequency of light determines colour & the amount of light determines intensity.

Characterization of Light :-

(1) Achromatic Light :- Light void of colour is called achromatic light. Its attributes are intensity. It is the light that viewers see on black & white TV.

(2) Chromatic Light :- This light spans over the visible from 400-700nm. The quality of chromatic light is defined by radiance, luminance & brightness.

a) Radiance :- It is the total amount of energy that flow from the light source (in watts)

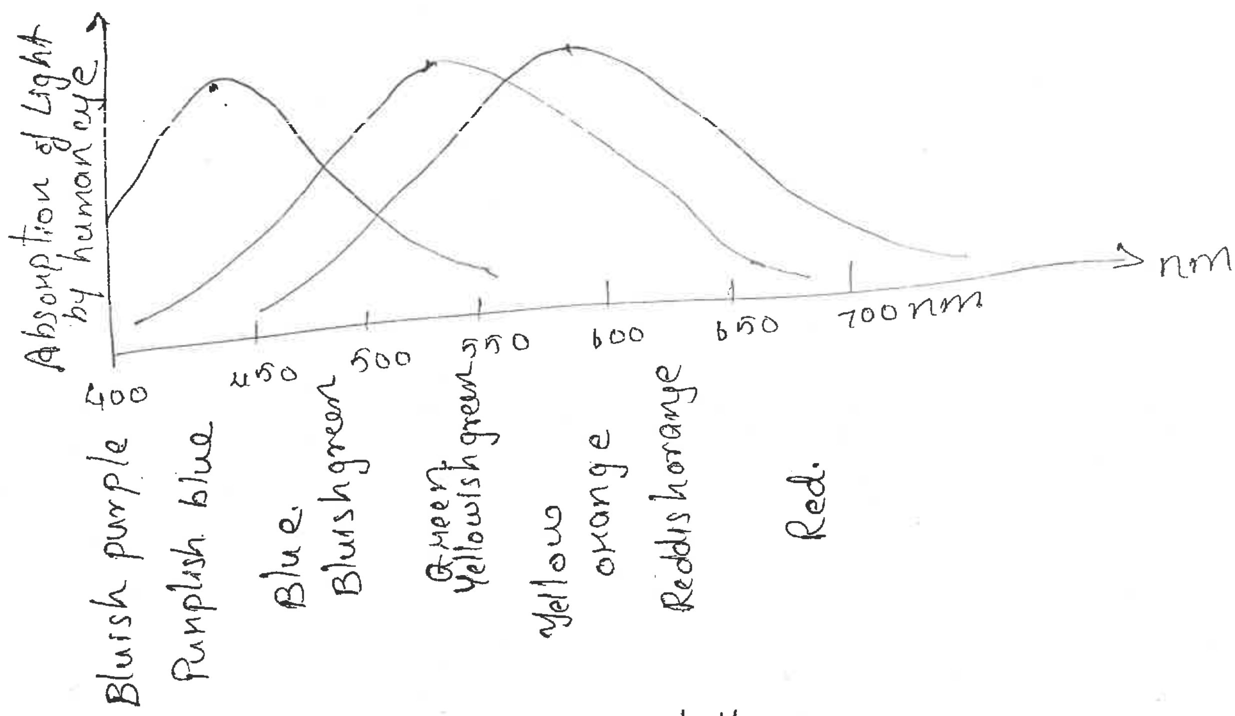
b) Luminance :- It is the measure of the amount of energy ~~of~~ an observer perceives from a light source (in lumens)

c) Brightness :- It is a subjective description that embodies the achromatic notion of intensity & is practically impossible to measure.

we know that in human eye we have receptor are called rods & cones. Among them rods are use for black & white through shades of gray where cones are used for colour vision in the eye.

Almost 65% cones are sensitive to Red light
35% cones are sensitive to ~~Blue~~ Green light
2% cones are sensitive to blue light

Due to absorption characteristics of eye there colours are seen as various combinations of Red, Green & Blue. Hence the three colours Red, Green & Blue are called Primary colours of light

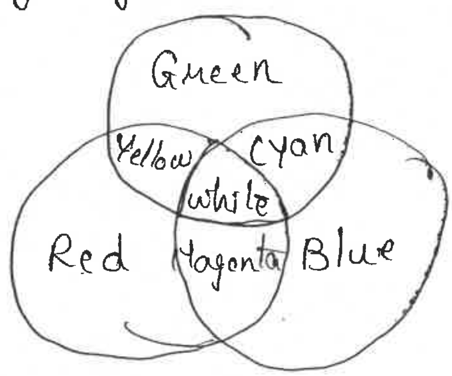


The standard wavelengths for

Blue = 435.8 nm
Green = 546.1 nm
Red = 700 nm.

Primary & Secondary colours of light:-

The primary colours of light are Red, Blue & Green. Adding these colours in certain proportions will produce secondary colours of light



Red + Green = Yellow

Green + Blue = ~~Magenta~~ Cyan

Red + Blue = Magenta

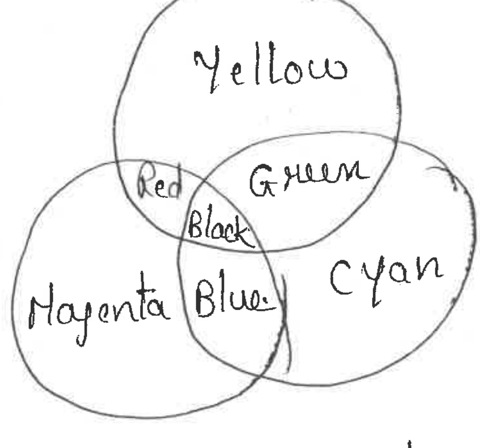
white light can be produced (a) by mixing of three primary colours Red, Green & Blue or (b) by mixing a secondary colour with opposite primary colour. Hence Red, Green & Blue are also called Additive Primaries. Yellow, cyan & magenta are obtained by mixing of two primaries & hence are called secondary colours of light.

Primary & Secondary Colours of Pigment:-

(or also called as colourants)

A primary colour of pigment is defined as the one that subtracts (or absorbs) primary colour of light & reflects (or transmits) the other two. Therefore, the primary colours of pigments are Magenta, Cyan & Yellow. The secondary colours of pigments are Red blue & Green. Fig shows the primary & secondary colour pigments.

A proper combination of three pigments primaries produces black.
Similarly, a proper combination of secondary pigment with opposite primary produces black.



Hue, Saturation & Intensity :- (write a short note on chromaticity)
These characteristics are used to distinguish one colour from another.

→ Intensity :- It represents the brightness or gray levels in an image.

→ Hue :- It represents the dominant colour as perceived by an observer (Ex :- Red, blue, orange e

→ Saturation :- It is the relative purity or the amou of white light mixed with hue
Ex (Pink - red + white ; lavender (violet + whit

Hue & Saturation together are called Chromaticity
Hence a colour can be characterized by Intensity & Chromaticity.

