UNIT I ILLUMINATION AND COLOUR MODELS

Light sources – basic illumination models – halftone patterns and dithering techniques; Properties of light – Standard primaries and chromaticity diagram; Intuitive colour concepts – RGB colour model – YIQ colour model – CMY colour model – HSV colour model – HLS colour model; Colour selection.

Color Models

Color Model is a method for explaining the properties or behavior of color within some particular context. No single color model can explain all aspects of color, so we make use of different models to help describe the different perceived characteristics of color.

Properties of Light

Light is a narrow frequency band within the electromagnetic system.

Other frequency bands within this spectrum are called radio waves, micro waves, infrared waves and x-rays. The below fig shows the frequency ranges for some of the electromagnetic bands.

Each frequency value within the visible band corresponds to a distinct color.

At the low frequency end is a red color ($4.3*10^{4}$ Hz) and the highest frequency is a violet color ($7.5 *10^{14}$Hz)

Spectral colors range from the reds through orange and yellow at the low frequency end to greens, blues and violet at the high end.

Since light is an electro magnetic wave, the various colors are described in terms of either the frequency for the wave length λ of the wave.

The wave length ad frequency of the monochromatic wave are inversely proportional to each other, with the proportionality constants as the speed of light
C where $C = \lambda f$
A light source such as the sun or a light bulb emits all frequencies within the visible range to produce white light. When white light is incident upon an object, some frequencies are reflected and some are absorbed by the object. The combination of frequencies present in the reflected light determines what we perceive as the color of the object.

If low frequencies are predominant in the reflected light, the object is described as red. In this case, the perceived light has the dominant frequency at the red end of the spectrum. The dominant frequency is also called the hue, or simply the color of the light.

AMSCE - 1101

Brightness is another property, which in the perceived intensity of the light.

Intensity in the radiant energy emitted per limit time, per unit solid angle, and per unit projected area of the source.

Radiant energy is related to the luminance of the source.

The next property in the purity or saturation of the light.

- Purity describes how washed out or how pure the color of the light appears.
- Pastels and Pale colors are described as less pure.

The term chromaticity is used to refer collectively to the two properties, purity and dominant frequency.

Two different color light sources with suitably chosen intensities can be used to produce a range of other colors.

If the 2 color sources combine to produce white light, they are called complementary colors. E.g., Red and Cyan, green and magenta, and blue and yellow.

Color models that are used to describe combinations of light in terms of dominant frequency use 3 colors to obtain a wide range of colors, called the color gamut.

The 2 or 3 colors used to produce other colors in a color model are called primary colors.

Standard Primaries

XYZ Color

The set of primaries is generally referred to as the XYZ or (X,Y,Z) color model

where X,Y and Z represent vectors in a 3D, additive color space.

Any color Cλ is expressed as

$$C\lambda = X\mathbf{X} + Y\mathbf{Y} + Z\mathbf{Z} \text{-------------} \qquad (1)$$

Where X,Y and Z designates the amounts of the standard primaries needed to match Cλ.

It is convenient to normalize the amount in equation (1) against luminance (X+ Y+ Z). Normalized amounts are calculated as,

$$x = X/(X+Y+Z), \qquad y = Y/(X+Y+Z), \qquad z = Z/(X+Y+Z)$$

AMSCE - 1101

with x + y + z = 1

Any color can be represented with just the x and y amounts. The parameters x and y are called the chromaticity values because they depend only on hue and purity.

If we specify colors only with x and y, we cannot obtain the amounts X, Y and Z. so, a complete description of a color in given with the 3 values x, y and Y.

$$X = (x/y)Y, \qquad Z = (z/y)Y$$

Where z = 1-x-y.

Intuitive Color Concepts

Color paintings can be created by mixing color pigments with white and black pigments to form the various shades, tints and tones.

Starting with the pigment for a „pure color" the color is added to black pigment to produce different shades. The more black pigment produces darker shades.

Different tints of the color are obtained by adding a white pigment to the original color, making it lighter as more white is added.

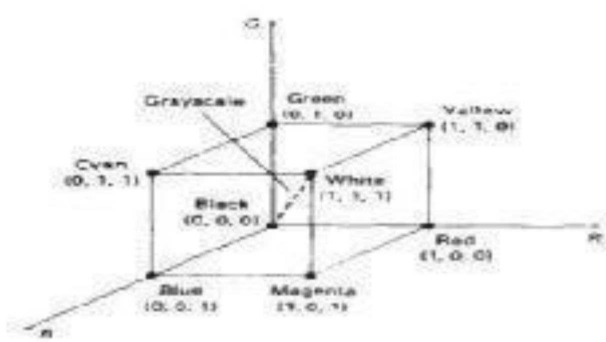
Tones of the color are produced by adding both black and white pigments.

RGB Color Model

Based on the tristimulus theory of version, our eyes perceive color through the stimulation of three visual pigments in the cones on the retina.

These visual pigments have a peak sensitivity at wavelengths of about 630 nm (red), 530 nm (green) and 450 nm (blue).
By comparing intensities in a light source, we perceive the color of the light.

This is the basis for displaying color output on a video monitor using the 3 color primaries, red, green, and blue referred to as the RGB color model.



AMSCE - 1101

Vertices of the cube on the axes represent the primary colors, the remaining vertices represents the complementary color for each of the primary colors.

The RGB color scheme is an additive model. (i.e.,) Intensities of the primary colors are added to produce other colors.

Each color point within the bounds of the cube can be represented as the triple (R,G,B) where values for R, G and B are assigned in the range from 0 to1.

The color $C\lambda$ is expressed in RGB component 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 (1,0,1) and white at (1,1,1) in 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.

YIQ Color Model

The National Television System Committee (NTSC) color model for forming the composite video signal in the YIQ model.

In the YIQ color model, luminance (brightness) information in contained in the Y parameter, chromaticity information (hue and purity) is contained into the I and Q parameters.

A combination of red, green and blue intensities are chosen for the Y parameter to yield the standard luminosity curve.

Since Y contains the luminance information, black and white TV monitors use only the Y signal.

Parameter I contain orange-cyan hue information that provides the flash-tone shading and occupies a bandwidth of 1.5 MHz.

Parameter Q carries green-magenta hue information in a bandwidth of about 0.6MHz.

CMY Color Model

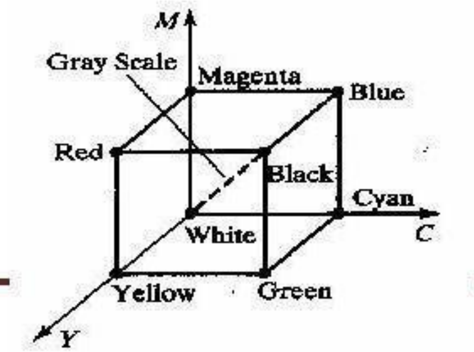
A color model defined with the primary colors cyan, magenta, and yellow (CMY) in useful for describing color output to hard copy devices.

It is a subtractive color model (i.e.,) cyan can be formed by adding green and blue light. When white light is reflected from cyan-colored ink, the reflected light must have no red

AMSCE - 1101

component. i.e., red light is absorbed or subtracted by the link.

Magenta ink subtracts the green component from incident light and yellow subtracts the blue component.



