14BECS606 **Graphics and Multimedia**

L	T	P	C
3	0	0	3

Course Objectives:

- To study the graphics techniques and algorithms.
- To study the multimedia concepts and various I/O technologies.
- To enable the students to develop their creativity
- To impart the fundamental concepts of Computer Graphics and Multimedia.

UNIT- I Output Primitives 10

Overviews of graphics system-Video display devices, Raster scan system-Random scan systems. Line, Circle Drawing Algorithms – Attributes – Two-Dimensional Geometric Transformations – Two-Dimensional Clipping and Viewing.

UNIT- II Three-Dimensional concepts 9

Three-Dimensional Object Representations – Three-Dimensional Geometric and Modeling Transformations – Three-Dimensional Viewing – Color models – Animation.

UNIT- III Multimedia systems design 9

An Introduction – Multimedia applications – Multimedia System Architecture – Evolving technologies for Multimedia – Defining objects for Multimedia systems – Multimedia Databases.

UNIT- IV Multimedia storage technologies 8

Compression & Decompression–Types of Compression-Data and file format standards-Multimedia I/O technologies - Storage and retrieval Technologies.

UNIT- V Hypermedia 9

Digital voice and audio – Video image and animation – Full motion video – Hypermedia messaging - Mobile Messaging – Hypermedia message component – Creating Hypermedia message – Integrated multimedia message standards-Distributed Multimedia Systems.

Total Hours: 45

Text Books:

1. Judith Jeffcoate, Multimedia in practice technology and Applications, PHI,2007
2. Foley, Vandam, Feiner, Huges, Computer Graphics: Principles & Practice, Pearson Education,2005

References:

1. Udit Agarwal, Computer Graphics & Multimedia, S.K. Kataria & Sons;
2. Donald Hearn and M.Pauline Baker, Computer Graphics C Version, Pearson Education,2013
3. Prabat K Andleigh and Kiran Thakrar, Multimedia Systems and Design,PHI,2011

KARPAGAM ACADEMY OF HIGHER EDUCATION
FACULTY OF ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

14BECS606 **Graphics and Multimedia** **Lecture plan**

Session No	Topics to be covered	Duration	Book & Page Nos. Used for teaching	Teaching Method
1.	Introduction to computer graphics Applications Overview of graphics and Multimedia systems	1 hr	Ref[1]	BB
2.	CRT and Color CRT	1 hr	Ref[1]	BB
3.	Raster Display System Random Display System	1 hr	Ref[1]	BB
4.	**Tutorial**	1 hr		
	UNIT- I OUTPUT PRIMITIVES Introduction - Line, Circle Drawing Algorithms – Attributes – Two-Dimensional Geometric Transformations – Two-Dimensional Clipping and Viewing.			
5.	DDA Line Drawing algorithm. Bresenham's line drawing algorithm.	1 hr	Ref[1] & Page no 107,108	BB
6.	Midpoint Circle algorithm Ellipse generating algorithm	1 hr	Ref[1] & Page no 117,123	BB
7.	**Attributes** 1..Line 1.5.2.Curve 1.5.3. Area fill 1.5.4. Character 1.5.5.Bundled	1 hr	Ref[1] & Page no 163-188	BB
8.	**Tutorial**	1 hr		
9.	1.6.Two Dimensional transformation 1.6.1.Translation 1.6.2.Rotation	1 hr	Ref[1] & Page no 204-208	BB

	1.6.3. Scaling 1.6.4. Matrix Representation			
10.	1.7. Composite transformation 1.8. Reflection 1.9.Shearing	1 hr	Ref[1] & Page no 211-223	BB
11.	1.10.Two Dimensional viewing 1.11.Viewing pipeline 1.12. Viewing Coordinate Reference frame.	1 hr	Ref[1] & Page no 236-239	BB
12.	**Tutorial**	1 hr		
13.	1.13. Window to view port coordinate transformation 1.14. Two Dimensional viewing functions	1 hr	Ref[1] & Page no 240-242	BB
14.	1.15. Clipping Operation 1.15.1. Point clipping 1.15.2. Text clipping 1.15.3. Line clipping algorithm	1 hr	Ref[1] & Page no 244-250	BB
15.	1.16.Polygon clipping algorithm	1 hr	Ref[1] & Page no 257-262	BB
16.	**Tutorial**	1 hr		
	Total : 12 hrs			
	UNIT- II THREE DIMENSIONAL CONCEPTS Three-Dimensional Object Representations – Three-Dimensional Geometric and Modeling Transformations – Three-Dimensional Viewing – Color models – Animation.			
17.	2.1. 3D concepts 2.2.Display methods 2.2.1.Parallel projections 2.2.2.Perspective projections	1 hr	Ref[1] & Page no 316-319	BB,PPT
18.	2.3. 3D geometric transformation 2.3.1.Translation 2.3.2.Scaling	1 hr	Ref[1] & Page no 427-428,440	BB,PPT
19.	2.4.Reflection 2.5.Shear 2.6. Modeling transformation. 2.7. Coordinate transformation	1 hr	Ref[1] & Page no 442-449	BB
20.	**Tutorial**	1 hr		
21.	2.8. 3D viewing 2.8.1. Viewing pipeline 2.8.2. Viewing coordinates	1 hr	Ref[1] & Page no 451-453	BB
22.	2.9. 3D Projections	1 hr	Ref[1] & Page no 458-467	BB
23.	2.10. Color models 2.11. Properties of light 2.12.Chromaticity diagram	1 hr	Ref[1] & Page no 584-588	BB
24.	**Tutorial**	1 hr		
25.	2.13.XYZ color model 2.14.RGB color model 2.15. YIQ color model	1 hr	Ref[1] & Page no 589-594	BB

26.	2.16. CMYcolor model 2.17.HSV color model 2.18. HLS color model	1 hr	Ref[1] & Page no 594-599	BB
27.	2.19. Animation Key frame systems 2.20. Morphing 2.21. Fractals	1 hr	Ref[1] & Page no 604-608	BB
28.	**Tutorial**	1 hr		
	Total : 12 hrs			
	UNIT- III MULTIMEDIA SYSTEMS DESIGN An Introduction – Multimedia applications – Multimedia System Architecture – Evolving technologies for Multimedia – Defining objects for Multimedia systems – Multimedia Databases.			
29.	I3.1.Introduction to multimedia 13.1.1.Definition 13.1.2.Benefits 13.1.3. Need 13.1.4. Multimedia Elements	1 hr	Ref[2] & Page no1-8	BB
30.	13.2.Multimedia applications 13.3.Document imaging 13.4. Image processing 13.5. Image recognition	1 hr	Ref[2] & Page no 8-13	BB
31.	13.6. Motion video application 13.7. Electronic messaging 13.8. Universal multimedia application.	1 hr	Ref[2] & Page no 13-16	BB
32.	**Tutorial**	1 hr		
33.	13.9.Multimedia System Architecture 13.10.HRGD 13.11.IMA framework	1 hr	Ref[2] & Page no19-22	BB
34.	13.12.Network architecture for multimedia system 13.13.Networking standards	1 hr	Ref[2] & Page no22-26	BB
35.	13.14.Evolving technologies for multimedia systems 13.5.Hypermedia documents 13.6. HDTV technologies 13.7.UDTV	1 hr	Ref[2] & Page no 26-28	BB
36.	13.8. 3D Holography 13.9.Fuzzy logic 13.10. DSP	1 hr	Ref[2] & Page no 29 -33	BB
37.	13.11.Defining objects for Multimedia systems 13.12.MM storage and retrieval	1 hr	Ref[2] & Page no 33,44 -47	BB
38.	13.13. DBMS for MM systems	1 hr	Ref[2] & Page no 47-50	
39.	**Tutorial**	1 hr		
	Total : 12 hrs	1 hr		

	UNIT- IV MULTIMEDIA STORAGE TECHNOLOGIES Compression & Decompression for bitmap images, still images and Movies– Multimedia I/O technologies -– Storage and retrieval Technologies.			
40.	4.1. Compression 4.1.1. Definition, Types 4.2. Binary image compression	1 hr	Ref[2] & Page no 52-67	BB
41.	4.3.Color still image compression(JPEG)	1 hr	Ref[2] & Page no 70-106	BB
42.	4.4.Moving Picture Compression (MPEG)	1 hr	Ref[2] & Page no 107-117	BB
43.	4.5.Multimedia I/O technologies 4.5.1.Issues 4.5.2.Digital Pen	1 hr	Ref[2] & Page no 189-199	BB
44.	4.6.Display systems	1 hr	Ref[2] & Page no 199-210	BB
45.	4.7. Types of Printers	1 hr	Ref[2] & Page no 211-216	BB
46.	**Tutorial**	1 hr		
47.	4.8. Types of scanners	1 hr	Ref[2] & Page no 216-220	BB
48.	4.9..Storage and retrieval technologies 4.9.1.Magnetic media	1 hr	Ref[2] & Page no 266-273	BB
49.	4.10.RAID levels 4.11.Optical Media 4.12. Cache management for storage systems	1 hr	Ref[2] & Page no 274-284,288,311	BB
50.	**Tutorial**	1 hr		
	Total : 12hrs			
	UNIT- V HYPERMEDIA Digital voice and audio – Video image and animation – Full motion video – Hypermedia messaging - Mobile Messaging – Hypermedia message component – Creating Hypermedia message – Integrated multimedia message standards			
51.	5.1.Voice recognition system 5.2.Digital camera 5.3.Video camera	1 hr	Ref[2] & Page no 224 -260	BB
52.	5.4.Hypermedia Messaging	1 hr	Ref[2] & Page	BB, PPT

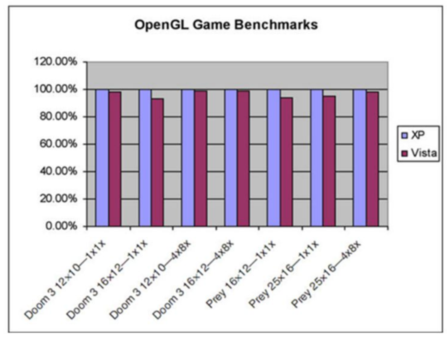
			no 472--473	
53.	5.5.Applications 5.6.Design consideration	1 hr	Ref[2] & Page no 448-455	BB
54.	**Tutorial**	1 hr		
55.	5.7.Display and Play back issues 5.8.Mobile messaging	1 hr	Ref[2] & Page no 465-468	BB, PPT
56.	5.9.Hypermedia message components 5.9.1.Text 5.9.2.Voice 5.9.3.Motion video	1 hr	Ref[2] & Page no 474-480	BB, PPT
57.	5.10. Full Motion Video	1 hr	Ref[2] & Page no 260-264	BB
58.	5.11.Linking in Hypertext documents	1 hr	Ref[2] & Page no 480-482	BB
59.	5.12.Creating Hypermedia messages	1 hr	Ref[2] & Page no 482	BB
60.	5.13.Integrated MM message standards.	1 hr	Ref[2] & Page no 482 483-488	BB
	Total : 12 hrs	1 hr		

Total : 60 Hrs

UNIT- I OUTPUT PRIMITIVES

Introduction - Line, Circle Drawing Algorithms – Attributes – Two-Dimensional Geometric Transformations – Two-Dimensional Clipping and Viewing

1. BASICS

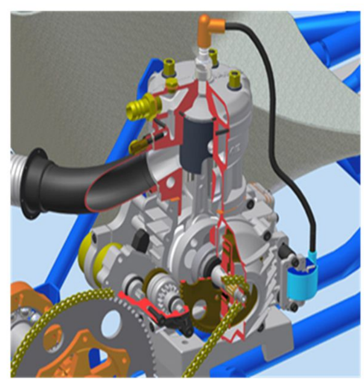


Computer graphics is a sub-field of computer science and is concerned with digitally synthesizing and manipulating visual content. Although the term often refers to three-dimensional computer graphics, it also encompasses two-dimensional graphics and image processing.

APPLICATIONS OF COMPUTER GRAPHICS

Computer graphics finds its application in various areas such as advertisements, art, education, entertainment, engineering, medicine, mathematics, business, scientific research and spatial, temporal applications.

A few application areas of computer graphics are listed below

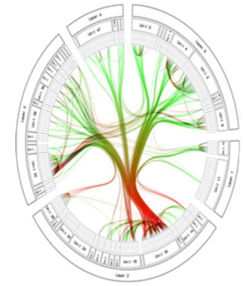


COMPUTER AIDED DESIGN

A major use of computer graphics is in design processes, particularly for engineering and architectural systems, but almost all products are now computer designed..

PRESENTATION GRAPHICS

Another major application area is presentation graphics, which is used to produce illustrations for reports or to generate 35-mm slides or transparencies for use with projectors. Presentation graphics is commonly used to summarize financial, statistical, mathematical, scientific, and economic data for research reports, managerial reports, consumer information bulletins, and other types of reports.



COMPUTER ART

Various painting packages are available. With cordless, pressure-sensitive stylus, artists can produce electronic paintings which simulate different brush strokes, brush widths, and colors. Photorealistic techniques, morphing and animations are very useful in commercial art. For films, 24 frames per second are required. For video monitor, 30 frames per second are required.

VISUALIZATION

Scientists, engineers, medical personnel, business analysts, and others often need to analyze large amounts of information or to study the behavior of certain processes. Numerical simulations carried out on supercomputers frequently produce data files containing thousands and even millions of data values.

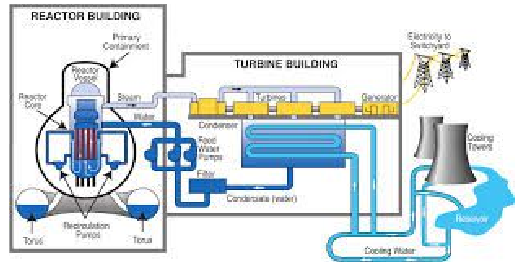
GRAPHICAL USER INTERFACES

It is common now for software packages to provide a graphical interface. A major component of a graphical interface is a window manager that allows a user to display multiple-window areas. Each window can contain a different process that can contain graphical or non graphical displays.

ENTERTAINMENT

Computer graphics are commonly used in making Motion pictures, Music videos, and TV shows, Computer games. Graphics objects can be combined with the live action or graphics and image processing techniques can be used to produce a transformation of one person or object into another (Morphing).

EDUCATION AND TRAINING



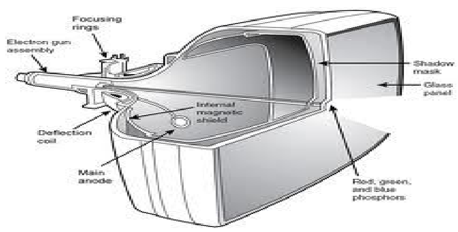
In Education, Computer graphics is used to produce computer-based training courses (popularly called CBTs) and reference books like encyclopedia and almanacs. A CBT lets the user go through a series of presentations, text about a particular topic, and associated illustrations in various information formats. Edutainment is an informal term used to describe combining education with entertainment, especially multimedia entertainment.

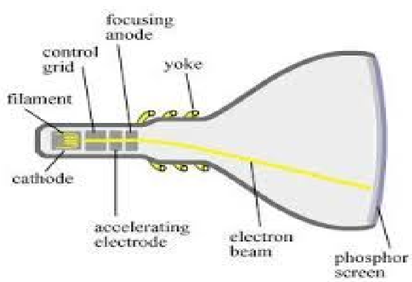
1.1 GRAPHIC SYSTEMS

VIDEO DISPLAY DEVICES

Typically, the primary output device in a graphics system is a video monitor the operation of most video monitors is based on the standard cathode-ray tube (CRT) design.

Cathode-Ray Tubes (CRT) – it is the most common video display device.





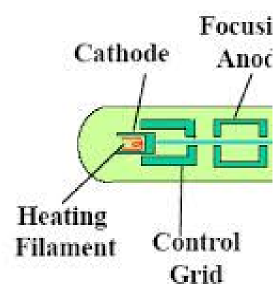


Figure 1. Cathode Ray Tube

The basic operation of a CRT:

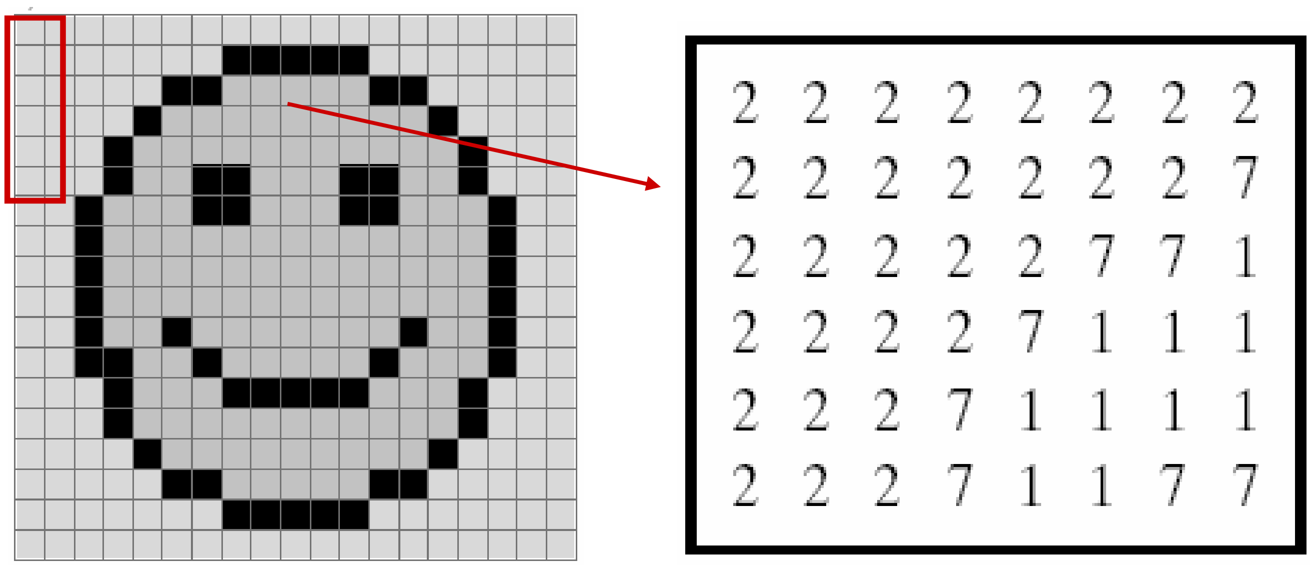
A beam of electrons (cathode rays) emitted by an electron gun, passes through focusing and deflection systems that direct the beam toward specified positions on the phosphor screen.

Figure 2. Illustration of a shadow-mask Color CRT

The light emitted by phosphor fades very rapidly, so it needs to redraw the picture repeatedly. There are 2 kinds of redrawing mechanisms: Raster-Scan and Random-Scan

Raster-Scan Displays

In a raster-scan system, the electron beam is swept across the screen, one row at a time from top to bottom. Stored intensity values are then retrieved from the refresh buffer and "painted" on the screen one row (scan line) at a time .Each screen point is referred to as a pixel or peel (shortened form of picture element). The capability of a raster-scan system to store intensity information for each screen point makes it well suited for the realistic display of scenes containing subtle shading and color patterns.



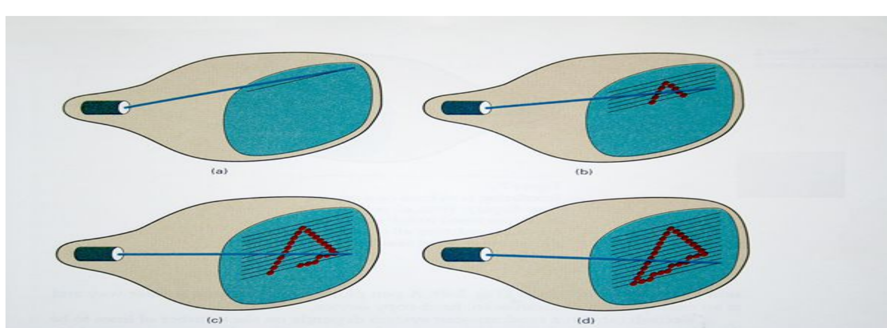


Figure 3. Raster Scan Display

Random-Scan Displays

When operated as a random-scan display unit, a CRT has the electron beam directed only to the parts of the screen where a picture is to be drawn. Random scan monitors draw a picture one line at a time and for this reason are also referred to as vector displays (or stroke-writing or calligraphic displays). The component lines of a picture can be drawn and refreshed by a random-scan system in any specified order

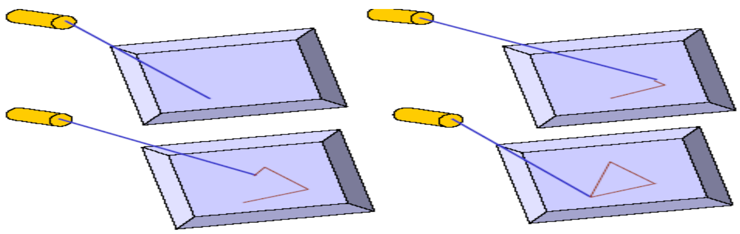


Figure 4. Random Scan Display

1.2 OUTPUT PRIMITIVES

Shapes and colors of the objects can be described internally with pixel arrays or with sets of basic geometric structures, such as straight line segments and polygon color areas. The scene is then displayed either by loading the pixel arrays into the frame buffer or by scan converting the basic geometric-structure specifications into pixel patterns.

POINTS AND LINES

Point plotting is accomplished by converting a single coordinate position furnished by an application program into appropriate operations for the output device in use.

The formula for the point to be plotted is put pixel(x, y, color)

1.3 LINE-DRAWING ALGORITHMS

The Cartesian slope-intercept equation for a straight line is

$y = m.x + c$ (1)

With m representing the slope of the line and c as the y intercept. Given that the two endpoints of a line segment are specified at positions (x_1 y_1,) and (x_2, y_2), we can determine values for the slope m and y intercept c with the following calculations:

$m = y_2 - y_1 / x_2 - x_1$ (2)

$c = y1 - m.x_1$ (3)

Algorithms for displaying straight lines are based on the line equation (1) and the calculations given in Eqs. (2) and (3).

For any given x interval Δx along a line, we can compute the corresponding y interval from y from Eq. (2) as

$\Delta y = m \, \Delta x$ (4)

$\Delta x = \Delta y/m$ (5)

These equations form the basis for determining deflection voltages in analog devices. For lines with slope magnitudes | m| < 1, Δx can be set proportional to a small horizontal deflection voltage and the corresponding vertical deflection is then set proportional to Δy as calculated from Eq. (4). For lines whose slopes have magnitudes |m|> 1, Δy can be set proportional to a small vertical deflection voltage with the corresponding horizontal deflection voltage set proportional to Δx, calculated from Eq. (5). For lines with m = 1, Δx = Δy and the horizontal and vertical deflections voltages are equal. In each case, a smooth line with slope m is generated between the specified endpoints.

1)DDA algorithm

Consider first a line with positive slope. If the slope is less than or equal to 1, we sample at unit x intervals (Δx = 1) and compute each successive y value as

if slope m > 1

then $\Delta y > \Delta x$ & $\Delta x = x_2 - x_1 = 1$

Since $\Delta y/\Delta x = m$

$\therefore \quad \Delta y/1 = m$

$\Delta y = m$ or $y_2 - y_1 = m$

$\therefore \quad y_2 = y_1 + m$

$\therefore \quad y_{k+1} = y_k + m$ and $x_{k+1} = x_k + 1$ (6)

Subscript k takes integer values starting from 1, for the first point, and increases by 1 until the final endpoint is reached. Since m can be any real number between 0 and 1, the calculated y values must be rounded to the nearest integer.

For lines with a positive slope greater than 1, we reverse the roles of x and y. That is, we sample at unit y intervals Δy = 1) and calculate each succeeding x value as

$$x_{k+1} = x_k + 1/m \quad \text{and} \quad y_{k+1} = y_k + 1 \qquad (7)$$

Equations (6) and (7) are based on the assumption that lines are to be processed from the left endpoint to the right endpoint . If this processing is reversed, so that the starting endpoint is at the right, then either we have $\Delta x = -1$ and

$$y_{k+1} = y_k - m \qquad (8)$$

Or (when the slope is greater than 1) we have $\Delta y = -1$ with

$$x_{k+1} = x_k - 1/m \qquad (9)$$

Equations (6) through (9) can also be used to calculate pixel positions along a line with negative slope. If the absolute value of the slope is less than 1 and the start endpoint is at the left, we set $\Delta x = 1$ and calculate y values with Eq. (6). When the start endpoint is at the right (for the same slope), we set $\Delta x = -1$ and obtain y positions from Eq. (8). Similarly, when the absolute value of a negative slope is water than 1, we use $\Delta y = -1$ and Eq. (9) or we use $\Delta y = 1$ and Eq. (7).

Advantage:

Faster method for calculating pixel position then the equation of a pixel position.

$y=mx+b$

Disadvantage:

The accumulation of round of error is successive addition of the floating point increments is used to find the pixel position but it take lot of time to compute the pixel position.

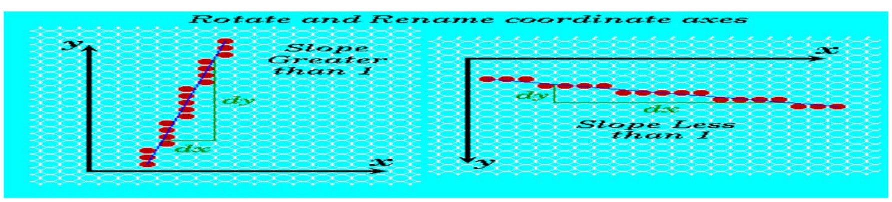
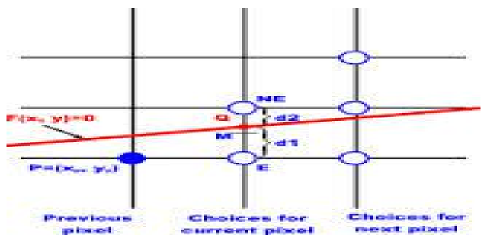
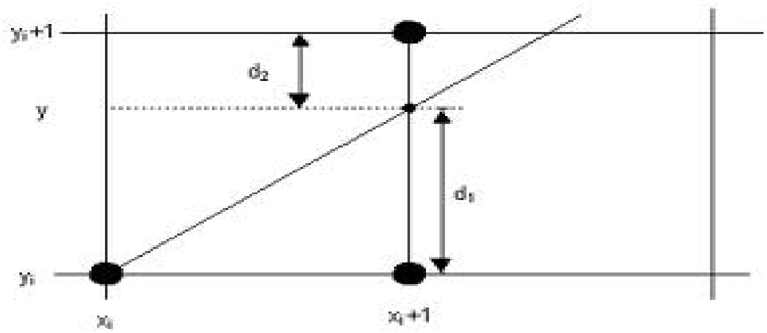


Figure 1.1 DDA Line Algorithms

2) Bresenham's line algorithm

The Bresenham line algorithm is an algorithm which determines which order to form a close approximation to a straight line between two given points. It is commonly used to draw lines on a computer screen, as it uses only integer addition, subtraction and bit shifting, all of which are very cheap operations in standard computer architectures. It is one of the earliest algorithms developed in the field of computer graphics





Decision Parameter Derivation in Midpoint Circle Algorithm:

The equation of a line at a point (x, y) is

$y=mx + b$---------> 1

The equation of a line at a point (x_{k+1}, y) is

$y=m (x_k + 1) + b$--------> 2

From the figure, the distance d_1 and d_2 are calculated as follows,

$d_1=y - y_k$-------->3

$d_2=y_{k+1} - y$-------->4

Sub 2 in 3

$d_1=m (x_{k+1}) +b - y_k$

Sub 2 in 4

$d_2=y_{k+1} - [m (x_k + 1) +b]$

$d_2=y_k + 1- m (x_k + 1)- b$

if $d1 - d2 < 0$ then $y_{i+1} \leftarrow y_i$
if $d1 - d2 > 0$ then $y_{i+1} \leftarrow y_i + 1$

We want integer calculations in the loop, but m is not an integer. Looking at definition of m ($m = \Delta y / \Delta x$) we see that if we multiply m by Δx, we shall remove the denominator and hence the floating point number.

For this purpose, let us multiply the difference (d1 - d2) by Δx and call it $\mathbf{P_k}$

$\mathbf{P_k=\Delta x (d_1-d_2)}$

$d_1-d_2=m (x_k + 1) + b - y_k - y_k -1 + m (x_k + 1) + b$
$=2m (x_k + 1) - 2y_k + 2b - 1$
$P_k=\Delta x [2m (x_k +1) - 2y_k + 2b - 1]$ ---------->5

$=\Delta x\ 2\ \Delta y/\Delta x\ (x_k+1)\ -\Delta x_k\ 2y_k + \Delta x\ 2b - \Delta x$
$=2\Delta y\ (x_k + 1) + 2\Delta x\ y_k + \Delta\ (2b - 1)$
$=2\Delta y\ x_k + 2\Delta y - 2\Delta x y_k + \Delta x\ (2b - 1)$
$P_k=2\Delta y x_k - 2\Delta x y_k + c$ where $c=2\Delta y + \Delta x\ (2b - 1)$

Because the sign of P_k is the same as the sign of d1 – d2, we could use it inside the loop to decide whether to select pixel at $(x_k + 1, y_k)$ or at $(x_k + 1, y_k+1)$. Note that the loop will only include integer arithmetic. There are now 6 multiplications, two additions and one selection in each turn of the loop.

However, we can do better than this, by defining P_k recursively.

$P_{k+1}=2\Delta y x_{k+1} - 2\Delta x y_{k+1} + c$

$P_{k+1}-P_k=2\Delta y\ (x_{k+1} - x_k) - 2\Delta x\ (y_{k+1} - y_k)$

$P_{k+1}-P_k=2\Delta y\ (x_k +1 - x_k)-2\Delta x\ (y_{k+1} - y_k)$

$\mathbf{P_{K+1}=P_k+2\Delta y - 2\Delta x\ (y_{k+1} - y_k)}$.

Recursive definition for P_k

$\mathbf{P_{K+1}=P_k + 2\Delta y - 2\Delta x\ (y_{k+1} - y_k)}$.

If you now recall the way we construct the line pixel by pixel, you will realize that the underlined expression: $y_{k+1} - y_k$ can be either 0 (when the next pixel is plotted at the same y-coordinate, i.e. d1 – d2 < 0); or 1 (when the next pixel is plotted at the next y-coordinate, i.e. d1 – d2 > 0). Therefore the final recursive definition for P_k will be based on choice, as follows (remember that the sign of P_k is the same as the sign of d1 – d2):

if $P_k < 0$, $\mathbf{P_{K+1} = P_k + 2\Delta y}$ because $2\Delta x \cdot (y_{k+1} - y_k) = 0$

if $P_k > 0$, $\mathbf{P_{K+1} = p_k + 2\Delta y - 2\Delta x}$ because $(y_{k+1} - y_k) = 1$

At this stage the basic algorithm is defined and the initial value for parameter p_0.

$\mathbf{P_0 = 2\Delta y - \Delta x}$

ALGORITHM

- Input line endpoints, (X1,Y1) and (X2, Y2)
- Calculate constants:

  $\Delta x = X2 - X1$

$\Delta y = Y2 - Y1$

$2\Delta y$

$2\Delta y - \Delta x$

- Assign value to the starting parameters: k = 0

  $P_0 = 2\Delta y - \Delta x$

- Plot the pixel at ((X1,Y1)
- For each integer x-coordinate, x_k, along the line

  if $P_k < 0$ plot pixel at $(x_k + 1, y_k)$

  $P_{k+1} = P_k + 2\Delta y$

  else plot pixel at $(x_k + 1, y_k + 1)$

  $P_{k+1} = P_k + 2\Delta y - 2 \Delta x$

  increment k while $x_k < X2$

1.4 CIRCLE / CURVE GENERATING ALGORITHMS

Properties of a circle**:**

- A circle is defined as a set of points that are all the given distance *(xc, yc).* This distance relationship is expressed by the Pythagorean theorem in Cartesian coordinates as

  $(x - xc) 2 + (y - yc) 2 = r2$

- We could use this equation to calculate the points on the circle circumference by stepping along x-axis in unit steps from *xc-r* to *xc+r* and calculate the corresponding y values at each position as

  $y = yc + (r2 - (x - xc) 2)1/2$

- This is not the best method:
  - *Considerable amount of computation*
  - *Spacing between plotted pixels is not uniform*

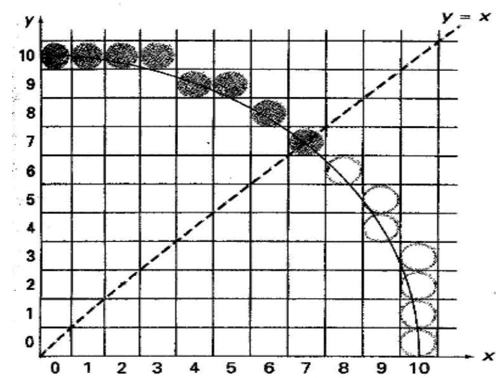
Polar co-ordinates for a circle

• We could use polar coordinates r and θ,

$x = x_c + r\cos\theta \quad y = y_c + r\sin\theta$

• A fixed angular step size can be used to plot equally spaced points along the circumference

• A step size of 1/r can be used to set pixel positions to approximately 1 unit apart for a continuous boundary • But, note that circle sections in adjacent octants within one quadrant are symmetric with respect to the 45 deg line dividing the to octants



• Thus we can generate all pixel positions around a circle by calculating just the points within the sector from x=0 to x=y

• This method is still computationally expensive

Midpoint Circle Algorithm

Jack E. Bresenham invented this algorithm in 1962. The objective was to optimize the graphic algorithms for basic objects, in a time when computers were not as powerful as they are today. This algorithm is called incremental, because the position of the next pixel is calculated on the basis of the last plotted one, instead of just calculating the pixels from a global formula. Such logic is faster for computers to work on and allows plotting circles without trigonometry. The algorithm use only integers, and that's where the strength is: floating point calculations slow down the processors.

All in all, incremental algorithms are 30% faster than the classical ones, based on floating points and trigonometry.

Decision Parameter in Midpoint Circle Algorithm:

A circle is defined as a set of points that are all at a given distance *r* from a center positioned at. (x_c, y_c).

This is represented mathematically by the equation

$$(x - x_c)^2 + (y - y_c)^2 = r^2 \quad \text{------------- (1)}$$

Using equation (1) we can calculate the value of y for each given value of x as

$$y = y_c \pm \sqrt{r^2 - (x_c - x)^2}$$ ---------------- (2)

Thus one could calculate different pairs by giving step increments to x and calculating the corresponding value of y. But this approach involves considerable computation at each step and also the resulting circle has its pixels sparsely plotted for areas with higher values of the slope of the curve.

Midpoint Circle Algorithm uses an alternative approach, wherein the pixel positions along the circle are determined on the basis of incremental calculations of a decision parameter.

Let

$$f(x,y) = (x - x_c)^2 + (y - y_c)^2 - r^2$$ -------------- (3)

Thus f(x, y) =0 represents the equation of a circle.

Further, we know from coordinate geometry, that for any point, the following holds:

1. $f(x,y) = 0 \quad \Rightarrow$ *The point lies on the circle.*
2. $f(x,y) < 0 \quad \Rightarrow$ *The point lies within the circle.*
3. $f(x,y) > 0 \quad \Rightarrow$ *The point lies outside the circle.*

In Midpoint Circle Algorithm, the decision parameter at the k^{th} step is the circle function evaluated using the coordinates of the midpoint of the two pixel centers which are the next possible pixel position to be plotted.

Let us assume that we are giving unit increments to x in the plotting process and determining the y position using this algorithm. Assuming we have just plotted the k^{th} pixel at (X_k,Y_k), we next need to determine whether the pixel at the position (X_{k+1},Y_k) or the one at (X_{k+1} , Y_{k-1}) is closer to the circle.

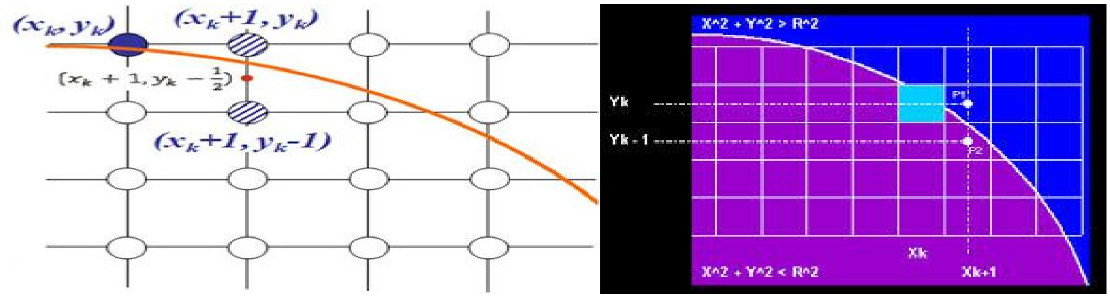


Figure 1.2 Bresenham's Circle Algorithm

Our decision parameter p_k at the k^{th} step is the circle function evaluated at the midpoint of these two pixels.

The coordinates of the midpoint of these two pixels are $(X_k+1,\ Y_k-1/2)$. Thus p_k

$$p_k = f\left(x_k + 1, y_k - \tfrac{1}{2}\right) = (x_k + 1)^2 + (y_k - \tfrac{1}{2})^2 - r^2 \quad \text{--------------- (4)}$$

Successive decision parameters are obtained using incremental calculations, thus avoiding a lot of computation at each step. We obtain a recursive expression for the next decision parameter i.e. at the $k+1^{th}$ step, in the following manner.