[Define brightness hue & saturation]

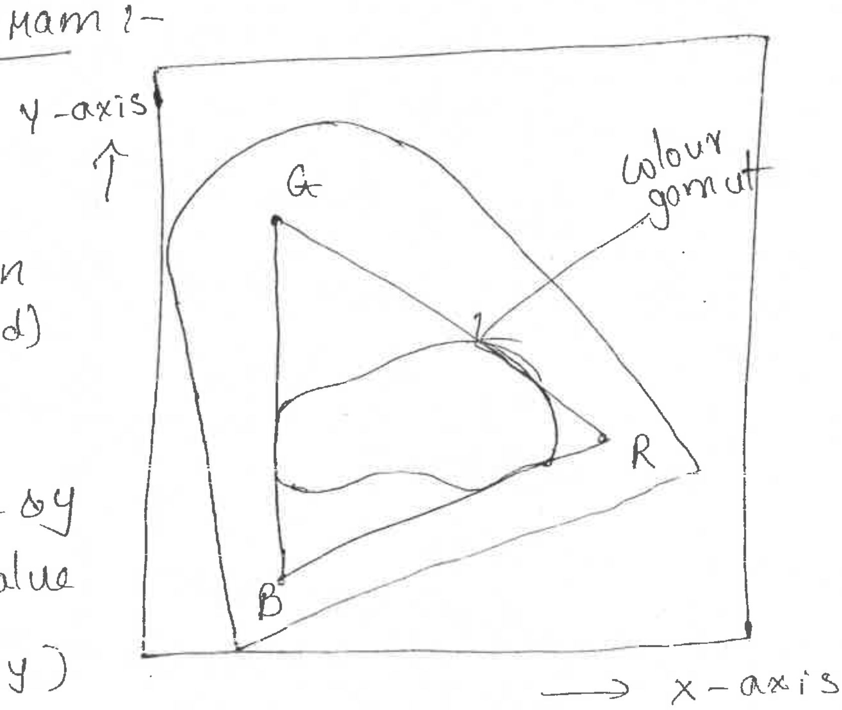
Tri-Stimulus Values :- (write a short note on tristimul val
The amount of Red, Green & Blue needed to form any particular colour ~~are~~ called Tri-stimulus values & are denoted by X, Y, Z

A colour can hence be specified by its tri-chromat coefficients defined as

$$x = \frac{X}{X+Y+Z} \; ; \; y = \frac{Y}{X+Y+Z} \; ; \; z = \frac{Z}{X+Y+Z}$$

& $x + y + z = 1$.

Chromaticity Diagram :-



The chromaticity diagram represents the colour composition as a function of x (red) & y (green). For any value of x & y the corresponding value of $z = 1-(x+y)$

Any point within the boundary of tongue shape represents some mixture of ~~colle~~ colours. As a point is more towards centre white light is added. This diagram is useful for colour mixing. The triangle shows typical range of colours called colour gamut produced by RGB monitors. The irregular region inside the triangle is representati of colour gamut of todays high quality colour printing devices.

COLOUR MODELS :- (what is the purpose of colour models)

The purpose of a colour model is to facilitate the specification of colours in some standard.

There are three colour models

(List out colour models)

→ RGB model [Red, Green, Blue] → Monitor, Video camera

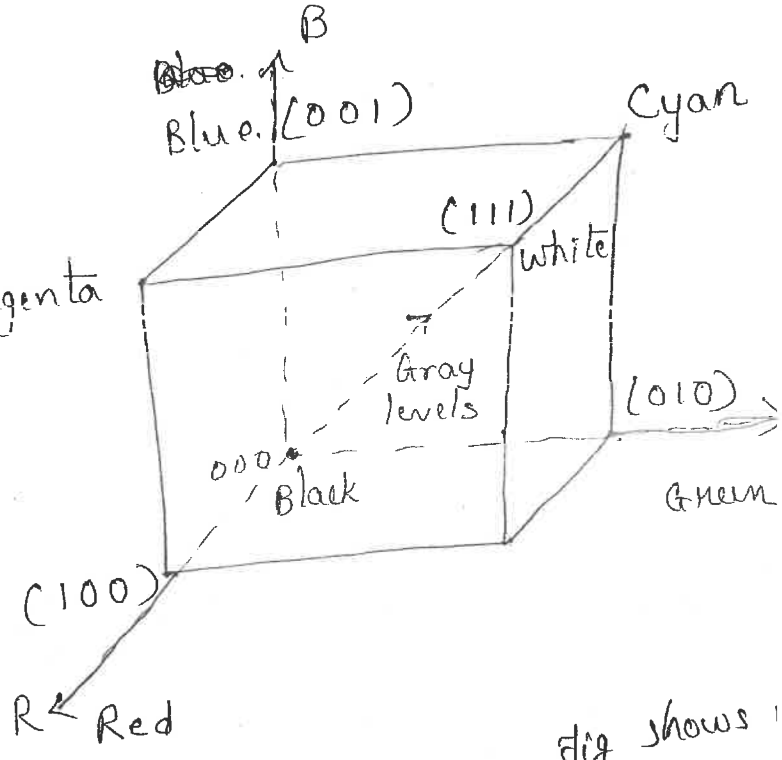
→ CMY model [Cyan, Magenta, yellow]

CMYK model [Cyan, Magenta, Yellow, Black] → colour printer

→ HSI model [Hue, Saturation, Intensity]

1) RGB colour model :-

This model is based on cartesian co-ordinate system. RGB primary colours are at three corners & CMY secondary colours are at other three corners. Black is at origin & white is in a corner far from origin. Fig shows the gray levels b/w black & white with dotted lines joining two points. If all colours inside the cube are normalized i.e all values of RGB colour are in range 0 - 1, then the cube is called a unit cube.



Explain about RGB colour model & mention its applications

The no. of bits used to represent each pixel in RGB colour space is called <u>pixel depth</u>.

Inorder to represent each pixel in RGB colour space 8 bits are required. To represent each RGB colour space in an image, no. of bits are required are

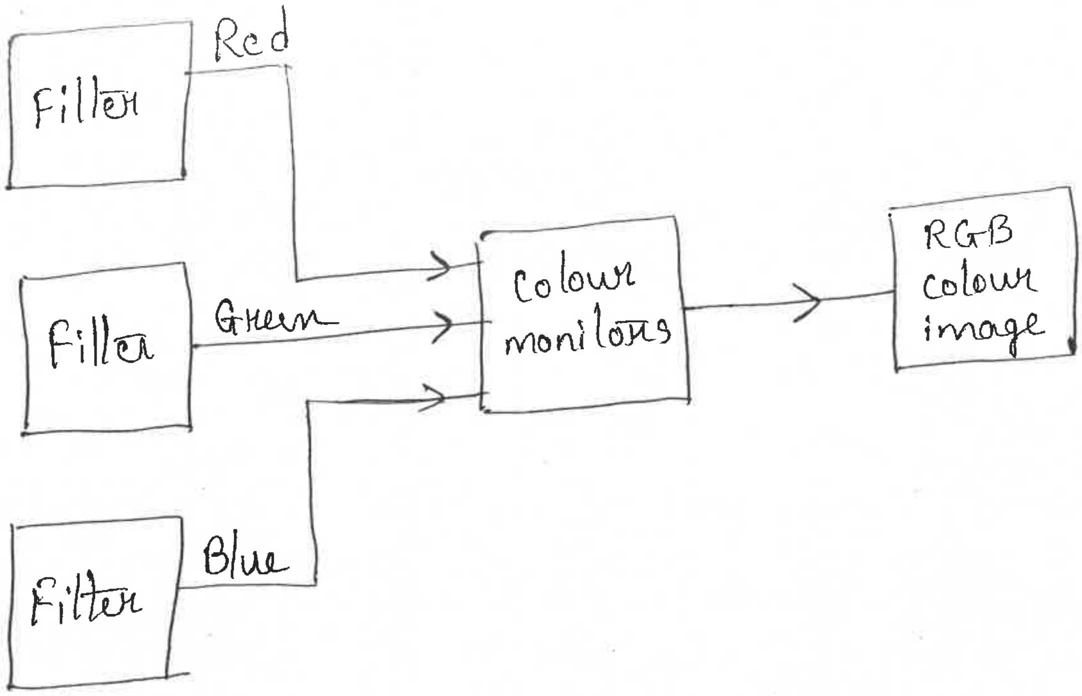
$8 \times 3 = 24$ bits

↓
RGB

Hence total no. of colours in a 24 bit RGB colour image is $(2^8)^3 \approx 1.6$ crores.

Practically it is not possible to use all the 1.6 crores colours. Hence we use only 256 colours called <u>safe RGB colours</u>.

A colour image can be acquired by using three filters sensitive to red, green & blue.



When we view a colour scene with a monochrome camera equipped with one of these filters, the result is a monochrome image whose intensity is proportional to the response of the filter. By repeating this process with each filter, it produces three monochrome images that are the RGB component images of colour scene.

The subset of colours is called the set of safe RGB colours or set of all systems safe colours. In internet they are called safe web colours or safe browser colours.

Among the 256 colours, 40 colours are processed differently by various operating systems leaving only 216 colours that are common to most systems for internet applications. These 216 values are represented in hexadecimal form. 000000 represents Black whereas FFFFFF represents white.