In CMY model, point (1,1,1) represents black because all components of the incident light are subtracted.

The origin represents white light.

Equal amounts of each of the primary colors produce grays along the main diagonal of the cube.

A combination of cyan and magenta ink produces blue light because the red and green components of the incident light are absorbed.

The printing process often used with the CMY model generates a color point with a collection of 4 ink dots; one dot is used for each of the primary colors (cyan, magenta and yellow) and one dot in black.

The conversion from an RGB representation to a CMY representation is expressed as

[ ] [ ] [ ]

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

Similarly the conversion of CMY to RGB representation is expressed as [ ] [ ] [ ]

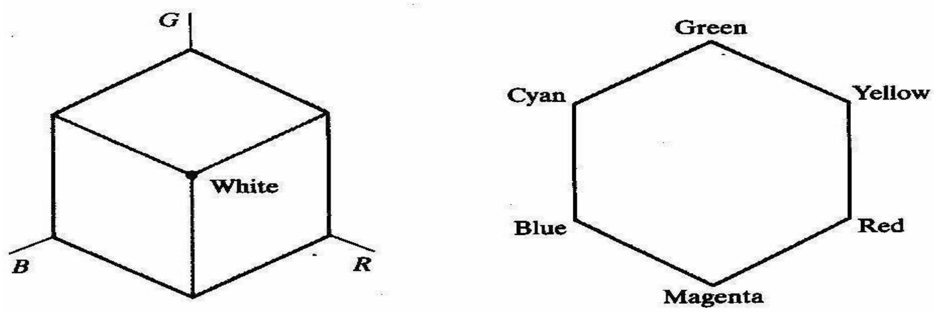
Where black is represented in the CMY system as the unit column vector.

HSV Color Model

The HSV model uses color descriptions that have a more interactive appeal to a user.

Color parameters in this model are hue (H), saturation (S), and value (V). The 3D representation of the HSV model is derived from the RGB cube. The outline of the cube has the hexagon shape.

AMSCE - 1101



RGB Color Cube
(a)

Color Hexagon
(b)

The boundary of the hexagon represents the various hues, and it is used as the top of the HSV hexcone.

In the hexcone, saturation is measured along a horizontal axis, and value is along a vertical axis through the center of the hexcone.

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

AMSCE - 1101

are 180 apart.

Saturation S varies from 0 to 1. the maximum purity at S = 1, at S = 0.25, the hue is said to be one quarter pure, at S = 0, we have the gray scale. Value V varies from 0 at the apex to 1 at the top.

the apex representation black.

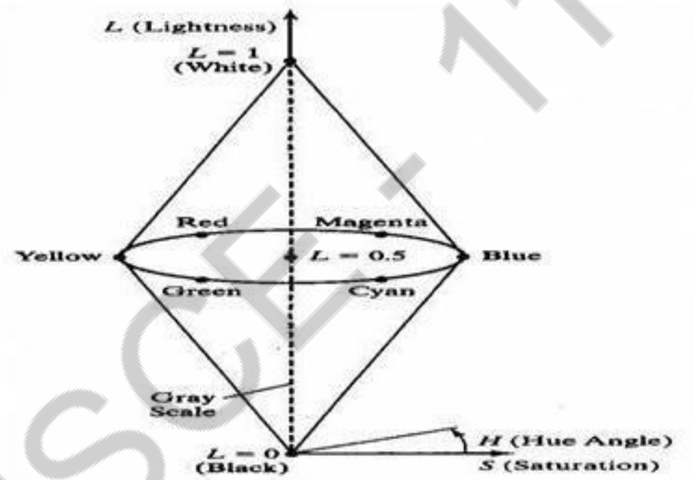
At the top of the hexcone, colors have their maximum intensity.
When V = 1 and S = 1 we have the „pure" hues.

White is the point at V = 1 and S = 0.

HLS Color Model

HLS model is based on intuitive color parameters used by Tektronix.

It has the double cone representation shown in the below figure. The 3 parameters in this model are called Hue (H), lightness (L) and saturation (s).
Hue specifies an angle about the vertical axis that locates a chosen hue.



In this model H = θ^0 corresponds to Blue.

The remaining colors are specified around the perimeter of the cone in the same order as in the HSV model.

Magenta is at 60^0,Red in at 120^0 and cyan in at H = 180^0 .

The vertical axis is called lightness (L). At L = 0, we have black, and white is at L

= 1 Gray scale in along the L axis and the "purehues" on the L = 0.5 plane.

Saturation parameter S specifies relative purity of a color. S varies from 0 to 1 pure hues are those for which S = 1 and L = 0.

As S decreases, the hues are said to be less pure.
At S= 0, it is said to be gray scale.

AMSCE - 1101

Line Drawing algorithm

Slope-intersept equation for a straight line is

$Y=mx+b$

m- slope of the line. b-constant

Two end points of the line segment are (x1,y1) , (x2,y2)
Slope m = y2-y1 / x2-x1 Δy/Δx

Δx- x interval – x2-x1
Δy – y interval – y2-y1

$\Delta y = m. \Delta x$

$\Delta x=\Delta y/m$

If the interval is known, we can find the next point.
xi+1= x i+ Δx xi+ Δy/m

yi+1= y i+ Δy yi+ m. Δx

we sample at unit x interval and y interval then this equation becomes xi+1= x i+ Δx xi+ (1/m)

yi+1= y i+ Δy yi+ m.1

The above equations are for the lines which are processed from left to right.

The equations are for the lines which are processed from right to left is xi+1= x i+ Δx xi- (1/m)

yi+1= y i+ Δy yi- m.1

Since m can be any real number between 0 and 1 the y value must be rounded to the nearest integer.

AMSCE - 1101

DDA Algorithm (Digital differential analyzer)

1. Get the two end points
2. Calculate the Horizontal and vertical difference between two end points (dx,dy)
3. The greatest difference value is taken as the length value.
4. Calculate the step value dx=dx/length, dy=dy / length
5. Start plotting the line from the first point.
6. Repeat through 9 step up to length value times
7. if dx>dy and xa<xb then increment x by 1 and increment y by m
8. if dx>dy and xa>xb then decrement x by 1 and decrement y by m
9. if dx<dy ya<yb then increment y by 1 and increment x by 1/m
10. if dx<dy and ya>yb then decrement y by 1 and decrement x by 1/m

DDA Line Algorithm

It generates lines from their differential equations.

Advantages

1. It is the faster method for calculating pixel positions.

2. It is simplest algorithm.It does not require special skills for implementation.

Disadvantages

Floating point arithmetic in DDA algorithm is still time consuming

The algorithm is orientation dependent.Hence end point accuracy is poor.

Bresenham's Line Algorithm

This algorithm uses only integer addition, subtraction, and multiplication by 2. So it is efficient for scan converting algorithms. This algorithm increments either x or y by one unit depending on the slope of the line. The increment in the other variable is determined by examine the distance between the actual line location and the nearest pixel. This distance is called **decision variable** or the **error.**

In matemetical terms the error or decision variable is defined as

e = Db – Da

If e>0 then the pixel above the line is closer to the true line.

Else the pixel below the line is closer to the true line.