Using Equ. (4), at the $k+1^{th}$ step, we have:

$$p_k = f\left(x_k + 1, y_k - \tfrac{1}{2}\right) = (x_k + 1)^2 + (y_k - \tfrac{1}{2})^2 - r^2$$

$$p_{k+1} = f\left(x_{k+1} + 1, y_{k+1} - \tfrac{1}{2}\right) = (x_{k+1} + 1)^2 + (y_{k+1} - \tfrac{1}{2})^2 - r^2$$

Or, $p_{k+1} = (x_k + 1 + 1)^2 + (y_{k+1} - \tfrac{1}{2})^2 - r^2$

Or, $p_{k+1} = (x_k + 2)^2 + (y_{k+1} - \tfrac{1}{2})^2 - r^2$ ---------------------------(5)

(5)-(4) gives

$$p_{k+1} - p_k = (x_k + 2)^2 - (x_k + 1)^2 + (y_{k+1} - \tfrac{1}{2})^2 - (y_k - \tfrac{1}{2})^2 - r^2 + r^2$$

Or, $p_{k+1} = p_k + (2x_k + 3).1 + (y_{k+1} + y_k - 1)(y_{k+1} - y_k)$ -------(6)

Now if $P_k <= 0$, then the midpoint of the two possible pixels lies within the circle, thus north pixel is nearer to the theoretical circle. Hence, $Y_{k+1} = Y_k$. Substituting this value of in Equ. (6), we have

$$p_{k+1} = p_k + (2x_k + 3) + (y_k + y_k - 1)(y_k - y_k)$$

If $p_k > 0$ then the midpoint of the two possible pixels lies outside the circle, thus south pixel is nearer to the theoretical circle. Hence, $Y_{k+1} = Y_k - 1$. Substituting this value of in Equ. (6), we have

$$p_{k+1} = p_k + (2x_k + 3) + (y_k - 1 + y_k - 1)(y_k - y_k - 1)$$

$$\text{Or, } p_{k+1} = p_k + 2(x_k - y_k) + 5$$

For the boundary condition, we have x=0, y=r. Substituting these values in (4), we have

For integer values of pixel coordinates, we can approximate $P_0 = 1 - r$,

Thus we have:

$$\text{If } p_k \leq 0: \quad y_{k+1} = y_k \text{ and } p_{k+1} = p_k + (2x_k + 3)$$

$$\text{If } p_k > 0: \quad y_{k+1} = y_k - 1 \text{ and } p_{k+1} = p_k + 2(x_k - y_k) + 5$$

$$\text{Also, } p_0 = 1 - r$$

Drawing the circle:

We can reduce our calculation drastically (8th fraction) by making use of the fact that a circle has 8 way symmetry. Thus after calculating a pixel position (x, y) to be plotted, we get 7 other points on the circle corresponding to it. These are:

(x, y); (x,-y); (-x, y); (-x,-y); (y, x); (y,-x); (-y, x); (-y,-x)

The Actual algorithm

1: Input radius r and circle center (xc,yc) and obtain the first point on the circumference of the circle centered on the origin as

(x0,y0) = (0, r)

2: Calculate the initial value of the decision parameter as

P0 = 5/4 - r

3: At each xk positions starting at k = 0, perform the following test:

If pk < 0 , the next point along the circle centered on (0,0) is (xk+1, yk) and

pk+1 = pk + 2xk+1 + 1

Otherwise the next point along the circle is (xk+1, yk-1) and

pk+1 = pk + 2xk+1 +1 -2yk+1

Where 2xk+1 = 2xk+2 and 2yk+1 = 2yk-2

4: Determine symmetry points in the other seven octants

5: Move each calculated pixel position (x,y) onto the circular path centered on (x,yc) and plot the coordinate values

x = x+ xc, y= y+ yc

6: Repeat steps 3 through 5 until x >= y

1.5 ELLIPSE-GENERATING ALGORITHMS

Properties of Ellipses

An ellipse is defined as the set of points such that the sum of the distances from two fixed positions (foci) is the same for all points (Fig. *b17).*

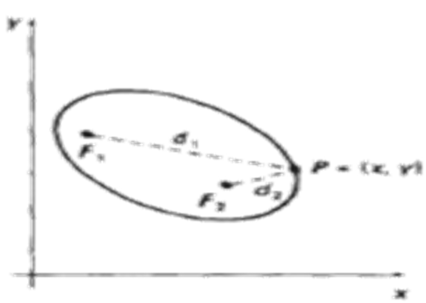


Figure 1.3 Ellipse generated about foci

If the distances to the two foci from any point P = ***(x, y)*** on the ellipse are labeled **dl** and **d2,** then the general equation of an ellipse can be stated as

d, + ***d,*** = constant ------------------------------- 1

Expressing distance d1 and d2 in terms of F1=(x1, y1) and F2=(x2, y2), we have

$$\sqrt{(x - x_1)^2 + (y - y_1)^2} + \sqrt{(x - x_2)^2 + (y - y_2)^2} = \text{constant}$$

By squaring this equation, isolating the remaining radical, and then squaring again, we can rewrite the general ellipse equation in the form

Ax2 + **By2** + ***Cxy*** + ***Dx*** + ***Ey*** + **F** = **0------------------------------------ (3)**

Where the coefficients *A,* B, C, *D, E,* and F are evaluated in terms of the focal coordinates and the dimensions of the major and minor axes of the ellipse.

The major axis $2r_y^2x \geq 2r_x^2y$ is the straight line segment extending from one side of the ellipse to the other through the foci. The minor axis spans the shorter dimension of the ellipse, bisecting the major axis at the halfway position (ellipse center) between the two foci. An interactive method for specifying an ellipse in an arbitrary orientation is to input the two foci and a point on the ellipse boundary.

With these three coordinate positions, we can evaluate the constant in Eq. 2

Then the coefficients in Eq. 3 can be evaluated and used to generate pixels along the Elliptical path.

Using polar coordinates r and 0**.** we can also describe the ellipse in standard position with the parametric equations:

$$x = x_c + r_x \cos\theta$$

$$y = y_c + r_y \sin\theta \qquad \text{------------ (5)}$$

The midpoint ellipse method is applied throughout the first quadrant in two parts.

$$\left(\frac{x - x_c}{r_x}\right)^2 + \left(\frac{y - y_c}{r_y}\right)^2 = 1$$

Figure 1.4 Regions in ellipse

Regions 1 and 2 can be processed in various ways.

We can start at position (0, r_y) and step clockwise along the elliptical path in the first quadrant, shifting from unit steps in *x* to unit steps in *y* when the slope becomes less than −1.0.

Alternatively, we could start at (r_x, 0) and select points in a counterclockwise order, shifting from unit steps in *y* to unit steps in *x* when the slope becomes greater than −1.

$$f_{\text{ellipse}}(x, y) = r_y^2x^2 + r_x^2y^2 - r_x^2r_y^2 \qquad 2r_y^2x = 2r_x^2y$$

- We define an ellipse function from Eq. 3-37 with $(x_c, y_c) = (0, 0)$ as which has the following properties:
- Starting at $(0, r_y)$, we take unit steps in the x direction until we reach the boundary between region 1 and region 2 .
- Then we switch to unit steps in the y direction over the remainder of the curve in the first quadrant.
- At each step we need to test the value of the slope of the curve.
- The ellipse slope is calculated as
- At the boundary between region 1 and region 2, $dy/dx = -1$ and

Assuming position (x_k, y_k) has been selected in the previous step; we determine the next position along the ellipse path by evaluating the decision parameter (that is, the ellipse function) at the midpoint:

$$p1_k = f_{\text{ellipse}}\left(x_k + 1, y_k - \frac{1}{2}\right.$$

$$= r_y^2(x_k + 1)^2 + r_x^2\left(y_k - \right.$$

$$\frac{dy}{dx} = -\frac{2r_y^2x}{2r_x^2y}$$

$$f_{\text{ellipse}}(x, y) \begin{cases} < 0, & \text{if } (x, y) \text{ is inside the ellipse boundary} \\ = 0, & \text{if } (x, y) \text{ is on the ellipse boundary} \\ > 0, & \text{if } (x, y) \text{ is outside the ellipse boundary} \end{cases}$$

- If $p1_k < 0$, the midpoint is inside the ellipse and the pixel on scan line y_k is closer to the ellipse boundary.
- Otherwise, the imposition is outside or on the ellipse boundary, and we select the pixel on scan line $y_k - 1$.
- At the next sampling position (x_k+1 + 1 = x_k + 2), the decision parameter for region 1 is evaluated as

$$p1_{k+1} = f_{\text{ellipse}}\left(x_{k+1} + 1, y_{k+1} - \frac{1}{2}\right)$$

$$= r_y^2[(x_k + 1) + 1]^2 + r_x^2\left(y_{k+1} - \frac{1}{2}\right)^2 - r_x^2r_y^2$$

or

$$p1_{k+1} = p1_k + 2r_y^2(x_k + 1) + r_y^2 + r_x^2\left[\left(y_{k+1} - \frac{1}{2}\right)^2 - \left(y_k - \frac{1}{2}\right)^2\right]$$

where y_{k+1} is either y_k or $y_k - 1$, depending on the sign of $p1_k$.

Decision parameters are incremented by the following amounts:

$$\text{increment} = \begin{cases} 2r_y^2 x_{k+1} + r_y^2, & \text{if } p1_k < 0 \\ 2r_y^2 x_{k+1} + r_y^2 - 2r_x^2 y_{k+1}, & \text{if } p1_k \geq 0 \end{cases}$$

At the initial position (0, r_y), these two terms evaluate to

$$2r_y^2 x = 0$$
$$2r_x^2 y = 2r_x^2 r_y$$

- In region 1, the initial value of the decision parameter is obtained by evaluating the ellipse function at the start position $(x_0, y_0) = (0, r_y)$:

$$p1_0 = f_{\text{ellipse}}\left(1, r_y - \frac{1}{2}\right)$$
$$= r_y^2 + r_x^2\left(r_y - \frac{1}{2}\right)^2 - r_x^2 r_y^2$$

$$p1_0 = r_y^2 - r_x^2 r_y + \frac{1}{4} r_x^2$$

- Over region 2, we sample at unit intervals in the negative y direction, and the midpoint is now taken between horizontal pixels at each step
- For this region, the decision parameter is evaluated as

$$p2_k = f_{\text{ellipse}}\left(x_k + \frac{1}{2}, y_k - 1\right)$$
$$= r_y^2\left(x_k + \frac{1}{2}\right)^2 + r_x^2(y_k - 1)^2 - r_x^2 r_y^2$$

- If $p2_k > 0$, the mid position is outside the ellipse boundary, and we select the pixel at x_k.
- If $p2_k <= 0$, the midpoint is inside or on the ellipse boundary, and we select
- Pixel position x_{k+1}.
- To determine the relationship between successive decision parameters in region 2,we evaluate the ellipse function at the next sampling step $y_k+1 -1 = y_k-2$:

Midpoint Ellipse Algorithm

1. Input r_x, r_y, and ellipse center (x_c, y_c), and obtain the first point on an ellipse centered on the origin as

$$(x_0, y_0) = (0, r_y)$$

2. Calculate the initial value of the decision parameter in region 1 as

$$p1_0 = r_y^2 - r_x^2 r_y + \frac{1}{4} r_x^2$$

$$p2_{k+1} = f_{\text{ellipse}}\left(x_{k+1} + \frac{1}{2}, y_{k+1} - 1\right)$$
$$= r_y^2\left(x_{k+1} + \frac{1}{2}\right)^2 + r_x^2[(y_k - 1) - 1]^2 - r_x^2 r_y^2$$

or

$$p2_{k+1} = p2_k - 2r_x^2(y_k - 1) + r_x^2 + r_y^2\left[\left(x_{k+1} + \frac{1}{2}\right)^2 - \left(x_k + \frac{1}{2}\right)^2\right]$$

with x_{k+1} set either to x_k or to $x_k + 1$, depending on the sign of $p2_k$.

- When we enter region 2, the initial position (x_0, y_0) is taken as the last position selected in region 1 and the initial decision parameter in region 2 is then

$$p2_0 = f_{\text{ellipse}}\left(x_0 + \frac{1}{2}, y_0 - 1\right)$$
$$= r_y^2\left(x_0 + \frac{1}{2}\right)^2 + r_x^2(y_0 - 1)^2 - r_x^2 r_y^2$$

3. At each x_k position in region 1, starting at $k = 0$, perform the following test. If $p1_k < 0$, the next point along the ellipse centered on $(0, 0)$ is (x_{k+1}, y_k) and

$$p1_{k+1} = p1_k + 2r_y^2 x_{k+1} + r_y^2$$

Otherwise, the next point along the ellipse is $(x_k + 1, y_k - 1)$ and

$$p1_{k+1} = p1_k + 2r_y^2 x_{k+1} - 2r_x^2 y_{k+1} + r_y^2$$

with

$$2r_y^2 x_{k+1} = 2r_y^2 x_k + 2r_y^2, \qquad 2r_x^2 y_{k+1} = 2r_x^2 y_k - 2r_x^2$$

and continue until $2r_y^2 x \geq 2r_x^2 y$.

4. Calculate the initial value of the decision parameter in region 2 as

$$p2_0 = r_y^2\left(x_0 + \frac{1}{2}\right)^2 + r_x^2(y_0 - 1)^2 - r_x^2 r_y^2$$

where (x_0, y_0) is the last position calculated in region 1.

5. At each y_k position in region 2, starting at $k = 0$, perform the following test. If $p2_k > 0$, the next point along the ellipse centered on $(0, 0)$ is $(x_k, y_k - 1)$ and

$$p2_{k+1} = p2_k - 2r_x^2 y_{k+1} + r_x^2$$

Otherwise, the next point along the ellipse is $(x_k + 1, y_k - 1)$ and

$$p2_{k+1} = p2_k + 2r_y^2 x_{k+1} - 2r_x^2 y_{k+1} + r_x^2$$

using the same incremental calculations for x and y as in region 1. Continue until $y = 0$.

6. For both regions, determine symmetry points in the other three quadrants.

7. Move each calculated pixel position (x, y) onto the elliptical path centered on (x_c, y_c) and plot the coordinate values:

$$x = x + x_c, \qquad y = y + y_c$$

1) LINE ATTRIBUTES

Basic attributes of a straight line segment are its type, its width, and its color.

In some graphics packages, lines can also be displayed using selected pen or brush options.

Line Type

Line-type attribute - solid lines, dashed lines, and dotted lines.

A dashed line could be displayed by generating an inter dash spacing that is equal to the length of the solid sections.

To set line type attributes in a PHICS application program, a user invokes the function setline type (lt)

Where parameter 'lt' is assigned a positive integer value of 1, 2, 3, or 4 to generate lines that are, respectively, solid, dashed, dotted, or dash-dotted.

Line Width

For lines with slope magnitude greater than 1, we can plot thick lines with horizontal spans, alternately picking up pixels to the right and left of the line path.

.

A third type of line cap is the projecting square cap. Here, we simply extend the line and add butt caps that are positioned one-half of the line width beyond the specified endpoints.

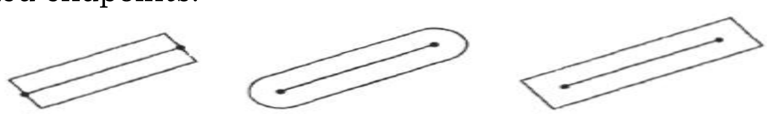


Figure 1.5 a) Butt caps b) Round caps and c) Projecting square caps

A miter join is accomplished by extending the outer boundaries of each of the two lines until they meet.

A round join is produced by capping the connection between the two segments with a circular boundary whose diameter is equal to the line width.

A bevel join is generated by displaying the line segments with butt caps and filling in the triangular gap where the segments meet.

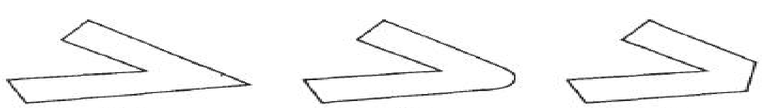


Figure 1.6 Thick line segments connected with a) Miter join b) Round join and c) Bevel join

Pen and Brush Options

These shapes can be stored in a pixel mask that identifies the array of pixel positions that are to be set along the line path. Lines generated with pen (or brush) shapes can be displayed in various widths by changing the size of the mask.

Line Color

When a system provides color (or intensity) options, a parameter giving the current color index is included in the list of system-attribute values.

2)CURVE ATTRIBUTES

Parameters for curve attributes are the same as those for line segments. We can display curves with varying colors, widths, dot dash patterns, and available pen or brush options.

3)CHARACTER ATTRIBUTES

The appearance of displayed characters is controlled by attributes such as font, size, color, and orientation. Attributes can be set for entire character strings (text) and for individual characters defined as marker symbols.

4)TEXT ATTRIBUTES

There are a great many text options that can be made available to graphics programmers. First of all, there is the choice of font (or typeface), London, 'Times Roman, and various special symbol groups.

1.7 TWO DIMENSIONAL TRANSFORMATIONS

2D Transformations are editing operations like change in position, shape, size, and orientation etc. of an object with respect to some point or its frame of reference. They are named so as the transformations take place in the same plane. Transformations can be broadly classified into the following:

Geometry transformations: Changing the position, shape, size, etc. of the object with respect to a frame of reference.

Coordinates transformations: Changing the frame of reference itself.

TYPES OF 2D GEOMETRIC TRANSFORMATIONS:

1.7.1 Basic Transformations

1)Translation: It refers to the transfer of a point from one position to another along a straight-line path in the same frame of reference. We translate a 2D point by adding translation distances, 'tx' and 'ty', to the original coordinate position. Transforming P(x, y) to P' (x', y'):

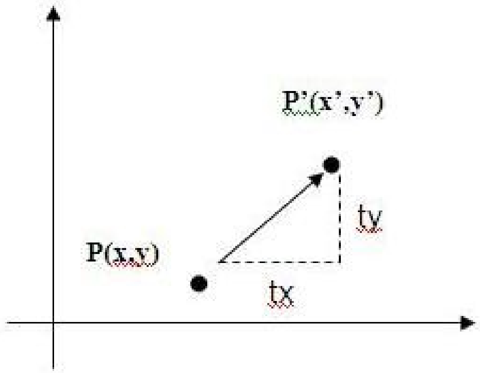


Figure 1.7 Translating a point from position P to position P' with translation vector T

From the above figure we get the coordinates of the new point P' (x', y') as:

$$x' = x + tx$$
$$y' = y + ty$$

Matrix representation of the above equations is as follows:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} tx \\ ty \end{bmatrix} + \begin{bmatrix} x \\ y \end{bmatrix}$$

We convert the above matrix representation to Homogeneous Coordinate System so that later on, 'n' processes or transformations can be combined into a single Composite transformation using only Multiplication as the operation.

Homogeneous Coordinate System:

A point P(x, y) in Cartesian coordinate system can be represented in Homogeneous System as:

P(x/h, y/h, h)

Where, 'h' is known as the Homogeneous Factor.

For e.g. (2,3) in Cartesian System can be written as (2,3,1) or (4,6,2) or (6,9,3) in Homogeneous System.

Hence, the translation matrix using the Homogeneous Coordinates System can be written as follows:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & tx \\ 0 & 1 & ty \\ 0 & 0 & 1 \end{pmatrix} * \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

2) **<u>Scaling:</u>** It alters the size of an object. This operation can be carried out for polygons by multiplying the coordinate values (x, y) of each vertex by the **Scaling Factors,** 'sx' and 'sy', to produce the transformed coordinates (x', y') for each vertex. Hence, all the vertices are repositioned.

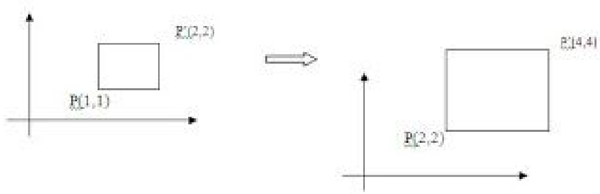


Figure 1.8 scaling a polygon with scaling factors

In the above Figure, the scaling factors are as follows:

sx = 2, sy = 2

If the 2 scaling factors are equal, then it is called **uniform scaling.** If the 2 are different then it is called **Differential scaling** (For e.g.: sx = 2, sx = 3).Hence, the scaling equations for each transformed vertex are:

x' = sx * x

y' = sy * y

The above can be represented in the matrix form as follows:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} sx & 0 \\ 0 & sy \end{pmatrix} * \begin{pmatrix} x \\ y \end{pmatrix}$$

As stated before, for composite transformations, we need to convert the above form to the Homogeneous Coordinates System. Hence the above can be written as:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} sx & 0 & 0 \\ 0 & sy & 0 \\ 0 & 0 & 1 \end{pmatrix} * \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

In case the 2 operations of translation and scaling are performed together, the result of the below matrix multiplication will give us the composite matrix:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} sx & 0 & 0 \\ 0 & sy & 0 \\ 0 & 0 & 1 \end{pmatrix} * \begin{pmatrix} 1 & 0 & tx \\ 1 & 0 & ty \\ 0 & 0 & 1 \end{pmatrix} * \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

3) Rotation: It is applied to an object by repositioning it along a circular path in the xy-plane about the rotation point or the pivot point. Considering rotation of a point about the origin we have,

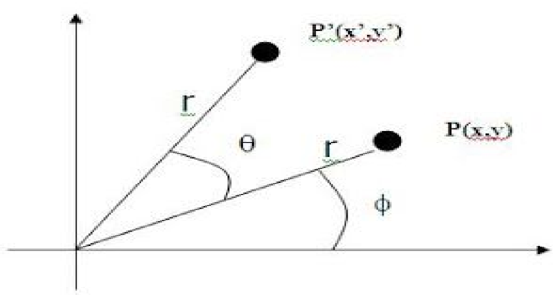


Figure 1.9. Rotating a point from position P to position P' with angle θ

In the above figure a point P(x, y), at an angle Φ, distant 'r' from the origin, has been rotated about the origin by angle θ in counter clockwise direction, hence giving the new point, P' (x', y'). For counter-clockwise rotation, θ is considered to be positive and negative for clockwise rotation.

For calculation for the coordinates of the new point P', from the figure we have the following:

$x' = r \cos(\Phi + \theta) = r \cos\Phi \cos\theta - r \sin\Phi \sin\theta$ (Using cos (a+b) = cos a.cos b – sin a.sin b)

$y' = r \sin(\Phi + \theta)$

$= r \cos\Phi \sin\theta + r \sin\Phi \cos\theta$ (Using sin (a+b) = sin a.cos b + cos a.sin b)

From figure, $x = r \cos\Phi$ and $y = r \sin\Phi$. Hence, we get the equations for rotation of a point about the origin as,

$$\mathbf{x' = x \cos\theta - y \sin\theta}$$

$$\mathbf{y' = x \sin\theta + y \cos\theta}$$

Converting the above to matrix multiplication form, we get,

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} * \begin{pmatrix} x \\ y \end{pmatrix}$$

Expressing the above in Homogeneous coordinate system, we get,

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} * \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

The above matrix is with respect to the origin.

General Pivot point Rotation:

For any other point as reference, we need to apply a combination of the 3 transformations to find out the actual result by using Composite matrix.

Consider the following transformation:

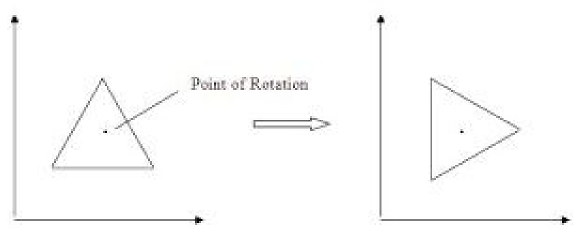
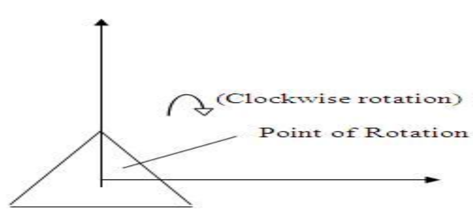


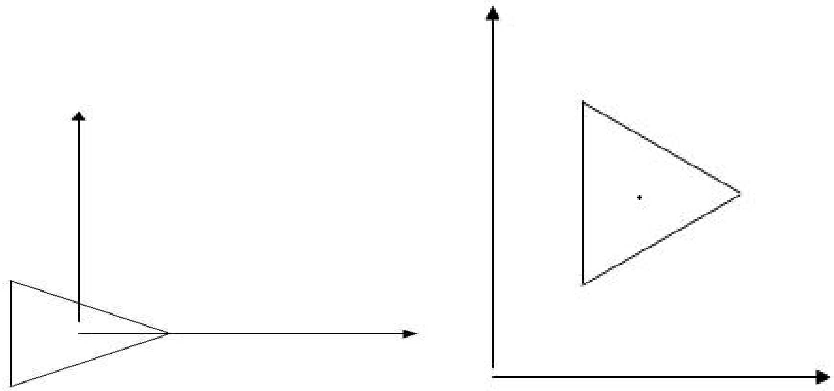
Figure 1.10 General pivot point rotation

We can perform the above transformation in 3 steps:

1. Shift the pivot of rotation to coincide with the origin using Translation.



- Rotate the triangle about the origin as the pivot of rotation.



- Inverse Translation to the previous points.

So we can combine the 3 steps into one by writing the composite matrix as a result of the following matrix multiplication,

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & tx \\ 0 & 1 & ty \\ 0 & 0 & 1 \end{pmatrix} * \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} * \begin{pmatrix} 1 & 0 & tx \\ 0 & 1 & ty \\ 0 & 0 & 1 \end{pmatrix} * \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

1.7.2 Other Transformations

1) 2D Reflection:

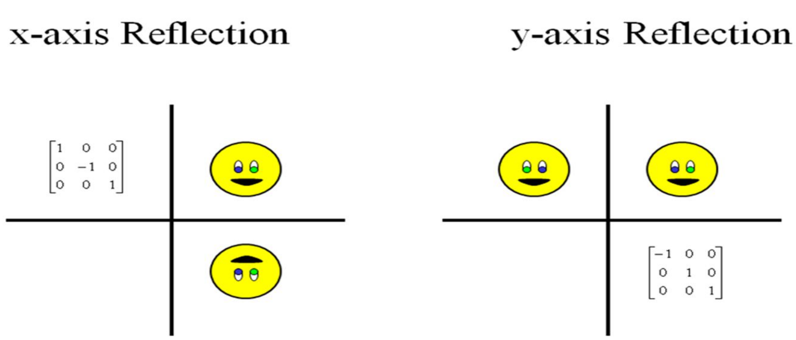
It is the transformation that produces the mirror image of an object. Mirror image is generated by rotating the object 180^0 around the axis of reflection.

The Cartesian coordinate equation for x axis reflection is

$x = x; \quad y = -y$

The Cartesian coordinate equation for y axis reflection is

$x = -x; \quad y = y$



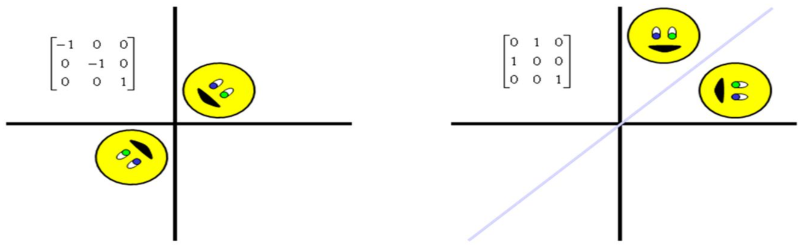
The Cartesian coordinate equation for reflection about the origin is

$x = x; \quad y = -y$

The Cartesian coordinate equation for reflection about the line y=x are

$x = y; \qquad y = x$

Reflection about the origin and the line x=y



2) 2D Shear

It is a transformation that distorts the shape of the object. It causes the object to be pushed on one side.

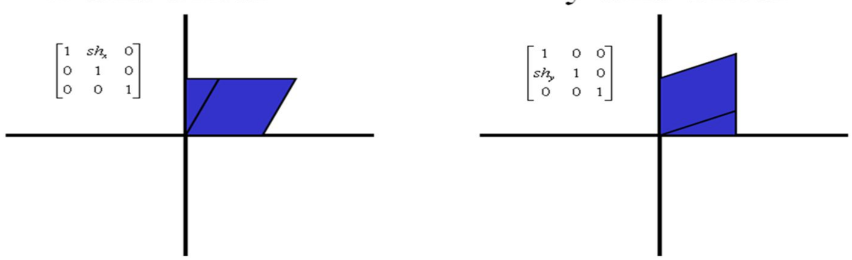
The Cartesian coordinate equation for x axis shear is

$x = x; \qquad y = y + sh_y$

The Cartesian coordinate equation for y axis shear is

$x = x + sh_x; \qquad y = y$

x-axis Shear   y-axis Shear



1.8 TWO DIMENSIONAL CLIPPING AND VIEWING

2D Viewing

• Clipping/View Window - The rectangle defining the part of the world we wish to display. [Rectangle specified in World Coordinates]

• Display/Screen/Device Coordinate System - The coordinate system of the frame buffer.

• Display Window - The graphics application’s window on the screen on computers.

• Viewport - The rectangle within the screen (or display window) defining where the image will appear. (Default is usually entire screen or display window.) [Viewport rectangle specified in Display Window or Screen Coordinates]

• Viewing Transformation - The process of going from a view window in world coordinates to a viewport in screen coordinates. Figure 8-1 A clipping window and associated viewport, specified as rectangles aligned with the coordinate axes.

Window to Viewport Transformation

Window - Area that defines what is to be displayed.

Viewport - Area that defines where it is to be displayed.

Whatever area is selected in the window is sent to the viewport. The size of the viewport might be smaller or larger than the window.

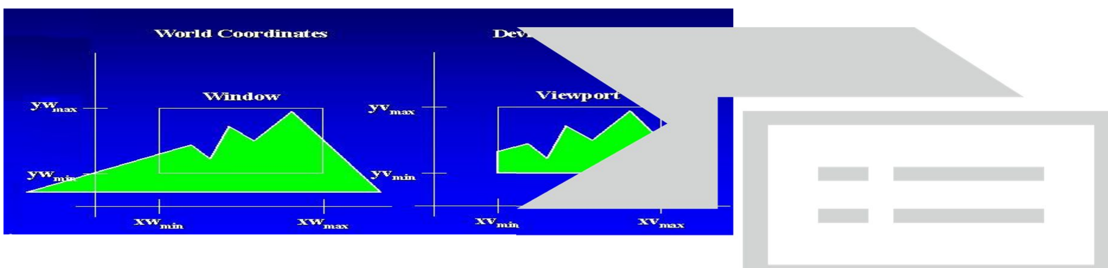
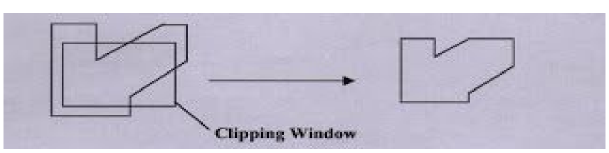
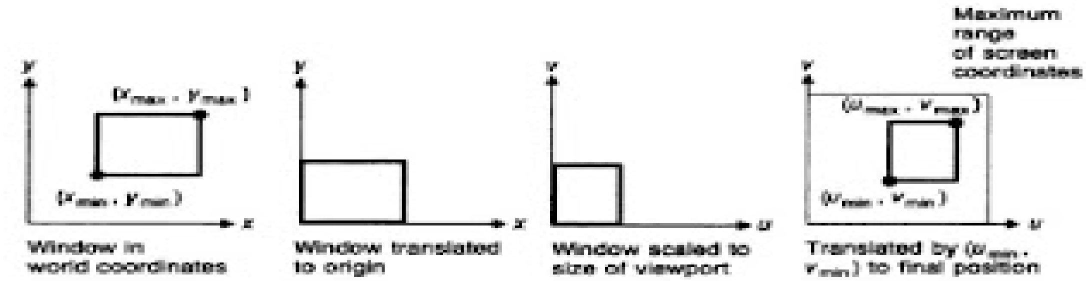


Figure 1.11 the window in world coordinates and the viewport in screen coordinates

The window-to-viewport transformation maintains the relative position of a point in window as well as in the viewport. A point at position (xw, yw) in the window is mapped into position (xv, yv) in the associated viewport.





The inputs to this transformation are window-size and viewport-size. The following steps are involved in the transformation:

1. Translation 2. Scaling 3. Inverse Translation

The transformation matrix that maps the window from world coordinates into the viewport in screen coordinates can be developed from the above equations. The overall matrix is:

$$\begin{bmatrix} xv \\ yv \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & xv_{min} \\ 0 & 1 & yv_{min} \\ 0 & 0 & 1 \end{bmatrix} . \begin{bmatrix} sx & 0 & 0 \\ 0 & sy & 0 \\ 0 & 0 & 1 \end{bmatrix} . \begin{bmatrix} 1 & 0 & -xw_{min} \\ 0 & 1 & -yw_{min} \\ 0 & 0 & 1 \end{bmatrix} . \begin{bmatrix} xw \\ yw \\ 1 \end{bmatrix}$$

1.8.1 CLIPPING OPERATIONS

A line clipping algorithms takes as input two endpoints of line segment and returns one (or more) line segments. A polygon clipper takes as input the vertices of a polygon and returns one (or more) polygons. The region against which an object is to be clipped is referred as clip window and it can be in any shape.

Types of clipping

- Point clipping
- Line clipping
- Area (Polygon) clipping
- Curve clipping
- Text clipping

Point clipping

If the x coordinate boundaries of the clipping rectangle are Xmin and Xmax, and the y coordinate boundaries are Ymin and Ymax, then the following inequalities must be satisfied for a point at (X, Y) to be inside the clipping rectangle:

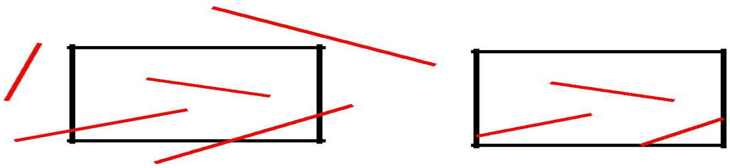
Xmin < X < Xmax and Ymin < Y < Ymax

If any of the four inequalities does not hold, the point is outside the clipping rectangle.

Line Clipping

It is the process of clipping the lines outside the clipping window.

Possible relationships between line positions and a rectangular clipping region



Before clipping After clipping

Figure 1.12.Line Clipping

Cohen-Sutherland Line Clipping

- Every line end-point in a picture is assigned a four-digit binary code, called *a region code that* identifies the location of the point relative to the point relative to the boundaries of the clipping rectangle. Regions are set up in reference to the boundaries as shown in fig.

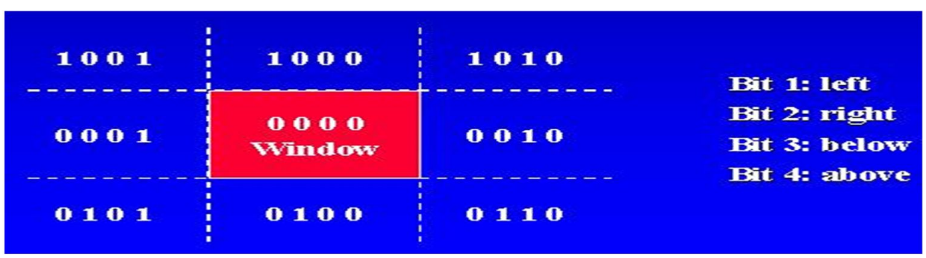


Figure 1.13 Region code

- Each bit position in the region code is used to indicate one of the four relative co-ordinates positions of the point with respect to the clip window: to the left, right, top or bottom.

  bit 1: left

  bit 2: right

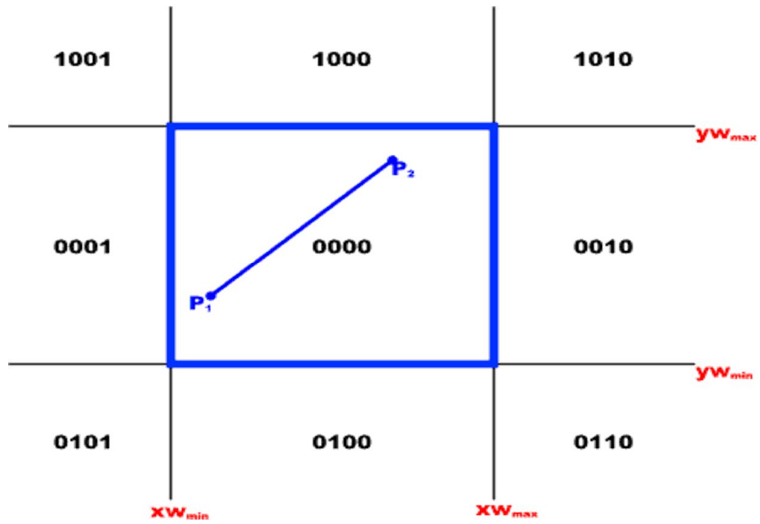
  bit 3: below

  bit 4: above

- A value of 1 in any bit position indicates that the point is in that relative position; otherwise, the bit position is set to 0.

- Bit values in the region code are determined by comparing endpoint coordinate values*(x, y)* to the clip boundaries.

  Bit 1 is set to 1 if $x<xw_{min}$. *Or the sign bit of x – xmin*
- Trivially Accept: Any lines that are completely contained within the window and boundaries have a region code of 0000 for both endpoints, and we trivially accept these lines.
- Trivially Reject: Any lines that have a 1 in the same bit position in the region codes for each end point are completely outside the clipping rectangle, and we trivially reject these lines. A method that can be used to *test* lines for total clipping is to perform the 'logical AND' operation with both region codes. If the result is not 0000, the line is completely outside the clipping region.
- Intersecting Lines: Lines that cannot be identified as completely inside or completely outside a clip window by these tests are checked for intersection with the window *boundaries*.
- Intersection points with a clipping boundary can be calculated using the slope-intercept form of the line equation. For a line with endpoint coordinates (x_1, y_1) and (x_2, y_2), the y coordinates of the intersection point with a vertical boundary can be obtained with the vertical boundary can be obtained with the calculation.



4 bit clip code:

A B R L

A = $y > yw_{max}$

B = $y < yw_{min}$

R = $x > xw_{max}$

L = $x < xw_{min}$

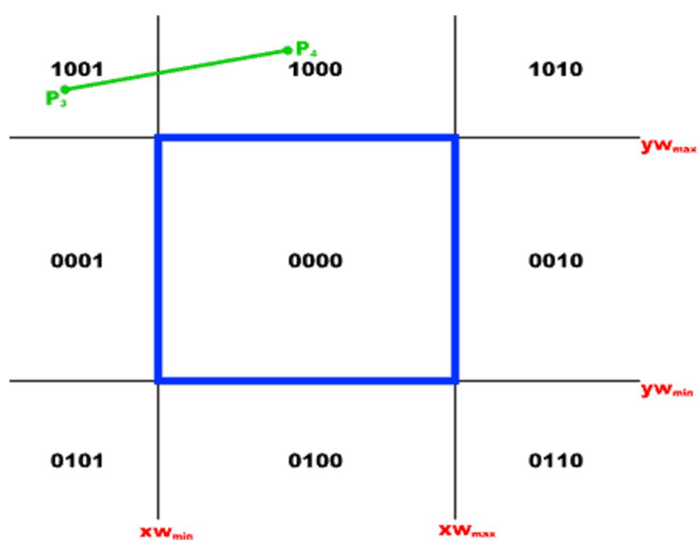
Calculate clip code

P_1 = 0 0 0

P_2 = 0 0 0 0

--- 0 0 0 0

=> accept segment P_1 P_2 (trivial accept since both end points are in the window



$$y = y_1 + m(x - x_1)$$

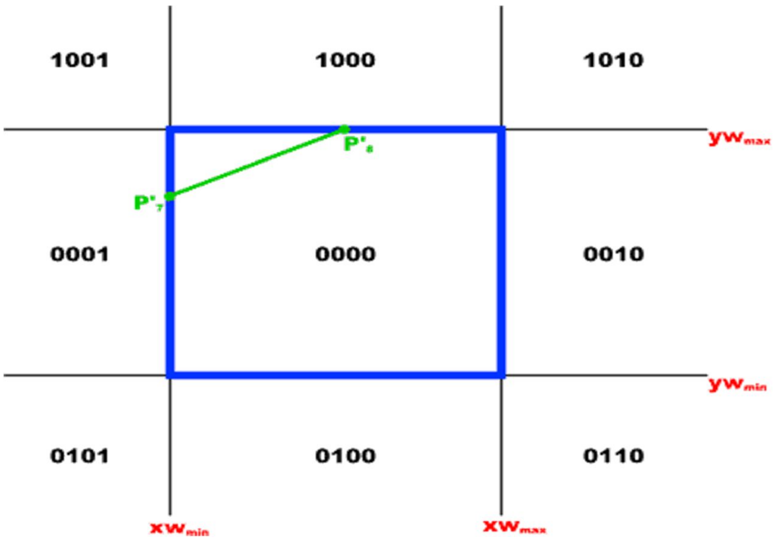
Where the x value is set either to xw_{min} or to xw_{max} , and the slope the line is calculated as

$$m = (y_2 - y_1) / (x_2 - x_1).$$

- Similarly, if we are looking for the intersection with a horizontal boundary ,the *x coordinate* can be calculated as

$$x = x_1 + (y - y_1 / m)$$

With y set as either to ym_{min} or to yw_{max}.