(write short notes on RGB to CMY conversion (Discuss about CMY & CMYK models)

CMY & CMYK Colour Model:-

Cyan, Magenta & Yellow are the secondary colours of light or primary colours of pigment. Most devices that deposit coloured pigments on paper, such as printers & copiers require CMY data i/p

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

All colours are normalized to range [0, 1]

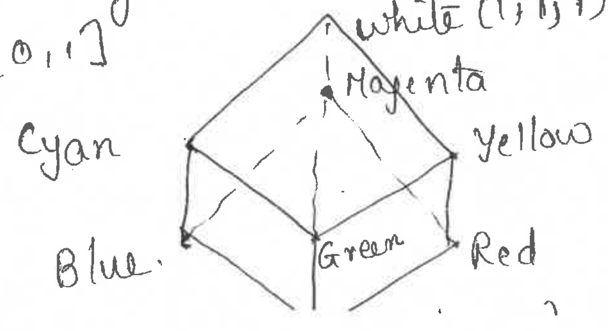
The o/p black colour obtained from colour printing is muddy black. (combination of cyan, magenta & Yellow) So in order to produce true black, a fourth colour, black is added, giving rise to the CMYK colour model.

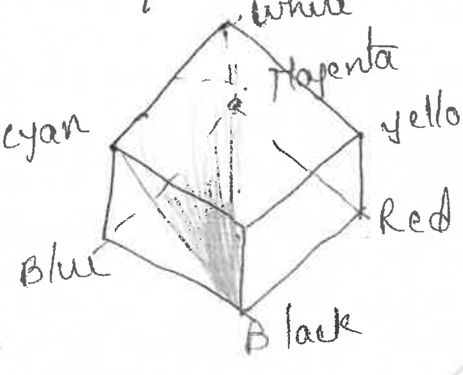
(3) The HSI Model :- (or HIS model)

The disadvantage of RGB, CMY & other colour models are not suitable for describing colours in terms that are practical for human interpretation. When humans describe a colour object it is done in terms of hue, saturation & intensity. Hence colour model called HSI model is used.

HSI model is an ideal tool for developing image processing algorithms. It decouples the colour & gray levels & provides information about image that is easily described & interpreted.

Figure shows a cube that has axis connecting black & white points called intensity axis. To calculate intensity of any colour, draw the plane ⊥ar to intensity axis & containing that colour point. Intersection of plane with intensity axis gives us a point with intensity value in rang [0,1]



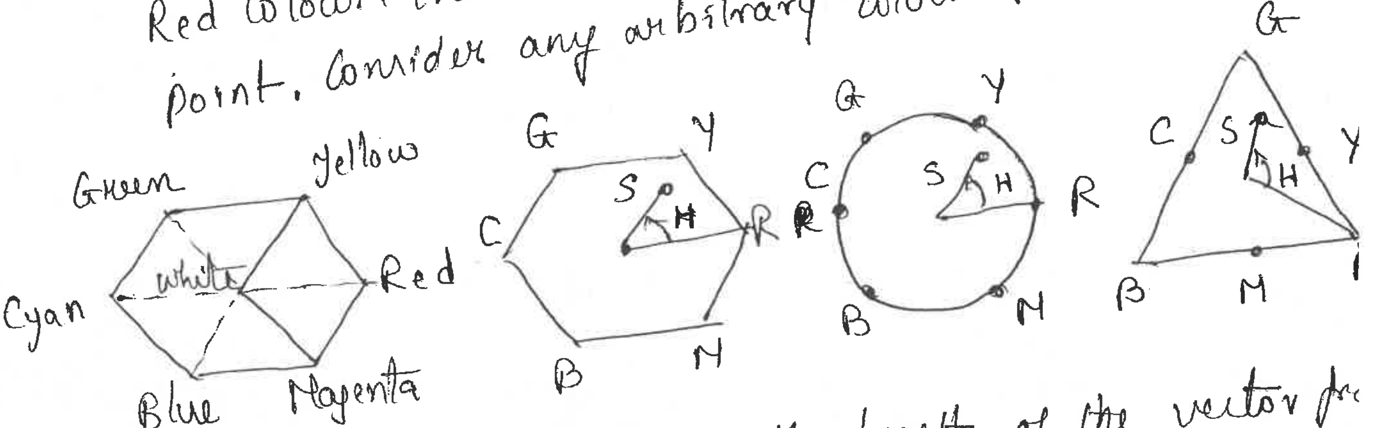


If two points are black & white & third is a colour point (ex: cyan) all points on triangle will have same hue because black & white cannot change hue. Any RGB model can be converted to HSI model.

Fig. demonstrates that saturation increases if distance from intensity axis has been increased. Hue is same at every point. In order to obtain different hue we have to rotate the plane at vertical axis.

As the plane moves up & down the intensity axis, the boundaries defined by intersection of cube have triangular or hexagonal shape.

Figure represents hexagonal shape. The primary & secondary colours are separated from each other by 120°. The angle b/w primary & secondary colour is 60°. Take the centre point & draw a line to Red colour. Hue can be calculated at any arbitrary point. Consider any arbitrary colour point. Any point.



The length of the vector from Red axis gives hue. The length of the vector from origin to point gives saturation. HSI planes can be defined in terms of circles & triangles also

(b) Discuss concept of converting colours RGB to HSI

Converting RGB to HSI model :-

Given an image in RGB colour format, the HSI components of each RGB pixel is obtained as follows:

$$H = \begin{cases} \theta & \text{if } B \le G \\ 360 - \theta & \text{if } B > G \end{cases}$$

where

$$\theta = \cos^{-1}\left\{ \frac{\frac{1}{2}\left[(R-G) + (R-B)\right]}{\left[(R-G)^2 + (R-B)(G-B)\right]^{1/2}} \right\}$$

Saturation Component is given by

$$S = 1 - \frac{3}{R+G+B}\left[\min R, G, B\right]$$

Intensity component is given by

$$I = \frac{1}{3}\left[R+G+B\right]$$

It is assumed that RGB values have been normalized to range [0, 1]

Angle θ has been measured w.r.t red axis of HIS space.

HSI COLOUR MODELS [Conversion from HSI to RGB]

Given values of HSI in the interval [0, 1] we find corresponding values of RGB in same range

There are three sectors of interest corresponding to $120°$ intervals in separation of primaries.

(1) RG Sector :- $0° \leq H < 120°$

when H is in this sector, RGB components are given in same range.

~~there are three section of interest correspon~~

~~to ro~~

$$B = I(1-S)$$

$$R = I\left[1 + \frac{S\cos H}{\cos(60°-H)}\right]$$

$$G = 3I - (R+B)$$

(2) GB sector :- $120° \leq H < 240°$

when H is in this sector,

$$H = H - 120°$$

$$R = I(1-S)$$

$$G = I\left[1 + \frac{S\cos H}{\cos(60°-H)}\right]$$

$$B = 3I - (R+G)$$

(3) BR sector :- $240° \leq H < 360°$

Given value of H in this sector

$$H = H - 240°$$

$$G = I(1-S)$$

$$B = I\left[1 + \frac{S\cos H}{\cos(60°-\right.$$

$$R = 3I - (G+B)$$

(what is pseudo colour image processing? Explain)

I PSEUDO COLOUR IMAGE PROCESSING :-

Pseudo colour (false colour) image processing consists of assigning colours to gray values based on a specified criterion. The principal use of pseudo colour is for human visualization & interpretation of gray scale events in an image or sequence of images.