Assume the current pixel is (xk,yk)

We have to find the next pixel position either ((xk+1,yk) and (xk+1,yk+1)

The y coordinate on the line at the pixel position xk+1 Is y = m (xk+1) + c Then the distance d1= y - yk = m (xk+1) + c-yk

d2= (yk+1-y)= yk+1-m (xk+1) + c

AMSCE - 1101

$d1-d2 = m (xk+1) + c-yk -yk--1-(yk+1)+m (xk+1) + c$
$= 2m (xk+1)-2 yk +2c-1$

The error term is initially set as $e = 2\Delta y-\Delta x$ where $\Delta y=y2-y1$, $\Delta x=x2-x1$

Bresenhams algorithm

1. Get the two end points
2. Calculate the values dx,dy,2dy and 2dy-2dx where dx=X2-X1 and dy=Y2-Y1
3. Calculate the starting decision parameter d=2dy-dx
4. plot the first point
5. At each Xk along the line, starting at k=0, perform the following test

If pk<0 , the next point to plot is (Xk+1,Yk) and Pk+1=pk+2dy
Otherwise the next point to plot is (Xk+1,Yk+1) Pk+1=pk+2dy-
2dx 6. Repeat step 5 for dx times.

The following program is used to for generalized bresenham algorithm, which will work for all the four coordinates.

Parallel Line algorithm

We can calculate the pixel position along the path simultaneously by partitioning the computations among the various processors available. One approach to the partitioning problem is to adapt an exsisting sequential algorithm to take advantage of multiple processors. Alternatively, we can look for other ways to set up the processing so that pixel positions can be calculated efficiently in parallel.

Circle Generating algorithm
Properties of the Circle

Circle is defined as a set of points that are all at a given distance r from a center position (Xc,Yc). This distance relationship is expressed by the Pythagorean theorem in Cartesian coordinates as

$$(X-Xc)2 +(Y-Yc)2=r2$$

Bresenham’s line algorithm for raster display is adapted to circle generation by setting up the decision parameters for finding the closest pixel for each sampling step.

A method for direct distances comparison is to test the halfway position between two pixels, to determine if this midpoint is inside or outside the circle boundary. This method is more easy . For an integer circle radius, the midpoint approach generates the same pixel position.

Midpoint Circle Algorithm

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

$$(x0,y0)=(0,r)$$

2.Calculate the initial value of the decision parameter

AMSCE - 1101

as P0=5/4 –r

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

if Pk<0 , then 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, 2yk+1=2yk-2

4. Determine the symmetry points in the other seven octants

5.Move each calculated position(x,y) onto the circular path centered
on (xc,yc) and plot the coordinate values x=x+xc,y=y+yc

6.Repeat steps 3 to 5 until x > = y

Ellipse generating Algorithm

Properties of the Ellipse

An ellipse is a set of points such that the sum of the distances from two fixed positions (foci) is the same for all points.If the distances to any two foci from any point P=(x,y) on the ellipse are labeled d1 and d2 then the general equation of an an ellipse can be stated as d1 + d2 is constant.

An ellipse in standard position is symmetric between quadrants. But it not symmetric between the two octants of the quadrant. So, we must calculate the pixel positions along the elliptical arc throughout one quadrant, then we obtained positions in the remaining three quadrants by symmetry.

Midpoint Ellipse Algorithm

1.Input rx,ry and ellipse center (xc,yc) and obtain the first point on
an ellipse centered on the origin as

(x0,y0)=(0,ry)

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

as p10 = r2y-r2xry + ¼ r2x

3.At each xk position in region 1,starting at k=0,perform the following test
if p1k<0, the next point along the ellipse centered on (0,0)is (xk+1,yk)
otherwise the next point along the circle is (xk+1,yk-1) and

p1k+1 = p1k + 2r2yxk+1 – 2r2yxk+1 + r2y

with

2r2yxk+1 = 2r2yxk + 2r2y

2r2xyk-2r2xyk – 2r2x

4.Calculate the initial value of the decision parameter in region 2 using
the last point (x0,y0) calculated in region as

p20 = r2y (x0+1/2)2 + r2x(y0-1)2-r2xr2y

5.At each yk position in region2 starting at k=0 ,perform the following test
if p2k>0 the next point along the ellipse centered on (0,0) is

(xk,yk-1) and p2k+1=p2k-
2r2xyk+1 + r2x

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

AMSCE - 1101

p2k+1 = p2k + 2r2yxk+1 – 2r2xyk+1 +r2x

using the same incremental calculations for x and y as in region1
6.Determine symmetry points in the other three quadrants.

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

X=x+xc,y=y+yc

8. Repeat the steps for region1 until 2r2yx > = 2r2xy

Attributes

Any parameter that affects the way a primitive is to be displayed referred to as an attribute parameter

Line Attributes
1.**Line Type**:

possible line type attribute include solid lines, dashed lines, dotted lines

2.Line Width

possible line width attribute include thicker line, thinner line and standard line. 3. **Line Color**

no of colors that can be provided to an output primitives depends upon the display device we used.

4. Pen and brush options

lines generated with pen or brush shapes can be displayed in various widths by changing the size of the mask. 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

Curve Attribute

Area Fill Attributes

Options of filling a defined region include a choice between a solid color or a patterned fill and choices for the particular colors and patterns.

Basic Fill Styles:
Hollow fill, Solid fill , Patterned fill

Character Attributes:

Text Attributes

A set of characters are affected by a particular attribute.(Font,size,style,color,alignment,hight,bold,italic,)

Marker Attributes
A particular character is affected by a particular attribute(marker type, marker precision).

Antialiasing

The distortion (or) deformation of information due to low frequency sampling is called aliasing. The aliasing effect can be reduced by adjusting intensities of pixels along the line to minimize the effect of alising is called antialiasing.

We can improve the appearance of displayed raster lines by applying antialisaing methods

AMSCE - 1101

that compensate for the under sampling process.

Methods of antialising

1. Increasing Resolution
2. Unweighted area sampling
3. Weighted area sampling
4. Thick line segments

1. Increasing Resolution

The alising effect can be minimized by increasing resolution of the raster display. By increasing resolution and making it twice the original one, the line passes through twice as many column of pixels and therefore has twice as many jags, but each jag is half as large in x and in y direction.

This improvement cause increase in cost of memory, bandwidth of memory and scan- conversion time. So it is a expensive method to reduce the aliasing method.

2. Unweighted area sampling

In general the line drawing algorithm select the pixels which is closer to the true line. In antialsing instead of picking closest pixel, both pixels are high lighted. However their intensity values may differ.

In unweighted area sampling, the intensity of pixel is proportional to the amount of line area occupied by the pixel. It produces better results than does setting pixels either to full intensity or to zero intensity.

3. Weighted area sampling

In weighted area sampling small area closer to the pixel center has greater intensity than does one at a greater distance. Thus in weighted area sampling the intensity of the pixel is dependent on the line area occupied and the distance of area from the pixel's center.

4. Thick line segment

In raster displays it is possible to draw lines with thickness greater than one pixel. To produce a thick line, we have to run two line drawing algorithms in parallel to find the pixels along the line edges, and while stepping along the line we have to turn on all the pixels which lie between the boundaries.