4 bit clip code:

A B R L

A = y > yw_{max}

B = y < yw_{min}

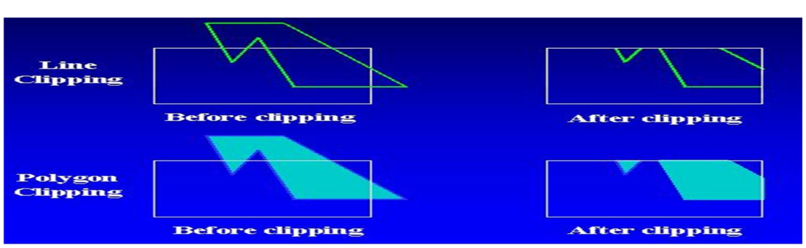
R = x > xw_{max}

L = x < xw_{min}

P_5 = 0 0 0 1 P_6 = 1 0 0 0	P_7 = 0 0 0 1 P_8 = 1 0 0 0	(1) Calculate clip code
0 0 0 0	0 0 0 0	
P'_5 = (xw_{min}, y_5 + m(xw_{min} - x_5))	P'_7 = (xw_{min}, y_7 + m(xw_{min} - x_7))	(2) Clip against left
P'_5 = 1 0 0 0 P_6 = 1 0 0 0	P'_7 = 0 0 0 0 P_8 = 1 0 0 0	(3) Update clip code
--- 1 0 0 0 => reject P'_5 P_6	--- 0 0 0 0	

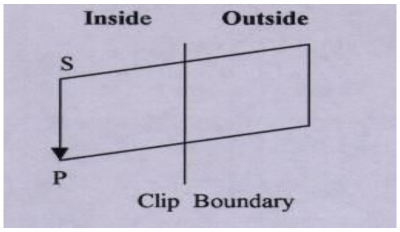
1.8.2 Polygon Clipping

In polygon clipping we clip a polygon by processing polygon boundary as a whole against each edge of the clipping window .

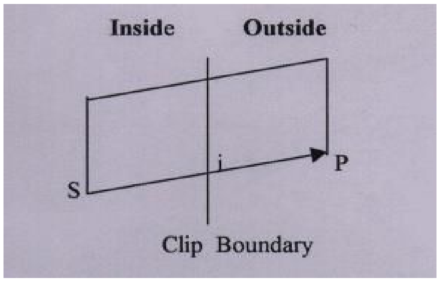


Sutherland Hodgeman Algorithm

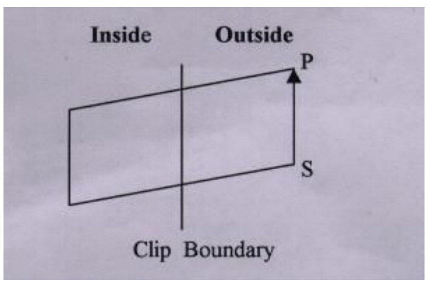
Case 1: If both endpoints (*s* and *p*) of the edge are inside the clipping edge, vertex *p* is added to the output list.

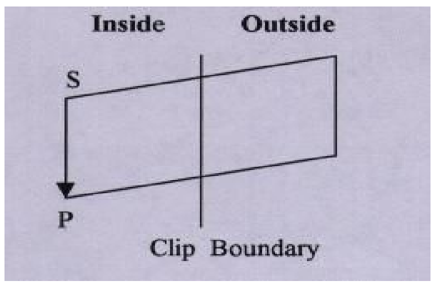


Case 2: If vertex s is inside the clipping edge and vertex p is outside of it, the intersection point, i, is output.

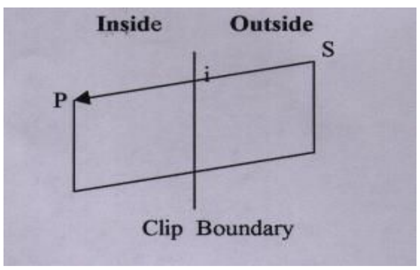


Case 3: If both endpoints (*s* and *p*) of the edge are outside of the clipping edge, there is no output.





Case 4: If vertex *p* is inside of the clipping edge and vertex *s* is outside of it, the intersection point, *i*, and vertex *p* are both added to the output list.



Processing the polygon boundary as a whole against each window edge is processing all polygon vertices against each clip rectangle boundary in turn.

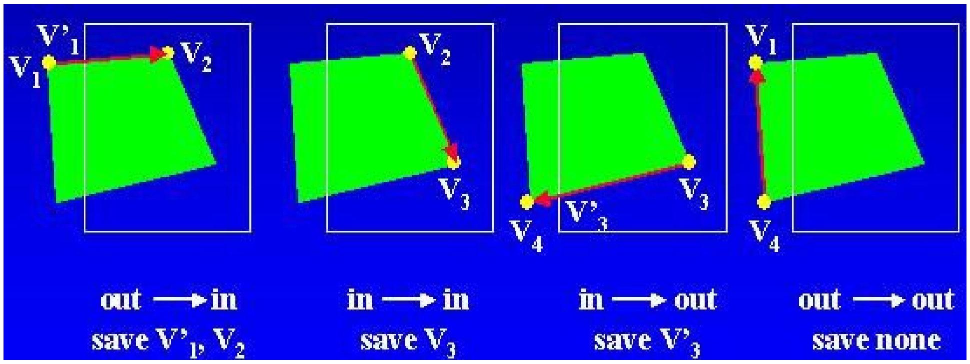


Figure 1.14. Sutherland Hodgeman Polygon clipping

Curve clipping

- Use bounding rectangle to test for overlap with a rectangular clip window.

Text clipping

- All-or-none string-clipping
- All-or-none character-clipping
- Clip the components of individual characters

PART C

1. (a) Digitize the line with end points (10,12) and (20,18) using DDA line drawing algorithm. (8)

(b) Using Bresenham's line drawing algorithm , find out which pixel would be turned on for the line with end points (4,4) to (12,9). (8)

2. Derive the Bresenham's line drawing algorithm and explain it with an example (16)

3. Consider the line from (1,1) to (6,4) . Use Bresenham's line drawing algorithm to rasterize this line and give output pixels.

5. Write down and explain midpoint circle drawing algorithm . Assume 10 cm as the radius and coordinate origin as the centre of the circle. with respect to 2D transformation

(i) scaling
(ii) Rotation
(iii) Translation

8. Explain window to viewport coordinate transformation

9. Explain the Sutherland Hodgeman polygon clipping algorithm
10. Write short notes on two dimensional viewing & anti aliasing

UNIT- II THREE-DIMENSIONAL CONCEPTS

Three-Dimensional Object Representations – Three-Dimensional Geometric and Modeling Transformations – Three-Dimensional Viewing – Color models – Animation.

2.1 THREE DIMENSIONAL OBJECT REPRESENTATIONS

To obtain display of a three-dimensional scene that has been modeled in world coordinates. we must first set up a coordinate reference for the "camera". This coordinate reference defines the position and orientation for the plane of the camera film which is the plane we want to us to display a view of the objects in the scene.

Boundary representations (B-reps) describe a three-dimensional object as a set of surfaces that separate the object interior from the environment. Typical examples of boundary representations are polygon facets and spline patches.

Space-partitioning representations are used to describe interior properties, by partitioning the spatial region containing an object into a set of small, non overlapping, contiguous solids (usually **cubes).**

3D Object representation Methods:

1. Polygon and Quadric surfaces: For simple Euclidean objects
2. Spline surfaces and construction: For curved surfaces
3. Procedural methods: Eg. Fractals, Particle systems
4. Physically based modeling methods
5. Octree Encoding

Objects may also associate with other properties such as mass, volume, so as to determine their response to stress and temperature etc.

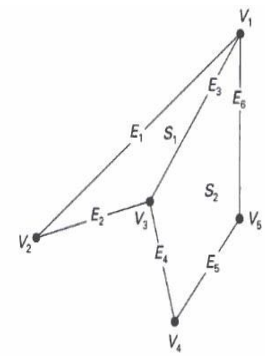
2.1.1 Polygon Surfaces

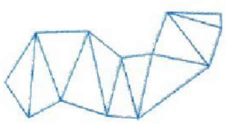
This method simplifies and speeds up the surface rendering and display of objects. For other 3D objection representations, they are often converted into polygon surfaces before rendering.

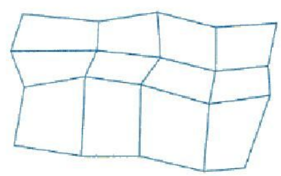
Polygon Mesh

Using a set of connected polygonally bounded planar surfaces to represent an object, which may have curved surfaces or curved edges.

- The wireframe display of such object can be displayed quickly to give general indication of the surface structure.
- Realistic renderings can be produced by interpolating shading patterns across the polygon surfaces to eliminate or reduce the presence of polygon edge boundaries.
- Common types of polygon meshes are triangle strip and quadrilateral mesh.







VERTEX TABLE
V_1: x_1, y_1, z_1
V_2: x_2, y_2, z_2
V_3: x_3, y_3, z_3
V_4: x_4, y_4, z_4
V_5: x_5, y_5, z_5

EDGE TABLE
E_1: V_1, V_2
E_2: V_2, V_3
E_3: V_3, V_1
E_4: V_3, V_4
E_5: V_4, V_5
E_6: V_5, V_1

POLYGON-SURFACE TABLE
S_1: E_1, E_2, E_3
S_2: E_3, E_4, E_5, E_6

Polygon Tables

This is the specification of polygon surfaces using vertex coordinates and other attributes:

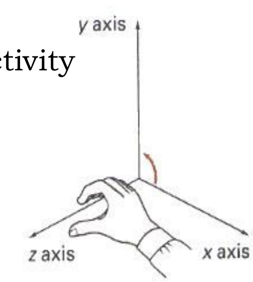
1. Geometric data table: vertices, edges, and polygon surfaces.
2. Attribute table: e.g. Degree of transparency and surface reflectivity

Some consistency checks of the geometric data table:

§ Every vertex is listed as an endpoint or at least 2 edges

§ Every edge is part of at least one polygon

§ Every polygon is closed



Plane equation and visible points

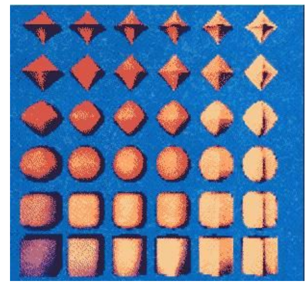
Consider a cube; each of the 6 planes has 2 sides: inside face and outside face. For each plane (in a right-handed coordinate system), if we look at its surface and take 3 points in counter-clockwise direction: (x_1, y1), (x_2, y2), and (x_3, y3), we can compute 4 values: A, B, C, D as

$$A = \begin{vmatrix} 1 & y_1 & z_1 \\ 1 & y_2 & z_2 \end{vmatrix} \qquad B = \begin{vmatrix} x_1 & 1 & z_1 \\ x_2 & 1 & z_2 \end{vmatrix} \qquad C = \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} \qquad D = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix}$$

$\begin{vmatrix} 1 & y_3 & z_3 \end{vmatrix}$ $\begin{vmatrix} x_3 & 1 & z_3 \end{vmatrix}$ $\begin{vmatrix} x_3 & y_3 & 1 \end{vmatrix}$ $\begin{vmatrix} x_3 & y_3 & z_3 \end{vmatrix}$

Then, the plane equation at the form: Ax+By+Cz+D=0 has the property that: If we substitute any arbitrary point (x, y) into this equation, then,Ax + By + Cz + D < 0 implies that the point (x, y) is inside the surface, and Ax + By + Cz + D < 1 implies that the point (x, y) is outside the surface.

2.1.2 Curved Surfaces



1. Regular curved surfaces can be generated as
-Quadric Surfaces, e.g. Sphere, Ellipsoid, or
-Super quadrics, e.g. Super ellipsoids

These surfaces can be represented by some simple parametric equations, e.g., for ellipsoid:

$x = r_x \cos^{s1}\phi \cos^{s2}\theta, \ -\pi/2 <= \phi <= \pi/2$

$y = r_y \cos^{s1}\phi \sin^{s2}\theta, \ -\pi <= \theta <= \pi$

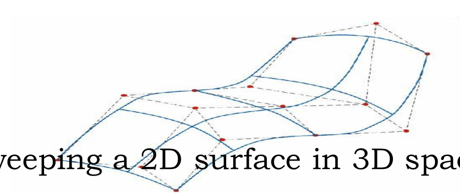
$z = r_z \sin^{s1} \phi$

Where s1, r_x, r_y, and r_x are constants. By varying the values of ϕ and θ, points on the surface can be computed.

2. Irregular surfaces can also be generated using some special formulating approach, to form a kind of blobby objects. The shapes showing a certain degree of fluidity.

3. Spline Representations

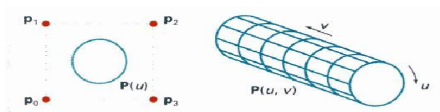
Spline means a flexible strip used to produce a smooth curve through a designated set of points. Several small weights are distributed along the length of the strip to hold it in position on the drafting table as the curve is

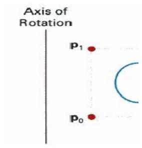


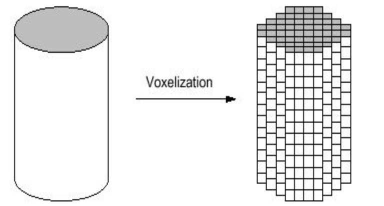
2.1.3 Sweep Representations

Sweep representations mean sweeping a 2D surface in 3D space to create an object. However, the objects created by this method are usually converted into polygon meshes and/or parametric surfaces before storing.

A Translational Sweep: A Rotational Sweep:







Other variations:

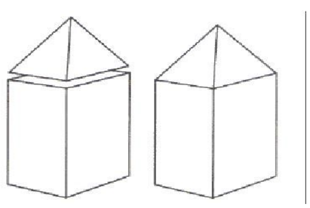
- We can specify a special path for the sweep as some curve function.
- We can vary the shape or size of the cross section along the sweep path.
- We can also vary the orientation of the cross section relative to the sweep path.

2.1.4 Constructive Solid-Geometry Methods

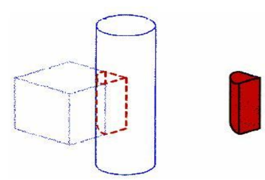
The Constructive Solid-Geometry Method (CSG) combines the volumes occupied by overlapping 3D objects using <u>set operations</u>:

- Union
- Intersection
- Difference

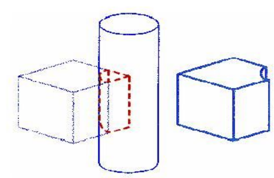
Object created by a union operation:



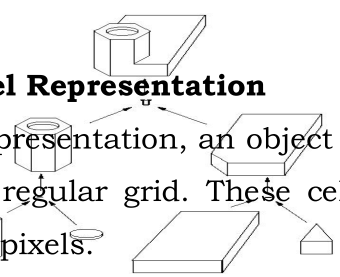
Object created by an intersection operation:



Object created by a difference operation:



A CSG object can be represented with a binary tree:



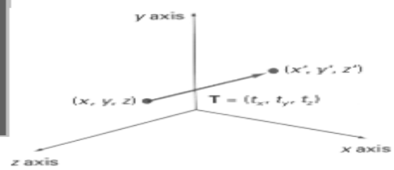
2.1.5 Voxel Representation

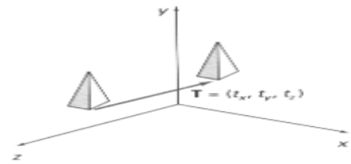
In voxel representation, an object is decomposed into identical cells arranged in a fixed regular grid. These cells are called voxels (volume elements), in analogy to pixels.

2.1.6 Octrees

Octrees are hierarchical tree structures that describe each region of 3D space as nodes. When compared with the basic voxel representation, octrees

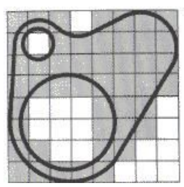
$$x' = x + t_x$$
$$y' = y + t_y$$
$$z' = z + t_z$$

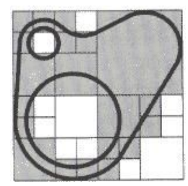




reduce storage requirements for 3D objects. It also provides a convenient representation for storing information about object interiors.Octree encoding procedure is an extension of the quad tree encoding of 2D images:

Bitmap representation of a 2D object: Quad tree encoding of a 2D object:





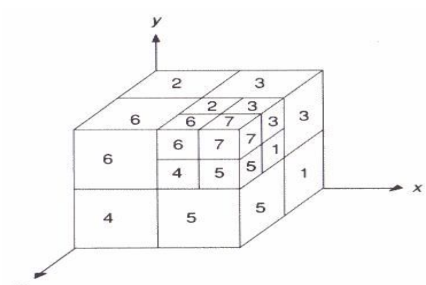
For octree representation of 3D data, the numbering method is as follows:

2.2 THREE DIMENSIONAL GEOMETRIC TRANSFORMATION:

1) 3D TRANSLATION

Figure 2.1.Translating a point with Translation vector (tx, ty, tz)

In a three-dimensional homogeneous coordinate representation, a point is translated from position P = ***(x, y, z)*** to position P' = ***(x', y', z'***) as shown in figure



$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Or

$$P' = T \cdot P$$

Parameters t_x, t_y, and t_z specifying translation distances for the coordinate directions x, y, and z, are assigned any real values.

An object is translated in three dimensions by transforming each of the defining points of the object.

2) 3D ROTATION

To generate a rotation transformation for an object, we must designate an axis of rotation (about which the object is to be rotated) and the amount of angular rotation.

Coordinate-Axes Rotations

The two-dimensional z-axis rotation equations are easily extended to three dimensions:

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 & (1-s_x)x_f \\ 0 & s_y & 0 & (1-s_y)y_f \\ 0 & 0 & s_z & (1-s_z)z_f \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$P' = S.P$$

Figure 2.2. 3D Rotation

Any coordinate position P on the object in this figure is transformed with the sequence shown as

$$\mathbf{P}' = \mathbf{T}^{-1} \cdot \mathbf{R}_z(\theta) \cdot \mathbf{T} \cdot \mathbf{P}$$

- *Z-axis Rotation*

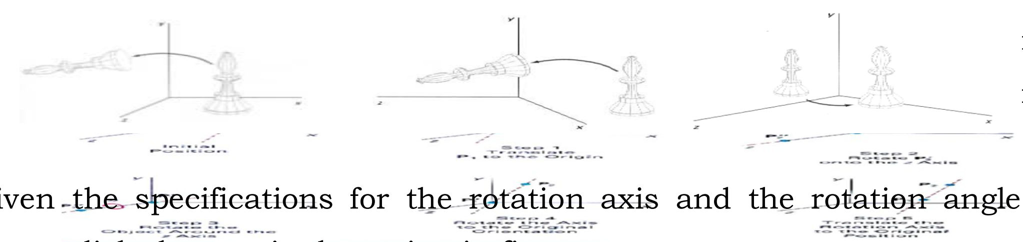
$$x' = xCos\theta - ySin\theta$$
$$y' = xSin\theta + yCos\theta$$
$$z' = z$$

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} Cos\theta & -Sin\theta & 0 & 0 \\ Sin\theta & Cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

- *X-axis Rotation*

$$y' = yCos\theta - zSin\theta$$
$$z' = xSos\theta + zCin\theta$$
$$x' = x$$

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & Cos\theta & -Sin\theta & 0 \\ 0 & Sin\theta & Cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$



Where the composite matrix for the transformation is

$$\mathbf{R}(\theta) = \mathbf{T}^{-1} \cdot \mathbf{R}_z(\theta) \cdot \mathbf{T}$$

Given the specifications for the rotation axis and the rotation angle, we can accomplish the required rotation in five steps

1 Translate the object so that the rotation axis pass= through the coordinate origin.

2.Rotate the object so that the axis of rotation coincides with one of the coordinate axes.

3. Perform the specified rotation about that coordinate axis.
4. Apply inverse rotations to bring the rotation axis back to original orientation.
5. Apply the inverse translation to bring the rotation axis back to its original position with one coordinate point and axes. We will assume that the axes is to be counter clockwise

The
$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & a & 0 \\ 0 & 1 & b & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

3) 3D SCALING

matrix expression tor the scaling transformation of a position P = (x, y, *z)* relative to the coordinate origin can be written as

$$x' = x \cdot s_x, \quad y' = y \cdot s_y, \quad z' = z \cdot s_z$$

Explicit expressions for the coordinate transformations for scaling relative to the origin are

1. Translate the fixed point to the origin.
2. Scale the object relative to the coordinate origin
3. Translate the fixed point back to its original position.

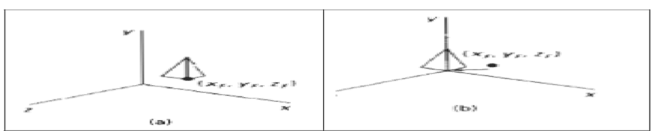


Figure 2.3. 3D Scaling

4) 3D Reflections

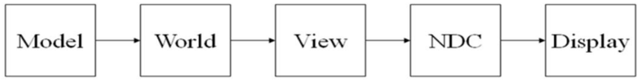
A three-dimensional reflection can be performed relative to a selected *reflection axis* or with respect to a selected *reception plane.*

Transformation matrices for inverting *x* and *y* values are defined similarly, as reflections relative to *yz* plane and *xz* plane, respectively.

3D Shear

The matrix expression for the shearing transformation of a position P = (*x*, *y*, *and z*), to produce *z*-axis shear, can be written as:

From model to image



Parameters a and b can be assigned any real values. The effect of this transformation is to alter x and y coordinate values by an amount that is proportional to the z value, while leaving the z coordinate unchanged. Shearing transformations for the x axis and y axis are defined similarly.

2.3 THREE-DIMENSIONAL VIEWING

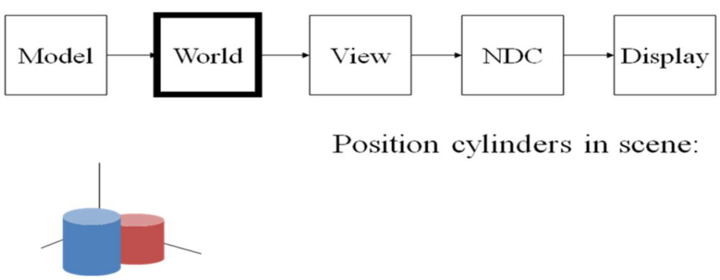
1. We can view an object from any spatial position, e.g. In front of an object,Behind the object,In the middle of a group of objects,Inside an object, etc.
2. 3D descriptions of objects must be projected onto the flat viewing surface of the output device.
3. The clipping boundaries enclose a volume of space.

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

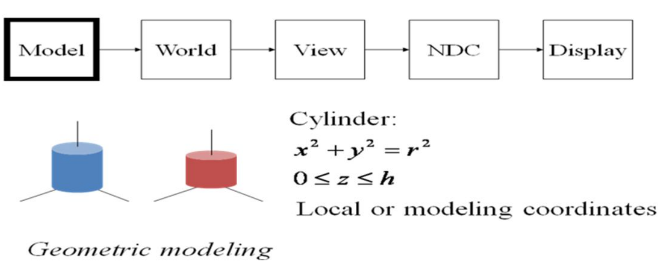
2.3.1 Viewing Pipeline

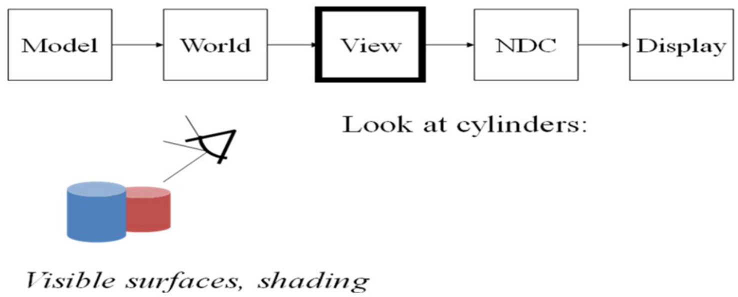
Step 1: Modeling Coordinates

Step 2: World Coordinates

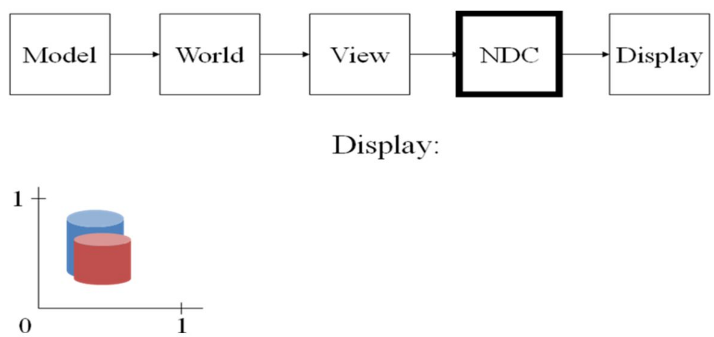


Step 3: Viewing Coordinates

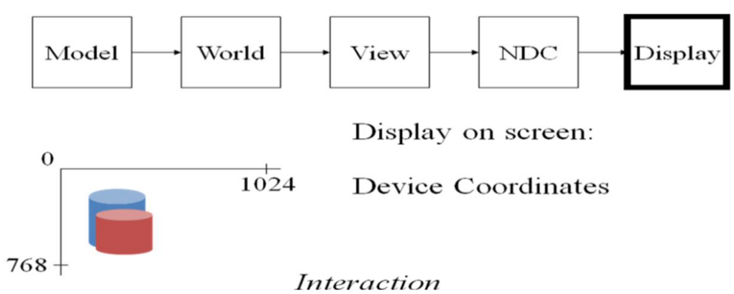




Step 4: Normalized Coordinates



Step 5: Device Coordinates



2.3.2 Projections

1. Parallel Projection transforms object positions to the view plane along parallel lines. A parallel projection preserves relative proportions of objects. Accurate views of the various sides of an object are obtained with a parallel projection. But not a realistic representation.

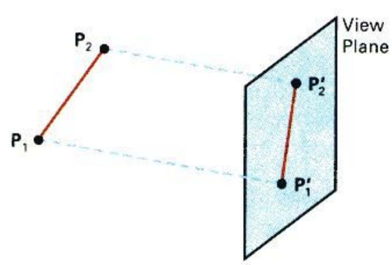


Figure 2.4 Parallel Projections

2.Perspective Projection transforms object positions to the view plane while converging to a center point of projection.

Perspective projection produces realistic views but does not preserve relative proportions. Projections of distant objects are smaller than the projections of objects of the same size that are closer to the projection plane.

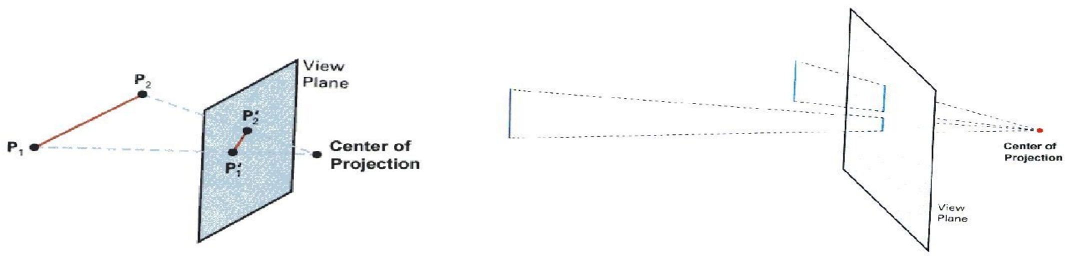
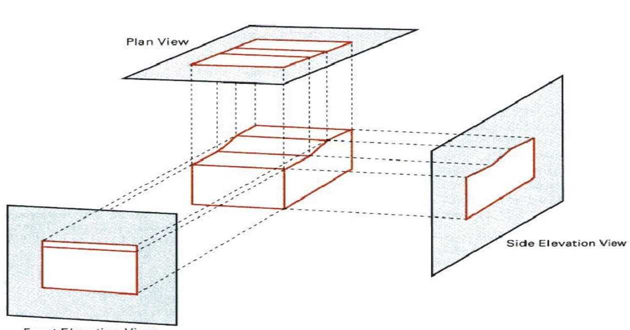


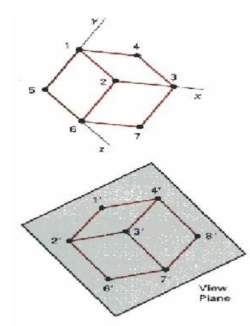
Figure 2.5 Perspective Projections

<u>Classification of Parallel Projection</u>

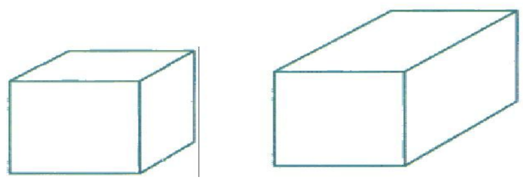
1. Orthographic Parallel Projection   2.Oblique Projection:

Orthographic parallel projections are done by projecting points along parallel lines that are perpendicular to the projection plane. Oblique projections are obtained by projecting along parallel lines that are NOT perpendicular to the projection plane. Some special Orthographic Parallel Projections involve Plan View (Top projection), Side Elevations, and Isometric Projection:



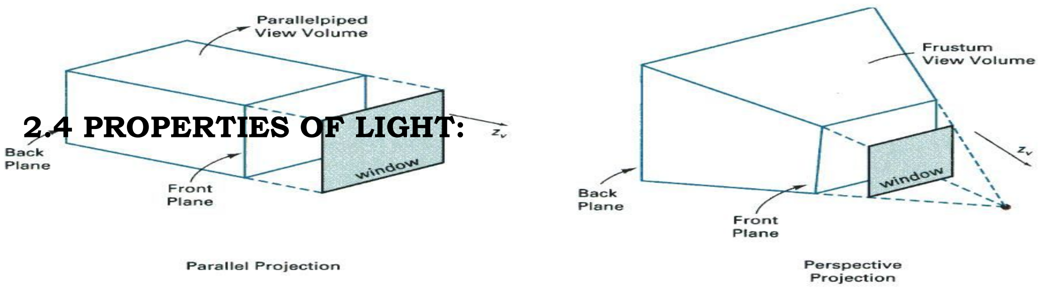


The following results can be obtained from oblique projections of a cube:



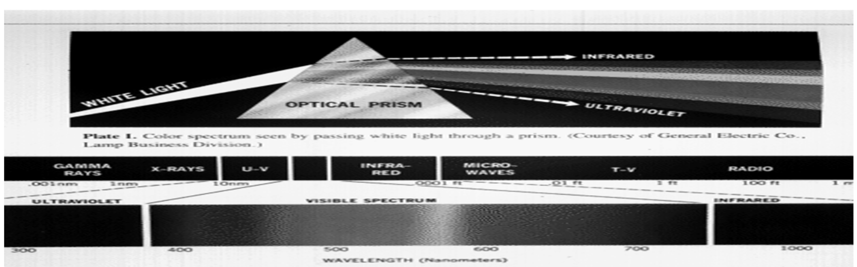
Perspective Projection

Perspective projection is done in 2 steps: Perspective transformation and Parallel projection. These steps are described in the following section.



2.4 PROPERTIES OF LIGHT:

The wavelength distribution of sunlight is relatively flat, indicating that sunlight contains an approximate equal intensity of each wavelength across the spectrum.



2.5 Color:-

Color is the product of the spectrum of light, as it is reflected or absorbed, as received by the human eye and processed by the human brain.

The Physics behind Colors:

The Light is an electromagnetic wave, so we can describe the various colors in terms of either frequency (f) or wavelength (λ) of the wave. Frequency and wavelength are inversely proportional to each other with proportionality constant as the speed of light

$C = \lambda * f$

BRIGHTNESS/ LUMINANCE:

It is the perceived intensity of light, i.e. the radiant energy emitted per unit time, per solid angle and per unit projected area of the source perceived by the eye.

PURITY/ SATURATION:

Describes how washed out or how "pure" the color of the light appears. It is the difference of a color against its own brightness.

2.6 *COLOR MODELS*

There are two broad categories of color models:

i) Hardware color model: models implemented at hardware level. Some of the models under this category are.

- RGB
- CMY
- CMYK
- YIQ

ii) Software color models: act as an interface between user and hardware color model. Models under this category are:

- HSV/ HSB
- HLS

Criteria for choosing the three primary colors

All the chosen three colors should be able to produce white color on mixing together.

No two colors on combining should produce the third primary color.

The three selected primary colors should be widely spread in visible spectrum.

Additive color model and Subtractive color model

the tools we use to describe color are different when the color is printed than from when it is projected. Projected color is additive. Printed color is subtractive.

Additive Color

In additive color model, no light or color is black. The additive color system starts with no light (black). Light sources add wavelength to make a color. The additive reproduction process usually uses red, green and blue light to produce the other colors. Combining one of these additive primary colors with another in equal amounts produces the additive secondary colors (cyan, magenta, and yellow). Combining all three primary lights (colors) in equal intensities produces white.

Subtractive Color

A subtractive color model explains the mixing of paints, dyes, inks, and natural colorants to create a range of colors, where each color is produced by the mixture, absorbing some wavelengths of light and reflecting others. The colors that an opaque object appears to have, is based on what parts of the electromagnetic spectrum are reflected by it, or by what parts of the spectrum are not absorbed.

2.7 CIE Chromaticity System

In the mid 1940's it was determined that color needed to be added to the existing monochrome video signal. The first step in this undertaking was to analyze and quantify the properties of the human perception. The Committee International d~Eclarge (CIE) was established to define an "average" human observer.

There are two parameters of color: saturation (how much color) and hue/tine (which shade of color)

As two properties are extremely subjective, the CIE tested thousands of subjects using a light comparison apparatus in order to define a "standard observer". The results are shown here, and are called "CIE color space". In this diagram the point labeled C is defined as standard pure white. The distance away from C represents saturation, and the position around the diagram represents hue (tint). Any point not actually on the solid outer line but within the area enclosed by the curve represents a mixture of colors. Pure spectrum colors are defined as being on the solid outer line.

2.7.1 The CIE XYZ Color Model

As mentioned in the preceding page, CIE considered the tristimulus values for red, green, and blue to be undesirable for creating a standardized color model. Instead, they used a mathematical formula to convert the RGB data to a system that uses only positive integers as values. The reformulated tristimulus values were indicated as *XYZ*. These values do not directly correspond to red, green, and blue, but are approximately so. The curve for the *Y* tristimulus value is equal to the curve that indicates the human eye's response to the total power of a light source. For this reason the value Y is called the luminance factor and the *XYZ* values have been normalized so that *Y* always has a value of 100.

Note: The tristumulus values *XYZ* are always indicated in upper case while the chromaticity coordinates, *xyz*, are always in lower case.

The chromaticity coordinates are used in conjunction with a chromaticity diagram, the most familiar one being CIE's 1931 *xyY* Chromaticity Diagram:

The third dimension is indicated by the tristimulus value *Y*. As previously mentioned, this value indicates the lightness or luminance of the color. The scale for Y extends from the white spot in a line perpendicular to the plane formed by x and y using a scale that runs from 0 to 100. The fullest range of color exists at 0 where the white point is equal to CIE Illuminant C. As the *Y*value increases and the color becomes lighter, the range of color, or gamut, decreases so that the color space at 100 is just a sliver of the original area:

2.7.2 The RGB Color Model

- The RGB color model was first described by Thomas Young and Herman Helmholtz in their Theory of Trichromatic Color vision (first half of the 19th century) and by James Maxwell's .
- Based on the tri stimulus theory of vision, our eyes perceive color through the stimulation of three visual pigments in the cones of the retina. These visual pigments have peak sensitivity at wavelengths of about 630 run (red), 530 nm (green), and 450 nm (blue).

- By comparing intensities in a light source, we perceive the color of the light. This theory of vision is the basis for displaying color output on a video monitor using the three color primaries, red, green, and blue, referred to as the RGB color model. The relationship between the colors can be seen in this illustration:
- In addition to providing a good description of human color perception, the RGB model is the basis for displaying colors in television and computer screens.
- The RGB model is also used for recording colors in digital cameras, including still image and video cameras.
- The RGB model is an additive system.
- The secondary colors of RGB, cyan, magenta, and yellow, are formed by the mixture of two of the primaries and the exclusion of the third. Red and green combine to make yellow, green and blue make cyan, blue and red make magenta.
- The combination of red, green, and blue in full intensity makes white. White light is created when all colors of the EM spectrum converge in full intensity.
- To represent colors for a television or computer screen, each pixel of the screen is recorded as the triple (r, g, b) of numbers. One popular system uses numbers that range from 0 to 255 for each color.
- Thus, a color C, is expressed in RGB components as

$$C_\lambda = R\mathbf{R} + G\mathbf{G} + B\mathbf{B}$$

- The magenta vertex is obtained by adding red and blue to produce the triple (1,0,l). And white at (1, l, 1) is the sum of the red, green, and blue vertices. Shades of gray are represented along the main diagonal of the cube from the origin (black) to the white vertex.

- Each point along this diagonal has an equal contribution from each primary color, so that a gray shade halfway between black and white is represented as (0.5, 0.5, and 0.5).

  We can represent this model with the unit cube defined on R, G, and B axes,

  The origin represents black, and the vertex with coordinates (1, 1, l) is white. Vertices of the cube on the axes represent the primary colors and the remaining vertices represent the complementary color for each of the primary colors.

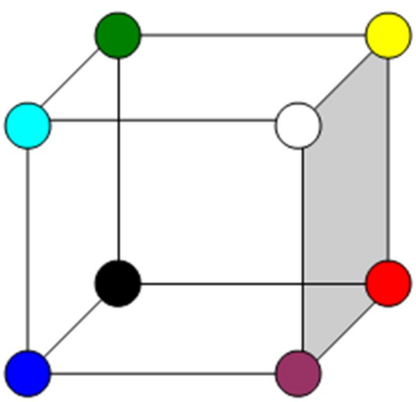


Figure 2.6 RGB Color Cube

Characteristics:

We cannot use RGB in printing devices because we use white paper and to produce black color we need to have no red, no green and no blue color, which

again produces white as background is white. Whereas to produce white we need full amount of all three primaries which is unnecessary wastage of color. Thus it is not economical and produces error.
The red-green-blue model is formed by a color cube {(R, G, B): 0 <= R, G, B <=1}. Conversion from (R,G,B) to (X,Y,Z) is given via the chromaticity's (Xr,Yr,Zr), (Xg,Yg,Zg) and (Xb,Yb,Zb) of the CRTs phosphors by matrix multiplication via:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} Xr & X_g & X_b \\ Yr & Y_g & Y_b \\ Zr & Z_g & Z_b \end{pmatrix} \cdot \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

Let Cr= Xr + Yr + Zr

Then Xr=xr.Cr, Yr = Yr.Cr, and Zr=zr.Cr = (1 – xr – yr).Cr

1.7.3 The CMY Color Model

This stands for cyan-magenta-yellow and is used for hardcopy devices. In contrast to color on the monitor, the color in printing acts subtractive and not additive. A printed color that looks red absorbs the other two components G and B and reflects R. Thus its (internal) color is G+B=CYAN. Similarly R+B=MAGENTA and R+G=YELLOW. Thus the C-M-Y coordinates are just the complements of the R-G-B coordinates:

Characteristics:

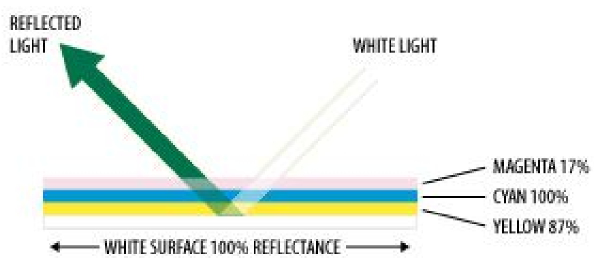
- Designed basically to be used for printing purpose.
- It is a Subtractive Model – ink and paint absorbs certain colors from image instead of adding to it.

  For E.g. Cyan has a property of absorbing red color form white. Magenta has a property of absorbing green color form white and Yellow has a property of absorbing blue color form white. Whereas Black absorbs all three colors red, green and blue from white.