Pseudo colour image processing is of two type

(1) Intensity slicing
(2) Gray level (or) Intensity to colour transformation

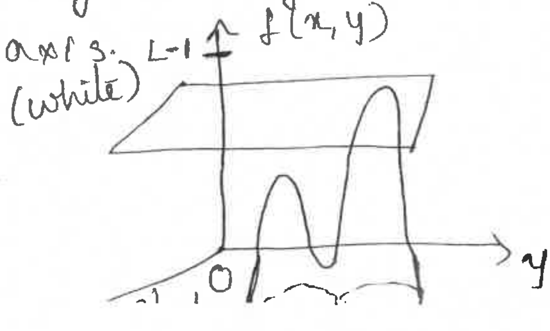
(1) Intensity slicing :- (Explain about intensity slicing & write its applications)

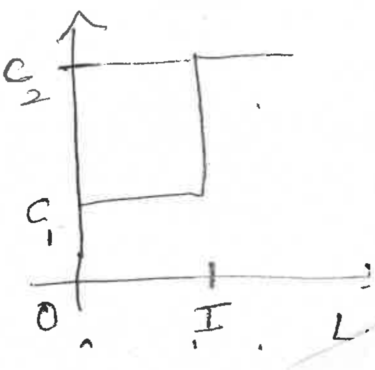
This is the simplest method of pseudo colour image processing. Take a 3D image function & place a parallel plane to the co-ordinate plane of the image

Assign two different colours.

→ If gray levels are above the plane, assign one colour
→ If gray levels are below the plane, assign second colour
→ The gray level at the plane can be assigned by any one of the two colours.

The result is a two colour image & can be controlled by moving slicing plane up & down the intensity axis.





Let $[0, L-1]$ represent gray scale. Let I_0 represent black $f(x,y) = 0$

level I_{K-1} represents white $f(x,y) = L-1$

Suppose planes ⊥lar to the intensity axis are defined at different levels. Assume $0 < P < L-1$

The 'p' planes partition gray scale into $P+1$ intervals, $V_1, V_2, V_3 \text{-----} V_{P+1}$

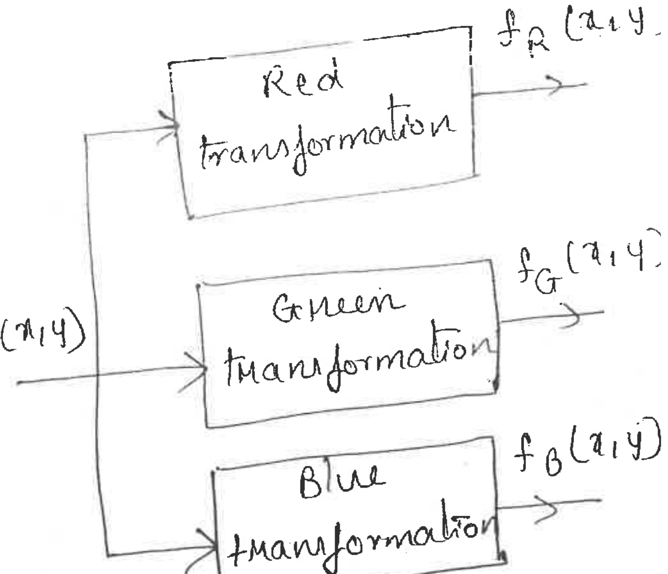
Intensity to colour assignments are made according to the relation

$f(x,y) = c_K$ if $f(x,y) \in V_K$

c_K - colour associated with kth interval

(2) Gray level or intensity to colour transformations :-

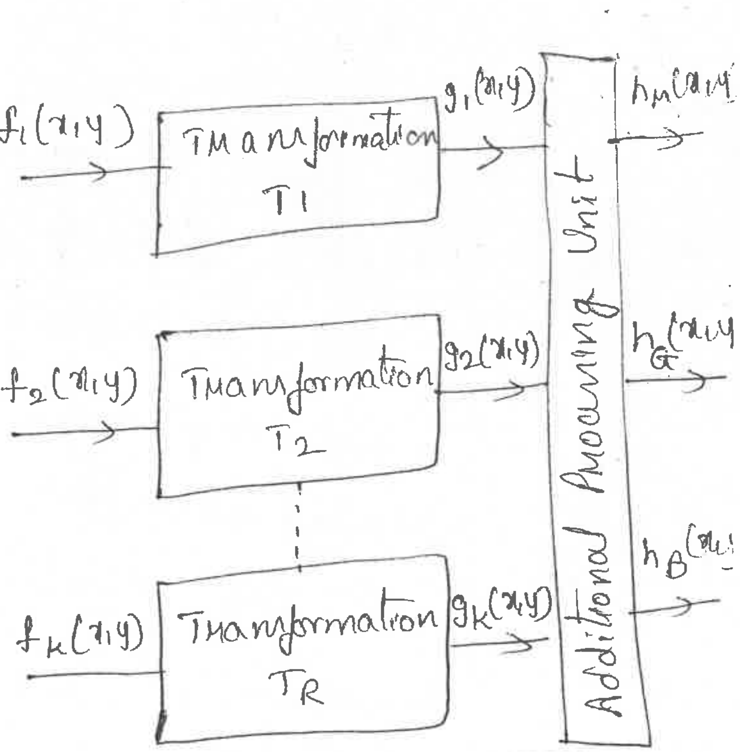
three transformations (Red, Green & Blue) are performed on the intensity of any i/p pixel. The three results are fed to Red, Green & Blue channels of colour television monitor



Transformation is a function on gray levels only, not the position of gray levels.

In order to combine several monochrome images into a single colour composite, an additional processing unit is used.

This approach is used in multi-spectral image processing. Different sensors produce individual monochrome images each in a different spectral band.



(2) FULL COLOUR IMAGE PROCESSING:

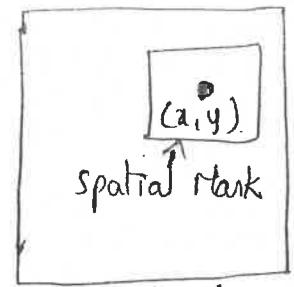
Full colour image processing approaches fall into two categories:

a) We process each colour component individually and then form a composite processed colour image

b) We work with colour pixels directly

Let 'c' represent an arbitrary vector in RGB colour space.

$$c = \begin{bmatrix} c_R \\ c_G \\ c_B \end{bmatrix} = \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$



Gray Scale Image

For an image of size $M \times N$

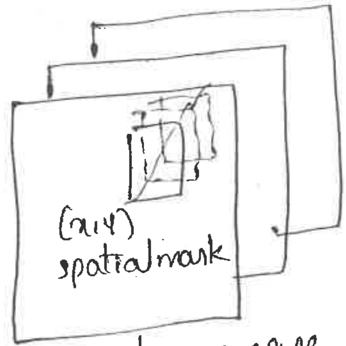
$$C(x,y) = \begin{bmatrix} C_R(x,y) \\ C_G(x,y) \\ C_B(x,y) \end{bmatrix} = \begin{bmatrix} R(x,y) \\ G(x,y) \\ B(x,y) \end{bmatrix}$$

where

$x = 0, 1, 2, 3 \ldots N-1$

$y = 0, 1, 2, 3 \ldots N-1$

The results of two methodologies are not the same.



RGB colour image

COLOUR TRANSFORMATIONS :-

Colour Transformations deal with processing of colour components of a colour image within the context of a single colour model, as opposed to the conversion of those components b/w models [like RGB-HSI, HSI-RGB]

The colour transformations are

- Formulation
- Colour complements
- Colour slicing
- Tone & colour corrections
- Histogram Processing

(1) Formulation :- Formulation means processing of a colour image within the context of a single colour model.

$$g(x,y) = T[f(x,y)]$$

$$s_i = T_i(r_1, r_2 \ldots r_n)$$

$$i = 1, 2, \ldots n$$

s_i = colour components of g.

$r_1, r_2 \ldots r_n$ = colour components of f (R, G, B)

T_i = colour mapping function

for HSI : $s_3 = kr_3$

RGB : $s_i = kr_i$; $i = 1, 2, 3$

CMY : $s_i = kr_i + (1-k)$ $i = 1, 2, 3$

(Briefly discuss about colour complements in colour circle.)