AMSCE - 1101

UNIT II **TWO DIMENSIONAL GRAPHICS**

Two dimensional geometric transformations – Matrix representations and homogeneous coordinates, composite transformations; Two dimensional viewing – viewing pipeline, viewing coordinate reference frame; widow-to-viewport coordinate transformation, Two dimensional viewing functions; clipping operations – point, line, and polygon clipping algorithms.

TWO DIMENSIONAL GRAPHICS TRANSFORMATIONS

Geometric Transformations

Changes in size, shape are accomplished with geometric transformation. It alter the coordinate descriptions of object.

The basic transformations are Translation, Roatation, Scaling. Other transformations are Reflection and shear. Basic transformations used to reposition and resize the two dimentional objects.

AMSCE - 1101

Two Dimensional Transformations Translation

A Translation is applied to an object by repositioning it along a straight line path from one co-ordinate location to another. We translate a two dimensional point by adding translation distances tx and ty to the original position (x,y) to move the point to a new location (x',y')

X'=x+tx Y'=y+ty

triangle = { p1=(1,0), p2=(2,0), p3=(1.5,2) }

The translation distance pair (tx,ty) is called a translation vector or shift vector.

$$P = \begin{bmatrix} X1 \\ X2 \end{bmatrix} \qquad P' = \begin{bmatrix} X1' \\ X2' \end{bmatrix} \qquad T = \begin{bmatrix} tx \\ ty \end{bmatrix}$$

$$P' = P + T \qquad p' = \begin{bmatrix} X1+tx \\ X2+ty \end{bmatrix}$$

It moves objects without deformation. (ie) Every point on the objet is translated by the same amount. It can be applied to lines, polygons.

Rotation

A two dimensional rotation is applied to an object by repositioning it along a circular path in the xy plane. To generate a rotation, we specify a rotation angle theta and the position (xr,yr) of the rotation point (or pivot point) about which the object is to be rotated.

Positive value of the rotation angle defines counter clock wise rotation. Negative value of the rotation angle defines the clock wise rotation.

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

Using column vector P'=P*R

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Rotation of an arbitary pivot point

Rotation of a point about any specified rotation position (xr,yr)

$$X' = X_r + (X - X_r)\cos\theta - (Y - Y_r)\sin\theta$$
$$Y' = Y_r + (X - X_r)\sin\theta + (Y - Y_r)\cos\theta$$

It moves objects without deformations. Every point on an object is rotated through the same angle.

Scaling

A scaling transformation alters the size of an object. This operation can be carried out for polygon by multiplying the coordinate values (x,y) of each vertex by scaling factors sx and sy to produce the transformed coordinates (x',y').

$$X' = x.sx$$

$$Y' = y.sy$$

$$P = \begin{matrix} X1 \\ X2 \end{matrix} \quad P' = \begin{matrix} X1' \\ X2' \end{matrix} \quad S = \begin{matrix} sx & 0 \\ 0 & sy \end{matrix}$$

$$P' = P*S$$

If sx=sy , then it produces the uniform scaling

Sx<> sy , different scaling.

If sx,sy<0, then it produces the reduced object size

If sx,sy > 0, then it produces the enlarged size objects.

By choosing the position called fixed point, we can control the location of the scaled object. This point is remain unchanged after the scaling transformation.

$$X' = X_f + (X - X_f)sx \Rightarrow X' = X.sx + (X_f(1-sx))$$
$$Y' = Y_f + (Y - Y_f)sy \Rightarrow Y' = Y.sy + Y_f(1-sy)$$

Matrix representations and homogeneous coordinates

Graphics applications involves sequences of geometric transformations. The basic transformations expressed in terms of

$$P' = M1 * P + M2$$

P, P' --> Column vectors.

M1 --> 2 x 2 array containing multiplicative factors

M2 --> 2 Element column matrix containing translation terms

For translation --> M1 is the identity matrix

For rotation or scaling -->M2 contains transnational terms associated with the pivot point or scaling fixed point.

For coordinate positions are scaled, then rotated then translated, these steps are combined together into one step, final coordinate positions are obtained directly from the initial coordinate values.

To do this expand the 2 x 2 matrix into 3 x 3 matrix.

AMSCE - 1101

To express 2 dimensional transformation as a multiplication, we represent each cartesion coordinate position (x,y) with the homogeneous co ordinate triple (Xh,Yh, h) where

X= xh/h, Y=Yh/h

So we can write (h.x, h.y,h), set h=1. Each two dimensional position is represented with homogeneous coordinates(x,y,1). Coordinates are represented with three element column vector. Transformation operations are written as 3 by 3 matrices.

For translation

X’	1	0	tx	X
Y’	0	1	ty	Y
1	0	0	1	1

P’=T(tx,ty)*P

Inverse of the translation matrix is obtained by replacing tx, ty by –tx, -ty

Similarly rotation about the origin

X’	Cosθ	-Sinθ	0	X
Y’	Sinθ	Cosθ	0	Y
1	0	0	1	1

P’= R(θ)*P

We get the inverse rotation matrix when θ is replaced with (-θ)

Similarly scaling about the origin

X’	Sx	0	0	X
Y’	0	Sy	0	Y
1	0	0	1	1

P’= S(sx,sy)*P

Composite transformations

- sequence of transformations is called as composite transformation. It is obtained by forming products of transformation matrices is referred as a concatenation (or) composition of matrices.

Tranlation : -

Two successive translations

1 0 tx	1 0 tx	1 0 tx1+tx2
0 1 ty	0 1 ty	0 1 ty1+ty2
0 0 1	0 0 1	0 0 1

T(tx1,ty1) + T(tx2,ty2) = T(tx1+tx2, ty1+ty2)

Two successive translations are additive.

AMSCE - 1101

Rotation

Two successive rotations are additive.

$R(\theta1)* R(\theta2)= R(\theta1+ \theta2)$

$P' = P. R(\theta1+ \theta2)$

Scaling

Sx1	0	0	Sx2	0	0	Sx1.x2	0	0
0	Sy1	0	0	Sy2	0	0	Sy1.y2	0
0	0	1	0	0	1	0	0	1

S(x1,y1).S(x2,y2) = S(sx1.sx2 , sy1.sy2)

1. the order we perform multiple transforms can matter

- eg. translate + scale can differ from scale + translate
- eg. rotate + translate can differ from translate + rotate
- eg. rotate + scale can differ from scale + rotate (when scale_x differs from scale_y)

2. When does M1 + M2 = M2 + M1?

M1	M2
translate	translate
scale	scale
rotate	rotate
scale (sx = sy)	rotate

General pivot point rotation

Rotation about any selected pivot point (xr,yr) by performing the following sequence of translate – rotate – translate operations.

1. Translate the object so that the pivot point is at the co-ordinate origin.
2. Rotate the object about the coordinate origin
3. Translate the object so that the pivot point is returned to its original position

1	0	xr	Cosθ	-Sinθ	0	1	0	-xr
0	1	yr	Sinθ	Cosθ	0	0	1	-yr
0	0	1	0	0	1	0	0	1

Concatenation properties

$T(xr,yr).R(\theta).T(-xr,-yr) = R(xr,yr, \theta)$

Matrix multiplication is associative. Transformation products may not be commutative. Combination of translations, roatations, and scaling can be expressed as

X'	rSxx	rSxy	trSx	X
Y'	rSyx	rSyy	trSy	Y
1	0	0	1	1

AMSCE - 1101

Other transformations
Besides basic transformations other transformations are reflection and shearing

Reflection :

Reflection is a transformation that produces the mirror image of an object relative to an axis of reflection. The mirror image is generated relative to an axis of reflection by rotating the object by 180 degree about the axis.