- The CMY model measures how much Cyan, Magenta, and Yellow component each image has.
- The conversion from RGB to CMY takes place at printout by

(C, M, Y) = (1, 1, and 1) − (R, G, and B)

The CMY model used in printing lays down overlapping layers of varying percentages of transparent cyan, magenta, and yellow inks. Light is transmitted through the inks and reflects off the surface below them (called the substrate). The percentages of CMY ink (which are applied as screens of halftone dots), subtract inverse percentages of RGB from the reflected light so that we see a particular color:



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

If we want to print a red looking color (i.e. with RGB coordinates (1,0,0)) we have to use CMY values of (0, 1,1). Note that M absorbs G, similarly Y absorbs B and hence M + Y absorbs all but R.

Black ((R,G,B) = (0,0,0)) corresponds to(C,M,Y)=(1,1,1) which should in principle absorb R, G and B. But in practice this will appear as some dark gray. So in order to be able to produce better contrast printers often use black as fourth color. This is the CMYK-model. Its coordinates are obtained from that of the CMY-model by K=max (C, M, Y), C=C-K, M=M-K and Y=Y-K.

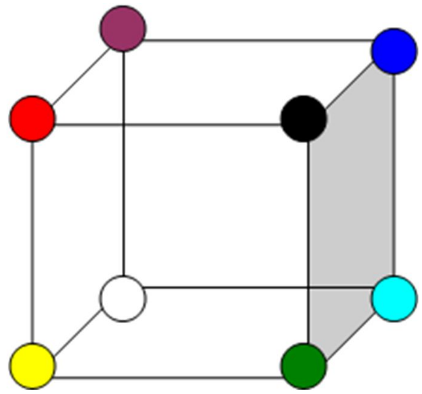


Figure 2.7. CMY Color Cube

we can express the conversion from an RGB representation to a CMY representation with the matrix transformation

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

Where the white is represented in the RGB system is the unit column vector. Similarly, we convert from a CMY color representation to an RGB representation with the matrix transformation where black is represented In the CMY system as the unit column vector.

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

2.7.4. YIQ COLOR MODEL

The National Television System Committee (NTSC) color model for forming the composite video signal is the YIQ model, which is based on concepts in the CIE XYZ model. In the YI Q color model, parameter Y is the same as in the XYZ model. Luminance (brightness) information is contained in the Y parameter, while chromaticity information (hue and purity) is incorporated into the 1 and Q parameters.

from RGB values to YIQ values is accomplished with the transformation

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.528 & 0.311 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

This transformation is based on the NTSC standard RGB phosphor, whose chromaticity coordinates were given in the preceding section. The larger proportions of red and green assigned to parameter Y indicate the relative importance of these hues in determining brightness, compared to blue.

An NTSC video signal can be converted to an RGB signal sing an NTSC decoder, which separates the video signal into the YIQ components, then converts to RGB values. We convert from YIQ space to RGB space with the inverse matrix transformation from Eq. 15-6:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.000 & 0.956 & 0.620 \\ 1.000 & -0.272 & -0.647 \\ 1.000 & -1.108 & 1.705 \end{bmatrix} \cdot \begin{bmatrix} Y \\ I \\ Q \end{bmatrix}$$

2.7.5. The HSV Color Model

Alvy Ray Smith in 1978 was the first to describe colors using hue, saturation, and value (HSV model).Instead of a set of color primaries, the HSV model uses color descriptions that have a more intuitive appeal to a user. Color parameters in this model are hue (H), saturation(S) and value

Hue is a saturated color on the outer rim of the HSV Color Wheel

Saturation is the amount of white added to the color. 0% means that the color (at V=100%) is totally white; 100% means totally saturated with no white added (a fully saturated color is a pure hue on the outer rim of the HSV color wheel).

Value is the brightness of the color. 0% means totally dark or black; 100% means full brightness, with the color is fully determined by the hue and saturation.

Hue is represented as an angle about the vertical axis ranging from 0 degree at red through 360 degrees. Vertices of the hexagon are separated by 60^0 intervals. Yellow is at 60^0, green at 120^0 and cyan opposite e red at H = 180^0 complementary colors are 1800 apart

Value V varies from 0 at the apex of the hex cone to I at the top. The apex represents black. At the top of the hex cone, colors have their maximum intensity. When V = 1 and S = 1, we have the "pure" hues. White is the point at V = 1 and S =0

Saturation S varies from 0 to 1. It is represented in this model as the ratio of the purity of a selected hue to its maximum purity at S = 1. A selected hue is said to be one-quarter pure at the value S = 0.25. At S = 0, we have the gray scale.

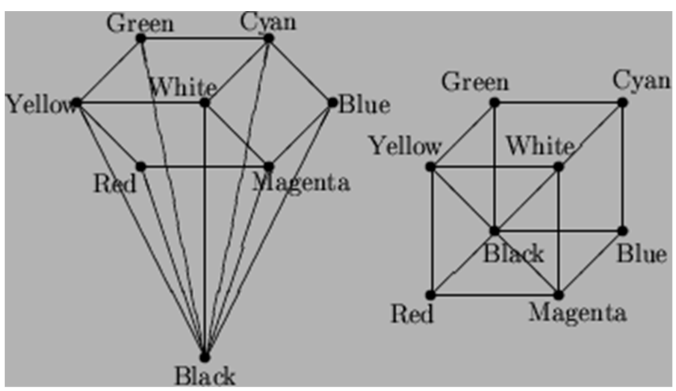


Figure 2.8. The HSV-model versus the RGB-model

The conversion from RGB to HSV is given by affine coordinate changes on each of the 3 four-sided sub-pyramids corresponding each to 1/3 of the color cube

2.7.6 The HLS Color Model

The HSL model describes colors in terms of hue, saturation, and lightness (also called luminance). (Note: the definition of saturation in HSL is substantially different from HSV, and lightness is not intensity.) The model has two prominent properties:

- The transition from black to a hue to white is symmetric and is controlled solely by increasing lightness

- Decreasing saturation transitions to a shade of gray dependent on the lightness, thus keeping the overall intensity relatively constant
- The angular relationship between tones around the color circle is easily identified
- Shades, tints, and tones can be generated easily without affecting the hue

Saturation, or the lack of it, produces tones of the reference hue that converge on the zero-saturation shade of gray, which is determined by the lightness. The following examples use the hues red, orange, and yellow at midpoint lightness with decreasing saturation. The resulting RGB value and the total intensity are shown.

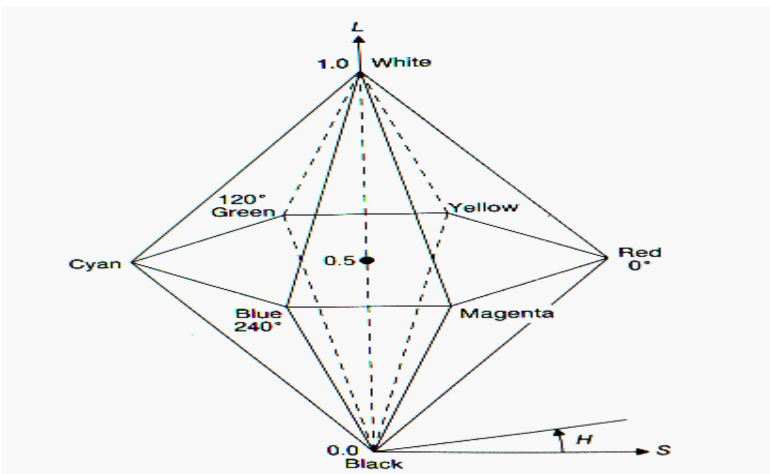


Figure2.9. HLS Color Model

2.8 ANIMATION

2.10DESIGN OF ANIMATION SEQUENCES

In general, an animation sequence is designed with the following steps:

1. Storyboard layout
2. Object definitions
3. Key-frame specifications
4. Generation of in-between frames

This standard approach for animated cartoons is applied to other animation applications as well, although there are many special applications that do not follow this sequence.

The Storyboard layout is an outline of the action. It defines the motion sequence as a set of basic events that are to take place. Depending on the type of animation to be produced, the storyboard could consist of a set of rough sketches or it could be a list of the basic ideas for the motion.

An object definition is given for each participant in the action. Objects can be defined in terms of basic shapes, such as polygons or splines. In addition, the associated movements for each object are speeded along with the shape.

A Key frame is a detailed drawing of the scene at a certain time in the animation sequence. Within each key frame, each object is positioned according to the time for that frame. Some key frames are chosen at extreme positions in the action; others are spaced so that the time interval between key frames is not too great. More key frames are specified for intricate motions than for simple, slowly varying motions.

In-between frames are the intermediate frames between the key frames. The number of in-betweens needed is determined by the media to be used to display the animation. Film requires 24 frames per second, and graphics terminals are refreshed at the rate of 30 to 60 frames per second. Typically, time intervals for the motion are set up so that there are from three to five in-betweens for each pair of key frames.

GENERAL COMPUTER-ANIMATION FUNCTIONS

Some steps in the development of an animation sequence are well-suited to computer solution. These include object manipulations and rendering, camera motions, and the generation of in-betweens. One function available in animation packages is provided to store and manage the object database. Object shapes and associated parameters are stored and updated in the database.

Raster Animations

On raster systems, we can generate real-time animation in limited applications using raster operations. Sequences of raster operations can be executed to produce real-time animation of either two-dimensional or three-dimensional objects, as long as we restrict the animation to motions in the projection plane.

Computer-Animation Languages

Design and control of animation sequences are handled with a set of animation routines. A general-purpose language, such as C, Lisp, Pascal, or FORTRAN, is often used to program the animation functions, but several specialized animation languages have been developed. Animation functions include a graphics editor, a key-frame generator, an in-between generator and standard graphics routines.

Key Frame Systems

We generate each set of in-betweens from the specification of two (or more) key frames. Motion paths can be given with a kinematic an s a set of spline curves, or the motions can be physically based by specifying the actions on the objects to be animated.

Morphing

Transformation of object shapes from one form to another is called morphing, which is a shortened form of metamorphosis. Morphing methods can be applied to any motion or transition involving a change in shape.

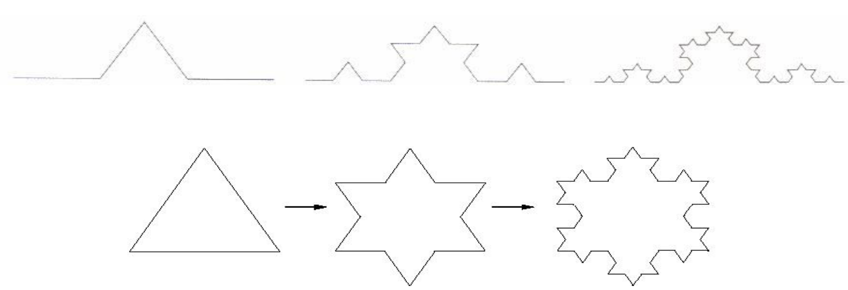
Motion Specifications There are several ways in which the motions of objects can be specified in an animation system. Direct Motion Specification is the most straightforward method for defining a motion sequence is direct specification of the motion parameters. Here, we explicitly give the rotation angles and translation vectors. Then the geometric transformation matrices are applied to transform coordinate positions

2.9 Fractals

Fractal objects refer to those objects which are self-similar at all resolutions.

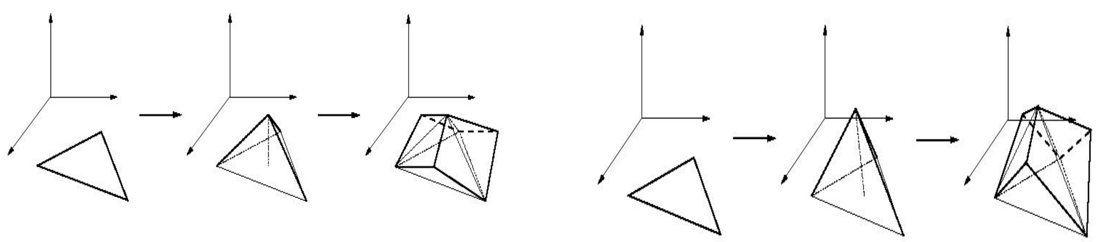
Most of the natural objects such as trees, mountains and coastlines are considered as fractal objects because no matter how far or how close one looks at them, they always appear to be somewhat similar.

For example, a Fractal Snowflake:



Euclidean dimensions and Fractal dimensions:

- A line segment is 1D. If we divide a line into N equal parts, the parts each look like the original line scaled down by a factor of $N = N^{1/1}$.
- A square is 2D. If we divide it into N parts, each part looks like the original scaled down by a factor of $N^{1/2}$.
- For the fractal snowflake, when it is divided into 4 pieces, each resulting piece looks like the original scaled down by a factor of 3, so it has the dimension ***d*** such that $4^{1/d}=3$. That is, d = 1.26.



UNIT- II THREE-DIMENSIONAL CONCEPTS

Three-Dimensional Object Representations – Three-Dimensional Geometric and Modeling Transformations – Three-Dimensional Viewing – Color models – Animation.

2.1 THREE DIMENSIONAL OBJECT REPRESENTATIONS

To obtain display of a three-dimensional scene that has been modeled in world coordinates. we must first set up a coordinate reference for the "camera". This coordinate reference defines the position and orientation for the plane of the camera film which is the plane we want to us to display a view of the objects in the scene.

Boundary representations (B-reps) describe a three-dimensional object as a set of surfaces that separate the object interior from the environment. Typical examples of boundary representations are polygon facets and spline patches.

Space-partitioning representations are used to describe interior properties, by partitioning the spatial region containing an object into a set of small, non overlapping, contiguous solids (usually **cubes).**

3D Object representation Methods:

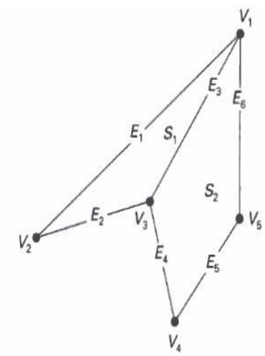
1. Polygon and Quadric surfaces: For simple Euclidean objects
2. Spline surfaces and construction: For curved surfaces

3. Procedural methods: Eg. Fractals, Particle systems
4. Physically based modeling methods
5. Octree Encoding

Objects may also associate with other properties such as mass, volume, so as to determine their response to stress and temperature etc.

2.1.1 Polygon Surfaces

This method simplifies and speeds up the surface rendering and display of objects. For other 3D objection representations, they are often converted into polygon surfaces before rendering.



VERTEX TABLE
V_1: x_1, y_1, z_1
V_2: x_2, y_2, z_2
V_3: x_3, y_3, z_3
V_4: x_4, y_4, z_4
V_5: x_5, y_5, z_5

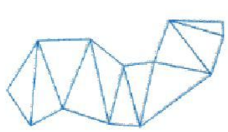
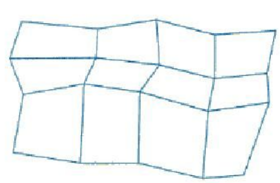
EDGE TABLE
E_1: V_1, V_2
E_2: V_2, V_3
E_3: V_3, V_1
E_4: V_3, V_4
E_5: V_4, V_5
E_6: V_5, V_1

POLYGON-SURFACE TABLE
S_1: E_1, E_2, E_3
S_2: E_3, E_4, E_5, E_6

Polygon Mesh

Using a set of connected polygonally bounded planar surfaces to represent an object, which may have curved surfaces or curved edges.

- The wireframe display of such object can be displayed quickly to give general indication of the surface structure.
- Realistic renderings can be produced by interpolating shading patterns across the polygon surfaces to eliminate or reduce the presence of polygon edge boundaries.
- Common types of polygon meshes are triangle strip and quadrilateral mesh.

Polygon Tables

This is the specification of polygon surfaces using vertex coordinates and other attributes:

3. Geometric data table: vertices, edges, and polygon surfaces.
4. Attribute table: e.g. Degree of transparency and surface reflectivity

Some consistency checks of the geometric data table:

§ Every vertex is listed as an endpoint or at least 2 edges

§ Every edge is part of at least one polygon

§ Every polygon is closed

<u>Plane equation and visible points</u>

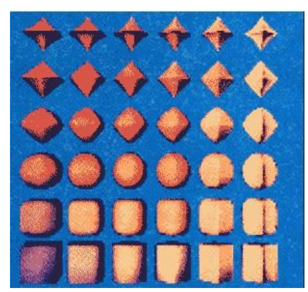
Consider a cube; each of the 6 planes has 2 sides: inside face and outside face. For each plane (in a right-handed coordinate system), if we look at its surface and take 3 points in counter-clockwise direction: $(x_1, y1)$, $(x_2, y2)$, and $(x_3, y3)$, we can compute 4 values: A, B, C, D as

$$A = \begin{vmatrix} 1 & y_1 & z_1 \\ 1 & y_2 & z_2 \\ 1 & y_3 & z_3 \end{vmatrix} \qquad B = \begin{vmatrix} x_1 & 1 & z_1 \\ x_2 & 1 & z_2 \\ x_3 & 1 & z_3 \end{vmatrix} \qquad C = \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} \qquad D = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$$

Then, the plane equation at the form: Ax+By+Cz+D=0 has the property that:

If we substitute any arbitrary point (x, y) into this equation, then,Ax + By + Cz + D < 0 implies that the point (x, y) is inside the surface, and Ax + By + Cz + D < 1 implies that the point (x, y) is outside the surface.

2.1.2 Curved Surfaces



1. Regular curved surfaces can be generated as

-Quadric Surfaces, e.g. Sphere, Ellipsoid, or

-Super quadrics, e.g. Super ellipsoids

These surfaces can be represented by some simple parametric equations, e.g., for ellipsoid:

$x = r_x \cos^{s1}\phi \cos^{s2}\theta, \ -\pi/2 <= \phi <= \pi/2$

$y = r_y \cos^{s1}\phi \sin^{s2}\theta, \ -\pi <= \theta <= \pi$

$z = r_z \sin^{s1}\phi$

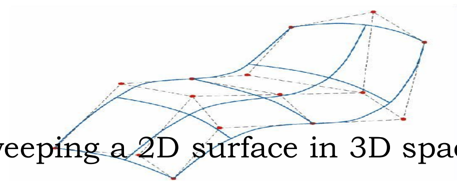
Where s1, r_x, r_y, and r_x are constants. By varying the values of ϕ and θ, points on the surface can be computed.

2. Irregular surfaces can also be generated using some special formulating approach, to form a kind of blobby objects. The shapes showing a certain

degree of fluidity.

3. Spline Representations

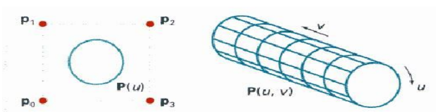
Spline means a flexible strip used to produce a smooth curve through a designated set of points. Several small weights are distributed along the length of the strip to hold it in position on the drafting table as the curve is



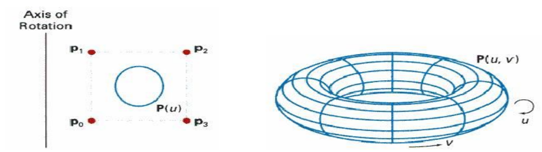
2.1.3 Sweep Representations

Sweep representations mean sweeping a 2D surface in 3D space to create an object. However, the objects created by this method are usually converted into polygon meshes and/or parametric surfaces before storing.

A Translational Sweep:



A Rotational Sweep:



Other variations:

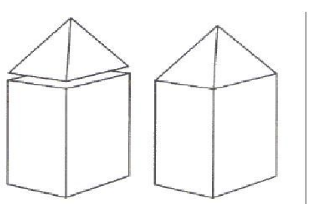
- We can specify a special path for the sweep as some curve function.
- We can vary the shape or size of the cross section along the sweep path.
- We can also vary the orientation of the cross section relative to the sweep path.

2.1.4 Constructive Solid-Geometry Methods

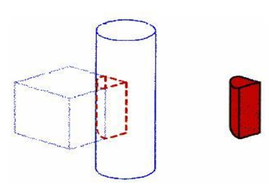
The Constructive Solid-Geometry Method (CSG) combines the volumes occupied by overlapping 3D objects using set operations:

- Union
- Intersection
- Difference

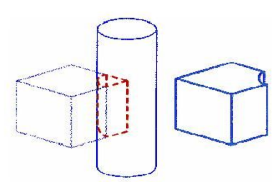
Object created by a union operation:



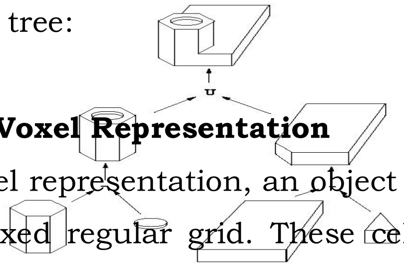
Object created by an intersection operation:

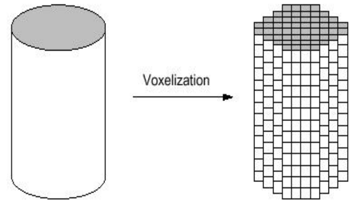


Object created by a difference operation:



A CSG object can be represented with a binary tree:





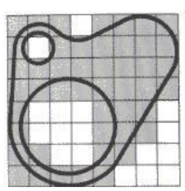
2.1.5 Voxel Representation

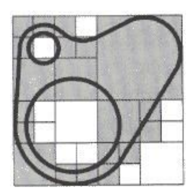
In voxel representation, an object is decomposed into identical cells arranged in a fixed regular grid. These cells are called voxels (volume elements), in analogy to pixels.

2.1.6 Octrees

Octrees are hierarchical tree structures that describe each region of 3D space as nodes. When compared with the basic voxel representation, octrees reduce storage requirements for 3D objects. It also provides a convenient representation for storing information about object interiors.Octree encoding procedure is an extension of the quad tree encoding of 2D images:

Bitmap representation of a 2D object: Quad tree encoding of a 2D object:





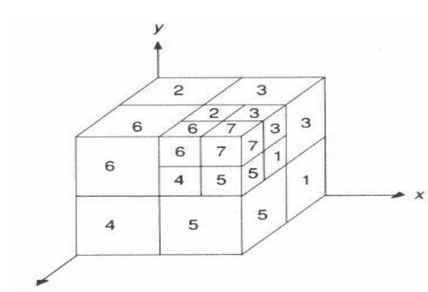
For octree representation of 3D data, the numbering method is as follows:

2.2 THREE DIMENSIONAL GEOMETRIC TRANSFORMATION:

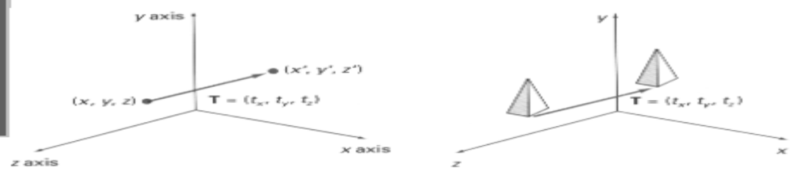
5) 3D TRANSLATION

Figure 2.1.Translating a point with Translation vector (tx, ty, tz)

In a three-dimensional homogeneous coordinate representation, a point is translated from position P = ***(x, y, z)*** to position P' = ***(x', y', z'***) as shown in figure



$$x' = x + t_x$$
$$y' = y + t_y$$
$$z' = z + t_z$$



$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Or

$$P' = T \cdot P$$

Parameters t_x, t_y, and t_z specifying translation distances for the coordinate directions x, y, and z, are assigned any real values.

An object is translated in three dimensions by transforming each of the defining points of the object.

6) 3D ROTATION

To generate a rotation transformation for an object, we must designate an axis of rotation (about which the object is to be rotated) and the amount of angular rotation.

Coordinate-Axes Rotations

The two-dimensional z-axis rotation equations are easily extended to three dimensions:

- *Z-axis Rotation*

$$x' = xCos\theta - ySin\theta$$
$$y' = xSin\theta + yCos\theta$$
$$z' = z$$

$$\Leftrightarrow \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} Cos\theta & -Sin\theta & 0 & 0 \\ Sin\theta & Cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

- *X-axis Rotation*

$$y' = yCos\theta - zSin\theta$$
$$z' = xSos\theta + zCin\theta$$
$$x' = x$$

$$\Leftrightarrow \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & Cos\theta & -Sin\theta & 0 \\ 0 & Sin\theta & Cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

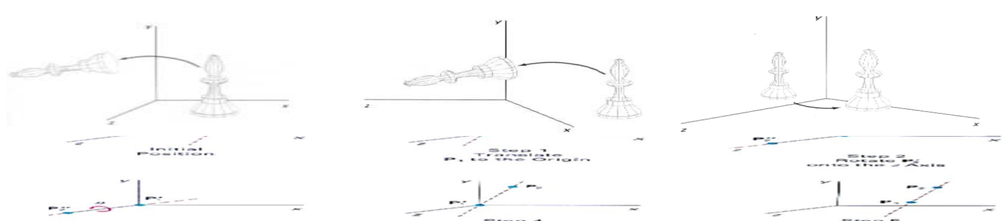


Figure 2.2. 3D Rotation

Any coordinate position P on the object in this figure is transformed with the sequence shown as

$$P' = T^{-1} \cdot R_x(\theta) \cdot T \cdot P$$

Where the composite matrix for the transformation is

$$\mathbf{R}(\theta) = \mathbf{T}^{-1} \cdot \mathbf{R}_x(\theta) \cdot \mathbf{T}$$

Given the specifications for the rotation axis and the rotation angle, we can accomplish the required rotation in five steps

1 Translate the object so that the rotation axis pass= through the coordinate origin.

6. Rotate the object so that the axis of rotation coincides with one of the coordinate axes.

7. Perform the specified rotation about that coordinate axis.

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 & (1-s_x)x_f \\ 0 & s_y & 0 & (1-s_y)y_f \\ 0 & 0 & s_z & (1-s_z)z_f \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$P' = S.P$$

8. Apply inverse rotations to bring the rotation axis back to original orientation.

9. Apply the inverse translation to bring the rotation axis back to its original position with one coordinate point and axes. We will assume that the axes is to be counter clockwise

7) 3D SCALING

The matrix expression tor the scaling transformation of a position P = (x, y, *z)* relative to the coordinate origin can be written as

$$x' = x \cdot s_x, \qquad y' = y \cdot s_y, \qquad z' = z \cdot s_z$$

Explicit expressions for the coordinate transformations for scaling relative to the origin are

1. Translate the fixed point to the origin.
2. Scale the object relative to the coordinate origin
3. Translate the fixed point back to its original position.

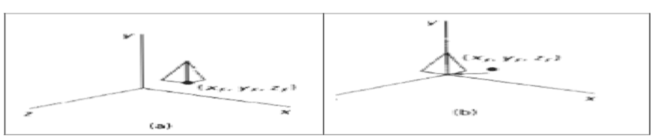
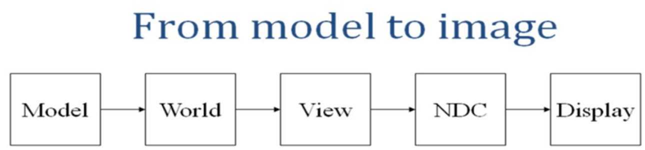


Figure 2.3. 3D Scaling

8) 3D Reflections

A three-dimensional reflection can be performed relative to a selected *reflection axis* or with respect to a selected *reception plane.*

Transformation matrices for inverting *x* and *y* values are defined similarly, as reflections relative to *yz* plane and *xz* plane, respectively.



3D Shear

The matrix expression for the shearing transformation of a position P = (x, y, z), to produce z-axis shear, can be written as:

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & a & 0 \\ 0 & 1 & b & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

and

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

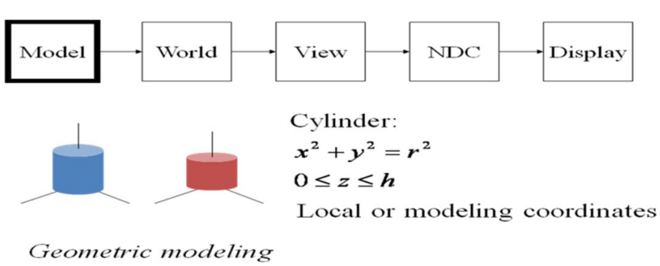
Parameters a and b can be assigned any real values. The effect of this transformation is to alter x and y coordinate values by an amount that is proportional to the z value, while leaving the z coordinate unchanged. Shearing transformations for the x axis and y axis are defined similarly.

2.3 THREE-DIMENSIONAL VIEWING

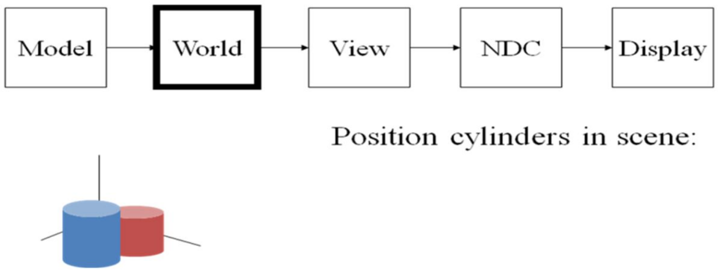
1. We can view an object from any spatial position, e.g. In front of an object,Behind the object,In the middle of a group of objects,Inside an object, etc.
3. 3D descriptions of objects must be projected onto the flat viewing surface of the output device.
3. The clipping boundaries enclose a volume of space.

2.3.1 Viewing Pipeline

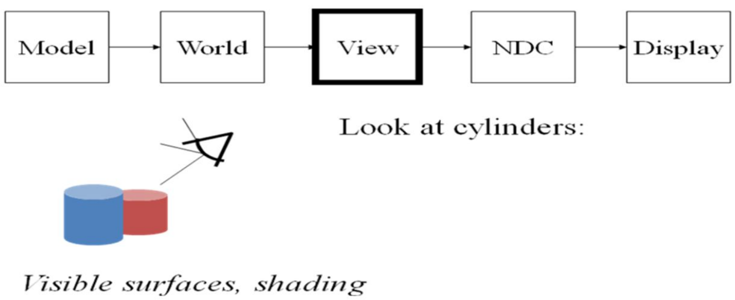
Step 1: Modeling Coordinates



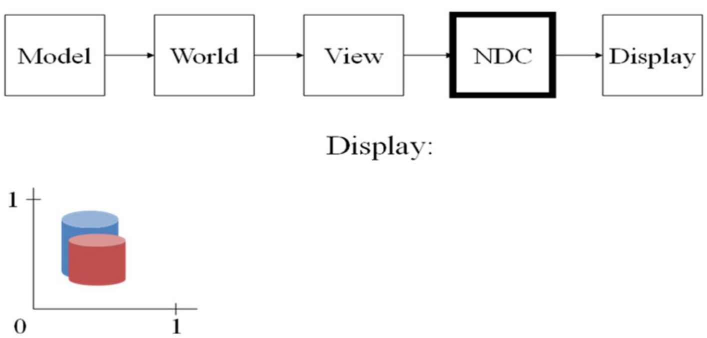
Step 2: World Coordinates



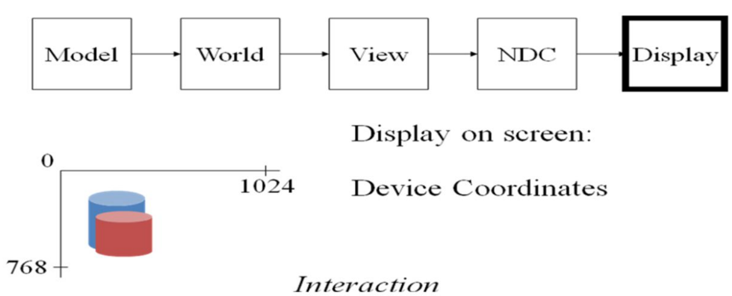
Step 3: Viewing Coordinates



Step 4: Normalized Coordinates



Step 5: Device Coordinates



2.3.2 Projections

1. Parallel Projection transforms object positions to the view plane along parallel

lines. A parallel projection preserves relative proportions of objects. Accurate views of the various sides of an object are obtained with a parallel projection. But not a realistic representation.

Figure 2.4 Parallel Projections

2.Perspective Projection transforms object positions to the view plane while converging to a center point of projection.

Perspective projection produces realistic views but does not preserve relative proportions. Projections of distant objects are smaller than the projections of objects of the same size that are closer to the projection plane.

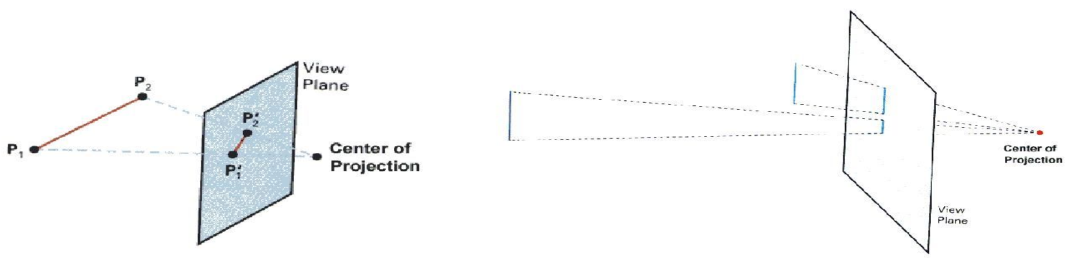
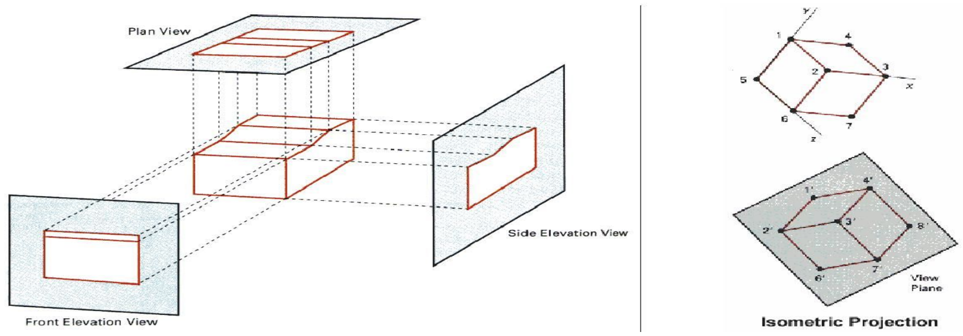


Figure 2.5 Perspective Projections

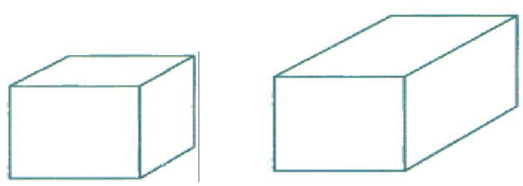
<u>Classification of Parallel Projection</u>

1. Orthographic Parallel Projection   2.Oblique Projection:

Orthographic parallel projections are done by projecting points along parallel lines that are perpendicular to the projection plane. Oblique projections are obtained by projecting along parallel lines that are NOT perpendicular to the projection plane. Some special Orthographic Parallel Projections involve Plan View (Top projection), Side Elevations, and Isometric Projection:

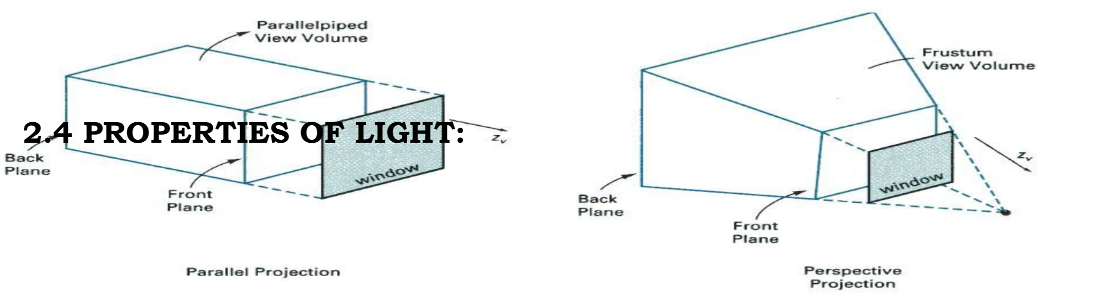


The following results can be obtained from oblique projections of a cube:



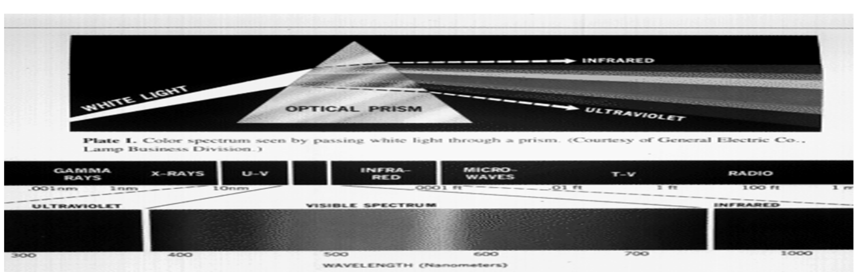
<u>Perspective Projection</u>

Perspective projection is done in 2 steps: Perspective transformation and Parallel projection. These steps are described in the following section.



2.4 PROPERTIES OF LIGHT:

The wavelength distribution of sunlight is relatively flat, indicating that sunlight contains an approximate equal intensity of each wavelength across the spectrum.



2.5 Color:-

Color is the product of the spectrum of light, as it is reflected or absorbed, as received by the human eye and processed by the human brain.

The Physics behind Colors:

The Light is an electromagnetic wave, so we can describe the various colors in terms of either frequency (f) or wavelength (λ) of the wave. Frequency and wavelength are inversely proportional to each other with proportionality constant as the speed of light

C= λ*f

BRIGHTNESS/ LUMINANCE:

It is the perceived intensity of light, i.e. the radiant energy emitted per unit time, per solid angle and per unit projected area of the source perceived by the eye.

PURITY/ SATURATION:

Describes how washed out or how "pure" the color of the light appears. It is the difference of a color against its own brightness.

2.6 *COLOR MODELS*

There are two broad categories of color models:

i) Hardware color model: models implemented at hardware level. Some of the models under this category are.

- RGB
- CMY
- CMYK
- YIQ

ii) Software color models: act as an interface between user and hardware color model. Models under this category are:

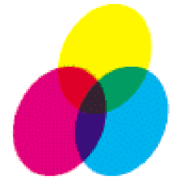
- HSV/ HSB
- HLS

Criteria for choosing the three primary colors

All the chosen three colors should be able to produce white color on mixing together.

No two colors on combining should produce the third primary color.

The three selected primary colors should be widely spread in visible spectrum.



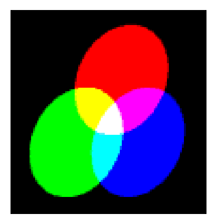
Additive color model and Subtractive color model the tools we use to describe color are different when the color is printed than from when it is projected. Projected color is additive. Printed color is subtractive.

Additive Color

In additive color model, no light or color is black. The additive color system starts with no light (black). Light sources add wavelength to make a color. The additive reproduction process usually uses red, green and blue light to produce the other colors. Combining one of these additive primary colors with another in equal amounts produces the additive secondary colors (cyan, magenta, and yellow). Combining all three primary lights (colors) in equal intensities produces white.