(2) Colour Complement :-

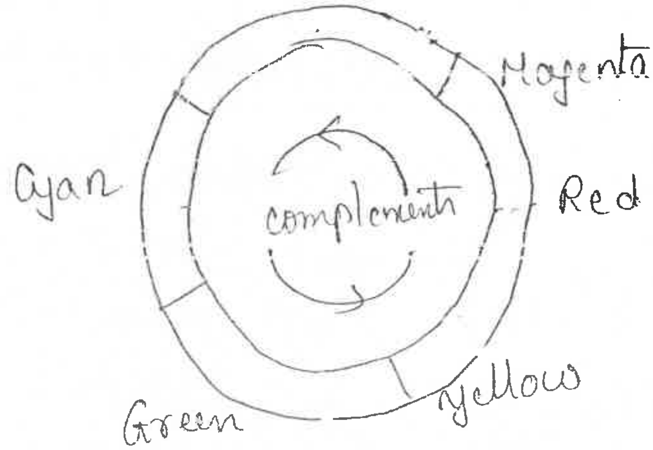


Figure represents a colour circle. Hues directly opposite to one another on the colour circle are called colour complements. These colour complements are useful for enhancing details embedded in dark regions of colour image, particularly when the regions are dominant in size.

(3) Colour Slicing :-

This transformation is used to highlight a specific range of colours in an image. It displays the colour of interest so that they stand out from the background. It uses the region defined by the colours as a mask for further processing.

The simplest way to slice an image is to map the colours outside the same range of interest to a non-prominent neutral colour.

If colours are enclosed by a cube or a hypercube ($n > 3$) of width w & centered at a average colour with components $(a_1, a_2, a_3 \ldots a_n)$ the necessary set of transformation is

$$S_i = \begin{cases} 0.5 & \text{if } \left[|r_j - a_j| > \frac{w}{2} \right]_{\text{any } 1 \le j \le} \\ r_i & \text{otherwise} \end{cases}$$

$i = 1, 2 \ldots n.$

If colours of interest are enclosed by a sphere

$$S_i = \begin{cases} 0.5 & \text{if } \sum_{j=1}^{n} (r_j - a_j)^2 > R_0^2 \quad i = 1, 2 \ldots \\ r_i, & \text{otherwise} \end{cases}$$

(4) Tone & Colour Correction:-

The tonal range of an image, also called its key-type, refers to its general distribution of colour intensities.

→ High key Images :- In these image most of the information is concentrated at high intensities

Low-key images :- In these images, most of the information is concentrated at low intensities

Middle-key images :- In these images, most of the information is concentrated at middle intensities

(5) Histogram Processing :- (Describe histogram processing in colour image) If the components of a colour image are equalized independently using histogram process, it results in error. Hence a logical approach is to spread the colour intensities uniformly leaving the colours themselves unchanged. { Ans:- Define histogram, process equalization from unit-II & then write for colour image

---

Colour Image Smoothing :- (Explain about colour image smoothing by averaging)

In gray scale smoothing we use scalar quantities to represent intensity values whereas in colour image smoothing, we use vector quantities to represent intensity values.

Averaging :-

Let S_{xy} denote set of co-ordinates defining a neighbourhood centered at (x,y) in an RGB colour image. The average of the RGB component vectors in this neighbourhood is

$$\bar{c}(x,y) = \frac{1}{K} \sum_{x,y \in S_{xy}} c(x,y)$$

From properties of vector addition $\left[c(x,y) = \begin{bmatrix} R(x,y) \\ G(x,y) \\ B(x,y) \end{bmatrix} \right]$

$$\bar{c}(x,y) = \begin{bmatrix} \frac{1}{K} \sum_{x,y \in S_{xy}} R(x,y) \\ \frac{1}{K} \sum_{x,y \in S_{xy}} G(x,y) \\ \frac{1}{K} \sum_{x,y \in S_{xy}} B(x,y) \end{bmatrix}$$

~~these are can~~ the components of this vector are recognized as the scalar images that would be obtained by independently smoothing each plane of the starting RGB image using gray scale processing. Hence smoothing can be done by neighbourhood averaging.

Colour Image Sharpening :- (Explain about colour image sharpening)

The principle objective of sharpening is to highlight transitions in intensity, Applications of image sharpening range from electronic printing & medical imaging to industrial inspection & autonomous guidance in military systems.

For gray scale images sharpening is done by using Laplacian. From vector analysis, the Laplacian of a vector is defined as a vector whose components are equal to the Laplacian of the individual scalar components of the i/p vector.

In RGB colour system the Laplacian of vector 'C' is

$$\nabla^2[c(x,y)] = \begin{bmatrix} \nabla^2 R(x,y) \\ \nabla^2 G(x,y) \\ \nabla^2 B(x,y) \end{bmatrix}$$

we can compute the Laplacian of a full colour image by computing the Laplacian of each component image separately.

(Discuss about segmentation in RGB colour space)

Explain about colour segmentation process

COLOUR IMAGE SEGMENTATION :-

Segmentation is a process that partitions an image into regions. One segment will not have any property that any other segment has. In a segment all pixels will have same qualities. There are three types of segmentation.

(1) Segmentation in HIS colour space :-

This method can be used to segment an image based on colour & process on individual planes.

Colour can be conveniently represented in hue of an image. Saturation is used to isolate regions of interest in the hue image. Intensity in image is used very less for segmentation of colour images because it carries no colour information.

(2) Segmentation in RGB vector space:-

Suppose we have to segment objects of a specified colour range in an RGB image. Given a set of sample colour points of interest in image, we estimate the 'average' colour we wish to segment. Let the avg. colour be denoted by RGB vector 'a'.

The objective of segmentation is to classify each RGB pixel in a given image as having a colour in the specified range or not.

To perform this comparison, it is necessary to have a measure of similarity. The simplest method is Euclidean Distance.

'z' is similar to 'a' if the distance b/w them is less than a specified threshold.

$$D(z,a) = \|z-a\| = \left[(z-a)^T(z-a)\right]^{1/2}$$

$$\Rightarrow D(z,a) = \left[(z_R-a_R)^2 + (z_G-a_z)^2 + (z_B-a_B)^2\right]^{\frac{1}{2}}$$

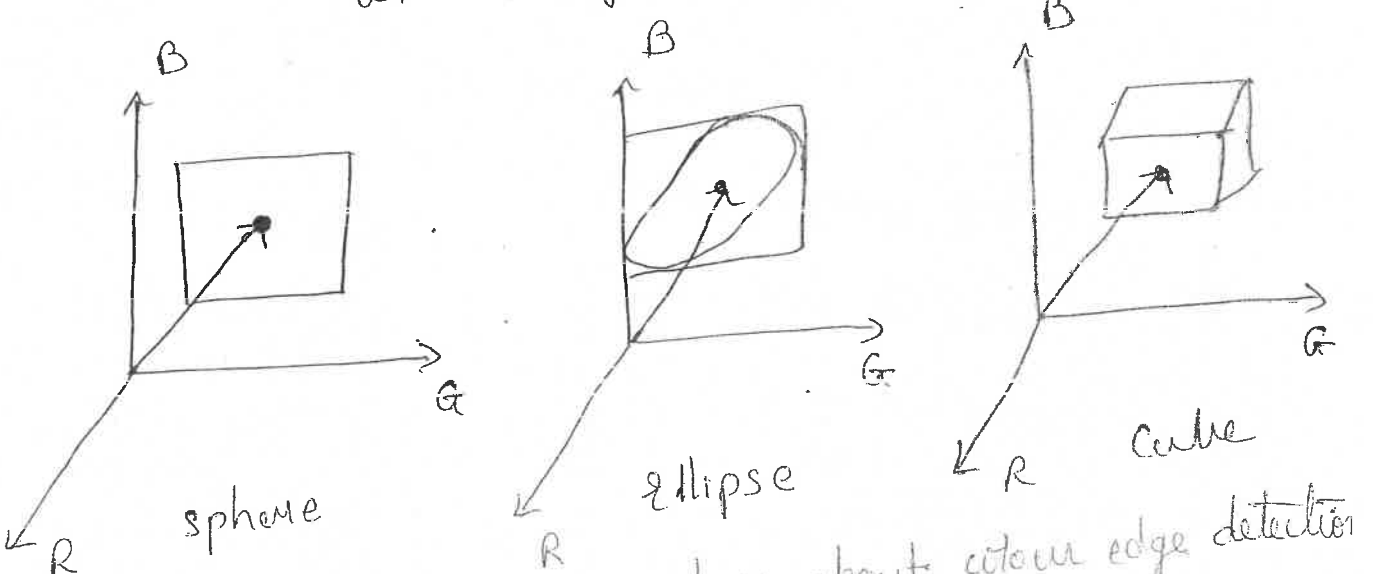
The subscripts R, G, B denote the Red, Green, Blue components of vectors a & z.