Reflection about the line y=0 (ie about the x axis), the x-axis is accomplished with the transformation matrix.

1 0 0

0 -1 0

0 0 1

It keeps the x values same and flips the y values of the coordinate positions. Reflection about the y-axis

-1 0 0
0 1 0
0 0 1

It keeps the y values same and flips the x values of the coordinate positions. Reflection relative to the coordinate origin.

-1 0 0
0 -1 0
0 0 1

Reflection relative to the diagonal line y=x , the matrix is 0 1 0 1 0 0 0 0 1

Reflection relative to the diagonal line y=x , the matrix is 0 -1 0 -1 0 0 0 0 1

Shear

A transformation that alter the shape of an object is called the shear transformation. Two shearing transformations

1. Shift x coordinate values (X- shear)
2. Shifts y coordinate values. (Y-shear)

In both cases only one coordinate (x or y) changes its coordinates and other preserves its values.

X –Shear

It preserves the y value and changes the x value which causes vertical lines to tilt right or left

	1	0 0
X-sh =	shx 1 0	
	0	0 1

X’= X+shx*y
Y’=Y

AMSCE - 1101

Y –Shear

It preserves the x value and changes the y value which causes vertical lines to tilt right or left 1 shy 0

Y-sh = 0 1
0 0 0 1

Y'= Y+shy*X
X'=X

Shearing Relative to other reference line

We can apply x and y shear transformations relative to other reference lines. In x shear transformation we can use y reference line and in y shear we can use x reference line.

The transformation matrices for both are given below.

	1	shx	-shx*yref
X shear with y reference line	0	1	0
	0	0	1

x'=x+shx(y-yref) , y'= y

	1	0	0
Y shear with x reference line	shy	1	-shy*xref
	0	0	1

which generates transformed coordinate positions.
x'=x , y'= shy(x-xref)+y

This transformation shifts a coordinate position vertically by an amount proposal to its distance from the reference line x=x ref.

Transformations between coordinate systems

Transformations between Cartesian coordinate systems are achieved with a sequence of translate-rotate transformations. One way to specify a new coordinate reference frame is to give the position of the new coordinate origin and the direction of the new y-axis. The direction of the new x-axis is then obtained by rotating the y direction vector 90 degree clockwise. The transformation matrix can be calculated as the concatenation of the translation that moves the new origin to the old co-ordinate origin and a rotation to align the two sets of axes. The rotation matrix is obtained from unit vectors in the x and y directions for the new system

TWO DIMENSIONAL GRAPHICS
TRANSFORMATIONS Composite transformations

- sequence of transformations is called as composite transformation. It is obtained by forming products of transformation matrices is referred as a concatenation (or) composition of matrices.

Tranlation : -

Two successive translations

1 0 tx	1 0 tx	1 0 tx1+tx2
0 1 ty	0 1 ty	0 1 ty1+ty2
0 0 1	0 0 1	0 0 1

T(tx1,ty1) + T(tx2,ty2) = T(tx1+tx2, ty1+ty2)
Two successive translations are additive.

AMSCE - 1101

Rotations

Two successive rotations are additive.

$R(\theta1)* R(\theta2)= R(\theta1+ \theta2)$

$P' = P. R(\theta1+ \theta2)$

Scaling

$$\begin{bmatrix} Sx1 & 0 & 0 \\ 0 & Sy1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} Sx2 & 0 & 0 \\ 0 & Sy2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} Sx1.x2 & 0 & 0 \\ 0 & Sy1.y2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$S(x1,y1).S(x2,y2) = S(sx1.sx2 , sy1.sy2)$

the order we perform multiple transforms can matter

eg. translate + scale can differ from scale + translate eg.
rotate + translate can differ from translate + rotate

eg. rotate + scale can differ from scale + rotate (when scale_x differs from scale_y)

When does M1 + M2 = M2 + M1?

M1	M2
translate	translate
Scale	Scale
Rotate	Rotate
Scale(sx=sy)	Rotate

General pivot point rotation

Rotation about any selected pivot point (xr,yr) by performing the following sequence of translate – rotate – translate operations.

$$\begin{bmatrix} 1 & 0 & xr \\ 0 & 1 & yr \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Cos\theta & -Sin\theta & 0 \\ Sin\theta & Cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -xr \\ 0 & 1 & -yr \\ 0 & 0 & 1 \end{bmatrix}$$

$T(xr,yr).R(\theta).T(-xr,-yr) = R(xr,yr, \theta)$

Concatenation properties

Matrix multiplication is associative. Transformation products may not be commutative. Combination of translations, roatations, and scaling can be expressed as

$$\begin{bmatrix} X' \\ Y' \\ 1 \end{bmatrix} \quad \begin{bmatrix} rSxx & rSxy & trSx \\ rSyx & rSyy & trSy \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

Affine transformations

Two dimensional geometric transformations are affine transformations. Ie they can be expressed as a linear function of co-ordinates x and y . Affine transformations transform the parallel lines to parallel lines and transform finite points to finite points . Geometric transformations that do not include scaling or shear also preserve angles and lengths.

AMSCE-1101

Raster methods for transformations

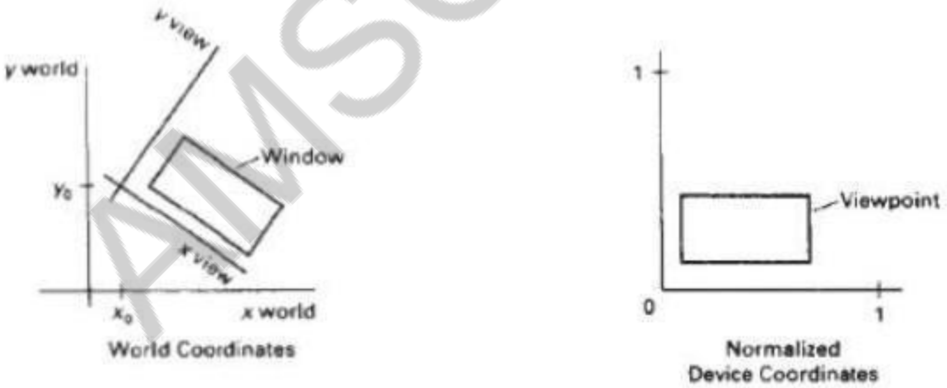
Moving blocks of pixels can perform fast raster transformations. This avoids calculating transformed coordinates for an object and applying scan conversion routines to display the object at the new position. Three common operations (bitBlts or pixBlts(bit block transfer)) are copy, read, and write. when a block of pixels are moved to a new position in the frame buffer(block transfer) , we can simply replace the old pixel values or we can combine the pixel values using Boolean or arithmetic operations. Copying a pixel block to a new location in the frame buffer carries out raster translations. Raster rotations in multiples of 90 degree are obtained by manipulating row and column positions of the pixel values in the block. Other rotations are performed by first mapping rotated pixel areas onto destination positions in the frame buffer, then calculate the overlap areas. Scaling in raster transformation is also accomplished by mapping transformed pixel areas to the frame buffer destination positions.