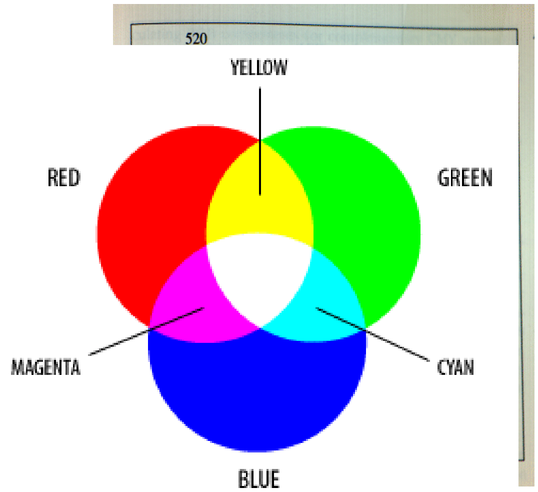
Subtractive Color

A subtractive color model explains the mixing of paints, dyes, inks, and natural colorants to create a range of colors, where each color is produced by the mixture, absorbing some wavelengths of light and reflecting others. The colors that an opaque object appears to have, is based on what parts of the electromagnetic spectrum are reflected by it, or by what parts of the spectrum are not absorbed.



2.7 CIE Chromaticity System

In the mid 1940's it was determined that color needed to be added to the existing monochrome video signal. The first step in this undertaking was to analyze and quantify the properties of the human perception. The Committee International d~Eclarge (CIE) was established to define an "average" human observer.

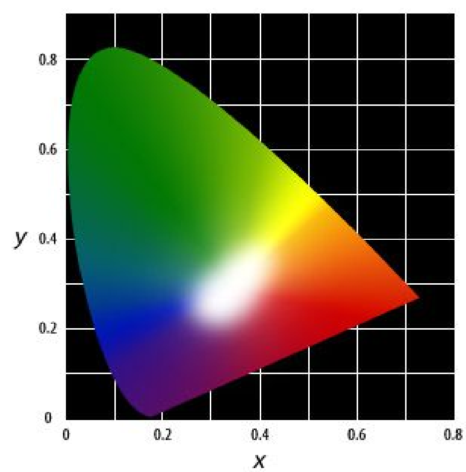


There are two parameters of color: saturation (how much color) and hue/tine (which shade of color)

As two properties are extremely subjective, the CIE tested thousands of subjects using a light comparison apparatus in order to define a "standard observer". The results are shown here, and are called "CIE color space". In this diagram the point labeled C is defined as standard pure white. The distance away from C represents saturation, and the position around the diagram represents hue (tint). Any point not actually on the solid outer line but within the area enclosed by the curve represents a mixture of colors. Pure spectrum colors are defined as being on the solid outer line.

2.7.1 The CIE XYZ Color Model

As mentioned in the preceding page, CIE considered the tristimulus values for red, green, and blue to be undesirable for creating a standardized color model. Instead, they used a mathematical formula to convert the RGB data to a system that uses only positive integers as values. The reformulated tristimulus values were indicated as *XYZ*. These values do not directly correspond to red, green, and blue, but are approximately so. The curve for the *Y* tristimulus value is equal to the curve that indicates the human eye's response to the total power of a light source. For this reason the value Y is called the luminance factor and the *XYZ* values have been normalized so that *Y* always has a value of 100.



Note: The tristumulus values *XYZ* are always indicated in upper case while the chromaticity coordinates, *xyz*, are always in lower case.

The chromaticity coordinates are used in conjunction with a chromaticity diagram, the most familiar one being CIE's 1931 *xyY* Chromaticity Diagram:

The third dimension is indicated by the tristimulus value *Y*. As previously mentioned, this value indicates the lightness or luminance of the color. The scale for Y extends from the white spot in a line perpendicular to the plane formed by x and y using a scale that runs from 0 to 100. The fullest range of color exists at 0 where the white point is equal to CIE Illuminant C. As the *Y*value increases and the color becomes lighter, the range of color, or gamut, decreases so that the color space at 100 is just a sliver of the original area:

2.7.3 The RGB Color Model

- The RGB color model was first described by Thomas Young and Herman Helmholtz in their Theory of Trichromatic Color vision (first half of the 19th century) and by James Maxwell's color triangle.
- Based on the tri stimulus theory of vision, our eyes perceive color through the stimulation of three visual pigments in the cones of the retina. These visual pigments have peak sensitivity at wavelengths of about 630 run (red), 530 nm (green), and 450 nm (blue).

- By comparing intensities in a light source, we perceive the color of the light. This theory of vision is the basis for displaying color output on a video monitor using the three color primaries, red, green, and blue, referred to as the RGB color model. The relationship between the colors can be seen in this illustration:
- In addition to providing a good description of human color perception, the RGB model is the basis for displaying colors in television and computer screens.
- The RGB model is also used for recording colors in digital cameras, including still image and video cameras.

- The RGB model is an additive system.
- The secondary colors of RGB, cyan, magenta, and yellow, are formed by the mixture of two of the primaries and the exclusion of the third. Red and green combine to make yellow, green and blue make cyan, blue and red make magenta.
- The combination of red, green, and blue in full intensity makes white. White light is created when all colors of the EM spectrum converge in full intensity.
- To represent colors for a television or computer screen, each pixel of the screen is recorded as the triple (r, g, b) of numbers. One popular system uses numbers that range from 0 to 255 for each color.
- Thus, a color C, is expressed in RGB components as

$$C_\lambda = R\mathbf{R} + G\mathbf{G} + B\mathbf{B}$$

- The magenta vertex is obtained by adding red and blue to produce the triple (1,0,l). And white at (1, l, 1) is the sum of the red, green, and blue vertices. Shades of gray are represented along the main diagonal of the cube from the origin (black) to the white vertex.

- Each point along this diagonal has an equal contribution from each primary color, so that a gray shade halfway between black and white is represented as (0.5, 0.5, and 0.5).

  We can represent this model with the unit cube defined on R, G, and B axes,

  The origin represents black, and the vertex with coordinates (1, 1, l) is white. Vertices of the cube on the axes represent the primary colors and the remaining vertices represent the complementary color for each of the primary colors.

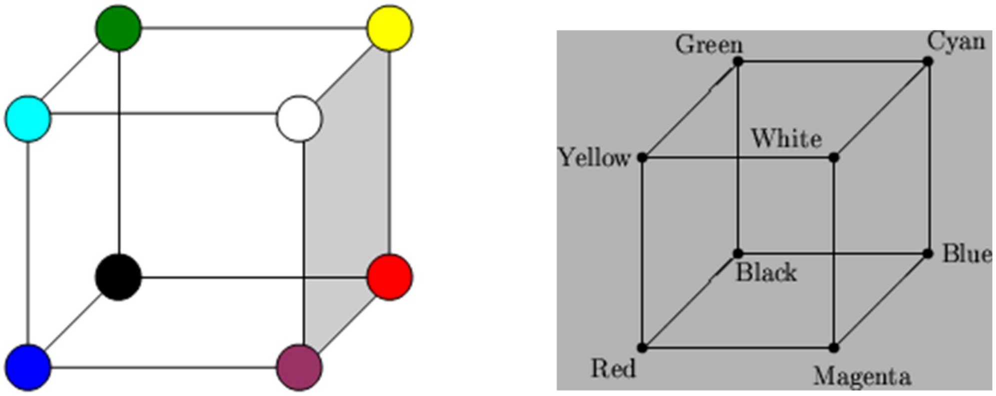


Figure 2.6 RGB Color Cube

Characteristics:

We cannot use RGB in printing devices because we use white paper and to produce black color we need to have no red, no green and no blue color, which again produces white as background is white. Whereas to produce white we need full amount of all three primaries which is unnecessary wastage of color. Thus it is not economical and produces error.

The red-green-blue model is formed by a color cube {(R, G, B): 0 <= R, G, B <=1}. Conversion from (R,G,B) to (X,Y,Z) is given via the chromaticity's (Xr,Yr,Zr), (Xg,Yg,Zg) and (Xb,Yb,Zb) of the CRTs phosphors by matrix multiplication via:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} Xr & X_g & X_b \\ Yr & Y_g & Y_b \\ Zr & Z_g & Z_b \end{pmatrix} \cdot \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

Let Cr= Xr + Yr + Zr

Then Xr=xr.Cr, Yr = Yr.Cr, and Zr=zr.Cr = (1 – xr – yr).Cr

1.7.3 The CMY Color Model

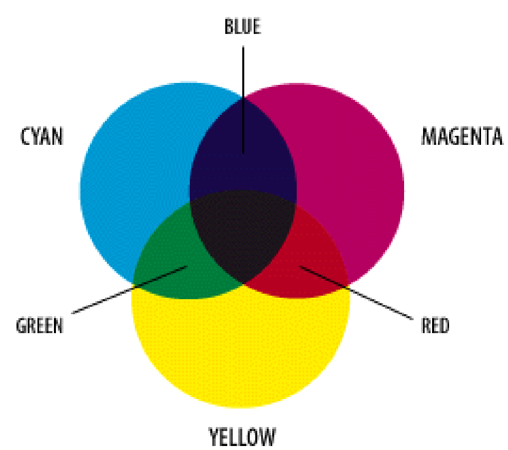
This stands for cyan-magenta-yellow and is used for hardcopy devices. In contrast to color on the monitor, the color in printing acts subtractive and not additive. A printed color that looks red absorbs the other two components G and B and reflects R. Thus its (internal) color is G+B=CYAN. Similarly R+B=MAGENTA and

R+G=YELLOW. Thus the C-M-Y coordinates are just the complements of the R-G-B coordinates:

Characteristics:

- Designed basically to be used for printing purpose.
- It is a Subtractive Model – ink and paint absorbs certain colors from image instead of adding to it.

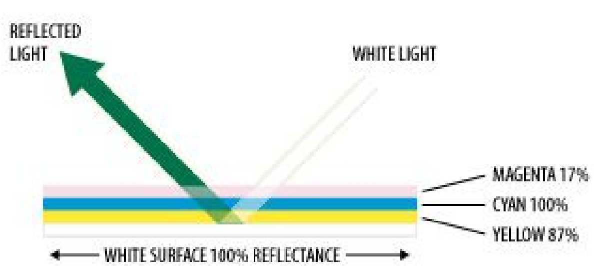
  For E.g. Cyan has a property of absorbing red color form white. Magenta has a property of absorbing green color form white and Yellow has a property of absorbing blue color form white. Whereas Black absorbs all three colors red, green and blue from white.



- The CMY model measures how much Cyan, Magenta, and Yellow component each image has.
- The conversion from RGB to CMY takes place at printout by

(C, M, Y) = (1, 1, and 1) – (R, G, and B)

The CMY model used in printing lays down overlapping layers of varying percentages of transparent cyan, magenta, and yellow inks. Light is transmitted through the inks and reflects off the surface below them (called the substrate). The percentages of CMY ink (which are applied as screens of halftone dots), subtract inverse percentages of RGB from the reflected light so that we see a particular color:



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

If we want to print a red looking color (i.e. with RGB coordinates (1,0,0)) we have to use CMY values of (0, 1,1). Note that M absorbs G, similarly Y absorbs B and hence M + Y absorbs all but R.

Black ((R,G,B) = (0,0,0)) corresponds to(C,M,Y)=(1,1,1) which should in principle absorb R, G and B. But in practice this will appear as some dark gray. So in order to be able to produce better contrast printers often use black as fourth color. This is the CMYK-model. Its coordinates are obtained from that of the CMY-model by K=max (C, M, Y), C=C-K, M=M-K and Y=Y-K.

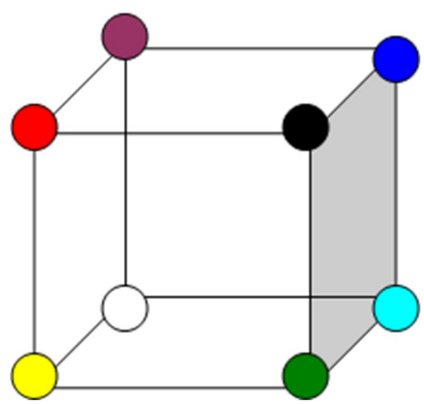


Figure 2.7. CMY Color Cube

we can express the conversion from an RGB representation to a CMY representation with the matrix transformation

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

Where the white is represented in the RGB system is the unit column vector.

Similarly, we convert from a CMY color representation to an RGB representation with the matrix transformation where black is represented In the CMY system as the unit column vector.

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

2.7.4. YIQ COLOR MODEL

The National Television System Committee (NTSC) color model for forming the composite video signal is the YIQ model, which is based on concepts in the CIE XYZ model. In the YI Q color model, parameter Y is the same as in the XYZ model. Luminance (brightness) information is contained in the Y parameter, while chromaticity information (hue and purity) is incorporated into the 1 and Q parameters.

from RGB values to YIQ values is accomplished with the transformation

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.528 & 0.311 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

This transformation is based on the NTSC standard RGB phosphor, whose chromaticity coordinates were given in the preceding section. The larger proportions of red and green assigned to parameter Y indicate the relative importance of these hues in determining brightness, compared to blue.

An NTSC video signal can be converted to an RGB signal sing an NTSC decoder, which separates the video signal into the YIQ components, then converts to RGB values. We convert from YIQ space to RGB space with the inverse matrix transformation from Eq. 15-6:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.000 & 0.956 & 0.620 \\ 1.000 & -0.272 & -0.647 \\ 1.000 & -1.108 & 1.705 \end{bmatrix} \cdot \begin{bmatrix} Y \\ I \\ Q \end{bmatrix}$$

2.7.5. The HSV Color Model

Alvy Ray Smith in 1978 was the first to describe colors using hue, saturation, and value (HSV model).Instead of a set of color primaries, the HSV model uses color descriptions that have a more intuitive appeal to a user. Color parameters in this model are hue (H), saturation(S) and value

Hue is a saturated color on the outer rim of the HSV Color Wheel

Saturation is the amount of white added to the color. 0% means that the color (at V=100%) is totally white; 100% means totally saturated with no white added (a fully saturated color is a pure hue on the outer rim of the HSV color wheel).

Value is the brightness of the color. 0% means totally dark or black; 100% means full brightness, with the color is fully determined by the hue and saturation.

Hue is represented as an angle about the vertical axis ranging from 0 degree at red through 360 degrees. Vertices of the hexagon are separated by 60^0 intervals. Yellow is at 60^0, green at 120^0 and cyan opposite e red at H = 180^0 complementary colors are 1800 apart

Value V varies from 0 at the apex of the hex cone to I at the top. The apex represents black. At the top of the hex cone, colors have their maximum intensity. When V = 1 and S = 1, we have the "pure" hues. White is the point at V = 1 and S =0

Saturation S varies from 0 to 1. It is represented in this model as the ratio of the purity of a selected hue to its maximum purity at S = 1. A selected hue is said to be one-quarter pure at the value S = 0.25. At S = 0, we have the gray scale.

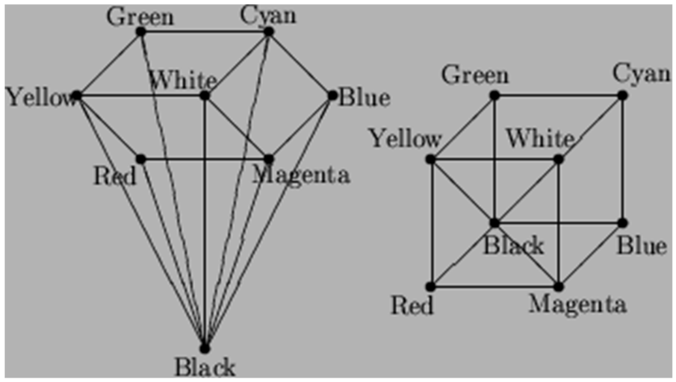


Figure 2.8. The HSV-model versus the RGB-model

The conversion from RGB to HSV is given by affine coordinate changes on each of the 3 four-sided sub-pyramids corresponding each to 1/3 of the color cube

2.7.7 The HLS Color Model

The HSL model describes colors in terms of hue, saturation, and lightness (also called luminance). (Note: the definition of saturation in HSL is substantially different from HSV, and lightness is not intensity.) The model has two prominent properties:

- The transition from black to a hue to white is symmetric and is controlled solely by increasing lightness
- Decreasing saturation transitions to a shade of gray dependent on the lightness, thus keeping the overall intensity relatively constant
- The angular relationship between tones around the color circle is easily identified
- Shades, tints, and tones can be generated easily without affecting the hue

Saturation, or the lack of it, produces tones of the reference hue that converge on the zero-saturation shade of gray, which is determined by the lightness. The following examples use the hues red, orange, and yellow at midpoint lightness with decreasing saturation. The resulting RGB value and the total intensity are shown.

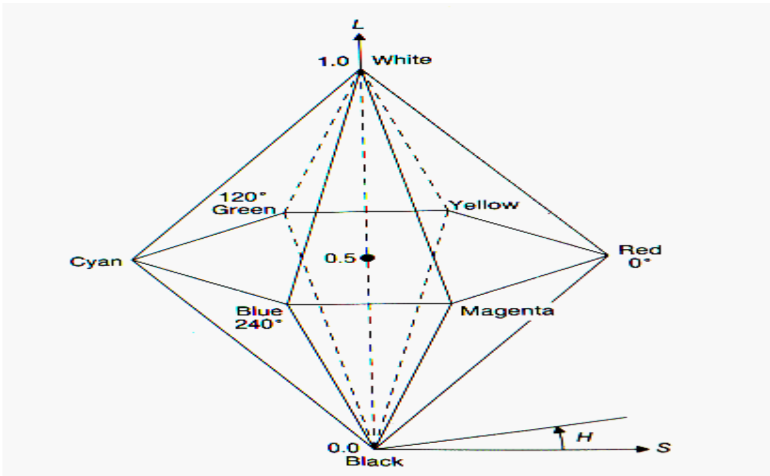


Figure2.9. HLS Color Model

2.8 ANIMATION

2.10DESIGN OF ANIMATION SEQUENCES

In general, an animation sequence is designed with the following steps:

5. Storyboard layout
6. Object definitions
7. Key-frame specifications
8. Generation of in-between frames

This standard approach for animated cartoons is applied to other animation applications as well, although there are many special applications that do not follow this sequence.

The Storyboard layout is an outline of the action. It defines the motion sequence as a set of basic events that are to take place. Depending on the type of animation to be produced, the storyboard could consist of a set of rough sketches or it could be a list of the basic ideas for the motion.

An object definition is given for each participant in the action. Objects can be defined in terms of basic shapes, such as polygons or splines. In addition, the associated movements for each object are speeded along with the shape.

A Key frame is a detailed drawing of the scene at a certain time in the animation sequence. Within each key frame, each object is positioned according to the time for that frame. Some key frames are chosen at extreme positions in the action; others are spaced so that the time interval between key frames is not too great. More key frames are specified for intricate motions than for simple, slowly varying motions.

In-between frames are the intermediate frames between the key frames. The number of in-betweens needed is determined by the media to be used to display the animation. Film requires 24 frames per second, and graphics terminals are refreshed at the rate of 30 to 60 frames per second. Typically, time intervals for the motion are set up so that there are from three to five in-betweens for each pair of key frames.

GENERAL COMPUTER-ANIMATION FUNCTIONS

Some steps in the development of an animation sequence are well-suited to computer solution. These include object manipulations and rendering, camera motions, and the generation of in-betweens. One function available in animation packages is provided to store and manage the object database. Object shapes and associated parameters are stored and updated in the database.

Raster Animations

On raster systems, we can generate real-time animation in limited applications using raster operations. Sequences of raster operations can be executed to produce real-time animation of either two-dimensional or three-dimensional objects, as long as we restrict the animation to motions in the projection plane.

Computer-Animation Languages

Design and control of animation sequences are handled with a set of animation routines. A general-purpose language, such as C, Lisp, Pascal, or FORTRAN, is often used to program the animation functions, but several specialized animation languages have been developed. Animation functions include a graphics editor, a key-frame generator, an in-between generator and standard graphics routines.

Key Frame Systems

We generate each set of in-betweens from the specification of two (or more) key frames. Motion paths can be given with a kinematic an s a set of spline curves, or the motions can be physically based by specifying the actions on the objects to be animated.

Morphing

Transformation of object shapes from one form to another is called morphing, which is a shortened form of metamorphosis. Morphing methods can be applied to any motion or transition involving a change in shape.

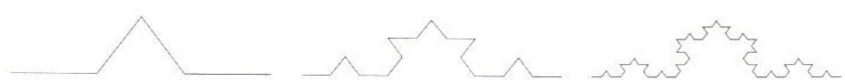
Motion Specifications There are several ways in which the motions of objects can be specified in an animation system. Direct Motion Specification is the most straightforward method for defining a motion sequence is direct specification of the motion parameters. Here, we explicitly give the rotation angles and translation vectors. Then the geometric transformation matrices are applied to transform coordinate positions

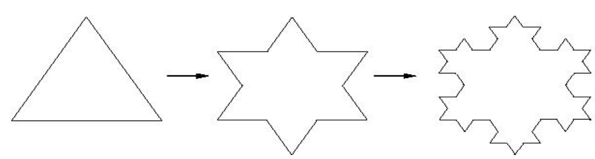
2.9 Fractals

Fractal objects refer to those objects which are self-similar at all resolutions.

Most of the natural objects such as trees, mountains and coastlines are considered as fractal objects because no matter how far or how close one looks at them, they always appear to be somewhat similar.

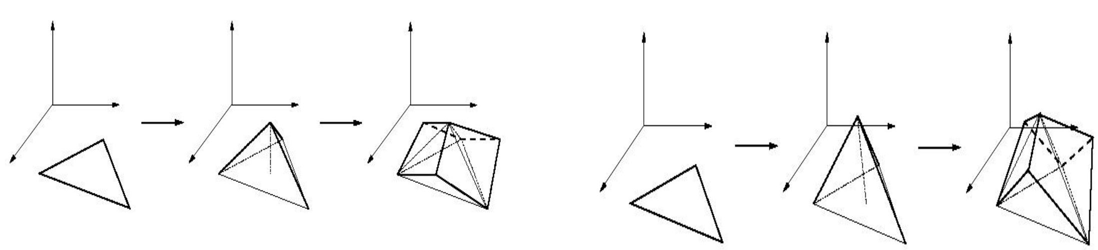
For example, a Fractal Snowflake:





Euclidean dimensions and Fractal dimensions:

- A line segment is 1D. If we divide a line into N equal parts, the parts each look like the original line scaled down by a factor of $N = N^{1/1}$.
- A square is 2D. If we divide it into N parts, each part looks like the original scaled down by a factor of $N^{1/2}$.
- For the fractal snowflake, when it is divided into 4 pieces, each resulting piece looks like the original scaled down by a factor of 3, so it has the dimension ***d*** such that $4^{1/d}=3$. That is, d = 1.26.



UNIT III MULTIMEDIA SYSTEMS DESIGN

Multimedia basics – Multimedia applications – Multimedia system architecture –Evolving technologies for multimedia – Defining objects for multimedia systems –Multimedia data interface standards – Multimedia databases.

3.1 Introduction

Multimedia has become an inevitable part of any presentation. It has found a variety of applications right from entertainment to education.

3.2 Categories of Multimedia

Multimedia may be broadly divided into linear and non-linear categories. Linear active content progresses without any navigation control for the viewer such as a cinema presentation.

Features of Multimedia

i) Multimedia presentations may be viewed in person on stage, projected, transmitted, or played locally with a media player.

ii) Multimedia games and simulations may be used in a physical environment with special effects, with multiple users in an online network, or locally with an offline computer, game system, or simulator.

iii) Enhanced levels of interactivity are made possible by combining multiple forms of media content But depending on what multimedia content you have it may vary Online multimedia is increasingly becoming object-oriented and data-driven, enabling applications with collaborative end-user innovation and personalization on multiple forms of content over time.

3.3 Elements of Multimedia System

The components that fall under our definition of multimedia are: '

1. **Facsimile:** Facsimile transmissions were the first practical means of transmitting document images over telephone lines.
2. **Document images:** Document images are used for storing business

documents that must be retained for long periods of time or may need to be accessed by a large number of people.

3. **Photographic images:** Photographic images are used for a wide range of applications.
4. **Geographic information systems map (GIS):** Map created in a GIS system is being used wildly for natural resources and wild life management as well as urban planning.
5. **Voice commands and voice synthesis:** Voice commands and voice synthesis are used for hands-free operations of a computer program.
6. **Audio message:** Annotated voice mail already uses audio or voice message as attachments to memos and documents such as maintenance manuals.
7. **Video messages:** Video messages are being used in a manner similar to annotated voice mail.
8. **Holographic images:** All of the technologies so for essentially present a flat view of information. Holographic images extend the concept of virtual reality by allowing the user to get "inside" a part, such as, an engine and view its operation from the inside.
9. **Fractals:** Fractals started as a technology in the early 1980s but have received serious attention only recently. This technology is based on synthesizing and storing algorithms that describes the information.

3.4 MULTIMEDIA APPLICATIONS

1. DOCUMENT IMAGING

Document Image Hardware requirements:

Real time image decompression and display place an important role on image processing hardware. Image decompression and display hardware supports 4 to 8 planes. 4 planes provide 16 colors and 8 planes provide 256 colors. The image planes are also called bit planes, because, they are addressed by a bit in a bytes. Images must be processed at the rate of tens to hundreds of pixels per nana-second.

2. Image processing and Image Recognition

Image processing involves image recognition, Image enhancement, image synthesis, and image reconstruction. An image processing system may actually alter the contents of the image itself.

Image enhancement: Most image display systems feature some level of image adjustment. Increasing the sensitivity and contrast makes the picture darker by making borderline pixels black or increasing the gray-scale level of pixels.

Capabilities built in the compression boards might include the following

I. ***Image calibration****:* The overall image density is calibrated, and the image pixels are adjusted to a predefined level.

II. ***Real time alignment****:* The image is aligned in real-time for skewing caused by improper feeding of paper.

III. ***Gray-Scale normalization****:* The overall gray level of an image or picture is evaluated to determine if it is skewed in one direction.

IV. ***RGB hue intensity adjustment****:* Too much color makes picture garish and fuzzy. Automatic hue intensity adjustment brings the hue intensity within pre-defined ranges.

V. ***Color Separation****:* A picture with very little color contrast can be dull and may not bring out the details. The hardware used can detect and adjust the range of color separation.

VI. ***Frame averaging****:* The intensity level of the frame is averaged to overcome the effects of very dark or very light areas by adjusting the middle tones.

Image Synthesis

Image Animation

Computers-created or scanned images can be displayed sequentially at controlled display speeds to provide image animation that simulates real processes.

Image annotation

It can be performed in one of two ways: as a text file stored along with the image or as a small image stored with the original image. The annotation is overplayed over the original image for display purposes.

Image Recognition:

Optical Character Recognition

Data entry is the most expensive component of data processing, because it requires extensive clerical staff work to enter data. Automating data entry, both typed and handwritten, is a significant application that can provide high returns. Optical Character Recognition (OCR) technology is used for data entry by scanning typed or printed words in a form.

Handwriting recognition

Research for Handwriting recognition was performed for *CADI* CAM systems for command recognition. Pen-based systems are designed to allow the user to write· commands on an electronic tablet.

Non-Textual Image Recognition

(i) 512 x 512 arrays of custom pixel processors that extract basic features such as lines and object boundaries.

(ii) The features of an object extracted by the first layer are tracked by the DSP array, and that information is fed into 5l2-M byte RAM. (iii) At the highest level, sophisticated AI algorithms perform the difficult task of object and scene recognition. .

3. Creative industries

Creative industries use multimedia for a variety of purposes ranging from fine arts, to entertainment, to commercial art, to journalism, to media and software services provided for any of the industries listed below.

4. Commercial

Much of the electronic old and new media utilized by commercial artists is multimedia. Exciting presentations are used to grab and keep attention in advertising.

Entertainment and Fine Arts

In addition, multimedia is heavily used in the entertainment industry, especially to develop special effects in movies and animations. Multimedia games are a popular pastime and are software programs available either as CD-ROMs or online.

6.Education

In Education, multimedia is used to produce computer-based training courses (popularly called CBTs) and reference books like encyclopedia and almanacs. A CBT lets the user go through a series of presentations, text about a particular topic, and associated illustrations in various information formats.

7.Engineering

Software engineers may use multimedia in Computer Simulations for anything from entertainment to training such as military or industrial training.

8.Industry

In the Industrial sector, multimedia is used as a way to help present information to shareholders, superiors and coworkers.

9.Mathematical and Scientific Research

In Mathematical and Scientific Research, multimedia is mainly used for modeling and simulation. For example, a scientist can look at a molecular model of a particular substance and manipulate it to arrive at a new substance.

10. Medicine

In Medicine, doctors can get trained by looking at a virtual surgery or they can simulate how the human body is affected by diseases spread by viruses and bacteria and then develop techniques to prevent it.

11. Multimedia in Public Places

In hotels, railway stations, shopping malls, museums, and grocery stores, multimedia will become available at stand-alone terminals or kiosks to provide information and help.

3.5 MULTIMEDIA SYSTEMS ARCHITECTURE

The following figure describes the architecture of a multimedia workstation environment.

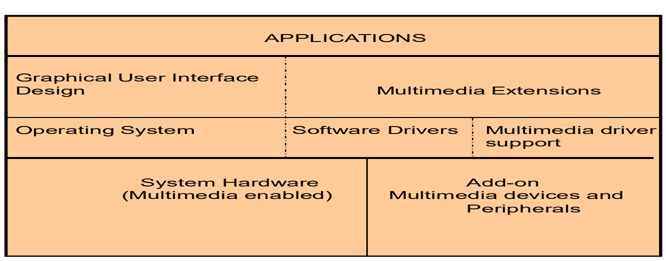


Figure 3.1 Multimedia System Architecture

In this diagram, the right side shows the new architectural entities required for supporting multimedia applications.

For each special devices such as scanners, video cameras, VCRs and sound equipments, a software device driver is need to provide the interface from an application to the device. The GUI require control extensions to support applications such as full motion video

High Resolution Graphics Display

(i) **VGA mixing:** In VGA mixing, the image acquisition memory serves as the display source memory, thereby fixing its position and size on screen:

(ii) **VGA mixing with scaling:** Use of scalar ICs allows sizing and positioning of images in pre-defined windows. Resizing the window causes the things to be retrieved again.

(iii)**Dual-buffered VGA/Mixing/Scaling:** Double buffer schemes maintain the original images in a decompression buffer and the resized image in a display buffer.

1) The IMA Architectural Framework

The architectural approach taken by IMA is based on defining interfaces to a multimedia interface bus. This bus would be the interface between systems and multimedia sources. It provides streaming I/O services; including filters and translators **Figure 3.5** describes the generalized architectural approach

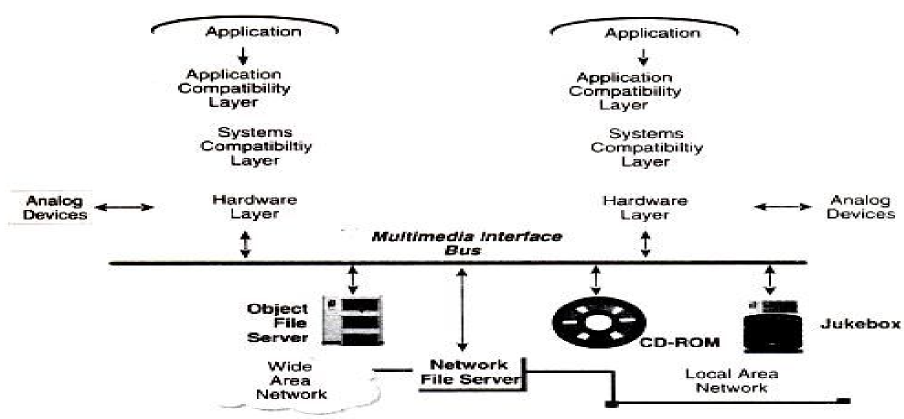


Figure 3.2 IMA Architecture

2) Network Architecture for Multimedia Systems:

Multimedia systems need special networks. Because large volumes of images and video messages are being transmitted. Asynchronous Transfer Mode technology (A TM) simplifies transfers across LANs and WANs.

3)Task based Multi level networking

Higher classes of service require more expensive components in the workstations as well as in the servers supporting the workstation applications.

Rather than impose this cost on all work stations, an alternate approach is to adjust the class of service to the specific requirement for the user. This approach is to adjust the class of services according to the type of data being handled at a time also.

We call this approach task-based multilevel networking.

1. **Duplication:** It is the process of duplicating an object that the user can manipulate. There is no requirement for the duplicated object to remain synchronized with the source (or master) object.
2. **Replication:** Replication is defined as the process of maintaining two or more copies of the same object in a network that periodically re-synchronize to provide the user faster and more reliable access to the data Replication is a complex process.

3. **Networking Standards:** The two well-known networking standards are Ethernet and token ring. ATM and FDDI are the two technologies which we are going to discuss in detail.

ATM: It is a high-speed networking standard designed to support both voice and data communications. ATM is normally utilized by Internet service providers on their private long-distance networks. ATM operates at the data link layer (Layer 2 in the OSI model) over either fiber or twisted-pair cable.

ATM differs from more common data link technologies like Ethernet in several ways. For example, ATM utilizes no routing. Hardware devices known as *ATM switches* establish point-to-point connections between endpoints and data flows directly from source to destination. Additionally, instead of using variable-length packets as Ethernet does, ATM utilizes fixed-sized cells. *ATM cells* are 53 bytes in length that includes 48 bytes of data and five (5) bytes of header information.

The performance of ATM is often expressed in the form of OC (Optical Carrier) levels, written as "OC-xxx." Performance levels as high as 10 Gbps (OC-192) are technically feasible with ATM. More common performance levels for ATM are 155 Mbps (OC-3) and 622 Mbps (OC-12).

ATM technology is designed to improve utilization and quality of service (QoS) on high-traffic networks. Without routing and with fixed-size cells, networks can much more easily manage bandwidth under ATM than under Ethernet, for example. The high cost of ATM relative to Ethernet is one factor that has limited its adoption to backbone and other high-performance, specialized networks.

ATM provides high capacity, low-latency switching fabric for data. It is independent of protocol and distances. ATM effectively manages a mix of data types, including text data, voice, images and full motion video. ATM was proposed as a means of transmitting multimedia applications over asynchronous networks.

FDDI: FDDI is an acronym of Fiber Distributed Data Interface. It is a standard for data transmission in a local area network. It uses optical fiber as its standard underlying physical medium. This FDDI network is an excellent candidate to act as the hub in a network configuration, or as a backbone that interconnects different types of LANs.

FDDI presents a potential for standardization for high speed networks. The ANSI standard for FDDI allows large-distance networking. It can be used as high-performance backbone networks to complement and extend current LANs. FDDI provides a 100 Mbit/s optical standard for data transmission in local area network that can extend in range up to 200 kilometers (120 mi). Although FDDI logical topology is a ring-based token network, it did not use the IEEE 802.5 token ring protocol as its basis; instead, its protocol was derived from the IEEE 802.4 token bus *timed token* protocol. In addition to covering large geographical areas, FDDI local area networks can support thousands of users. FDDI offers both a Dual-Attached Station (DAS), counter-rotating token ring topology and a Single Attached Station (SAS), token bus passing ring topology.

FDDI, as a product of American National Standards Institute X3T9.5 (now X3T12), conforms to the Open Systems Interconnection (OSI) model of functional layering using other protocols. The standards process started in the mid 1980s. FDDI-II, a version of FDDI described in 1989, added circuit-switched service capability to the network so that it could also handle voice and video signals.[3] Work started to connect FDDI networks to synchronous optical networking (SONET) technology.

A FDDI network contains two rings, one as a secondary backup in case the primary ring fails. The primary ring offers up to 100 Mbit/s capacities. When a network has no requirement for the secondary ring to do backup, it can also carry data, extending capacity to 200 Mbit/s. The single ring can extend the maximum distance; a dual ring can extend 100 km (62 mi). FDDI had a larger maximum-frame size (4,352 bytes) than the standard Ethernet family, which only supports a maximum-frame size of 1,500 bytes, allowing better effective data rates in some cases.

3.6 EVOLVING TECHNOLOGIES FOR MULTIMEDIA SYSTEMS

Multimedia applications use a number of technologies generated for both commercial business application as well as the video game industry.
Let us review some of these technologies in this section.

I) Hypertext

Hypertext systems allow authors to link information together; create information paths through a large volume of related text in documents. It also allows annotating existing text, and appending notes. It allows fast and easy searching and reading of selected excerpts.

II) Hypermedia

It is an extension of hypertext. Hypermedia documents are documents which have text, embedded or linked multimedia objects such as image, audio, hologram, or full-motion video. Hypermedia documents used for electronic mail and work flow applications provide a rich functionality for exchanging a variety of information types. The hypermedia document is a definition of a document and a set of pointers to help locate the various elements of the document on the network.

III) Hyper Speech

Multimedia stimulated the development of general-purpose speech interfaces. Speech synthesis and speech recognition are fundamental requirement for hyper speech systems. Speech recognition is nothing but converting the analog speech into a computer action and into ASCII text. Speech-recognition systems cannot segment a stream of sounds without breaks into meaningful units. The user must speak in a stilted fashion. He should make sure to interpose silence between each word.

IV) HDTV AND UDTV

HDTV is an acronym of High-Definition Television. The broadcasting standards such as NTSC, PAL, SECAM, and NHK have an idea of bringing the world together on a single high-definition Television broadcasting standard. The Japanese broadcasting services developed an 1125-line, along MUSE system. A competing standard in the U.S. changed direction from analog to digital technology: An 1125-line digital HDTV has been developed and is being commercialized. NHK of Japan is trying to leapfrog the digital technology to develop ultra definition television (digital UDTV) featuring approximately 3000 lines

High Definition Television (HDTV) provides high resolution in a 16:9 aspect ratio (see following Figure). This aspect ratio allows the viewing of Cinemascope and Panavision movies. There is contention between the broadcast and computer industries about whether to use interlacing or progressive-scan technologies.

V) 3D Technologies and Holography

Three-dimensional technologies are concerned with two areas: pointing devices and displays. 3-D pointing devices are essential to manipulate object in a 3-D display system. 3-D displays are achieved using holography techniques. The techniques developed for holography have been adapted for direct computer use.

VI) Fuzzy Logic

Fuzzy logic is logic which is used for low-level process controllers.Use of fuzzy logic in multimedia chips is the key to the emerging graphical interfaces of the future. It is expected to become an integral part of multimedia hardware. Fuzzy logic has mathematical principles. Hence, the application of multimedia can benefit those principles.

VII) Digital Signal Processing

Digital Signal Processing is used in applications such as digital servos in hard disk drives, and fax/modems. DSP technology is used in Digital wireless communications, such as personal communication networks (pens), wireless local area networks and digital cordless phones.

DSP Architectures and Applications

A typical DSP operating system architecture would contain the following subsystems:

I. *Memory Management:* DSP architectures provide dynamic allocation of arrays from multiple segments, including RAM, SRAM and DRAM.

II. *Hardware-Interrupt handling:* A DSP operating system must be designed to minimize hardware-interrupt latency to ensure fast response to real time events for applications, such as servo systems.

III. *Multitasking:* DSPs need real-time kernels that provide pre-emptive multitasking and user-defined and dynamic task prioritization

IV. *Interties Synchronization and Communication***:** Mechanisms for interties communication include message queues, semaphores, shared memory, and quick response event flags. Multiple timer services: The ability for the developer to set system clock interrupt managed timers to control and synchronize tasks is needed for most real-time applications.

Device-Independent I/O: DSP operating system should supports

i) Asynchronous data stream

ii) Synchronous message passing.

Use of DSP' s has evolved from traditional general purpose digital signal processors to application-specific and customizable DSPs. DSPs were conceived as math engines with a system architecture that was like that of a mini-computer with an array processor.

3.7 DEFINING OBJECTS FOR MULTIMEDIA SYSTEMS

The basic data types of object using in multimedia include text, image, audio, holograms and full-motion video.