The locus of points $D(z, a) \le D_0$ is a solid sphere of radius D_0. Points within the sphere satisfy the specified colour criterion. Plots outside the sphere do not.

<u>Generalized</u> form :-

$$D(z, a) = \left[(z-a)^T C^{-1} (z-a) \right]^{1/2}$$

where C - covariance matrix of the samples we wish to segment



explain about colour edge detection (with necessary eqns)

(3) <u>Colour Edge Detection</u> :-

It is an important tool for segmentation.

Let r, g, b be unit vectors along the R, G, B axis of the RGB colour space & define the vector

$$u = \frac{\partial R}{\partial x} r + \frac{\partial G}{\partial x} g + \frac{\partial B}{\partial x} b$$

$$v = \frac{\partial R}{\partial y} r + \frac{\partial G}{\partial y} g + \frac{\partial B}{\partial y} b$$

Let the quantities g_{xx}, g_{yy}, g_{xy} be defined in terms of dot product of these vectors

$$g_{xx} = u^T u = \left|\frac{\partial R}{\partial x}\right|^2 + \left|\frac{\partial G}{\partial x}\right|^2 + \left|\frac{\partial B}{\partial x}\right|^2$$

$$g_{yy} = v^T v = \left|\frac{\partial R}{\partial y}\right|^2 + \left|\frac{\partial G}{\partial y}\right|^2 + \left|\frac{\partial B}{\partial y}\right|^2$$

$$g_{xy} = u^T v = \frac{\partial R}{\partial x}\frac{\partial R}{\partial y} + \frac{\partial G}{\partial x}\frac{\partial G}{\partial y} + \frac{\partial B}{\partial x}\frac{\partial B}{\partial y}$$

The direction of maximum rate of change $c(x,y)$ is given by the angle

$$\theta = \frac{1}{2}\tan^{-1}\left[\frac{2g_{xy}}{(g_{xx} - g_{yy})}\right]$$

The value of the rate of change at (x,y) in the direction of $\theta(x,y)$ is given by the angle

$$F(\theta) = \left\{\frac{1}{2}\left[(g_{xx} + g_{yy}) + (g_{xx} - g_{yy})\cos 2\theta + 2g_{xy}\sin 2\theta\right]\right\}^{1/2}$$

Noise in Colour Images :- (Discuss about noise in colour images)

[Note :- For essay define noise & noise models from unit - 3]

The noise content of a colour image has same characteristics in each colour channel. It is possible for colour channels to be affected by noise. The fine grain noise is less noticable in a colour image compared to monochrome images. Noise models used for colour images are same as that used for gray scale images. Depending on the strength of illumination present in a particular colour channel, the effect of noise level will vary.

Ex :- The strength of illumination present in the green sensor of CCD camera is reduced by green fill. Thus green component of an RGB image will be more noisy than other component images at lower levels of illumination.

COLOUR IMAGE COMPRESSION

Compression means to compress the size of image by reducing the no. of bits. It place a major role in transmission & storage of information. The colour information is 3 to 4 times greater compared to gray level images.

WAVELETS & MULTI RESOLUTION PROCESSING

Wavelet transforms :- These transforms are based on "small waves" called wavelets.

wavelet :- A wavelet is a wave of varying frequencies limited duration.

Multi Resolution Analysis (MRA) :- It is a technique of storing & processing of images at multiple resolutions.

Multiple Resolutions :-

A small sized or low contrast objects in an image may require higher resolution or a large sized or high contrast objects may require low resolution for analysis.

[Ex :- Observing a tree from a distance gives low resoluti & observing closely gives high resolution]

Multiresolution analysis is implemented using wave transforms.

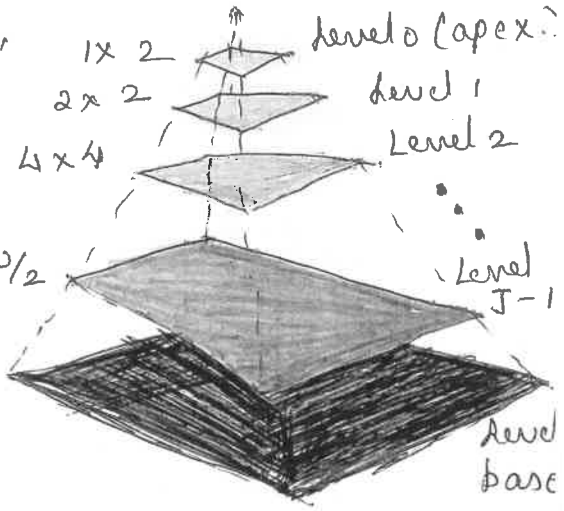
Image Pyramids:-

→ These are structures used to represent images at more than one resolution.

→ They are a collection of decreasing resolution images arranged in the shape of a pyramid.

→ Consider a high resolution image at the pyramid base

→ As we move up the pyramid, both size & resolution decrease.



→ Base level is of size $2^J \times 2^J$

→ General level j of size $2^j \times 2^j$, $0 \le j \le J$

→ Pyramid may get truncated at level P, $1 \le P \le J$

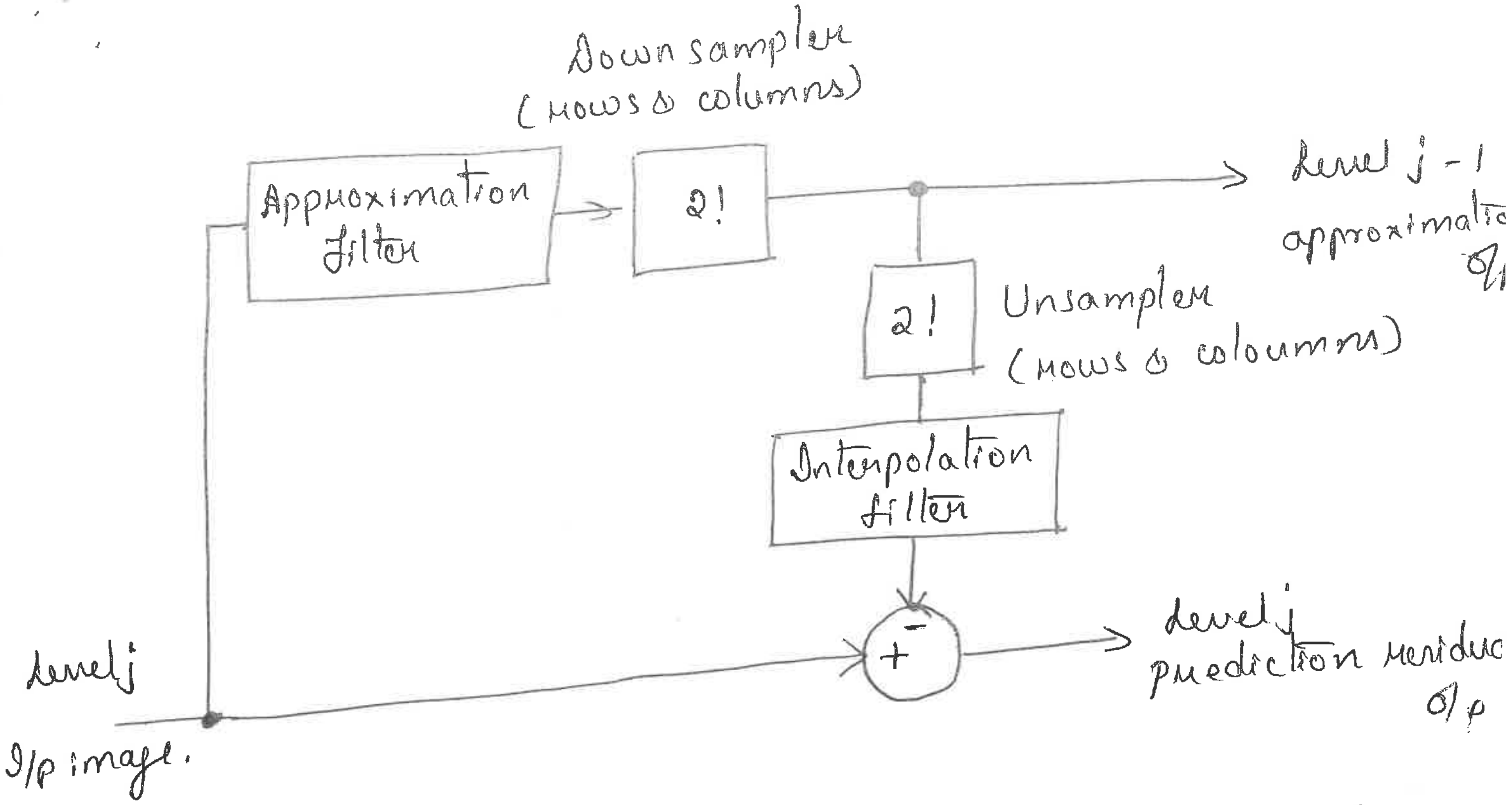
→ No. of pixels in a pyramid with $P+1$ levels ($P>0$)