Two dimensional viewing

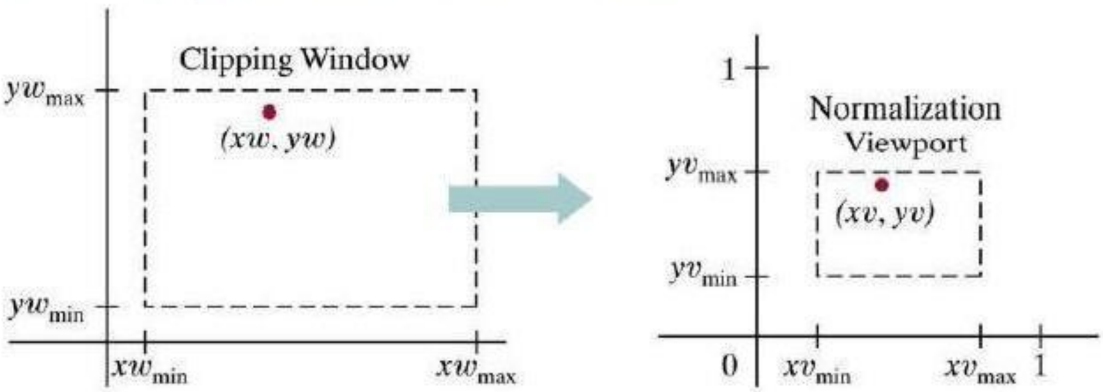
Two dimensional viewing The viewing pipeline A world coordinate area selected for display is called a window. An area on a display device to which a window is mapped is called a view port. The window defines what is to be viewed the view port defines where it is to be displayed. The mapping of a part of a world coordinate scene to device coordinate is referred to as viewing transformation. The two d imensional viewing transformation is referred to as window to view port transformation of windowing transformation.

A viewing transformation using standard rectangles for the window and viewport

The viewing transformation in several steps, as indicated in Fig. First, we construct the scene in world coordinates using the output primitives. Next to obtain a particular orientation for the window, we can set up a two-dimensional viewing-coordinate system in the world coordinate plane, and define a window in the viewing-coordinate system. ctangular windows.



Window to view port coordinate transformation:



AMSCE - 1101

At the final step all parts of the picture that lie outside the viewport are clipped, and the contents of the viewport are transferred to device coordinates. By changing the position of the viewport, we can view objects at different positions on the display area of an output device.

A point at position (xw,yw) in a designated window is mapped to viewport coordinates (xv,yv) so that relative positions in the two areas are the same. The figure illustrates the window to view port mapping. A point at position (xw,yw) in the window is mapped into position (xv,yv) in the associated view port. To maintain the same relative placement in view port as in window

The conversion is performed with the following sequence of transformations.

1. Perform a scaling transformation using point position of (xw min, yw min) that scales the window area to the size of view port.
2. Translate the scaled window area to the position of view port. Relative proportions of objects are maintained if scaling factor are the same(Sx=Sy).

Otherwise world objects will be stretched or contracted in either the x or y direction when displayed on output device. For normalized coordinates, object descriptions are mapped to various display devices. Any number of output devices can be open in particular application and another window view port transformation can be performed for each open output device. This mapping called the work station transformation is accomplished by selecting a window area in normalized apace and a view port are in coordinates of display device.

Window to Viewport transformation

The window defined in world coordinates is first transformed into the normalized device coordinates. The normalized window is then transformed into the viewport coordinate. The window to viewport coordinate transformation is known as workstation transformation. It is achieved by the following steps

1. The object together with its window is translated until the lower left corner of the window is at the orgin.

2. Object and window are scaled until the window has the dimensions of the viewport

3. Translate the viewport to its correct position on the screen.

The relation of the window and viewport display is expressed as

$$\frac{XV-XVmin}{XVmax-XVmin} = \frac{XW-XWmin}{XWmax-XWmin}$$

$$\frac{YV-Yvmin}{YVmax-YVmin} = \frac{YW-YWmin}{YWmax-YWmin}$$

$$XV=XVmin + (XW-XWwmin)Sx$$
$$YV=YVmin + (YW-YWmin)Sy$$

$$Sx= \frac{XVmax-XVmin}{XWmax-Xwmin}$$

AMSCE - 1101

$$Sy = \frac{YVmax - YVmin}{YWmax - YWmin}$$

2D Clipping

The procedure that identifies the portion of a picture that are either inside or outside of a specified regin of space is referred to as clipping. The regin against which an object is to be clipped is called a clip window or clipping window.

The clipping algorithm determines which points, lines or portions of lines lie within the clipping window. These points, lines or portions of lines are retained for display. All other are discarded.

Possible clipping are

1. Point clipping
2. Line clipping
3. Area clipping
4. Curve Clipping
5. Text Clipping

Point Clipping:

The points are said to be interior to the clipping
if XWmin <= X <=XW max

YWmin <= Y <=YW max

The equal sign indicates that points on the window boundary are included within the window.

Line Clipping:

- The lines are said to be interior to the clipping window, if the two end points of the lines are interior to the window.

- If the lines are completely right of, completely to the left of, completely above, or completely below the window, then it is discarded.

- Both end points of the line are exterior to the window, then the line is partially inside and partially outside the window.The lines which across one or more clipping boundaries requires calculation of multiple intersection points to decide the visible portion of them.To minimize the intersection calculation and increase the efficiency of the clipping algorithm, initially completely visible and invisible lines are identified and then intersection points are calculated for remaining lines.

There are many clipping algorithms. They are

1.Sutherland and cohen subdivision line clipping algorithm

It is developed by Dan Cohen and Ivan Sutharland. To speed up the processing this algorithm performs initial tests that reduces the number of intersections that must be calculated.

given a line segment, repeatedly:

AMSCE - 1101

1. check for trival acceptance
   both
2. check for trivial rejection

both endpoints of the same side of clip rectangle
3. both endpoints outside clip rectangle

divide segment in two where one part can be trivially rejected

Clip rectangle extended into a plane divided into 9 regions . Each region is defined by a unique 4-bit string

- left bit = 1: above top edge (Y > Ymax)
- 2nd bit = 1: below bottom edge (Y < Ymin)
- 3rd bit = 1: right of right edge (X > Xmax)
- right bit = 1: left of left edge (X < Xmin)
- left bit = sign bit of (Ymax - Y)
- 2nd bit = sign bit of (Y - Ymin)
- 3rd bit = sign bit of (Xmax - X)
- right bit = sign bit of (X - Xmin)

(the sign bit being the most significant bit in the binary representation of the value. This bit is '1' if the number is negative, and '0' if the number is positive.)

The frame buffer itself, in the center, has code 0000.
1001 | 1000 | 1010

0001 | 0000 | 0010

0101 | 0100 | 0110 For
each line segment:

1. each end point is given the 4-bit code of its region
2. repeat until acceptance or rejection
1. if both codes are 0000 -> trivial acceptance
2. if logical AND of codes is not 0000 -> trivial rejection
3. divide line into 2 segments using edge of clip rectangle
1. find an endpoint with code not equal to 0000
2. lines that cannot be identified as completely inside or outside are checked for the intersection with two boundaries.
3. break the line segment into 2 line segments at the crossed edge
4. forget about the new line segment lying completely outside the clip rectangle
5. draw the line segment which lies within the boundary regin.