1. TEXT

It is the simplest of data types and requires the least amount of storage. Text is the base element of a relational database. It is also the basic building of a document. The major attributes of text include paragraph styling, character styling, font families and sizes, and relative location in a document

2. HYPERTEXT

It is an application of indexing text to provide a rapid search of specific text strings in one or more documents. It is an integral component of hypermedia documents. A hypermedia document is the basic complex object of which text is a sub object. Sub objects include images, sound and full motion video.A hypermedia document always has text and has one or more other types of sub-objects

3. IMAGES

Image object is an object that is represented in graphics or encoded form. Image object is a sub object of the hypermedia document object. In this object, there is no direct relationship between successive representations in time.The image object includes all data types that are not coded text. It does not have a temporal property associated with them.

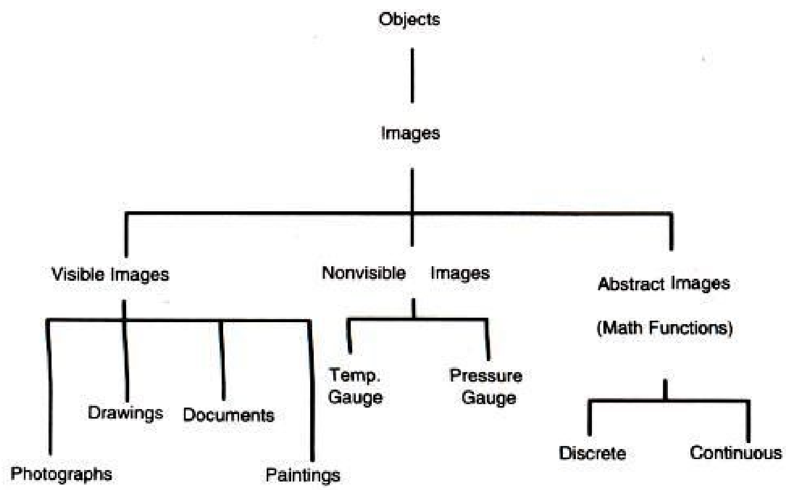


Figure 3.3 Image Classification

The data types such as document images, facsimile systems, fractals, bitmaps, Meta files, and still pictures or still video frames are grouped together. Figure describes a hierarchy of the object classes

i) Non-Visible: This type of images is not stored as images. But they are displayed as images. Example: Pressure gauges, and temperature gauges.

ii) Abstract: Abstract images are computer-generated images based on some arithmetic calculations. They are really not images that ever existed as real-world objects. Example of these images is fractals.

4. AUDIO AND VOICE

Stored-Audio and Video objects contain compressed audio information. This can consist of music, speech, telephone conversation and voice commands. An Audio object needs to store information about the sound clip.

Information here means length of the sound clip, its compression algorithm, playback characteristics, and any annotations associated with the original clip.

5. FULL MOTION AND LIVE VIDEO

Full motion video refers to pre-stored video clips. Live video refers to live and it must be processed while it is being captured by the camera. . From a storage perspective, we should have the information about the coding algorithm used for compression. It needs decoding also. From a processing perspective, video should be presented to user with smooth and there should not be any unexpected breaks.

Hence, video object and its associated audio object must be transferred over the network to the decompression unit. It should be then played at the fixed rate specified for it. For successful playback of compressed video, there are number of technologies. They are database storage, network media and protocols, decompression engines and display engines.

3.8 MULTIMEDIA DATABASES

Images, sounds and movies can be stored, retrieved and played by many databases. In future, multimedia databases will become a main source of interaction between users and multimedia elements.

Multimedia storage is characterized by a number of considerations. They are:

(i) massive storage volumes

(ii) large object sizes

(iii) multiple related objects

(iv) temporal requirements for retrieval

A single multimedia document may be a combination of different media. Hence indexing of documents, films and tapes is more complex. Locating massive data volumes requires searching through massive storage files. Locating and indexing systems can be understood only by a few key staff personnel. Hence it requires a major organizational effort to ensure that they are returned in proper sequence to their original storage location. There are two major mass storage technologies used currently for storage of multimedia documents. i)

Optical disk storage systems. (ii) High-speed magnetic storage.

Multimedia object storage

Multimedia object storage in an optical medium serves its original purpose, only if it can be located fast and automatically. A key issue here is random keyed Access the various components of hypermedia database record. Optical media provides very dense storage. Speed of retrieval is another consideration.

Retrieval speed is a direct result of the storage latency, size of the data relative to display resolution, transmission media and speed, and decompression efficiency. Indexing is important for fast retrieval of information. Indexing can be at multiple levels.

Multimedia document retrieval

The simplest form of identifying a multimedia document is by storage platter identification and its relative position on the platter (file number). These objects can then be grouped using a database in folders (replicating the concept of paper storage in file folders) or within complex objects representing hypermedia documents.

The capability to access objects using identifiers stored in a database requires capability in the database to perform the required multimedia object directory functions. Another important application for sound and full motion video is the ability to clip parts of it and combine them with another set. Indexing of sound and full-motion video is the subject of intense debate and a number of approaches have been used.

3.8.1 DATABASE MANAGEMENT SYSTEMS FOR MULTIMEDIA

Since most multimedia applications are based primarily on communications technologies, such as electronic mail, the database system must be fully distributed. A number of database storage choices are available.

The choices available are:

i) Extending the existing relational database management systems, (RDBMSs) to support the various objects for multimedia as binary objects.

ii) Converting to a fully fledged object oriented database that supports the standard SQL language.

Multimedia applications combine numerical and textual data, graphics from GUI front-ends, CAD/CAM systems and GIS applications, still video, audio and full-motion video with recorded audio and annotated voice components. Relational databases, the dominant database paradigm, have lacked the ability to support multimedia databases. Key limitations of relational database systems for implementing multimedia applications stem from two areas: the relational data model and the relational computational model.

RDBMSs have been designed to manage only tabular alphanumeric forms of data (along with some additional data types stored in binary form such as dates).

1. RDBMS EXTENSIONS FOR MULTIMEDIA

Binary Large Object (BLOB) is a data type which has been adapted by most of the leading relational databases. BLOBs are used for objects such as images or other binary data types.

The relational database is extended to access these BLOBs to present the user 'with a complete' data set. Extended relational databases provide a gradual migration path to a more object-oriented environment. Relational database tables include location information for the BLOBs which may be stored outside the database on separate image or video servers. Relational databases have the strength of rigorous set management for maintaining the integrity of the database

2. OBJECT-ORIENTED DATABASES FOR MULTIMEDIA

In object databases, data remains in RMS or flat files. Object databases can provide the fastest route to multimedia support. Object programming embodies the principles of reusable code and modularity. This will ease future maintenance of these databases. Object database capabilities such as message passing, extensibility, and the support of hierarchical structures, are important for multimedia systems. We can develop the application fastest class definitions. ODBMSs are extensible. They allow incremental changes to the database applications.

Features of ODBMS:

Extensibility: Extensibility means that the set of operations, structures and constraints that are available to operations are not fixed, and developers can define new operations, which can then be added as needed to their application.

Association: It is the ability to define a software entity in terms of its differences from another entity.

Classification: It is the ability to represent with a single software entity a number of data items that all have the same behavior and the same state attributes. Object orientation helps to organize the software in a more, modular and re-usable manner.

Encapsulation allows for the development of open systems where one part of the application does not need to know the functioning of other part.

Autonomy means we can interface to a variety of external programs can be built in one class of objects and the storage of the data in another class of objects.

3.8.2 Database Organization for Multimedia Applications

Data organization for multimedia systems has some key issues. They are:

(l) Data independence

(2)Distributed database servers

(3)Multimedia object management.

1) Data Independence

Flexible access by a number of databases requires that the data be independent from the application so that future applications can access the data without constraints related to a previous application.

Key features of data independent designs are:

1. Storage design in independent of specific applications.
2. Explicit data definitions are independent of application program.
3. Users need not know data formats or physical storage structures.
4. Integrity assurance in independent of application programs.
5. Recovery in independent of application programs.

2) Distributed Database Servers

Distributed database servers are a dedicated resource on a network accessible to a number of applications. The database server is built for growth and enhancement, and the network provides the opportunity for the growth of applications and distributed access to the data.

3) Multimedia Object Management

The object management system must be capable of indexing, grouping and storing multimedia objects in distributed hierarchical optional storage systems, and accessing these objects on or keyed basis.

The design of the object management system should be capable indexing objects in such a manner that there is no need to maintain multiple storage copies. Multimedia transactions are very complex transactions. We define a multimedia transaction as the sequence of events that starts when a user makes a request to display, edit, or print a hyper media document. The transaction is complete when the user releases the hypermedia document and stores back the edited versions or discards the copy in memory (including virtual memory) or local storage.

UNIT IV -MULTIMEDIA FILE HANDLING

Compression and decompression − Data and file format standards − Multimedia I/Technologies − Digital voice and audio − Video image and animation − Full motion video – Storage and retrieval technologies.

4.1 COMPRESSION AND DECOMPRESSION

Compression is the way of making files to take up less space. In multimedia systems, in order to manage large multimedia data objects efficiently, these data objects need to be compressed to reduce the file size for storage of these objects. Compression tries to eliminate redundancies in the pattern of data.

For example, if a black pixel is followed by 20 white pixels, there is no need to store all 20 white pixels. A coding mechanism can be used so that only the count of the white pixels is stored. Once such redundancies are removed, the data object requires less time for transmission over a network. This in turn significantly reduces storage and transmission costs.

Compression and decompression techniques are utilized for a number of applications, such as facsimile system, printer systems, document storage and retrieval systems, video teleconferencing systems, and electronic multimedia messaging systems. An important standardization of compression algorithm was achieved by the CCITT when it specified Group 2 compression for facsimile system. CCITT Comitee Consultatif International Telephonique et Telegraphique, is part of the ITU International Telecommunication Union, one of the specialized agencies of the United Nations

When information is compressed, the redundancies are removed. Sometimes removing redundancies is not sufficient to reduce the size of the data object to manageable levels. In such cases, some real information is also removed. The primary criterion is that removal of the real information should not perfectly affect the quality of the result. In the case of video, compression causes some information to be lost; some information at a delete level is considered not essential for a reasonable reproduction of the scene. This type of compression is called loss compression. Audio compression, on the other hand, is not loss. It is called lossless compression.

4.2 TYPES OF COMPRESSION

Lossless Compression:

In lossless compression, data is not altered or lost in the process of compression or decompression. Decompression generates an exact replica of the original object. Text compression is a good example of lossless compression. The repetitive nature of text, sound and graphic images allows replacement of repeated strings of characters or bits by codes. Lossless compression techniques are good for text data and for repetitive data in images all like binary images and gray-scale images.

Some of the commonly accepted lossless standards are given below:

- **Pack pits encoding (Run-length encoding)**
- **CCITT Group 3 I D**
- **CCITT Group 3 2D**
- **CCITT Group 4**

1. Binary Image compression schemes

Binary Image Compression Scheme is a scheme by which a binary image containing black and white pixel is generated when a document is scanned in a binary mode. The schemes are applicable in office/business documents, handwritten text, line graphics, engineering drawings, and so on. Let us view the scanning process. A scanner scans a document as sequential scan lines, starting from the top of the page.

A scan line is complete line of pixels, of height equal to one pixel, running across the page. It scans the first line of pixels (Scan Line), then scans second "line, and works its way up to the last scan line of the page. Each scan line is scanned from left to right of the page generating black and white pixels for that scan line. This uncompressed image consists of a single bit per pixel containing black and white pixels. Binary 1 represents a black pixel, binary 0 a white pixel. Several schemes have been standardized and used to achieve various levels of compressions. Let us review the more commonly used schemes.

Pack pits Encoding (Run-Length Encoding):

It is a scheme in which a consecutive repeated string of characters is replaced by two bytes. It is the simple, earliest of the data compression scheme developed. It is used to compress black and white (binary) images. Among two bytes which are being replaced, the first byte contains a number representing the number of times the character is repeated, and the second byte contains the character itself.

In some cases, one byte is used to represent the pixel value and the other byte is to represents the run length.

Example 3w 4b 9w 2b 2w 6b 5w 2b 5w...

2. CCITT Group 3 1-D Compression

Group 3 One-Dimensional coding (G3 1D) is called *Modified Huffman* (MH) as it encodes run lengths using a predefined Huffman code. This scheme is based on run length encoding and assumes that a typical scan line has long runs of the same color. This scheme was designed for black and white images only, not for gray scale or color images. The primary application of this scheme is in facsimile and early document imaging system.

Huffman Encoding:

A modified version of run-length encoding is Huffman encoding. It is variable-length encoding. It generates the shortest code for frequently occurring run lengths and longer code for less frequently occurring run lengths.

Mathematical Algorithm for Huffman encoding:

It is constructed based on the probability of occurrence of white pixels or black pixels in the run length or bit stream.

n	white runs	black runs
0	00110101	0000110111
1	000111	010
2	0111	11
3	1000	10
4	1011	011
5	1100	0011
6	1110	0010
7	1111	00011
8	10011	000101
9	10100	000100
10	00111	0000100
11	01000	0000101
12	001000	0000111

Table below shows the CCITT Group 3 tables showing codes or white run lengths and black run lengths

In order to maintain black/white syncronization, each line begins with a white run, eventually of zero length.The run-length code words are taken from a predefined table of values representing runs of black or white pixels. This table is used to encode and decode all Group 3 data.

The size of the code words were originally determined by the CCITT, based statistically on the average frequency of black-and-white runs occurring in typical type and handwritten documents. The documents included line art and were written in several different languages. Run lengths that occur more frequently are assigned smaller code words while run lengths that occur less frequently are assigned larger code words.

In printed and handwritten documents, short runs occur more frequently than long runs. Two- to 4-pixel black runs are the most frequent in occurrence. The maximum size of a run length is bounded by the maximum width of a Group 3 scan line.

Run lengths are represented by two types of code words: *makeup* and *terminating*. An encoded pixel run is made up of zero or more makeup code words and a terminating code word. Terminating code words represent shorter runs, and makeup codes represent longer runs. There are separate terminating and makeup code words for both black and white runs.

Pixel runs with a length of 0 to 63 are encoded using a single terminating code. Runs of 64 to 2623 pixels are encoded by a single makeup code and a terminating code. Run lengths greater than 2623 pixels are encoded using one or more makeup codes and a terminating code. The run length is the sum of the length values represented by each code word.

Here are some examples of several different encoded runs:

- A run of 20 black pixels would be represented by the terminating code for a black run length of 20. This reduces a 20-bit run to the size of an 11-bit code word, a compression ratio of nearly 2:1. This is illustrated in Figure a.
- A white run of 100 pixels would be encoded using the makeup code for a white run length of 64 pixels followed by the terminating code for a white run length of 36 pixels (64 + 36 = 100). This encoding reduces 100 bits to 13 bits, or a compression ratio of over 7:1. This is illustrated in Figure b.
- A run of 8800 black pixels would be encoded as three makeup codes of 2560 black pixels (7680 pixels), a makeup code of 1088 black pixels, followed by the terminating code for 32 black pixels (2560 + 2560 + 2560 + 1088 + 32 = 8800). In this case, we will have encoded 8800 run-length bits into five code words with a total length of 61 bits, for an approximate compression ratio of 144:1. This is illustrated in Figure c.

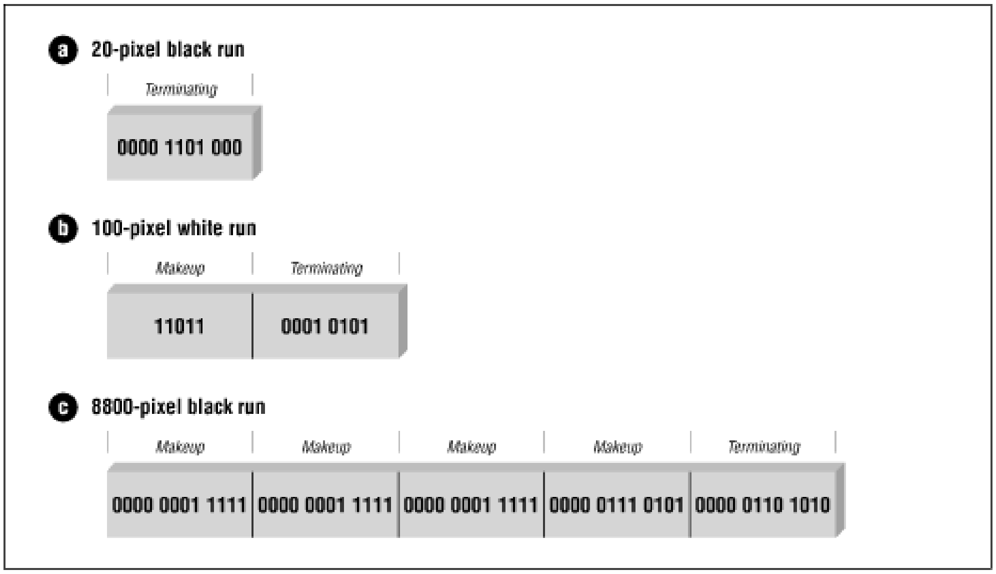


Figure 4.1 CCITT Group 3 1D encoding

- The use of run lengths encoded with multiple makeup codes has become a de facto extension to Group 3, because such encoders are necessary for images with higher resolutions. And while most Group 3 decoders do support this extension, do not expect them to do so in all cases.
- Decoding Group 3 data requires methods different from most other compression schemes. Because each code word varies in length, the encoded data stream must be read one bit at a time until a code word is recognized. This can be a slow and tedious process at best. To make this job easier, a state table can be used to process the encoded data one byte at a time. This is the quickest and most efficient way to implement a CCITT decoder.
- All scan lines are encoded to begin with a white run-length code word (most document image scan lines begin with white run lengths). If an actual scan line begins with a black run, a zero-length white run-length code word will be pretended to the scan line.
- A decoder keeps track of the color of the run it is decoding. Comparing the current bit pattern to values in the opposite color bit table is

wasteful. That is, if a black run is being decoded, there is no reason to check the table for white run-length codes.

- Several special code words are also defined in a Group 3-encoded data stream. These codes are used to provide synchronization in the event that a phone transmission experiences a burst of noise. By recognizing this special code, a CCITT decoder may identify transmission errors and attempt to apply a recovery algorithm that approximates the lost data.
- The EOL code is a 12-bit code word that begins each line in a Group 3 transmission. This unique code word is used to detect the start/end of a scan line during the image transmission. If a burst of noise temporarily corrupts the signal, a Group 3 decoder throws away the unrecognized data it receives until it encounters an EOL code. The decoder would then start receiving the transmission as normal again, assuming that the data following the EOL is the beginning of the next scan line. The decoder might also replace the bad line with a predefined set of data, such as a white scan line.
- A decoder also uses EOL codes for several purposes. It uses them to keep track of the width of a decoded scan line. (An incorrect scan-line width may be an error, or it may be an indication to pad with white pixels to the EOL.) In addition, it uses EOL codes to keep track of the number of scan lines in an image, in order to detect a short image. If it finds one, it pads the remaining length with scan lines of all white pixels.

Advantages of CCITT Group 3 ID

It is simple to implement in both hardware and software.

It is a worldwide standard for facsimile which is accepted for document imaging application. This allows document imaging applications to incorporate fax documents easily.

4.3 CCITT GROUP 3 2D COMPRESSION

It is also known as modified run length encoding. It is used for software based imaging system and facsimile. It is easier to decompress in software than CCITT Group 4. The CCITT Group 3 2D scheme uses a "k" factor where the image is divided into several group of k lines. This scheme is based on the statistical nature of images; the image data across the adjacent scan line is redundant. If black and white transition occurs on a given scan line, chances are the same transition will occur within + or - 3 pixels in the next scan line.

a. Necessity of k factor·

When CCITT Group 3 2D compression is used, the algorithm embeds Group 3 1 D coding between every k groups of Group 3 2D coding, allowing the Group 3 1 D coding to be the synchronizing line in the event of a transmission error. Therefore when a transmission error occurs due to a bad communication link, the group 3 I D can be used to synchronize and correct the error.

b. Data formatting for CCITT Group 3 2D

The 2D scheme uses a combination of additional codes called vertical code, pass code, and horizontal code to encode every line in the group of k lines. Group 3 Two-Dimensional coding (G3 2D) is called *Modified READ* (MR) as it is a variant of a previously defined code, called READ (Relative Element Address Designate). Many images have a high degree of *vertical coherence* between consecutive lines.C*hanging elements* are coded w.r.t. a "nearby" change position of the same color in the previous (reference) line .

- Nearby means within an interval of radius 3 pixels
- If there are *changing elements* in the current line without correspondents in the reference line à switch to *horizontal mode* (1D)
- On the opposite if the ref line has a run with no counterpart in the current line à special *pass code*

Advantage of CCITT Group 3 2D

The implementation of the k factor allows error-free transmission.
Compression ratio achieved is better than CCITT Group 3 1 D.
It is accepted for document imaging applications.

4.3.1 Lossy Compression:

Lossy compression is that some loss would occur while compressing information objects. Loss compression is used for compressing audio, gray-scale or color images, and video objects in which absolute data accuracy is not necessary. The idea behind the loss compression is that, the human eye fills in the missing information in the case of video. But, an important consideration is how much information can be lost so that the result should not affect. For example, in a grayscale image, if several bits are missing, the information is still perceived in an acceptable manner as the eye fills in the gaps in the shading gradient.

Lossy compression is applicable in medical screening systems, video tale conferencing, and multimedia electronic messaging systems. Loss compressions techniques can be used along with other compression methods in a multimedia object consisting of audio, color images, and video as well as other specialized data types.

The following lists some of the commonly used loss compression mechanisms:

- Joint Photographic Experts Group (JPEG)
- Moving Picture Experts Group (MPEG)

4.3.2 JOINT PHOTOGRAPHIC EXPERTS GROUP COMPRES5ION (JPEG)

ISO and CCITT working committee joint together and formed Joint Photographic Experts Group. It is focused exclusively on still image compression. Another joint committee, known as the Motion Picture Experts Group (MPEG), is concerned with full motion video standards. JPEG is a compression standard for still color images and grayscale images, otherwise known as continuous tone images. The **JPEG** image compression standard has become an important tool in the creation and manipulation of digital images. The primary algorithm underlying this standard is executed in several stages.

Stages in JPEG Compression:

- In the first stage, the image is converted from **RGB** format to a video-based encoding format in which the grayscale (**luminance**) and color (**chrominance**) information are separated. Such a distribution is desirable because grayscale information contributes more to perceptual image quality than does color information, due to the fact that the human eye uses grayscale information to detect boundaries. Color information can be dispersed across boundaries without noticeable loss of image quality. Thus, from a visual standpoint, it is acceptable to discard more of the color information than grayscale information, allowing for a greater compression of digital images.

- In the second stage, the luminance and chrominance information are each transformed from the spatial domain into the frequency domain. This process consists of dividing the luminance and chrominance information into square (typically 8 x 8) blocks and applying a two-dimensional **Discrete Cosine Transform** (**DCT**) to each block. The cosine transforms converts each block of spatial information into an efficient frequency space representation that is better suited for compression. Specifically, the transform produces an array of coefficients for real-valued basis functions that represent each block of data in frequency space. The magnitude of the DCT coefficients exhibits a distinct pattern within the array, where transform coefficients corresponding to the lowest frequency basis functions usually have the highest magnitude and are the most perceptually significant. Similarly, cosine transform coefficients corresponding to the highest frequency basis functions usually have the lowest magnitude and are the least perceptually significant.

The DCT Expression is

where

$$B(k_1, k_2) = \sum_{i=0}^{N_1-1} \sum_{j=0}^{N_2-1} 4 \cdot A(i,j) \cdot \cos\left[\frac{\pi \cdot k_1}{2 \cdot N_1} \cdot (2 \cdot i + 1)\right] \cdot \cos\left[\frac{\pi \cdot k_2}{2 \cdot N_2} \cdot (2 \cdot j + 1)\right]$$

block is pixels (in JPEG, 8x8)

- A(i,j) is the value of pixel of position (i,j)
- $B(k_1,k_2)$ is the DCT coefficient of position
- low values for k_1 corresponds to low vertical frequencies, low values for to k_2 low horizontal frequencies
- Generally higher frequencies have very low values

- In the third stage, each block of DCT coefficients is subjected to a process of **quantization**, wherein grayscale and color information are discarded. Each cosine transform coefficient is divided by its corresponding element in a scaled **quantization matrix**, and the resulting numerical value is rounded. The default quantization matrices for luminance and chrominance are specified in the JPEG standard, and were designed in accordance with a model of human perception. The scale factor of the quantization matrix directly affects the amount of image compression, and the **lossy** quality of JPEG compression arises as a direct result of this quantization process. Quantizing the array of cosine transform coefficients is designed to eliminate the influence of less perceptually significant basis functions.

The transform coefficients corresponding to these less significant basis functions are typically very small to begin with, and the quantization process reduces them to zeros in the resulting quantized coefficient array. As a result, the array of quantized DCT coefficients will contain a large number of zeros, a factor that is employed in the next stage to deliver significant data compression.

In practice the matrix of factor is usually

$$\begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

- In the fourth stage, a process of **run-length encoding** is applied to each block of quantized cosine transform coefficients. A zigzag pattern is employed in the run-length encoding scheme to exploit the number of consecutive zeros that occur in each block. The zigzag pattern progresses from low-frequency to high-frequency terms. Because the high-frequency terms are the ones most likely to be eliminated in the quantization stage, any run-length encoded block will typically contain at least one large run of zeros at the end. Thus, the amount of space required to represent each block can be substantially reduced by representing a run of zeros as (0, n), where n is the number of zeros occurring in the run. The process of run length encoding of an 8 x 8 DCT block is illustrated in Figure 1. In the fifth and final stage, the resulting data may be further compressed through a **loss-less** process of **Huffman coding**. The resulting compressed data may then be written to the computer hard drive in a file for efficient storage and transfer.

4.4 VIDEO IMAGE COMPRESSION

The development of digital video technology has made it possible to use digital video compression for a variety of telecommunications applications. Standardization of compression algorithms for video was first initiated by CCITT for teleconferencing and video telephony.MPEG (Moving Picture Experts Group) Compression is one of the widely used video image compression. Applications using MPEG standards can be symmetric or asymmetric. Symmetric applications are applications that require essentially equal use of compression and decompression. Asymmetric applications require frequent decompression.

Symmetric applications require on-line input devices such as video cameras, scanners and microphones for digitized sound. In addition to video and audio compression, this standards activity is concerned with a number of other Issues concerned with playback of video clips and sound clips. The MPEG standards consist of a number of different standards.The main profile is designed to cover the largest number of applications. It supports digital video compression in the range of 2 to 15 M bits/sec. It also provides a generic solution for television worldwide, including cable, direct broadcast satellite, fiber optic media, and optical storage media (including digital VCRs).

MPEG Coding Methodology

The above said requirements can be achieved only by incremental coding of successive frames. It is known as interface coding. If we access information randomly by frame requires coding confined to a specific frame, and then it is known as intraframe coding. The MPEG standard addresses these two requirements by providing a balance between interface coding and intraframe coding. The MPEG standard also provides for recursive and non-recursive temporal redundancy reduction.

The MPEG video compression standard provides two basic schemes: discrete-transform-based compression for the reduction of' spatial redundancy and block-based motion compensation for the reduction of temporal (motion) redundancy. During the initial stages of DCT compression, both the full motion MPEG and still image JPEG algorithms are essentially identical. First an image is converted to the YUV color space (a luminance/chrominance color space similar to that used for color television). The pixel data is then fed into a discrete cosine transform, which creates a scalar quantization (a two-dimensional array representing various frequency ranges represented in the image) of the pixel data.

Following quantization, a number of compression algorithms are applied, including run-length and Huffman encoding. For full motion video (MPEG I and 2), several more levels of block based motion-compensated techniques are applied to reduce temporal redundancy with both causal and non causal coding to further reduce spatial redundancy.

The MPEG algorithm for spatial reduction is loss and is defined as a hybrid which employs motion compensation, forward discrete cosine transform (DCF), a uniform quantize, and Huffman coding. Block-based motion compensation is *utilized* for *reducing temporal* redundancy (i.e. to reduce the amount of data needed to represent each picture in a video sequence). Motion-compensated reduction is a key feature of MPEG.

4.4.1 Moving Picture Types

Moving pictures consist of sequences of video pictures that are played back a fixed number of frames per second. To achieve the requirement of random access, a set of pictures can be defined to form a group of pictures (GOP) consisting of one or more of the following three types of pictures.

Intra pictures (I)

Unidirectional predicted pictures (U)

Bidirectional predicted pictures (B)

A GOP consists of consecutive pictures that begin with an intrapicture. The intrapicture is coded without any reference to any other picture in the group. Predicted pictures are coded with a reference to a past picture, either an intrapicture or a unidirectional predicted picture. Bidirectional predicted picture is never used as references Motion Compensation for Coding MPEG. Let us review the concept of Macro blocks and understand the role they play in compression

MACRO BLOCKS

For the video coding algorithm recommended by CCITT, CIF and QCIF are divided into a hierarchical block structure consisting of pictures, groups of blocks (GOPs), Macro Blocks (MBs), and blocks. Each picture frame is divided into 16 x 16 blocks. Each Macro block is composed of four 8 x 8 (Y) luminance blocks and two 8 x 8 (C_b and Ch) chrominance blocks. This set of six blocks, called a macro block; is the basic hierarchical component used for achieved a high level of compression.

Motion compensation

Motion compensation is the basis for most compression algorithms for visual telephony and full-motion video. Motion compensation assumes that the current picture is some translation of a previous picture. This creates the opportunity for using prediction and interpolation. Prediction requires only the current frame and the reference frame.

Based on motion vectors values generated, the prediction approach attempts to find the relative new position of the object and confirms it by comparing some block exhaustively. In the interpolation approach, the motion vectors are generated in relation to two reference frames, one from the past and the next predicted frame. The best-matching blocks in both reference frames are searched, and the average is taken as the position of the block in the current frame. The motion vectors for the two references, frames are averaged.

Picture Coding Method

In this coding method, motion compensation is applied bidirectional. In MPEG terminology, the motion-compensated units are called macro blocks (MBs).MBs are 16 x 16 blocks that contain a number of 8 x 8 luminance and chrominance blocks. Each 16 x 16 macro block can be of type intrapicture, forward-predicted, backward predicted, or average.

MPEG Encoder

Figure below shows the architecture of an MPEG encoder. It contains DCT quantize, Huffman coder and Motion compensation. These represent the key modules in the encoder.

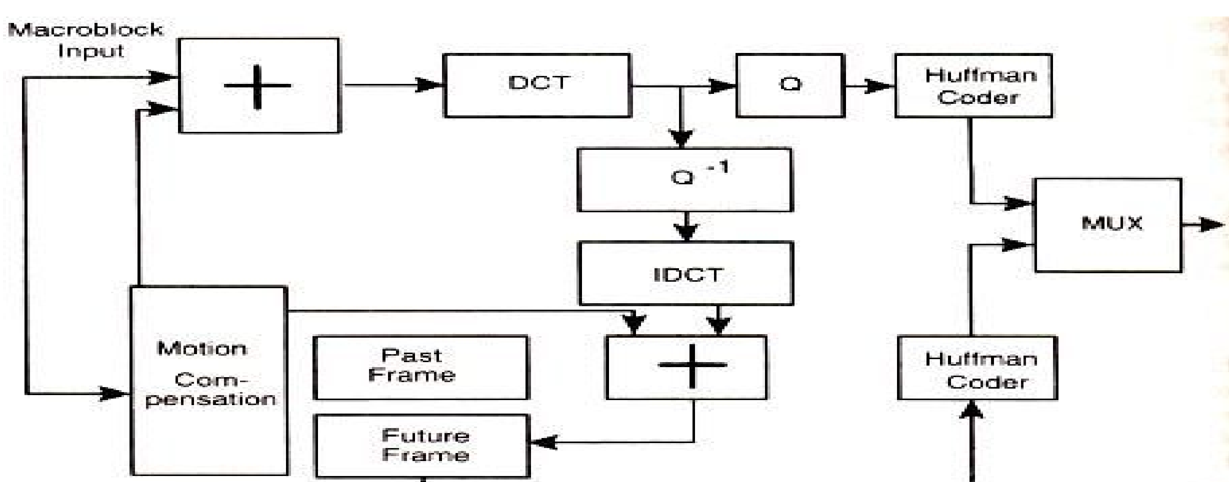


Figure 4.2 Architecture of MPEG Encoder

The Sequence of events for MPEG

- First an image is converted to the YUV color space.
- The pixel data is then fed into a DCT, which creates a scalar quantization of the pixel data.
- Following quantization, a number of compression algorithms are applied, including run-length and Huffman encoding. For full-motion video, several more levels of motion compensation compression and coding are applied.

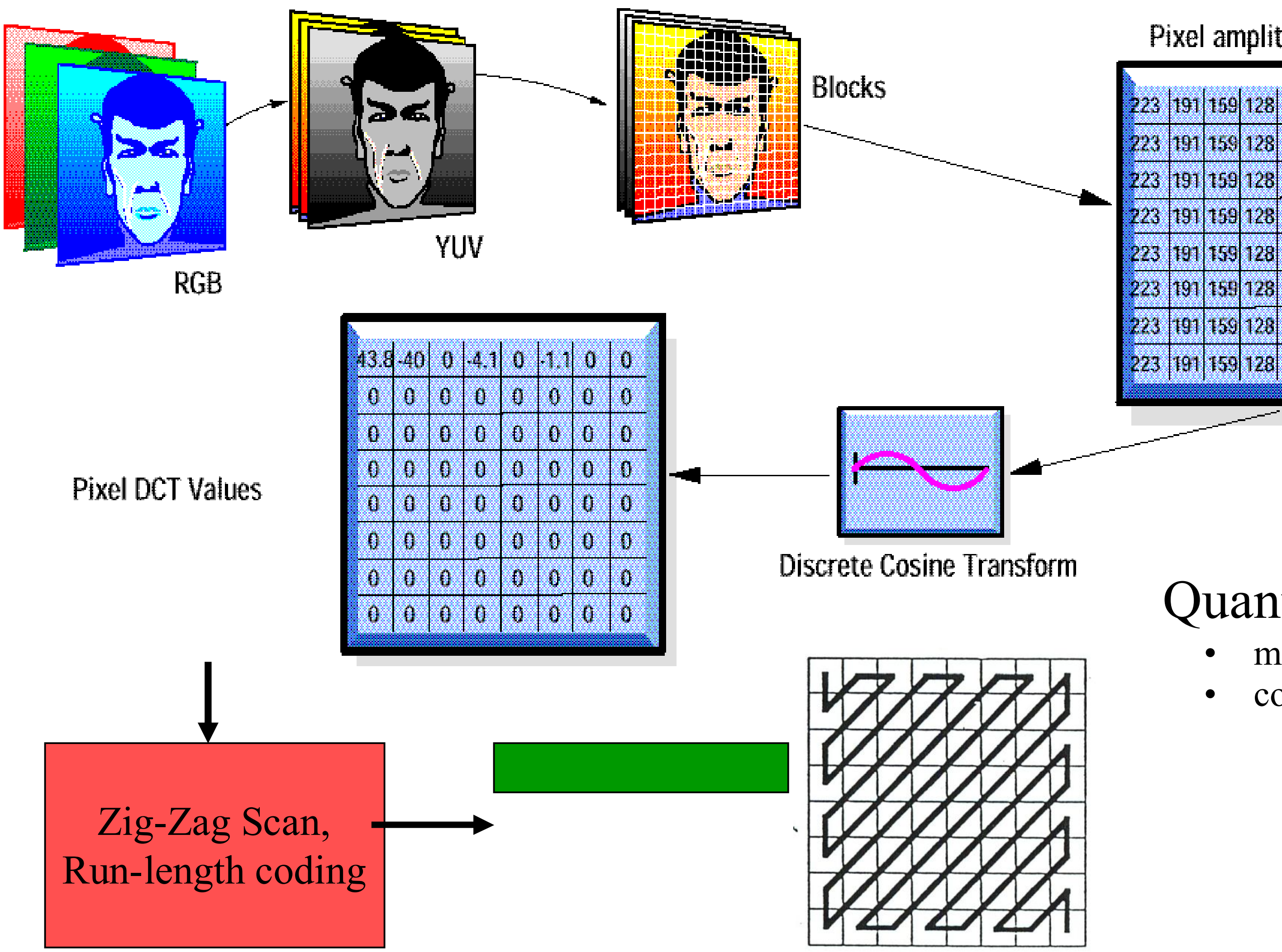


Figure 4.3. Sequence of Events for MPEG

MPEG -2

It is defined to include current television broadcasting compression and decompression needs, and attempts to include hooks for HDTV broadcasting.

The MPEG-2 Standard Supports:

1. Video Coding: * MPEG-2 profiles and levels.

2. Audio Coding:*MPEG-l audio standard for backward compatibility.

* Layer-2 audio definitions for MPEG-2 and stereo sound.
* Multichannel sound.
* Multiplexing: MPEG-2 definitions

4.5 MULTIMEDIA INPUT/OUTPUT TECHNOLOGIES

Input devices

Often, input devices are under direct control by a human user, who uses them to communicate commands or other information to be processed by the computer, which may then transmit feedback to the user through an output device. Input and output devices together make up the hardware interface between a computer and the user or external world. Typical examples of input devices include keyboards and mice. However, there are others which provide many more degrees of freedom. In general, any sensor which monitors, scans for and accepts information from the external world can be considered an input device, whether or not the information is under the direct control of a user.

4.5.1 Classification of Input Devices

Input devices can be classified according to the modality of input (e.g. mechanical motion, audio, visual, sound, etc.), the input is discrete or continuous (e.g. key presses or a mouse's position), the number of degrees of freedom involved (e.g. 2D positional input, or 3D input)

Pointing devices, which are input devices used to specify a position in space, can further be classified, is direct or indirect input devices. With direct input, the input space coincides with the display space, i.e. pointing is done in the space where visual feedback or the cursor appears. Touch screens and light pens involve direct input. Examples involving indirect input include the mouse and trackball.

4.5.2 Keyboards

A keyboard is the most common method of interaction with a computer. Keyboards provide various tactile responses (from firm to mushy) and have various layouts depending upon your computer system and keyboard model. Keyboards are typically rated for at least 50 million cycles (the number of times a key can be pressed before it might suffer breakdown). The most common keyboard for PCs is the 101 style (which provides 101 keys), although many styles are available with more are fewer special keys, LEDs, and others features, such as a plastic membrane cover for industrial or food service applications or flexible "ergonomic" styles. Macintosh keyboards connect to the Apple Desktop Bus (ADB), which manages all forms of user input from digitizing tablets to mice.

4.5.3 Pointing devices

A pointing device is any computer hardware component that allows a user to input spatial (i.e., continuous and multi-dimensional) data to a computer. CAD systems and graphical user interfaces (GUI) allow the user to control and provide data to the computer using physical gestures - point, click, and drag - typically by moving a hand-held mouse across the surface of the physical desktop and activating switches on the mouse.

While the most common pointing device by far is the mouse, many more devices have been developed. However, mouse is commonly used as a metaphor for devices that move the cursor. A mouse is the standard tool for interacting with a graphical user interface (GUI). All Macintosh computers require a mouse; on PCs, mice are not required but recommended. Even though the Windows environment accepts keyboard entry in lieu of mouse point-and-click actions, your multimedia project should typically be designed with the mouse or touch screen in mind. The buttons the mouse provide additional user input, such as pointing and double-clicking to open a document, or the click-and-drag operation, in which the mouse button is pressed and held down to drag (move) an object, or to move to and select an item on a pull-down menu, or to access context-sensitive help. The Apple mouse has one button; PC mice may have as many as three.