is $$N^2\left(1 + \frac{1}{(4)^1} + \frac{1}{(4)^2} + \cdots + \frac{1}{4^P}\right) \le \frac{4}{3} N^2$$

→ Fig below shows a simple system for constructing two intimately related image pyramids

→ Level $j-1$ approximation o/p provides the images needed to build an approximation pyramid.

→ Level j prediction residual o/p is used to build a complementary prediction residual pyramid.



→ Level j prediction residual contains only reduced-resolution approximation of the i/p image at the top level. All other levels contain prediction residuals where level j prediction residual is difference b/w level j approximation & an estimate of the level j-1 approximation.

→ Both approximation & prediction residual pyramids are computed in an iterative fashion.

<u>Step 1</u>:- Compute a reduced resolution approximation of level j input image done by filtering & downsampling the filtered result by a factor of 2. Place the resulting approximation at level j-1 of the approximation pyramid.

Step 2 :- Create an estimate of level j input image from the reduced resolution approximation generated in step 1; done by unsampling & filtering the generated approximation; resulting prediction image will have same dimensions as level j input image.

Step 3 :- Compute the difference b/w the prediction image of step 2 & i/p to step 1; place the result in level j of prediction residual pyramid.

→ A variety of approximation & interpolation filters can be incorporated into the system of fig (b).

Useful approximation filtering techniques include

- → Neighbourhood averaging producing mean pyramids.
- → Low pass Gaussian filtering producing Gaussian pyramids
- → No filtering producing subsampling pyramids.
- → Interpolation filter can be based on nearest neighbour, bilinear & bicubic.

→ Up Sampling :-

- → Doubles the spatial dimensions of approximation images
- → Given an integer n & 1D sequence of samples f(n), upsampled sequence is given by

$$f_{2\uparrow}(n) = \begin{cases} f(n/2) & \text{if } n \text{ is even} \\ 0 & \text{otherwise} \end{cases}$$

Down sampling :-

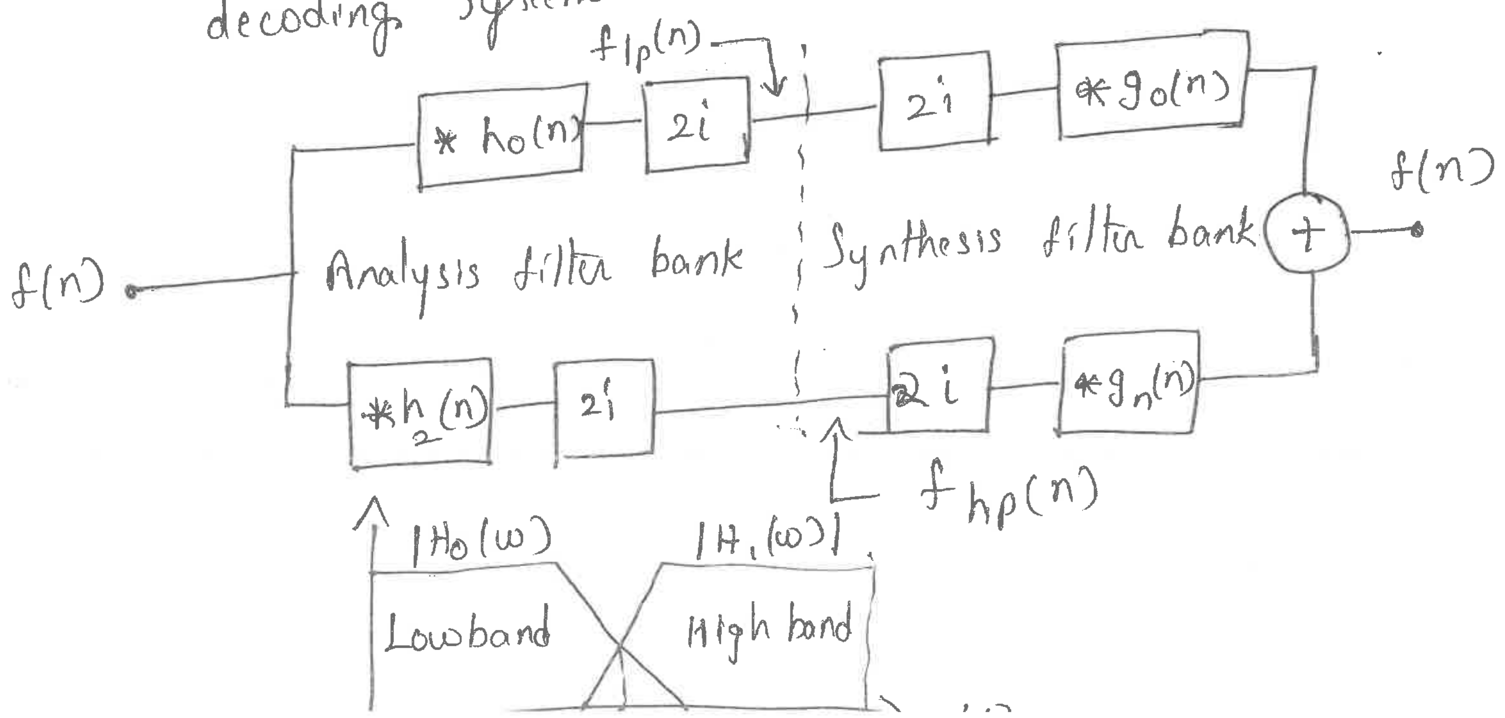
→ Halves the spatial dimensions of prediction images & is given by

$$f_{2\downarrow}(n) = f(2n)$$

→ Discard every other sample.

SUB BAND CODING :-

In subband coding image is decomposed into a set of bandlimited components as a result of decomposing an image. The decomposition is performed such that subbands can be reassembled to construct the original image without error. Consider the two-band sub band coding & decoding system.



The system is composed of two filter banks, each containing two FIR filters as shown in fig.

(1) Analysis of filter bank:

→ Uses filters $h_0(n)$ & $h_1(n)$ to split i/p sequence $f(n)$ into two downsampled sequences $f_{lp}(n)$ & $f_{hp}(n)$

→ $f_{lp}(n)$ & $f_{hp}(n)$ are two subbands to represent the i/p

→ $h_0(n)$ & $h_1(n)$ are two half-band filters whose idealized transfer characteristics H_0 & H_1 are shown in fig.

→ $h_0(n)$ is a LPF whose o/p subband is called an approximation of $f(n)$

→ $h_1(n)$ is a high pass filter whose o/p subband is called the detail part of $f(n)$.

(2) Synthesis filter bank:-

→ Filters $g_0(n)$ & $g_1(n)$ combine the o/p of analysis to produce $\hat{f}(n)$

The goal of subband coding is to select the four filters $h_0(n)$, $h_1(n)$, $g_0(n)$, $g_1(n)$ such that $f(n) = \hat{f}(n)$. [perfect reconstruction filters]

For perfect reconstruction, the impulse responses of the synthesis & analysis filters must be related in one of the following two ways:

$$g_0(n) = (-1)^n h_1(n)$$
$$g_1(n) = (-1)^{n+1} h_0(n)$$

or

$$g_0(n) = (-1)^{n+1} h_1(n)$$
$$g_1(n) = (-1)^n h_0(n)$$

Filters $h_0(n), h_1(n), g_0(n)$ & $g_1(n)$ are said to be cross modulated because diagonally opposed filters in the block diagram are related by modulation.
They satisfy biorthogonality condition.