2. Mid point subdivision algorithm

If the line partially visible then it is subdivided in two equal parts. The visibility tests are then applied to each half. This subdivision process is repeated until we get completely visible and

AMSCE - 1101

completely invisible line segments.

Mid point sub division algorithm

1. Read two end points of the line P1(x1,x2), P2(x2,y2)

2. Read two corners (left top and right bottom) of the window, say (Wx1,Wy1 and Wx2, Wy2)

3. Assign region codes for two end points using following steps

Initialize code with bits 0000
Set Bit 1 – if (x < Wx1) Set
Bit 2 – if (x > Wx1) Set Bit
3 – if (y < Wy1)
Set Bit 4 – if (y > Wy2)

4. Check for visibility of line

a. If region codes for both endpoints are zero then the line is completely visible. Hence draw the line and go to step 6.

b. If the region codes for endpoints are not zero and the logical ANDing of them is also nonzero then the line is completely invisible, so reject the line and go to step6

c. If region codes for two end points do not satisfy the condition in 4a and 4b the line is partially visible.

5. Divide the partially visible line segments in equal parts and repeat steps 3 through 5 for both subdivided line segments until you get completely visible and completely invisible line segments.

6. Stop.

This algorithm requires repeated subdivision of line segments and hence many times it is slower than using direct calculation of the intersection of the line with the clipping window edge.

3. Liang-Barsky line clipping algorithm

The cohen Sutherland clip algorithm requires the large no of intesection calculations.here this is reduced. The update parameter requires only one division and windows intersection lines are computed only once.

The parameter equations are given as

X=x1+u x, Y=Y1 + u y

0<=u<=1, where x =x2-x1 , u y=y2-y1

Algorithm

1. Read the two end points of the line p1(x,y),p2(x2,y2)

2. Read the corners of the window (xwmin,ywmax), (xwmax,ywmin)

3. Calculate the values of the parameter p1,p2,p3,p4 and q1,q2,q3,q4m such that

4. p1= x q1=x1-xwmin

p2= **-** x q2=xwmax-x1
p3= y q3=y1-ywmin
p4= **-** y q4=ywmax-y1

5. If pi=0 then that line is parallel to the ith boundary. if qi<0 then the line is completely outside the boundary. So discard the linesegment and and goto stop.

AMSCE - 1101

Else

{

Check whether the line is horizontal or vertical and check the line endpoint with the corresponding boundaries. If it is within the boundary area then use them to draw a line. Otherwise use boundary coordinate to draw a line. Goto stop.

}

6. initialize values for U1 and U2 as U1=0,U2=1

7. Calculate the values forU= qi/pi for I=1,2,3,4

8. Select values of qi/pi where pi<0 and assign maximum out of them as u1

9. If (U1<U2)

{

Calculate the endpoints of the clipped line as follows

XX1=X1+u1 x

XX2=X1+u 2 x

YY1=Y1+u1 y

YY2=Y1+u 2 y

}

10.Stop.

4. Nicholl-lee Nicholl line clipping

It Creates more regions around the clip window. It avoids multiple clipping of an individual line segment. Compare with the previous algorithms it perform few comparisons and divisions . It is applied only 2 dimensional clipping. The previous algorithms can be extended to 3 dimensional clipping.

1. For the line with two end points p1,p2 determine the positions of a point for 9 regions. Only three regions need to be considered (left,within boundary, left upper corner).

2. If p1 appears any other regions except this, move that point into this region using some reflection method.

3. Now determine the position of p2 relative to p1. To do this depends on p1 creates some new region.

a. If both points are inside the region save both points.

b. If p1 inside , p2 outside setup 4 regions. Intersection of appropriate boundary is calculated depends on the position of p2.

c. If p1 is left of the window, setup 4 regions . L, Lt,Lb,Lr

1. If p2 is in region L, clip the line at the left boundary and save this intersection to p2.

2. If p2 is in region Lt, save the left boundary and save the top boundary.

3. If not any of the 4 regions clip the entire line.

d. If p1 is left above the clip window, setup 4 regions . T, Tr,Lr,Lb

1. If p2 inside the region save point.

2. else determine a unique clip window edge for the intersection calculation.

e. To determine the region of p2 compare the slope of the line to the slope of the boundaries of the clip regions.

AMSCE - 1101

Line clipping using non rectangular clip window

Circles and other curved boundaries clipped regions are possible, but less commonly used. Clipping algorithm for those curve are slower.

1. Lines clipped against the bounding rectangle of the curved clipping region. Lines outside the region is completely discarded.

2. End points of the line with circle center distance is calculated . If the squre of the 2 points less than or equal to the radious then save the line else calculate the intersection point of the line.

Polygon clipping

Splitting the concave polygon

It uses the vector method , that calculate the edge vector cross products in a counter clock wise order and note the sign of the z component of the cross products. If any z component turns out to be negative, the polygon is concave and we can split it along the line of the first edge vector in the cross product pair.

Sutherland – Hodgeman polygon Clipping Algorithm

1. Read the coordinates of all vertices of the polygon.
2. Read the coordinates of the clipping window.
3. Consider the left edge of the window.
4. Compare the vertices of each edge of the polygon, Individually with the clipping plane.
5. Save the resulting intersections and vertices in the new list of vertices according to four possible relationships between the edge and the clipping boundary discussed earlier.
6. Repeats the steps 4 and 5 for remaining edges of the clipping window. Each time the resultant vertices is successively passed the next edge of the clipping window.
7. Stop.

The Sutherland –Hodgeman polygon clipping algorithm clips convex polygons correctly, But in case of concave polygons clipped polygon may be displayed with extraneous lines. It can be solved by separating concave polygon into two or more convex polygons and processing each convex polygons separately.
The following example illustrate a simple case of polygon clipping.

WEILER –Atherton Algorithm

Instead of proceding around the polygon edges as vertices are processed, we sometime wants to follow the window boundaries.For clockwise processing of polygon vertices, we use the following rules.
- For an outside to inside pair of vertices, follow the polygon boundary.
- For an inside to outside pair of vertices, follow a window boundary in a clockwise direction.

Curve Clipping

It involves non linear equations. The boundary rectangle is used to test for overlap with a rectangular clipwindow. If the boundary rectangle for the object is completely inside the window , then save the object (or) discard the object.If it fails we can use the coordinate extends of individual quadrants and then octants for preliminary testing before calculating

AMSCE - 1101

curve window intersection.

Text Clipping

The simplest method for processing character strings relative to a window boundary is to use the all or none string clipping strategy. If all the string is inside then accept it else omit it.

We discard only those character that are not completely inside the window. Here the boundary limits of individual characters are compared to the window.

Exterior clipping

The picture part to be saved are those that are outside the region. This is referred to as exterior clipping. An application of exterior clipping is in multiple window systems.

Objects within a window are clipped to the interior of that window. When other higher priority windows overlap these objects , the ojects are also clipped to the exterior of the overlapping window.