Examples of common pointing devices include

- mouse
- trackball
- touchpad
- space Ball - 6 degrees-of-freedom controller
- touch screen
- graphics tablets (or digitizing tablet) that use a stylus light pen

4.5.4 Scanners

A scanner may be the most useful piece of equipment used in the course of producing a multimedia project; there are flat-bed and handheld scanners. Most commonly available are gray-scale and color flat-bed scanners that provide a resolution of 300 or 600 dots per inch (dpi). Professional graphics houses may use even higher resolution units. Handheld scanners can be useful for scanning small images and columns of text, but they may prove inadequate for the multimedia development.

The scanned images, particularly those at high resolution and in color, demand an extremely large amount of storage space on the hard disk, no matter what instrument is used to do the scanning. Also remember that the final monitor display resolution for your multimedia project will probably be just 72 or 95 dpi-leave the very expensive ultra-high-resolution scanners for the desktop publishers. Most expensive flat-bed scanners offer at least 300 dpi resolution, and most scanners allow to set the scanning resolution.

Types of Scanners

1. Flatbed Scanners

The most commonly used scanner is a flatbed scanner also known as desktop scanner. It has a glass plate on which the picture or the document is placed. The scanner head placed beneath the glass plate moves across the picture and the result is a good quality scanned image. For scanning large maps or top sheets wide format flatbed scanners can be used.

2. Sheet Fed Scanners

Sheet fed scanners work on a principle similar to that of a fax machine. In this, the document to be scanned is moved past the scanning head and the digital form of the image is obtained. The disadvantage of this type of scanner is that it can only scan loose sheets and the scanned image can easily become distorted if the document is not handled properly while scanning.

3. Handheld Scanners

Hand-held scanners although portable, can only scan images up to about four inches wide. They require a very steady hand for moving the scan head over the document. They are useful for scanning small logos or signatures and are virtually of no use for scanning maps and photographs.

Overview of components

i. Glass Plate and Cover

The glass plate is the transparent plate wherein the original is placed so that the scanner can scan it and the cover keeps out stray light that can affect the accuracy of the scan

ii. Scanning head

Scanning head is the most important component because it is the one which does actual scanning. It contains components like

iii. Light source and mirror: It is the bright white light that is used to illuminate the original as it is being scanned and which bounces off the original and reflected off several mirrors

iv. Stabilizer bar: It is a long stainless steel rod that is securely fastened to the case of the scanner and it provides a smooth ride as the scanner scans down the page

v.CCD (Charge Coupled Device) or CIS (Contact Image Sensor): A CCD array is a device that converts photons into electricity. Any scanner that uses CCD use lens to focus the light coming from the mirrors within the scanning head. All scanners use charge-coupled devices as their photo sensors. CCDs consist of cells arranged in a fixed array on a small square or rectangular solid state surface. Light source moves across a document. The intensity of the light reflected by the mirror charges those cells. The amount of charge is depending upon intensity of the reflected light, which depends on the pixel shade in the document.

vi. Stepper Motor

The stepper motor in a scanner moves the scan head down the page during scan cycle and this is often located either on the scan head itself or attached to a belt to drive the scanner head.

The Scanning Process

Here are the steps that a scanner goes through when it scans a document:

- The document is placed on the glass plate and the cover is closed. The inside of the cover in most scanners is flat white, although a few are black. The cover provides a uniform background that the scanner software can use as a reference point for determining the size of the document being scanned. Most flatbed scanners allow the cover to be removed for scanning a bulky object, such as a page in a thick book.

- A lamp is used to illuminate the document. The lamp in newer scanners is either a cold cathode fluorescent lamp (CCFL) or a xenon lamp, while older scanners may have a standard fluorescent lamp.
- The entire mechanism (mirrors, lens, filter and CCD array) make up the scan head. The scan head is moved slowly across the document by a belt that is attached to a stepper motor. The scan head is attached to a stabilizer bar to ensure that there is no wobble or deviation in the pass. Pass means that the scan head has completed a single complete scan of the document.
- The image of the document is reflected by an angled mirror onto a lens. The lens focuses the image through a filter on the CCD array.
- The filter and lens arrangement vary based on the scanner. Some scanners use a three pass scanning method. Each pass uses a different color filter (red, green or blue) between the lens and CCD array. After the three passes are completed, the scanner software assembles the three filtered images into a single full-color image.

Touch screens

Touch screens are monitors that usually have a textured coating across the glass face. This coating is sensitive to pressure and registers the location of the user's finger when it touches the screen. The Touch Mate System, which has no coating, actually measures the pitch, roll, and yaw rotation of the monitor when pressed by a finger, and determines how much force was exerted and the location where the force was applied. Other touch screens use invisible beams of infrared light that crisscross the front of the monitor to calculate where a finger was pressed. Pressing twice on the screen in quick and dragging the finger, without lifting it, to another location simulates a mouse click-and-drag. A keyboard is sometimes simulated using an onscreen representation so users can input names, numbers, and other text by pressing "keys".

Touch screen recommended for day-to-day computer work, but are excellent for multimedia applications in a kiosk, at a trade show, or in a museum delivery system-anything involving public input and simple tasks. When your project is designed to use a touch screen, the monitor is the only input device required, so you can secure all other system hardware behind locked doors to prevent theft or tampering.

Digital Pen

- **Pen Driver**: It is a pen device driver that interacts with the digitizer to receive all digitized information about the pen location and builds pen packets for the recognition context manager. Recognition context manager: It is the main part of the pen system. It is responsible for co-coordinating windows pen applications with the pen. It works with Recognizer, dictionary, and display driver to recognize and display pen drawn objects.
- **Recognizer**: It recognizes hand written characters and converts them to ASCII.
- **Dictionary**: A dictionary is a dynamic link library (DLL); the windows form pen computing system uses this dictionary to validate the recognition results.
- **Display Driver**: It interacts with the graphics device interface' and display hardware. When a user starts writing or drawing, the display driver paints the ink trace on the screen.

OUTPUT DEVICES

Presentation of the audio and visual components of the multimedia project requires hardware that may or may not be included with the computer itself-speakers, amplifiers, monitors, motion video devices, and capable storage systems. The better the equipment, of course, the better the presentation. There is no greater test of the benefits of good output hardware than to feed the audio output of your computer into an external amplifier system: suddenly the bass sounds become deeper and richer, and even music sampled at low quality may seem to be acceptable.

1. Amplifiers and Speakers

Often the speakers used during a project's development will not be adequate for its presentation. Speakers with built-in amplifiers or attached to an external amplifier are important when the project will be presented to a large audience or in a noisy setting.

2. Monitors

The monitor needed for development of multimedia projects depends on the type of multimedia application created, as well as what computer is being used. A wide variety of monitors is available for both Macintoshes and PCs. High-end, large-screen graphics monitors are available for both, and they are expensive.

Serious multimedia developers will often attach more than one monitor to their computers, using add-on graphic board. This is because many authoring systems allow working with several open windows at a time, so we can dedicate one monitor to viewing the work we are creating or designing, and we can perform various editing tasks in windows on other monitors that do not block the view of your work. Editing windows that overlap a work view when developing with Macromedia's authoring environment, director, on one monitor. Developing in director is best with at least two monitors, one to view the work the other two view the "score". A third monitor is often added by director developers to display the "Cast".

4.5.7 Projectors

When it is necessary to show a material to more viewers than can huddle around a computer monitor, it will be necessary to project it on to large screen or even a white-painted wall. Cathode-ray tube (CRT) projectors, liquid crystal display (LCD) panels attached to an overhead projector, stand-alone LCD projectors, and light-valve projectors are available to splash the work on to big-screen surfaces.

CRT projectors have been around for quite a while- they are the original "big-screen" televisions. They use three separate projection tubes and lenses (red, green, and blue), and three color channels of light must “converge” accurately on the screen. Setup, focusing, and aligning are important to getting a clear and crisp picture. CRT projectors are compatible with the output of most computers as well as televisions.

LCD panels are portable devices that fit in a briefcase. The panel is placed on the glass surface of a standard overhead projector available in most schools, conference rooms, and meeting halls. While the overhead projectors do the projection work, the panel is connected to the computer and provides the image, in thousands of colors and, with active-matrix technology, at speeds that allow full-motion video and animation. Because LCD panels are small, they are popular for on-the-road presentations, often connected to a laptop computer and using a locally available overhead projector.

More complete LCD projection panels contain a projection lamp and lenses and do not recover a separate overheads projector. They typically produce an image brighter and shaper than the simple panel model, but they are somewhat large and cannot travel in a briefcase.

Light-valves complete with high-end CRT projectors and use a liquid crystal technology in which a low-intensity color image modulates a high-intensity light beam. These units are expensive, but the image from a light-valve projector is very bright and color saturated can be projected onto screen as wide as 10 meters.

4.5.8 Printers

With the advent of reasonably priced color printers, hard-copy output has entered the multimedia scene. From storyboards to presentation to production of collateral marketing material, color printers have become an important part of the multimedia development environment. Color helps clarify concepts, improve understanding and retention of information, and organize complex data. As multimedia designers already know intelligent use of colors is critical to the success of a project. Tektronix offers both solid ink and laser options, and either Phases 560 will print more than 10000 pages at a rate of 5 color pages or 14 monochrome pages per minute before requiring new toner. Epson provides lower-cost and lower-performance solutions for home and small business users; Hewlett Packard's Color LaserJet line competes with both. Most printer manufactures offer a color model-just as all computers once used monochrome monitors but are now color, all printers will became color printers.

PRINTING TECHNOLOGIES

There are various printers available namely Dot matrix, inkjet, laser print s and ink jet color. But, laser printing technology is the most common for multimedia systems. To explain this technology, let us take Hewlett Packard Laser jet-III laser printer as an example. The basic components of the laser printer are

Paper feed mechanism**,** Paper Guide, Laser Assembly, Fuser**,** Toner cartridge.

Working: The paper feed mechanism moves the paper from a paper tray through the paper path in the printer. The paper passes over a set of corona wires that induce a change in the paper. The charged paper passes over a drum coated with fine-grain carbon (toner), and the toner attaches itself to the paper as a thin film of carbon .The paper is then struck by a scanning laser beam that follows the pattern of the text on graphics to be printed. The carbon particles attach themselves to the pixels traced by the laser beam. The fuser assembly then binds the carbon particles to the paper.

Role of Software in the printing mechanism:

The software package sends information to the printer to select and control printing features. Printer drivers (files) are controlling the actual operation of the printer and allow the application

4.6 Video and Image Display Systems Technologies

There is variety of display system technologies employed for decoding compressed data for displaying. Mixing and scaling technology: For VGA screen, these technologies are used.

(i) **VGA mixing**: Images from multiple sources are mixed in the image acquisition memory.

(ii) **VGA mixing with scaling**: Scalar ICs are used to sizing and positioning of images in predefined windows.

(iii) **Dual buffered VGA mixing/Scaling**: If we provide dual buffering, the original image is prevented from loss. In this technology, a separate buffer is used to maintain the original image.

4.7 STORAGE AND RETRIVAL TECHNOLOGY

Multimedia systems require storage for large capacity objects such as video, audio and images. Another requirement is delivery of audio and video objects. Storage technologies include battery powered RAM, Nonvolatile flash, rotating magnetic disk drives, and rotating optical disk drives: Let us discuss these technologies in detail.

MAGNETIC MEDIA TECHNOLOGY

Magnetic hard disk drive storage is a mass storage medium. It has advantages of it continual reduction in the price per mega byte of high-capacity storage. It has high capacity and available in low cost. In this section let us concentrate on magnetic disk I/O subsystems most applicable to multimedia uses. Magnetic hard disk storage remains a much faster mass storage to play an important role~ in multimedia systems. It remains a much faster mass storage medium than any other mass storage medium.

Magnetic Storage Densities and Latencies

The Latency is divided into two categories: seek latency and rotational latency. Data management provides the command queuing mechanism to minimize latencies and also set-up the scatter-gather process to gather scattered data in CPU main memory.

- **Seek Latencies**: There are three seek latencies available. They are·

overlapped seek latency, Mid-transfer seek and Elevator seek.

- **Rotational Latencies**: To reduce latency, we use two methods. They are:

(i) Zero latency read/write: Zero latency reads allow transferring data immediately after the head settles. It does not wait for disk revolution to sector property.
ii) Interleaving factor: It keeps up with the data stream without skipping succors. It determines the organization of sectors.

4.8 RAID (Redundant Array of Inexpensive Disks)

RAID stands for redundant array of independent disks. The name indicates that the disk drives are independent, and are multiple in numbers. How the data is distributed between these drives depends on the RAID level used. The main advantage of RAID is the fact that, to the operating system the array of disks can be presented as a single disk.
It is an alternative to mass storage for multimedia systems that combines throughput speed and reliability improvements. RAID is an array of multiple disks. In RAID the data is spread across the drives. It achieves fault tolerance, large storage capacity and performance improvement.

Hard Disk drives are one of the most complex devices that are attached to a computer system (or a server machine). The complexity of a hard disk is due to the fact that it is a mechanical device, used for storage. Most of the internal parts of a disk drive are moving parts that moves its head to fetch data for the user.
Due to this movement there is a high chance of failure of the disk drive. Advancement in disk drives has resulted in removing the mechanical parts to make a solid state drive, normally called as an SSD. However there are yet some shortcomings in SSD drives, due to which, we cannot completely replace mechanical disk drives. Using a single hard disk for a server machine is not at all advisable because that will be a single point of failure (a heavy

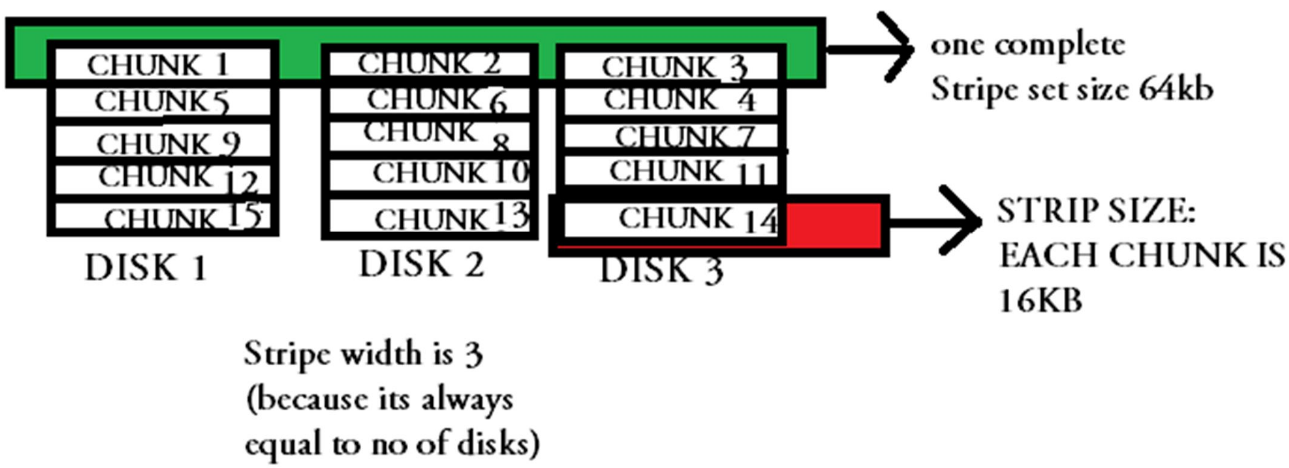
risk of data loss). Raid addresses most of these problems, because it's fast, fault tolerant, and a high performing solution.

Let's understand different terminologies that are used in RAID before getting inside different levels of RAID.

STRIPPING IN RAID

Writing data on a single disk is slower, but writing data by spreading it on multiple disks is faster (because data is written in small chunks to different disks, and also fetched in small chunks by different disks).When data is fetched from different disks, the CPU does not have to wait, because the throughput will be a combined one of all the disks.

Each and every disk drive is partitioned in small chunks (ranges from 4kb to 512kb sometimes).

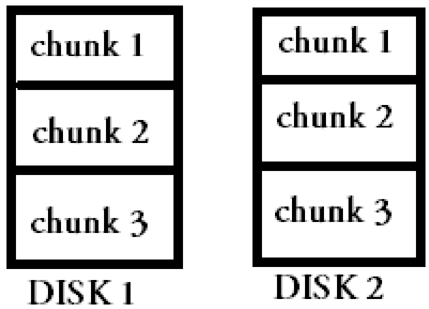


In the above shown example diagram you can see that all three disks contain different data (data that needs to be stored will be striped in chunks and will be spreader across different disks).

There is no redundancy in this method, but is better known for high performance.

Mirroring in Raid:

Mirroring is a mechanism in which the same data is written to another disk drive. The main advantage of mirroring (multiple sets of same data on two disks), is that it provides 100 percent redundancy.Suppose there are two drives in mirroring mode, then both of them will contain an exact same copy of data. So even if one disk fails, the data is safe on the other.



Parity in Raid

Parity is an interesting method used to rebuild data in case of failure of one of the disks. Although it's interesting to understand, how parity works, you will find less documentation about it on the internet. Parity makes use of a very famous mathematical binary operation called as "**XOR**".

XOR is a mathematical operation that's done to produce one output from two inputs. Some examples of XOR operations are as below.

1'st operator	2'nd operator	XOR OUPUT
1	1	0
1	0	1
0	1	1
0	0	0

You can simply make a rule while performing XOR binary operation, that if there is a difference in the operator then the XOR output is 1.

In the above shown example table consider the columns "**1'st operator**" and "**2nd operator**" as hard disks in a RAID array, and the third column "**XOR OUPUT**" as a parity disk.

And now if one of the disks fails, you can easily construct the data on the failed disk with the help of the parity disk and the other disk which is not failed.

Parity in raid can be of two types.

I.Dedicated Parity (XOR of data bits on a dedicated parity disk)

Ii.Distributed Parity (XOR of data bits distributed across all data disks)

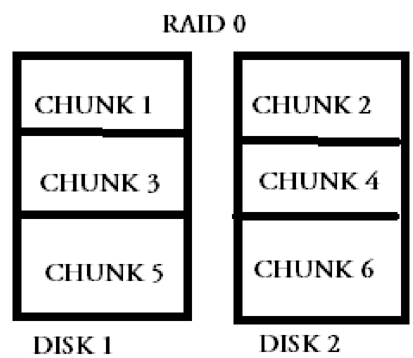
Hot spares in Raid

Hot Spare is an extra drive added to the disk array, to increase the fault tolerance. If you have a hot spare in your Raid disk array, the Raid controller will automatically start rebuilding data on to that hot spare drive, if one of the disks from the array fails.

Which means the hot spare will automatically take the role of the failed drive once data rebuilding is complete. You can later on replace your failed drive. RAID management software's will provide you with a mechanism to specify the hot spare drive for your array.

Hot swap in Raid

Hot swapping is a term used to describe the ability to replace a failed disk drive without rebooting the machine. In other words, hot swapping enabled you to replace a component without interrupting the normal operation of a server machine.



Different Levels of Raid

RAID (redundant array of independent disks), can be classified to different levels based on its operation and level of redundancy provided. There is no "One size fits all" solution as far as raid levels are concerned. Selecting the suitable raid level for your application depends on the following things.

Raid type is selected based on the performance that it provides, based on the level of redundancy it provides and based on read and writes operations.

Let's discuss some of the widely used raid levels.

I) RAID 0 or No RAID

If performance is given priority, then raid 0 fits right. An important fact that should be kept in mind is that, RAID 0 does not provide any kind of redundancy which means even if one drive fails the data is at risk.

It is simply striping done on your disk array. Data is broken into smaller chunks and are spread across the number of disks you have. It has no mirroring, no parity (which means no redundancy)

In fact raid level 0 is not RAID, because raid was primarily build for redundancy, and raid 0 does not provide any kind of redundancy, although it provides high performance.

II) Raid level 1(RAID 1)

RAID 1 implements heavy use of mirroring. All data in the drive is duplicated to another drive. It can be used in a situation where fault tolerance is of primary importance.

Maximum number of drives in RAID 1 can be 32, from a starting number of 2(even number of disks are required.)Striping and parity are not used in RAID 1.

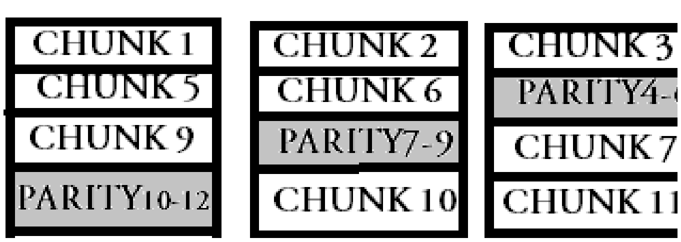
III) Raid Level 5(RAID 5)

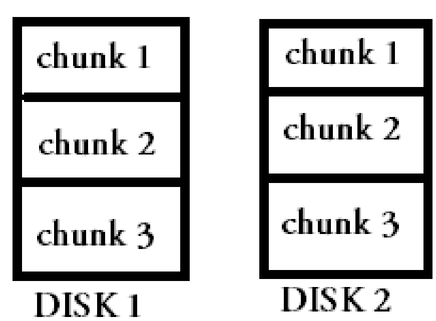
RAID level 5 uses striping, so data is spread across number of disks used in the array, and also provides redundancy with the help of parity.

RAID 5 is a best cost effective solution for both performance and redundancy. Striped method of storing data always improves performance, and parity used in this level of raid is distributed parity.

Minimum number of disks required for raid 5 is 3, and maximum can go up to 32(depending on the RAID controller used.)

One important fact to note is that, reading rate in raid 5 is much better than writing. This is because reading can be done, by a combined rate of all disks used.





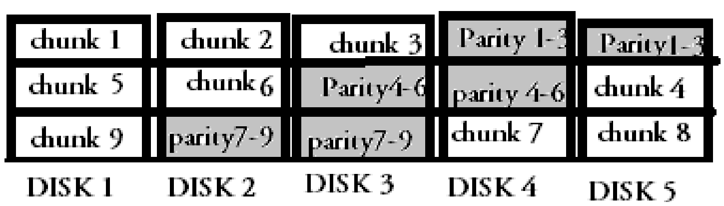
RAID 5 does a distributed parity so it can survive one disk failure. RAID 5 does a double distributed parity so it can service two disk failures.

It is noticed from the above picture that depicts distributed parity, if suppose one disk is failed, you can build the data in it with the parity data of other disk. Which means one stripe on one of the disk, while storing data will be used for parity of the data in other disks. So if a disk fails, data can be reconstructed from other disk parity, and parity blocks can also be reconstructed from other disk's data.

IV) Raid level 6 (RAID 6)

Raid level 6 is very much similar to raid level 5, but it has got one more added advantage.

The added advantage is that it can sustain 2 drive failures instead of 1. This is achieved again with the help of parity. In raid level 6, double distributed parity is used to achieve this level of redundancy.



You can Cleary see in the above diagram; each and every stripe set contains two parity on multiple disks.

So even if two disks get failed at one time, data can still be recreated.

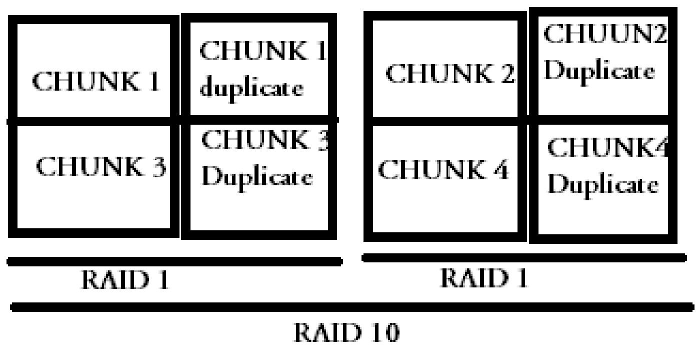
Performance of raid 6 is very much similar to performance of raid 5, it's much well suited for reads that writes.

V) Raid 10 (Combination of Raid 0 and Raid 1)

Raid 10 is a good solution that will give you both the performance advantage of raid 0 and also the redundancy of raid 1 mirroring.

Raid 10 was made by a combination of raid 0 and raid 1. And hence you get qualities of both the raid levels.

Let's understand how data is stored in raid 10 array.



If you see the above diagram data is redundant with a duplicate set in raid 1, and also is stripped across multiple raid 1 groups to achieve performance.

This is best suitable for heavy IO usage & also provides 100 percent redundancy. Minimum number of drives required is 4. It is quite expensive, as you can clearly see that you are dedicating one disk per raid 1 array for redundancy.

But is an excellent choice for both performance and redundancy.

Summary of different RAID levels

RAID 0 uses striping for high performance. Raid 0 cannot be considered as RAID as it does not provide fault tolerance.

RAID 1 uses mirroring for redundancy.

RAID 5 uses striping as well as parity for redundancy. It is well suited for heavy read and low writes operations.

RAID 6 uses striping and double parity for redundancy.

RAID 10 is a combination of raid 1 and raid 0. It also provides heavy redundancy because of mirroring, and also provides performance as the data is striped across multiple raid 1 groups.

4.9 OPTICAL MEDIA

Optical storage devices have become the order of the day. The high storage capacity available in the optical storage devices has influenced it as storage for multimedia content. Apart from the high storage capacity the optical storage devices have higher data transfer rate.

CD-ROM

A Compact Disc or CD is an optical disc used to store digital data, originally developed for storing digital audio. The CD, available on the market since late 1982, remains the standard playback medium for commercial audio recordings to the present day, though it has lost ground in recent years to MP3 players.

An audio CD consists of one or more stereo tracks stored using 16-bit PCM coding at a sampling rate of 44.1 kHz. Standard CDs have a diameter of 120 mm and can hold approximately 80 minutes of audio. There are also 80 mm discs, sometimes used for CD singles, which hold approximately 20 minutes of audio. The technology was later adapted for use as a data storage device, known as a CD-ROM, and to include record-once and re-writable media (CD-R and CD-RW respectively).

Physical details of CD-ROM

A Compact Disc is made from a 1.2 mm thick disc of almost pure polycarbonate plastic and weighs approximately 16 grams. A thin layer of aluminum (or, more rarely, gold, used for its longevity, such as in some limited-edition audiophile CDs) is applied to the surface to make it reflective, and is protected by a film of lacquer. CD data is stored as a series of tiny indentations (*pits*), encoded in a tightly packed spiral track molded into the top of the polycarbonate layer. The areas between pits are known as "lands". Each pit is approximately 100 nm deep by 500 nm wide, and varies from 850 nm to 3.5 µm in length.

The spacing between the tracks, the pitch, is 1.6 µm. A CD is read by focusing a 780 nm wavelength semiconductor laser through the bottom of the polycarbonate layer.

While CDs are significantly more durable than earlier audio formats, they are susceptible to damage from daily usage and environmental factors. Pits are much closer to the label side of a disc, so that defects and dirt on the clear side can be out of focus during playback. Discs consequently suffer more damage because of defects such as scratches on the label side, whereas clear-side scratches can be repaired by refilling them with plastic of similar index of refraction, or by careful polishing.

Types of CD

Audio CD

The logical format of an audio CD (officially Compact Disc Digital Audio or CD-DA) is described in a document produced in 1980 by the format's joint creators, Sony and Philips. The document is known colloquially as the "Red Book" after the color of its cover. The format is a two-channel 16-bit PCM encoding at a 44.1 kHz sampling rate. Four-channel sound is an allowed option within the Red Book format, but has never been implemented.

CD-Text

CD-Text is an extension of the Red Book specification for audio CD that allows for storage of additional text information (e.g., album name, song name, and artist) on a standards-compliant audio CD. The information is stored either in the lead-in area of the CD, where there is roughly five kilobytes of space available, or in the sub code channels R to W on the disc, which can store about 31 megabytes.

CD + Graphics

Compact Disc + Graphics (CD+G) are a special audio compact disc that contains graphics data in addition to the audio data on the disc. The disc can be played on a regular audio CD player, but when played on a special CD+G player, can output a graphics signal (typically, the CD+G player is hooked up to a television set or a computer monitor); these graphics are almost exclusively used to display lyrics on a television set for karaoke performers to sing along with.

CD + Extended Graphics

Compact Disc + Extended Graphics (CD+EG, also known as CD+XG) is an improved variant of the *Compact Disc + Graphics* (CD+G) format. Like CD+G, CD+EG utilizes basic CD-ROM features to display text and video information in addition to the music being played. This extra data is stored in sub code channels R-W.

CD-MIDI

Compact Disc MIDI or CD-MIDI is a type of audio CD where sound is recorded in MIDI format, rather than the PCM format of Red Book audio CD. This provides much greater capacity in terms of playback duration, but MIDI playback is typically less realistic than PCM playback.

Video CD

Video CD (aka VCD, View CD, Compact Disc digital video) is a standard digital format for storing video on a Compact Disc. VCDs are playable in dedicated VCD players, most modern DVD-Video players, and some video game consoles.

Recordable CD

Recordable compact discs, CD-Rs, are injection molded with a "blank" data spiral. A photosensitive dye is then applied, after which the discs are metalized and lacquer coated. The write laser of the CD recorder changes the color of the dye to allow the read laser of a standard CD player to see the data as it would an injection molded compact disc. The resulting discs can be read by *most* (but not all) CD-ROM drives and played in *most* (but not all) audio CD players.

CD-R recordings are designed to be permanent. Over time the dye's physical characteristics may change, however, causing read errors and data loss until the reading device cannot recover with error correction methods. The design life is from 20 to 100 years depending on the quality of the discs, the quality of the writing drive, and storage conditions. However, testing has demonstrated such degradation of *some* discs in as little as 18 months under normal storage conditions. This process is known as CD rot.

Rewritable CD

CD-RW is a re-recordable medium that uses a metallic alloy instead of a dye. The write laser in this case is used to heat and alter the properties (amorphous vs. crystalline) of the alloy, and hence change its reflectivity. A CD-RW does not have as great a difference in reflectivity as a pressed CD or a CD-R, and so many earlier CD audio players *cannot* read discs, although *later* CD audio players and stand-alone DVD players can.

4.10 DVD

DVD (also known as "**Digital Versatile Disc**" or "**Digital Video Disc**") is a popular optical disc storage media format. Its main uses are video and data storage. Most DVDs are of the same dimensions as compact discs (CDs) but store more than 6 times the data.

Variations of the term DVD often describe the way data is stored on the discs: DVD-ROM has data which can only be read and not written, DVD-R can be written once and then functions as a DVD-ROM, and DVD-RAM or DVD-RW holds data that can be re-written multiple times.

DVD-Video and DVD-Audio discs respectively refer to properly formatted and structured video and audio content. Other types of DVD discs, including those with video content, may be referred to as DVD-Data discs. The term "DVD" is commonly misused to refer to high density optical disc formats in general, such as Blue-ray and HD DVD.

"DVD" was originally used as an initialize for the unofficial term "digital video disc". It was reported in 1995, at the time of the specification finalization, that the letters officially stood for "digital versatile disc" (due to non-video applications), however, the text of the press release announcing the specification finalization only refers to the technology as "DVD", making no mention of what (if anything) the letters stood for. Usage in the present day varies, with "DVD", "Digital Video Disc", and "Digital Versatile Disc" all being common.

DVD-Video is a standard for storing video content on DVD media. Though many resolutions and formats are supported, most consumer DVD-Video discs use either 4:3 or anamorphic 16:9 aspect ratio MPEG-2 video, stored at a resolution Of 720×480 (NTSC) or 720×576 (PAL) at 24, 30, or 60 FPS. Audio is commonly stored using the Dolby Digital (AC-3) or Digital Theater System (DTS) formats, ranging from 16-bits/48kHz to 24-bits/96kHz format with monaural to 7.1 channel "Surround Sound" presentation, and/or MPEG-1 Layer 2. Although the specifications for video and audio requirements vary by global region and television system, many DVD players support all possible formats. DVD-Video also supports features like menus, selectable subtitles, multiple camera angles, and multiple audio tracks.

DVD-Audio is a format for delivering high-fidelity audio content on a DVD. It offers many channel configuration options (from mono to 7.1 surround sounds) at various sampling frequencies (up to 24-bits/192kHz versus CDDA's 16-bits/44.1kHz). Compared with the CD format, the much higher capacity DVD format enables the inclusion of considerably more music (with respect to total running time and quantity of songs) and/or far higher audio quality (reflected by higher linear sampling rates and higher vertical bit-rates, and/or additional channels for spatial sound reproduction).

DVD uses 650 nm wavelength laser diode light as opposed to 780 nm for CD. This permits a smaller spot on the media surface that is 1.32 µm for DVD while it was 2.11 µm for CD.

Writing speeds for DVD were 1x, that is 1350 kobo/s (1318 Kid/s), in first drives and media models. More recent models at 18 xs or 20 xs have 18 or 20 times that speed. Note that for CD drives, 1x means 153.6 kobo/s (150 Kid/s), 9 times slower. DVD FAQ

UNIT V HYPERMEDIA

Digital voice and audio – Video image and animation – Full motion video – Hypermedia messaging - Mobile Messaging – Hypermedia message component – Creating Hypermedia message – Integrated multimedia message standards

5.1Introduction

Sound is perhaps the most important element of multimedia. It is meaningful "speech" in any language, from a whisper to a scream. It can provide the listening pleasure of music, the startling accent of special effects or the ambience of a mood-setting background. Sound is the terminology used in the analog form, and the digitized form of sound is called as audio.

When something vibrates in the air is moving back and forth it creates wave of pressure. These waves spread like ripples from pebble tossed into a still pool and when it reaches the eardrums, the change of pressure or vibration is experienced as sound. Acoustics is the branch of physics that studies sound. Sound pressure levels are measured in decibels (db); a decibel measurement is actually the ratio between a chosen reference point on a logarithmic scale and the level that is actually experienced.

5.2 DIGITAL VOICE AND AUDIO

Sound is made up of continuous analog sine waves that tend to repeat depending on the music or voice. The analog waveforms are converted into digital format by analog-to-digital converter (ADC) using sampling process

Digital Audio

Digital audio is created when a sound wave is converted into numbers – a process referred to as digitizing. It is possible to digitize sound from a microphone, a synthesizer, existing tape recordings, live radio and television broadcasts, and popular CDs. You can digitize sounds from a natural source or prerecorded.

Digitized sound is sampled sound. Ever n^{th} fraction of a second, a sample of sound is taken and stored as digital information in bits and bytes. The quality of this digital recording depends upon how often the samples are taken.

Formula for determining the size of the digital audio

Monophonic = Sampling rate * duration of recording in seconds * (bit resolution / 8) * 1

Stereo = Sampling rate * duration of recording in seconds * (bit resolution / 8) * 2

- The sampling rate is how often the samples are taken.
- The sample size is the amount of information stored. This is called as bit resolution.
- The number of channels is 2 for stereo and 1 for monophonic.

- The time span of the recording is measured in seconds.

Editing Digital Recordings

Once a recording has been made, it will almost certainly need to be edited. The basic sound editing operations that most multimedia procedures needed are described in the paragraphs that follow

1. **Multiple Tasks**: Able to edit and combine multiple tracks and then merge the tracks and export them in a final mix to a single audio file.
2. **Trimming**: Removing dead air or blank space from the front of a recording and an unnecessary extra time off the end is your first sound editing task.
3. **Splicing and Assembly**: Using the same tools mentioned for trimming, you will probably want to remove the extraneous noises that inevitably creep into recording.
4. **Volume Adjustments**: If you are trying to assemble ten different recordings into a single track there is a little chance that all the segments have the same volume.
5. **Format Conversion**: In some cases your digital audio editing software might read a format different from that read by your presentation or authoring program.
6. **Resampling or down sampling**: If you have recorded and edited your sounds at 16 bit sampling rates but are using lower rates you must resample or down sample the file.
7. **Equalization:** Some programs offer digital equalization capabilities that allow you to modify recording frequency content so that it sounds brighter or darker.
8. **Digital Signal Processing**: Some programs allow you to process the signal with reverberation, multitap delay, and other special effects using DSP routines.
9. **Reversing Sounds**: Another simple manipulation is to reverse all or a portion of a digital audio recording.
10. **Time Stretching**: Advanced programs let you alter the length of a sound file without changing its pitch. This feature can be very useful but watch out:

most time stretching algorithms will severely degrade the audio quality.

Digital Voice

Speech is analog in nature and is converted to digital form by an analog-to-digital converter (ADC). An ADC takes an input signal from a microphone and converts the amplitude of the sampled analog signal to an 8, 16 or 32 bit digital value.

5.3 VOICE Recognition System

Voice Recognition Systems can be classified into three types.

1. Isolated-word Speech Recognition.
2. Connected-word Speech Recognition.
3. Continuous Speech Recognition.

1. Isolated-word Speech Recognition.

It provides recognition of a single word at a time. The user must separate every word by a pause. The pause marks the end of one word and the beginning of the next word.

Stage 1: Normalization

The recognizer's first task is to carry out amplitude and noise normalization to minimize the variation in speech due to ambient noise, the speaker's voice, the speaker's distance from and position relative to the microphone, and the speaker's breath noise.

Stage2: Parametric Analysis

It is a preprocessing stage that extracts relevant time-varying sequences of speech parameters. This stage serves two purposes: (i) It extracts time-varying speech parameters. (ii) It reduces the amount of data of extracting the relevant speech parameters.

Training mode: In training mode of the recognizer, the new frames are added to the reference list. **Recognizer mode:** If the recognizer is in Recognizer mode, then dynamic time warping is applied to the unknown patterns to average out the phoneme (smallest distinguishable sound, and spoken words are constructed by concatenate basic phonemes) time duration. The unknown pattern is then compared with the reference patterns.A speaker independent isolated word recognizer can be achieved by grouping a large number of samples corresponding to a word into a single cluster.

2. Connected-Word Speech Recognition: Connected-word speech consists of spoken phrase consisting of a sequence of words. It may not contain long pauses between words.

The method using Word spotting technique

It recognizes words in a connected-word phrase. In this technique, Recognition is carried out by compensating for rate of speech variations by the process called dynamic time warping (this process is used to expand or compress the time duration of the word), and sliding the adjusted connected-word phrase representation in time past a stored word template for a likely match.

3. Continuous Speech Recognition

This system can be divided into three sections:

(i) A section consisting of digitization, amplitude normalization, time normalization and parametric representation.

(ii) Second section consisting of segmentation and labeling of the speech segment into a symbolic string based on a knowledge based or rule-based systems.

(iii) The final section is to match speech segments to recognize word sequences.

Voice Recognition Applications

Database Input and Query Applications

A number of applications are developed around the voice recognition and voice synthesis function. The following lists a few applications which use Voice recognition.

- The voice-mail message can be integrated with e-mail messages to create an integrated message.
- Application such as order entry and tracking

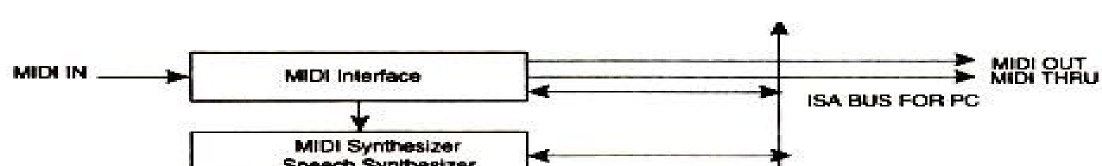
It is a server function; it is centralized; Remote users can dial into the system to enter an order or to track the order by making a Voice query.

- Voice-activated rolodex or address book

When a user speaks the name of the person, the rolodex application searches the name and address and voice-synthesizes the name, address, telephone numbers and fax numbers of a selected person. In medical emergency, ambulance technicians can dial in and register patients by speaking into the hospital's centralized system.

- Police can make a voice query through central database to take follow-up action to catch any suspect.
- Language-teaching systems are an obvious use for this technology. The system can ask the student to spell or speak a word. When the student speaks or spells the word, the systems performs voice recognition and measures the student's ability to spell. Based on the student's ability, the system can adjust the level of the course. This creates a self-adjustable learning system to follow the individual's pace.
- Foreign language learning is another good application where an individual student can input words and sentences in the system. The system can then correct for pronunciation or grammar.