Inner product $\langle h_i(2n-k), g_j(k) \rangle = \delta(i-j)\delta(n)$
$i, j = \{0, 1\}$.

$i \neq j \Rightarrow$ inner product $= 0$
$i = j \Rightarrow$ " " $=$ unit discrete impulse fn $\delta($

$$\langle g_i(n), g_j(n+2m) \rangle = \delta(i-j)\delta(m), \quad i, j = \{0, 1\}$$

which defines orthonormality for perfect reconstruction filter banks.
In addition to above eq, orthogonal filters can be shown to satisfy the following two conditions

$$g_i(n) = (-1)^n g_0(K_{even} - 1 - n)$$

$$h_i(n) = g_i(K_{even} - 1 - n), \quad i = \{0, 1\}$$

K_{even} indicates that the no. of filter coefficients must be divisible by 2. (an even no.)

HAAR WAVELET TRANSFORM

It is the oldest & simplest orthogonal wavelets

It is expressed in matrix form as

$$T = HFH^T$$

F is an $N \times N$ image matrix, $N = 2^n$

H is an $N \times N$ Haar transformation & contains the basis function $h_k(z)$ from the wavelet

T is resulting $N \times N$ transform

Transform is required because H is not symmetric

H is generated by defining the integer $k = 2^p + q - 1$

where $0 \leq p \leq n-1$, $q = 0$ or 1 for $p = 0$ &

$1 \leq q \leq 2^p$ for $p \neq 0$.

Haar basis functions are

$$h_0(z) = h_{00}(z) = \frac{1}{\sqrt{N}}, \quad z \in [0, 1]$$

$$h_k(z) = h_{pq}(z) = \frac{1}{\sqrt{N}} \begin{cases} 2^{p/2} & (q-1)/2^p \leq z < \frac{q-0.5}{2^p} \\ -2^{p/2} & (q-1)/2^p \leq z \leq q/2^p \\ \text{otherwise}, & z \in [0, 1] \end{cases}$$

The ith row of an $N \times N$ Haar transform matrix contains the elements of $h_i(z)$ for $z = \frac{0}{N}, \frac{1}{N}, \ldots \frac{(N-}{N}$

For $N = 2$, first row of a 2×2 Haar matrix is computed using $h_0(z)$ with $z = \frac{0}{2}, \frac{1}{2}$

From above eq $h_0(z) = \frac{1}{\sqrt{2}}$ independent of z.

First row of H_2 is $\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}$

The second row is computed by $h_1(z)$ for $\frac{0}{2}, \frac{1}{2}$

$k = 2^p + q - 1$ when $k = 1, p = 0, q = 1$

$$h_1(0) = \frac{2^0}{\sqrt{2}} = \frac{1}{\sqrt{2}}$$

$$h_i(1/2) = \frac{-2^0}{\sqrt{2}} = \frac{-1}{\sqrt{2}}$$

The 2×2 Haar matrix is

$$H_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

If $N = 4$, k, q & p assume the values

k	p	q
0	0	0
1	0	1
2	1	1
3	1	2

The 4×4 transformation matrix H_4 is

$$H_4 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix}$$

MULTI RESOLUTION EXPANSIONS

In multiresolution analysis, a scaling function is used to create a series of approximations of a function or image, each differing by a factor of 2 in resolution from its nearest neighbouring approximations. Additional functions called wavelets, are then used to encode the difference in information b/w adjacent approximations

Series Expansions:-

A signal or function $f(x)$ can often be better analyzed as a linear combination of expansion fns.

$$f(x) = \sum_k a_k \varphi_k(x) \qquad -(1)$$

where k is an integer index of a finite or infinite sum.

α_k - real valued expansion coefficients

$\varphi(x)$ - real-valued expansion functions.

If expansion is unique $\psi_k(x)$ are called basis functions the expressible functions form a function space that is referred to as the closed span of the expansion set, denoted

$$V = \overline{\text{span}\{\psi_k(x)\}} \qquad -(2)$$

To say that $f(x) \in V$ means that $f(x)$ is in the closed span of $\{\varphi_k(x)\}$

For any function space V & corresponding expansion set $\{\varphi_k(x)\}$, there is a set of dual functions denoted $\{\tilde{\varphi}_k(x)\}$ that can be used to compute the α_k coefficients for any $f(x) \in V$. These coefficients are computed by taking the integral inner products of the dual $\tilde{\varphi}_k(x)$ & function $f(x)$.

i.e., $\alpha_k = \langle \hat{\varphi}_k(x), f(x) \rangle = \int \tilde{\varphi}_k^*(x) f(x)\, dx$ — (3)

where $*$ denotes the complex conjugate operation.

Case 1:- If the expansion functions form an orthonormal basis for V, meaning that

$$\langle \varphi_j(x), \varphi_k(x) \rangle = \delta_{jk} = \begin{cases} 0 & j \neq k \\ 1 & j = k \end{cases} \quad \text{— (4)}$$

the basis & its dual are equivalent.

i.e $\varphi_k(x) = \tilde{\varphi}_k(x)$

Hence eq (3) becomes

$$\alpha_k = \langle \varphi_k(x), f(x) \rangle$$

Case 2:- If the expansion functions are not orthonormal but are an orthogonal basis for V, then

$$\langle \varphi_j(x), \varphi_k(x) \rangle = 0 \quad j \neq k$$

and the basis functions & their duals are called biorthogonal. The α_k are computed using eq (3) & the biorthogonal basis & its dual are such that

$$< \psi_j(x), \tilde{\psi}_k(x) > = \delta_{jk} = \begin{cases} 0 & j \neq k \\ 1 & j = k \end{cases}$$

Case 3:- If the expansion set is not a basis for for V, but supports the expansion defined by eq (1), it is a spanning set in which there is more than one set of α_k for any $f(x) \in V$. The expansion functions & their duals are said to be over complete or redundant. They form a frame in which

$$A \| f(x) \|^2 \leq \sum_k | \langle \psi_k(x), f(x) \rangle |^2 \leq B \| f(x) \|^2$$

for some $A > 0$, $B < \infty$ & all $f(x) \in V$.

Dividing this eq by the norm squared of $f(x)$, we see that A & B " frame" the normalized inner products of the expansion coefficients & the function. Eqns (3) & (5) can be used to find the expansion coefficients for frames. If $A = B$, the expansion set is called a tight frame

$$f(x) = \frac{1}{A} \sum_k (\psi_k(x), f(x)) \psi_k(x)$$

Scaling Functions:-

Consider the set of expansion functions composed of integer translations & binary scalings of the real, square-integrable function $\varphi(x)$; this is the set $\{\varphi_{j,k}(x)\}$, where

$$\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k)$$

$$\varphi(x) \in L^2(R)$$

for all $j, k \in Z$ &

Here, k determines the position of $\varphi_{j,k}(x)$ along the x-axis & j determines the width of $\varphi_{j,k}(x)$ i.e how broad or narrow it is along the x-axis.

$2^{j/2}$ - controls the amplitude of the function.

Because the shape of $\varphi_{j,k}(x)$ changes with j, $\varphi(x)$ is called a scaling function.

If we restrict j to a specific value, say $j = j_0$, the resulting expansion set $\{\varphi_{j_0,k}(x)\}$ is a subset of $\{\varphi_{j,k}(x)\}$ that spans a subspace of $L^2(R)$

Hence a subspace $V_{j_0} = \overline{\text{Span}_k\{\varphi_{j_0,k}(x)\}}$

V_{j_0} is the span of $\varphi_{j_0,k}(x)$ over k. If $f(x) \in V_{j_0}$

$$f(x) = \sum_k \alpha_k \varphi_{j_0,k}(x)$$

$V_j = \overline{\text{Span}\{\varphi_{j,k}(x)\}}$