5.4 SOUND BOARD ARCHITECTURE: A sound card consists of the following components:MIDI Input/output Circuitry, MIDI Synthesizer Chip, input mixture circuitry to mix CD audio input with LINE IN input and microphone input, analog-to-digital converter with a pulse code modulation circuit to convert analog signals to digital to create WAV files, a decompression and compression chip to compress and decompress audio files, a speech synthesizer to synthesize speech output, a speech recognition circuitry to recognize speech input and output circuitry to output stereo audio OUT or LINEOUT. The audio mixer component of the sound card typically has external inputs for stereo CD audio, stereo LINE IN, and stereo microphone MICIN. These are analog inputs, and they go through analog-to-digital conversion in conjunction with PCM or ADPCM to generate digitized



samples.

Figure 5.1 Sound Board Architecture:

Analog-to-Digital Converters: The ADC gets its input from the audio mixer and converts the amplitude of a sampled analog signal to either an 8-bit or 16-bit digital value.

Digital-to-Analog Converter (DAC): A DAC converts digital input in the form of W AVE files, MIDI output and CD audio to analog output signals.

Sound Compression and Decompression: Most sound boards include a codec for sound compression and decompression. ADPCM for windows provides algorithms for sound compression.

CD-ROM Interface: The CD-ROM interface allows connecting the CD ROM drive to the sound board.

5.5 VIDEO IMAGES AND ANIMATION

A video frame grabber is used to capture, manipulate and enhance video images. A video frame grabber card consists of video channel multiplexer, Video ADC, Input look-up table with arithmetic logic unit, image frame buffer, compression-decompression circuitry, output color look-up table, video DAC and synchronizing circuitry.

Components used in Video Board Architectures:

Video Channel Multiplexer:

A video channel multiplexer has multiple inputs for different video inputs. The video channel multiplexer allows the video channel to be selected under program control and switches to the control circuitry appropriate for the selected channel in a TV with multi – system inputs. Analog to Digital Converter: The ADC takes inputs from video multiplexer and converts the amplitude of a sampled analog signal to either an 8-bit digital value for monochrome or a 24 bit digital value for color.

Input lookup table: The input lookup table along with the arithmetic logic unit (ALU) allows performing image processing functions on a pixel basis and an image frame basis. The pixel image-processing functions ate histogram stretching or histogram shrinking for image brightness and contrast, and histogram sliding to brighten or darken the image. The frame-basis image-processing functions perform logical and arithmetic operations.

Image Frame Buffer Memory: The image frame buffer is organized as an l024 x 1024 x 24 storage buffer to store image for image processing and display.

Video Compression-Decompression: The video compression decompression Processor is used to compress and decompress still image data and video data.

Frame Buffer Output Lookup Table: The frame buffer data represents the pixel data and is used to index into the output look uptable. The output lookup table generates either an 8 bit pixel value for monochrome or a 24 bit pixel value for color.

SVGA Interface: This is an optional interface for the frame grabber. The frame grabber can be designed to include an SVGA frame buffer with its own output lookup table and digital-to-analog converter.

Analog Output Mixer: The output from the SVGA DAC and the output from image frame buffer DAC are mixed to generate overlay output signals. The primary components involved include the display image frame buffer and the display SVGA buffer. The display SVGA frame buffer is overlaid on the image frame buffer or live video; this allows SVGA to display live video.

Video and Still Image Processing

Video image processing is defined as the process of manipulating a bit map image so that the image can be enhanced, restored, distorted, or analyzed.

Let us discuss about some of the terms using in video and still image processing.

Pixel point to point processing: In pixel point-to-point processing, operations are carried out on individual pixels one at a time.

Histogram Sliding: It is used to change the overall visible effect of brightening or darkening of the image. Histogram sliding is implemented by modifying the input look-up table values and using the input lookup table in conjunction with arithmetic logic unit.

Histogram Stretching and Shrinking: It is to increase or decrease the contrast.

In histogram shrinking, the brighter pixels are made less bright and the darker pixels are made less dark.

Pixel Threshold: Setting pixel threshold levels set a limit on the bright or dark areas of a picture. Pixel threshold setting is also achieved through the input lookup table.

Inter- frame image processing: Inter- frame image processing is the same as point-to-point image processing, except that the image processor operates on two images at the same time. The equation of the image operations is as follows:

Pixel output (x, y) = (Image l(x, y) Operator (Image 2(x, y)

Image Averaging: Image averaging minimizes or cancels the effects of random noise.

Image Subtraction: Image subtraction is used to determine the change from one frame to the next .for image comparisons for key frame detection or motion detection.

Logical Image Operation: Logical image processing operations are useful for comparing image frames and masking a block in an image frame.

***Spatial Filter Processing*:** The rate of change of shades of gray or colors is called spatial frequency. The process of generating images with either low-spatial frequency-components or high frequency components is called spatial filter processing. A low pass filter causes blurring of the image and appears to cause a reduction in noise. The high-pass filter causes edges to be emphasized. The high-pass filter attenuates low-spatial frequency components, thereby enhancing edges and sharpening the image. The Laplacian Filter sharply attenuates low-spatial-frequency components without affecting and high-spatial frequency components, thereby enhancing edges sharply.

Frame Processing: Frame processing operations are most commonly for geometric operations, image transformation, and image data compression and decompression Frame processing operations are very compute intensive many multiply and add operations, similar to spatial filter convolution operations.

Image scaling: Image scaling allows enlarging or shrinking the whole or part of an image.

Image rotation*:* Image rotation allows the image to be rotated about a center point. The operation can be used to rotate the image orthogonally to reorient the image if it was scanned incorrectly. The operation can also be used for animation. The rotation formula is:

Pixel output-(x, *y)* = pixel input *(x,* cos Q + *y* sin Q, - *x* sin + Y cos Q) where, Q is the orientation angle

x, yare the spatial co-ordinates of the original pixel.

Image translation*:* Image translation allows the image to be moved up and down or side to side. Again, this function can be used for animation.

The translation formula is:

Pixel output *(x, y)* =Pixel Input *(x + Tx, y + Ty)* where *Tx* and *Ty* are the horizontal and vertical coordinates. *x,* yare the spatial coordinates of the original pixel.

Image transformation: An image contains varying degrees of brightness or colors defined by the spatial frequency. The image can be transformed from spatial domain to the frequency domain by using frequency transform.

Video and Image Animation Techniques

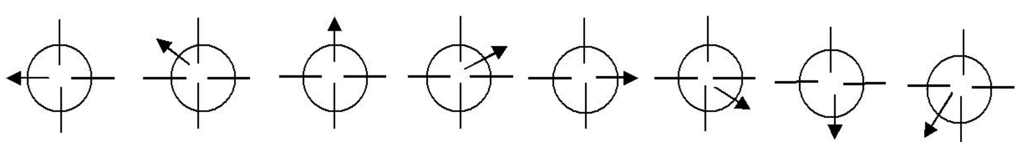
ANIMATION

Animation makes static presentations come alive. It is visual change over time and can add great power to our multimedia projects. Carefully planned, well-executed video clips can make a dramatic difference in a multimedia project. Animation is created from drawn pictures and video is created using real time visuals.

Principles of Animation

Animation is the rapid display of a sequence of images of 2-D artwork or model positions in order to create an illusion of movement. It is an optical illusion of motion due to the phenomenon of persistence of vision, and can be created and demonstrated in a number of ways. The most common method of presenting animation is as a motion picture or video program, although several other forms of presenting animation also exist

Animation is possible because of a biological phenomenon known as *persistence of vision* and a psychological phenomenon called *phi.* An object seen by the human eye remains chemically mapped on the eye's retina for a brief time after viewing. Combined with the human mind's need to conceptually complete a perceived action, this makes it possible for a series of images that are changed very slightly and very rapidly, one after the other, to seemingly blend together into a visual illusion of movement. The following shows a few cells or frames of a rotating logo. When the images are progressively and rapidly changed, the arrow of the compass is perceived to be spinning.



Television video builds entire frames or pictures every second; the speed with which each frame is replaced by the next one makes the images appear to blend smoothly into movement. To make an object travel across the screen while it changes its shape, just change the shape and also move or *translate* it a few pixels for each frame.

Animation Techniques

When you create an animation, organize its execution into a series of logical steps. First, gather up in your mind all the activities you wish to provide in the animation; if it is complicated, you may wish to create a written script with a list of activities and required objects. Choose the animation tool best suited for the job. Then build and tweak your sequences; experiment with lighting effects. Allow plenty of time for this phase when you are experimenting and testing. Finally, post-process your animation, doing any special rendering and adding sound effects.

1) Cel Animation

The term *cel* derives from the clear celluloid sheets that were used for drawing each frame, which have been replaced today by acetate or plastic. Cels of famous animated cartoons have become sought-after, suitable-for-framing collector's items.

Cel animation artwork begins with *key frames* (the first and last frame of an action). For example, when an animated figure of a man walks across the screen, he balances the weight of his entire body on one foot and then the other in a series of falls and recoveries, with the opposite foot and leg catching up to support the body.

- The animation techniques made famous by Disney use a series of progressively different on each frame of movie film which plays at 24 frames per second.
- A minute of animation may thus require as many as 1,440 separate frames.
- The term cel derives from the clear celluloid sheets that were used for drawing each frame, which is been replaced today by acetate or plastic.
- Cel animation artwork begins with key frames.

2) Computer Animation

Computer animation programs typically employ the same logic and procedural concepts as cel animation, using layer, key frame, and tweening techniques, and even borrowing from the vocabulary of classic animators. On the computer, paint is most often filled or drawn with tools using features such as gradients and anti-aliasing. The word *links*, in computer animation terminology, usually means special methods for computing RGB pixel values, providing edge detection, and layering so that images can blend or otherwise mix their colors to produce special transparencies, inversions, and effects.Computer Animation is same as that of the logic and procedural concepts as cel animation and use the vocabulary of classic cel animation – terms such as layer, Key frame, and tweening. The primary difference between the animation software programs is in how much must be drawn by the animator and how much is automatically generated by the software

In 2D animation the animator creates an object and describes a path for the object to follow. The software takes over, actually creating the animation on the fly as the program is being viewed by your user. In 3D animation the animator puts his effort in creating the models of individual and designing the characteristic of their shapes and surfaces. Paint is most often filled or drawn with tools using features such as gradients and anti- aliasing.

3) Kinematics

It is the study of the movement and motion of structures that have joints, such as a walking man. Inverse Kinematics is in high-end 3D programs, it is the process by which you link objects such as hands to arms and define their relationships and limits. Once those relationships are set you can drag these parts around and let the computer calculate the result.

4) Morphing

Morphing is popular effect in which one image transforms into another. Morphing application and other modeling tools that offer this effect can perform transition not only between still images but often between moving images as well.

- The morphed images were built at a rate of 8 frames per second, with each transition taking a total of 4 seconds.

- Some product that uses the morphing features are as follows
  - Black Belt's Easy Morph and Win Images,
  - Human Software's Squizz
  - Valis Group's Flo, MetaFlo, and MovieFlo.

5.6 FULL MOTION VIDEO

Most modem cameras use a CCD for capturing the image. HDTV video cameras will be all-digital, and the capture method will be significantly different based on the new NTSC HDTV Standard.

Full-Motion Video Controller Requirements

A full-motion video capture board is a circuit card in the computer that consists of the following components:

(i) Video input to accept video input signals.

(ii) S- Video input to accept RS 170 input.

(iii) Video compression-decompression processor to handle different video compression-decompression algorithms for video data.

(iv) Audio compression-decompression processor to compress and decompress audio data.

(v) Analog to digital converter.

(vi) Digital to analog converter.

(vii) Audio input for stereo audio LINE IN, CD IN.

(viii) (viii) Microphone.

A video capture board can handle a variety of different audio and video input signals and convert them from analog to digital or digital to analog.

Video Channel Multiplexer: It is similar to the video grabber's video channel multiplexer.

Video Compression and Decompression: A video compression and decompression processor is used to compress and decompress video data.The video compression and decompression processor contains multiple stages for compression and decompression. The stages include forward discrete cosine transformation and inverse discrete cosine transformation, quantization and inverse quantization, Zigzags and Zero run-length encoding and decoding, and motion estimation and compensation.

Audio Compression: MPEG-2 uses adaptive pulse code modulation (ADPCM) to sample the audio signal. The method takes a difference between the actual sample value and predicted sample value. The difference is then encoded by a 4-bit value or 8-bit value depending upon the sample rate

Analog to Digital Converter: The ADC takes inputs from the video switch and converts the amplitude of a sampled analog signal to either an 8-bit or 16-bit digital value.

Special features for video playback: Before seeing the features of video playback let us learn what is isochronous playback. The playback at a constant rate to ensure proper cadence (the rise and fall in pitch of a person's voice) is known as isochronous playback. But isochronous playback is more complex with video than it is for sound.

If video consists of multiple clips of video and multiple soundtracks being retrieved from different servers and combined for playback by accurately synchronizing them, the problem becomes more complex. To achieve isochronous playback, most video storage systems use frame interleaving concepts. Video Frame Interleaving: Frame interleaving defines the structure of the video file in terms of the layout of sound and video components.

VIDEO SCALING, PANNING AND ZOOMING:

Scaling:

Scaling is a feature since users are used in changing window sizes. When the size of the video window is changed, scaling take place.

Panning: Panning allows the user to move to other parts of the window. Panning is useful in combination with zooming. Only if the video is being displayed at full resolution and the video window is not capable of displaying the entire window then panning is useful. Therefore panning is useful only for video captured using very high resolution cameras.

Zooming:

Zooming implies that the stored number of pixels is greater than the number that can be displayed in the video window. In that case, a video scaled to show the complete image in the video window can be paused and an area selected to be shown in a higher resolution within the same video window. The video can be played again from that point either in the zoomed mode or in scaled to fit window mode.

5.7 HYPER MEDIA MESSAGING

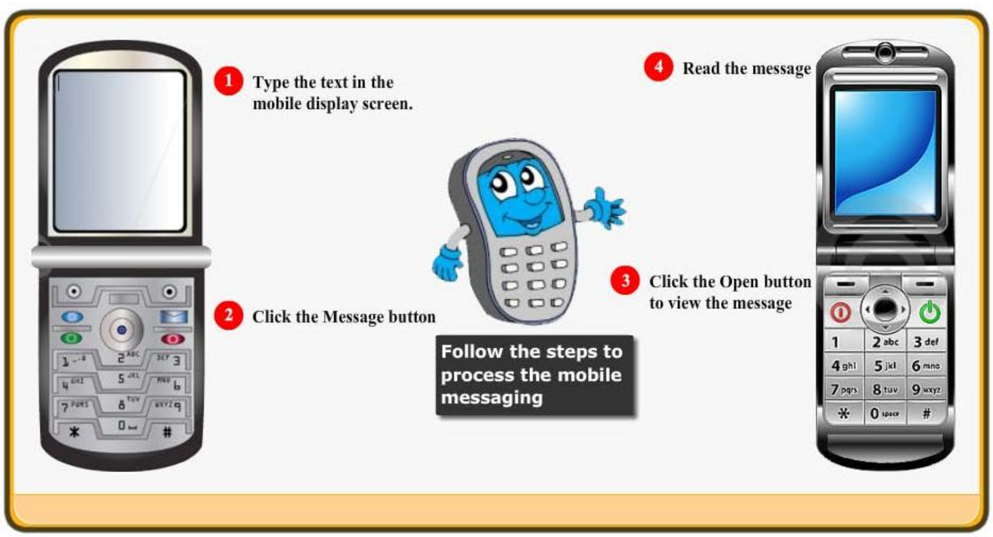
Messaging is one of the major multimedia applications. Messaging started out as a simple text-based electronic mail application. Multimedia components have made messaging much more complex. We see how these components are added to messages. Communication among members of a workgroup, among workgroups and departments, and among divisions is crucial for the success of any business enterprise. The communication can be face-to-face, via telephone and phone mail, via written communications – such as letter or electronic mail (e-mail), or via recorded sound and pictures. E-mail has emerged as a very well understood and used medium for communication in a business setting due to its speed and its close equivalence to written communications. An organization's objectives in deploying e-mail-based document interchange services should be to reduce paper output and in-house paper-handling costs and to streamline work processes by eliminating redundant functions. E-mail-based document interchange, generally known as messaging services, contributes to corporate productivity in the following important ways:

1. It strengthens the automation of the document life cycle (creation, transport, interchange, and output).
2. It allows document sharing (for viewing or revising) without forcing an organization to standardize on a particular word processor solely to achieve document interchange.
3. It cuts down on the paper output generated by organizations that must share documents.

Mobile Messaging

Mobile messaging represents a major new dimension in the user's interaction with the messaging system. With the emergence of remote access from users using personal digital assistants and notebook computers, made possible by wireless communications developments supporting wide ranging access using wireless modems and cellular telephone links, mobile messaging has significantly influence messaging paradigms.

Hypermedia messaging is not restricted to the desktops; it is increasingly being used on the road through mobile communications in metaphors very different from the traditional desktop metaphors. Short Message Service (SMS) is a communication service standardized in the GSM mobile communication system, using standardized communications protocols allowing the interchange of short text messages between mobile telephone devices.



- SMS text messaging is the most widely used data application on the planet, ith 2.4 billion active users, or 74% of all mobile phone subscribers sendingand receiving text messages on their phones.

• The SMS technology has facilitated the development and growth of text messaging. The connection between the phenomenon of text messaging and the underlying technology is so great that in parts of the world the term "SMS" is used as a synonym for a text message or the act of sending a text message, even when a different protocol is being used.

• SMS as used on modern handsets was originally defined as part of the GSM series of standards in 1985[1] as a means of sending messages of up to 160 characters (including spaces), to and from GSM mobile handsets.

• Since then, support for the service has expanded to include other mobile technologies such as ANSI CDMA networks and Digital AMPS, as well as satellite and landline networks. Most SMS messages are mobile-to-mobile text messages, though the standard supports other types of broadcast messaging as well.

5.7.1 HYPERMEDIA MESSAGE COMPONENTS

A hypermedia message may be a simple message in the form of text with an embedded graphics, sound track, or video clip, or it may be the result of analysis of material based books, CD ROMs, and other on-line applications. An authoring sequence for a message based on such analysis may consist of the following components.

1. The user may have watched some video presentation on the material and may want to attach a part of that clip in the message. While watching it, the user marks possible quotes and saves an annotated copy.
2. Some pages of the book are scanned as images. The images provide an illustration or a clearer analysis of the topic
3. The user writes the text of the message using a word processor. The text summarizes the highlights of the analysis and presents conclusions.

These three components must be combined in a message using an authoring tool provided by the messaging system. The messaging system must prompt the user to enter the name of the addressee for the message.

The message system looks up the name in an online directory and converts it to electronic addresses well as routing information before sending the message. The user is now ready to compose the message. The first step is to copy the word processed text report prepared in step 3 above in the body area of the message or use the text editor provided by the messaging system. The user then marks the spots where the images are referenced and uses the link and embed facilitates of the authoring tool to link in references to the images. The user also marks one or more spots for video clips and again uses the link and embeds facilities to add the video clips to the message when the message is fully composed, the user signs it (electronic signature) and mails to the message to the addressee (recipient). The addressing system must ensure that the images and video clips referenced in the message are also transferred to a server "local' to the recipient.

Text Messages

In earlier days, messaging systems used a limited subset of plain ASCII text. Later, messaging systems were designed to allow users to communicate using short messages. Then, new messaging standards have added on new capabilities to simple messages. They provide various classes of service and delivery reports.

Pratap

To : Karan

Copy to: Madhan Date: 01 Jan'07

Subject: WISHING A

HAPPY NEW

YEAR Hai Karan,

I wish you a very bright and prosperous new year.

Pratap Delivery notification: Normal Priority: High

Electronic mail message

Other capabilities of messaging systems includes a name and address directory of all users accessible to the messaging system.

Rich-Text Messages

Microsoft defined a standard for exporting and importing text data that included character set, font table, section and paragraph formatting, document formatting, and color information-called Rich Text Format (RTF), this standard is used for storage as well as Import and export of text files across a variety of word-processing and messaging systems.

When sections of this document are cut and pasted into another application, the font and formatting information is .retained. This allows the target application to display the text in the nearest equivalent fonts and formats.

Rich-text messages based on the RTF formats provide the capability to create messages in one word processor and edit in another at the recipient end. Most messaging systems provide rich text capability for the field of a message.

Voice Messages

Voice mail systems answer telephones using recorded messages and direct the caller through a sequence of touch tone key operations until the caller is connected to the desired party or is able to leave a recorded message. Voice message refers to a message that could be sent to a destination using voice media.

- Voice itself could be 'packaged' and sent through the IP backbone so that it reaches its marked 'address'.
- In a technical sense, the process of sending 'voice packets' is a semi passive way of communication.
- However, given the speed at which it could be delivered can make the communication sound seamless.

Full Motion Video Message:

- Those live-action video clips showing the computer's desktop, menus and files in motion are called screen casts. To take your own videos, you need a screen casting application. Programs like Cam Studio are available free, while for about $800 you can buy Adobe Captivate, software aimed at education and training professionals. You can also record screen casts right in your browser on sites like Screen Castle and Screen cast-O-Matic. When looking for screen casting software to suit your needs, read the program specifications. Some less-expensive options may not record sound along with the video or may limit recording time.

Full-Motion Video Management

Use of full-motion video for information repositories and memos are more informative. More information can be 'conveyed and explained in a short full-motion video clip than can be conveyed in a long text document. Because a picture is equivalent to thousand words.

Full Motion video Authoring System

An authoring system is an important component of a multimedia messaging system. A good authoring system must provide a number of tools for the creation and editing of multimedia objects. The subsets of tools that are necessary are listed below:

1. A video capture program - to allow fast and simple capture of digital video from analog sources such as a video camera or a video tape. .
2. Compression and decompression Interfaces for compressing the captured video as it is being captured.
3. A video editor with the ability to decompress, combines, edits, and compresses digital video clips.
4. Video indexing and annotating software for marking sections of a video clip and recording annotations.

Full-Motion Video Playback Systems

The playback system allows the recipient to detach the embedded video reference object, Interpret its contents and retrieve the actual video clip from a specialized video server and launch the Playback application. A number of factors are involved in playing back the video correctly.

They are:

1. How the compression format used for the storage of the video clip relates to the available hardware and software facilities for decompression.
2. Resolution of the screen and the system facilities available for managing display windows. The display resolution may be higher or lower

than the resolution of the source of the video clip.

3. The CPU processing power and the expected level of degradation as well as managing the degraded output on the fly.

4. Ability to determine hardware and software facilities of the recipient's system, and adjusting playback, parameters to provide the best resolution and performance on playback.

5.7.2 HYPERMEDIA LINKING AND EMBEDDING

Linking and embedding are two methods for associating multimedia objects with documents.

Linking Objects

When an object is linked, the source data object, called the link source, continues to stay whenever it was at the time the link was created. This may be at the object server where it was created, or where it has been copied.

Only reference is required in the hypermedia document. The' reference is also known as link. This link reference includes information about the multimedia object storage, its presentation parameters, and the server application that is needed to display/play or edit it. When this document is copied, the link reference is transferred. But the actual multimedia document remains in its original location. A linked object is not a part of the hypermedia document and it does not take up storage space within the hypermedia document. If the creator, or authorized user edits the original stored multimedia object, subsequent calls to the linked object bring the copy.

Embedded Objects

If a copy of the object is physically stored in the hypermedia document, then the multimedia objects: 3 said to be embedded. Any changes to the original copy of t4at object are not reflected in the embedded copy. When the hypermedia document is copied, the multimedia object is transferred with it to the new locations. Graphics and images can be inserted in a rich-text document on embedded using such techniques as OLE Voice and audio components can be included in a text message; or they can be part of a full voice-recorded message that has embedded text and other components.

CREATING HYPERMEDIA MESSAGES

Hypermedia message is a complex collection of a variety of objects. It is an integrated message consisting of text, rich text, binary files, images, and bitmaps. Voice, sound, and full motion video. Creating of a hypermedia message requires some preparation. A hypermedia report is more complex. It requires the following steps:

1. Planning
2. Creating each component
3. Integrating components

The planning phase for preparing the hypermedia message consists of determining the various sources of input. These can include any of the following:

1. A text report prepared in a word-processing system.
2. A spreadsheet in a spreadsheet program.
3. Some diagrams from a graphics program.
4. Images of documents.
5. Sound dips.
6. Video clips.

We should determine which components are required for the message, in what sequence should they be, and where in the text report they should be referenced. The length of each component should be determined. Careful planning is necessary to ensure that the capabilities of the messaging system are used appropriately.

Each component must be created using the authoring tool provided by the application used for creating it. All applications Involved in creating various components must have common formats to allow combining these various components. The various components must be authored, reviewed, and edited as needed, checked for smooth flow when the user launches an embedded object and stored in the final format in which it will become a part of the hypermedia message. The final step in this process is mailing the hypermedia message.

5.10 INTEGRATED MULTIMEDIA MESSAGE STANDARDS

Let us review some of the Integrated Multimedia Message Standards in detail.

1)Vendor Independent Messaging (VIM):

VIM interface is designed to facilitate messaging between VIM. Enabled electronic mail systems as well as other applications. The VIM interface makes mail and messages services available through a well defined interface. A messaging service enables its clients to communicate with each other in a store-and-forward manner. VIM-aware applications may also use one-or-more address books. Address books are used to store information about users, groups, applications, and so on.

VIM Messages:

VIM defines messaging as a stored-and-forward method of application-to-application all program-to-program data exchange. The objects transported by a messaging system are called messages. The message, along with the address is sent to the messaging system. The messaging system providing VIM services accept the responsibility for routing and delivering the message to the message container of the recipient.

Message Definition:

Each message has a message type. The message type defines the syntax of the message and the type of information that can be contained in the message.

A VIM message consists of message header. It may contain one or more message items. The message header consists of header attributes: recipient address, originator address, time/date prior. A message item is a block of arbitrary-sized (means any size) data of a defined type. The contents of the data block are defined by the data-item type.

The actual items in a message and its syntax and semantics are defined by the message type. The message may also contain file attachments. VIM allows the nesting of messages; means one message may be enclosed in another message. A VIM message can be digitally signed so that we can ensure that the message 'received is without any modification during the transit.

Mail Message: It is a message of a well-defamed type that must include a message header and may include note parts, attachments, and other application-defined components. End users can see their mail messages through their mail programs.

***Message Delivery**:* If message is delivered successfully, a delivery report is generated and sends to the sender of the message if the sender requested the deliver-y report. If a message is not delivered, a non-delivered report is sent to the sender. A message that delivered will be in a message container will be marked as 'unread', until the recipient open and read it.

***Message Container**:* Multiple users or applications can access one message container. Each message in a message container has a reference number associated with it for as long as the message remains stored in the message container.

VIM Services: The VIM interface provides a number of services for creating and mailing a message. Some of them are:

.:. Electronic message composition and submission.

.:. Electronic message sending and receiving.

.:. Message extraction from mail system.

.:. Address book services.

2) MAPI (Multimedia Application Programmable Interface)

Messaging Application Programming Interface (MAPI) is a messaging architecture and a Component Object Model based API for Microsoft Windows. MAPI allows client programs to become (e-mail) messaging-enabled, -aware, or -based by calling MAPI subsystem routines that interface with certain messaging servers.

MAPI provides a layer of functionality between applications and underlying messaging systems. The primary goals of MAPI are: Separate client applications from the underlying messaging services. Make basic mail enabling a standard feature for all applications. Support message-reliant workgroup applications.

• While MAPI is designed to be independent of the protocol, it is usually used with MAPI/RPC, the proprietary protocol that Microsoft Outlook uses to communicate with Microsoft Exchange.

• Simple MAPI is a subset of 12 functions which enable developers to add basic messaging functionality.

• Extended MAPI allows complete control over the messaging system on the client computer, creation and management of messages, management of the client mailbox, service providers, and so forth.

• Simple MAPI ships with Microsoft Windows as part of Outlook Express/Windows Mail while the full Extended MAPI ships with Office Outlook and Exchange.

• In addition to the Extended MAPI client interface, programming calls can be made indirectly through the Simple MAPI API client interface, through the Common Messaging Calls (CMC) API client interface, or by the object-based CDO Library interface.

• These three methods are easier to use and designed for less complex messaging-enabled and -aware applications. (Simple MAPI and CMC were removed from Exchange 2003.)

MAPI Architecture: MAPI Architecture provides two perspectives (i) a client API

(ii) A service provider interface. The Client API provides the link between the client applications andMAPI. The service provider interface links MAPI to the messaging system.

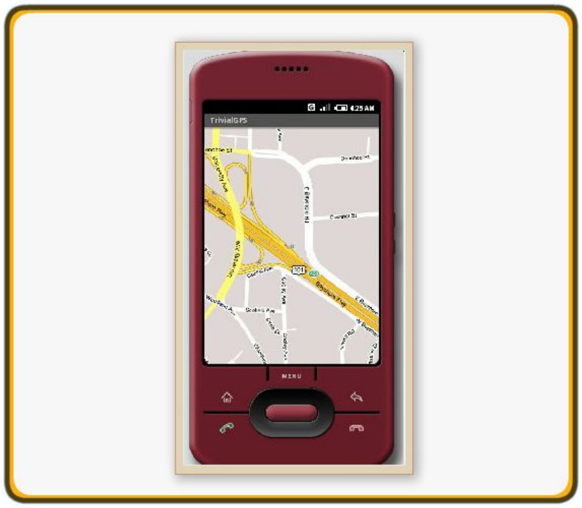
The two interfaces combine to provide an open architecture such that any messaging application can use any messaging service that has a MAPI driver. MAPI drivers are provided by Microsoft or third party developers.

3)TELEPHONY API (TAPI)

TAPI standard has been defined by Microsoft and Intel. The telephone can be used for reading e-mail as well as for entering e-mail messages remotely. The CG 6565e is supported by NMS's Natural Access Fusion VoIP service for IP-based enhanced services applications. This development environment offers significant advantages in development time, performance, and most importantly, scalability. Fusion provides a common software development environment that can be used to create IP telephony gateways, IP-enabled enhanced services platforms, and wire- less IP telephony and video gateways.

Telephony API

The CG 6565e is supported by NMS's Natural Access Fusion VoIP service for IP-based enhanced services applications. This development environment offers significant advantages in development time, performance, and most importantly, scalability. Fusion provides a common software development environment that can be used to create IP telephony gateways, IP-enabled enhanced services platforms, and wire- less IP telephony and video gateways.



4) X 400 MESSAGE HANDLING SERVICE

The CCITT X 400 series recommendations define the OSI message handling system, (MHS).The MHS describes a functional model that provides end users the ability t6 send and receive electronic messages. In the system end user is an originator. He composes and sends messages.

Receiver is the one who receives messages. A User Agent (UA) is an entity that provides the end user function for composing and sending messages and for delivering messages. Most user agent implementations provide storage of mail, sorting directories, and forwarding.

A Message Transfer Agent (MTA) forwards messages from the originator UA to another MT A. A number of MTAs are combining to form Message transfer System (MTS).The MTAs in an MTS provide message routing services at intermediate nodes in a WAN. Figure below shows the overall X 400 architecture and the relationships between the components.

Figure 5.2 X.400 Architecture

X·500 5.11DIRECTORY SYSTEM STANDARDS

The X·500 is the joint International Standard Organization. CCITT standard for a distributed directory system that lets users store information such as addresses and databases on a local server and easily query, exchange, and update that information in an interoperable networked environment. The X 500 directory structure is described in the CCITT standard known as Data Communications Network Directory, Recommendations X·500-X·521, 1988.

X·500 Directory System Architecture

Directory System Agents carryout updates and management operations. X ·500 defines a structured information model, an object oriented model and database schema. The X ·500 architecture is based on a number of models, as follow'>:

The information model: It specifies the contents of directory entries, how they are identified, and the way in which they are organized to form the directory information base.

The Directory model: It describes the directory and its users, the functional model for directory operation, and the organization of the directory. .

The security model: It specifies the way in which the contents of the directory are protected from unauthorized access and authentication methods for updates. The X 500 directory system is designed to be capable of spanning national and corporate boundaries.

X 500 Directory System Components: All information in an X 500 database is organized as entries in the Directory-Information Base(DlB). The directory system provides agents to manipulate entries in the DIB.

X 500 directories consist of the following basic components:

1. **Directory Information Base (DIB**); The DIB contains information about users, applications, resources and the configuration of the directory that enables servers to locate one another.
2. **Directory User Agents (DUA**): A DUA issues inquiry and update requests, and accesses directory information through the directory access protocol.
3. **Directory Service Agents (DSAs**): DSAs cooperate with one another to resolve user requests over a distributed network. They interact through a specialized protocol called a directory system protocol.

GRAPHICS AND MULTIMEDIA – ONLINE QUESTIONS

Questions	opt1	opt2	opt3	opt4	answer
Which technique gives a line drawing a more realistic appearance	Hidden line elimination	Animation	3D Imaging	All of these	Hidden line elimination
Which of the following is a combination of geometry and topology	3D image	3D scene	3D view parameter	All of these	3D scene
Morphing is	A type of animation	It is an anti aliasing technique	It is a technique creating special sound effect	All of these	A type of animation
Which application requires extreme realism including moving images	Simulation	Animation	CAD	All of these	Simulation
The process of displaying 3D objects on a 2D display is called	Animation	Projection	Transformation	Morphing	Projection
________ color model which uses color descriptions that have a more inituitive appeal to a user.	HSV model	CMV model	HLS model	RGB model	HSV model
________ color model useful for describing color output to hard copy devices	CMV	RGB	HLS	HSV	CMV
NTSC stands for	National Telegraph System Committee.	National Television System Commission.	National Television System Committee.	National Telegraph System Commission	National Television System Committee.

SHADES is combination of	Original Color + Black Pigment.	Original Color + White Pigment.	Original Color + Black Pigment + White Pigment.	Black Pigment + White Pigment.	Original Color + Black Pigment.
The 3 color parameters in HLS color model is	Hue, Luminance, Saturation.	Hue, Luminance, Saturation.	Hue, Brightness, Saturation.	All of these	Hue, Luminance, Saturation.
The color parameter in the YIQ model is	Luminance, Hue, Purity.	Saturation, Hue, Value.	Saturation, Hue, Lightness.	Cyan, Magneta, Yellow.	Luminance, Hue, Purity.
If the 2 color sources combine to produce white light, they are called	Primary colors.	complementary colors	Spectral colors.	Color jammet.	complementary colors
The 2 or 3 colors used to produce other colors in a color model are called	Complementary colors.	Primary colors.	Spectral colors.	Color jammet	Primary colors.
. The term chromaticity is used to refer collectively two properties	Purity & Saturation.	Saturation & Luminance.	Purity & Luminance.	Purity & dominant Frequency	Purity & dominant Frequency
The dominant frequency is also called	Saturation.	Hue	Brightness.	Luminance.	Hue.
____________ refers to any time sequence of visual changes in a scene	Morphing.	Computer Animation.	Transformation.	Framing.	Computer Animation.
The combination of cyan and Magneta produces ___________ light.	Blue.	Yellow.	Red.	Green.	Yellow.

Animation Functions are	Object manipulation & rendering.	Camera motion.	Generation of inbetween.	All of these	All of these
Another method to display the non visible lines as	Highlighted lines.	Dashed lines.	Hidden lines.	All of these	Dashed lines.
Polygon Facets, Spline patches are the example for	Boundary Representation.	Space Partitioning Representation	Non-Boundary Representation.	Spline Representation.	Boundary Representation.
___________ is a doughnet shaped object	Spline.	Torus.	Ellipsoid	Sphere.	Torus.
Red and Cyan, Green and Magneta, blue and yellow are example for	Primary colors.	Complementary colors.	Spectral colors.	Color gammet	Complementary colors.
Types of computer graphics are	Vector and raster	Scalar and raster	Vector and scalar	None of these	Vector and raster
Vector graphics is composed of	Pixels	Paths	Palette	None of these	Paths
Raster graphics are composed of	Pixels	Paths	Palette	None of these	Pixels
Raster images are more commonly called	Pix map	bitmap	both a & b	none of these	bitmap
Pixel can be arranged in a regular	One dimensional grid	Two dimensional grid	Three dimensional grid	None of these	Two dimensional grid
The brightness of each pixel is	Compatible	Incompatible	Both a & b	None of these	Incompatible
Each pixel has _________basic color components	Two or three	One or two	Three or four	None of these	Three or four

The quantity of an image depend on	No. of pixel used by image	No. of line used by image	No. of resolution used by image	None	No. of pixel used by image
Higher the number 0f pixels,________ the image quality	Bad	Better	Smaller	None of above	Better
A palette can be defined as a finite set of colors for managing the	Analog images	Digital images	Both a & b	None of these	Digital images
Display card are	VGA	EGA	Both a & b	None of above	Both a & b
Display card is used for the purpose of	Sending graphics data to input unit	Sending graphics data to output unit	Receiving graphics data from output unit	None of these	Sending graphics data to output unit
Several graphics image file formats that are used by most of graphics system are	GIF	JPEG	TIFF	All of these	All of these
The GIF format is much __________to be downloaded or uploaded over the	Slower	Faster	Medium	None of these	Faster
Once a file is saved in JPEG format ,some data is lost	Temporarily	Permanently	Both a & b	None	Permanently
EPS image file format is used for	Vector graphics	Bitmap	Both a & b	None of these	Both a & b
TIFF (tagged image file format)are used for	Vector graphics	Bitmap	Both a & b	None of these	Bitmap

EPS means	Entire post script	Entire post scale	Encapsulated post script	None of these	Encapsulated post script
The additive color models use the concept of	Printing ink	Light to display color	Printing line	None of these	Light to display color
The subtractive color model use the concept of	Printing ink	Light to display color	Printing line	None of these	Printing ink
Color apparent in additive model are the result of	Reflected light	Transmission of light	Flow of light	None of these	Transmission of light
Color apparent in subtractive model are the result of	Amount of Reflected light	Transmission of light	Flow of light	None of these	Amount of Reflected light
Two dimensional color model are	RGB and CMKY	RBG and CYMK	RGB and CMYK	None	RGB and CMYK
RGB model are used for	Computer display	Printing	Painting	None of these	Computer display
CMYK model are used for	Computer display	Printing	Painting	None of these	Printing
The intersection of three primary RGB color produces	White color	Black color	Magenta color	Blue color	White color
The intersection of primary CMYK color produces	White color	Black color	Cyan color	Magenta color	Black color
The RGB model display a much ________ percentage of the visible band as compared to CMYK	Lesser	Larger	Medium	None of these	Larger

Color depth can be defined by ________ which can be displayed on a display unit	Bits per pixel	Bytes per pixel	Megabyte per pixel	None of these	Bits per pixel
Each bit represent	One color	Two color	Three color	None	Two color
RGB true color model has _______ color depth	24bit	32bit	64bit	None	24bit
CMYK true color model has _______ color depth	24bit	32bit	64bit	None	32bit
Grey scale images have a maximum color depth of	8bit	16bit	24bit	32bit	8bit
Graphics with limited features is known as	Active graphics	Passive graphics	Grayscale image	None of these	Passive graphics
Computer of present time have much higher memory and ________ storage capacity	Much smaller	Much bigger	Much slower	None	Much bigger
CRT means	Common ray tube	Cathode ray tube	Common ray tube	None	Cathode ray tube
Refresh CRT consist of	Glass wrapper	The phosphor viewing surface	The electron gun assembly	All of above	All of above
The amount of time the phosphor produce light or shine is controlled by chemical composition of the phosphor. This is known as	Persistence	Resistance	Generators	None	Persistence