AMSCE - 1101

UNIT III THREE DIMENSIONAL GRAPHICS

Three dimensional concepts; Three dimensional object representations – Polygon surfaces- Polygon tables- Plane equations – Polygon meshes; Curved Lines and surfaces, Quadratic surfaces; Blobby objects; Spline representations – Bezier curves and surfaces -B-Spline curves and surfaces. TRANSFORMATION AND VIEWING: Three dimensional geometric and modeling transformations – Translation, Rotation, Scaling, composite transformations; Three dimensional viewing – viewing pipeline, viewing coordinates, Projections, Clipping; Visible surface detection methods.

LP 1: Three Dimensional Object Representations

Representation schemes for solid objects are divided into two categories as follows: 1. Boundary Representation (B-reps)

It describes a three dimensional object as a set of surfaces that separate the object interior from the environment. Examples are polygon facets and spline patches.

2. Space Partitioning representation

It describes the interior properties, by partitioning the spatial region containing an object into a set of small, nonoverlapping, contiguous solids(usually cubes). Eg: Octree Representation.

Polygon Surfaces

Polygon surfaces are boundary representations for a 3D graphics object is a set of polygons that enclose the object interior.

Polygon Tables

The polygon surface is specified with a set of vertex coordinates and associated attribute parameters.

For each polygon input, the data are placed into tables that are to be used in the subsequent processing.

Polygon data tables can be organized into two groups: Geometric tables and attribute tables.

Geometric Tables Contain vertex coordinates and parameters to identify the spatial orientation of the polygon surfaces.

Attribute tables Contain attribute information for an object such as parameters specifying the degree of transparency of the object and its surface reflectivity and texture characteristics. A convenient organization for storing geometric data is to create three lists:

1. The Vertex Table

Coordinate values for each vertex in the object are stored in this table. 2. The Edge Table

It contains pointers back into the vertex table to identify the vertices for each polygon edge.

3. The Polygon Table

It contains pointers back into the edge table to identify the edges for each polygon.

AMSCE - 1101

Listing the geometric data in three tables provides a convenient reference to the individual components (vertices, edges and polygons) of each object.

The object can be displayed efficiently by using data from the edge table to draw the component lines.

Extra information can be added to the data tables for faster information extraction. For instance, edge table can be expanded to include forward points into the polygon table so that common edges between polygons can be identified more rapidly.

E1 : V1, V2, S1

E2 : V2, V3, S1

E3 : V3, V1, S1, S2

E4 : V3, V4, S2

E5 : V4, V5, S2

E6 : V5, V1, S2

is useful for the rendering procedure that must vary surface shading smoothly across the edges from one polygon to the next. Similarly, the vertex table can be expanded so that vertices are cross-referenced to corresponding edges.

Additional geometric information that is stored in the data tables includes the slope for each edge and the coordinate extends for each polygon. As vertices are input, we can calculate edge slopes and we can scan the coordinate values to identify the minimum and maximum x, y and z values for individual polygons.

The more information included in the data tables will be easier to check for errors. Some of the tests that could be performed by a graphics package are:

1. That every vertex is listed as an endpoint for at least two edges.
2. That every edge is part of at least one polygon.
3. That every polygon is closed.
4. That each polygon has at least one shared edge.
5. That if the edge table contains pointers to polygons, every edge referenced by a polygon pointer has a reciprocal pointer back to the polygon.

Plane Equations:

To produce a display of a 3D object, we must process the input data representation for the object through several procedures such as,

- Transformation of the modeling and world coordinate descriptions to viewing coordinates.
- Then to device coordinates:
- Identification of visible surfaces
- The application of surface-rendering procedures.

For these processes, we need information about the spatial orientation of the individual surface components of the object. This information is obtained from the vertex coordinate value and the equations that describe the polygon planes.

The equation for a plane surface is

$Ax + By + Cz + D = 0$ ----(1)

AMSCE - 1101

Where (x, y, z) is any point on the plane, and the coefficients A,B,C and D are constants describing the spatial properties of the plane.

We can obtain the values of A, B,C and D by solving a set of three plane equations using the coordinate values for three non collinear points in the plane.

For that, we can select three successive polygon vertices (x_1, y_1, z_1), (x_2, y_2, z_2) and (x_3, y_3, z_3) and solve the following set of simultaneous linear plane equations for the ratios A/D, B/D and C/D.

$(A/D)x_k + (B/D)y_k + (c/D)z_k = -1, \quad k=1,2,3$ -----(2)

The solution for this set of equations can be obtained in determinant form, using Cramer's rule as

$$A=\begin{vmatrix}1 & y_1 & z_1\\ 1 & y_2 & z_2\\ 1 & y_3 & z_3\end{vmatrix} \quad B=\begin{vmatrix}x_1 & 1 & z_1\\ x_2 & 1 & z_2\\ x_3 & 1 & z_3\end{vmatrix} \quad C=\begin{vmatrix}x_1 & y_1 & 1\\ x_2 & y_2 & 1\\ x_3 & y_3 & 1\end{vmatrix} \quad D=\begin{vmatrix}x_1 & y_1 & z_1\\ x_2 & y_2 & z_2\\ x_3 & y_3 & z_3\end{vmatrix} \quad \text{----(3)}$$

Expanding the determinants , we can write the calculations for the plane coefficients in the form: $A = y_1 (z_2 - z_3) + y_2(z_3 - z_1) + y_3 (z_1 - z_2)$

$B = z_1 (x_2 - x_3) + z_2 (x_3 - x_1) + z_3 (x_1 - x_2)$

$C = x_1 (y_2 - y_3) + x_2 (y_3 - y_1) + x_3 (y_1 - y_2)$

$D = -x_1 (y_2 z_3 - y_3 z_2) - x_2 (y_3 z_1 - y_1 z_3) - x_3 (y_1 z_2 - y_2 z_1)$ ------(4)

As vertex values and other information are entered into the polygon data structure, values for A, B, C and D are computed for each polygon and stored with the other polygon data.

Plane equations are used also to identify the position of spatial points relative to the plane surfaces of an object. For any point (x, y, z) hot on a plane with parameters A,B,C,D, we have $Ax + By + Cz + D \neq 0$

We can identify the point as either inside or outside the plane surface according o the sigh (negative or positive) of $Ax + By + Cz + D$:

If $Ax + By + Cz + D < 0$, the point (x, y, z) is inside the surface. If $Ax + By + Cz + D > 0$, the point (x, y, z) is outside the surface.

These inequality tests are valid in a right handed Cartesian system, provided the plane parmeters A,B,C and D were calculated using vertices selected in a counter clockwise order when viewing the surface in an outside-to-inside direction.

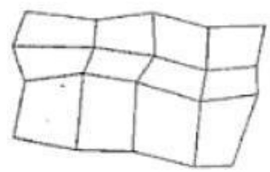
AMSCE - 1101

Polygon Meshes

A single plane surface can be specified with a function such as **fillArea**. But when object surfaces are to be tiled, it is more convenient to specify the surface facets with a mesh function. One type of polygon mesh is the **triangle strip**.A triangle strip formed with 11 triangles connecting 13 vertices.

This function produces n-2 connected triangles given the coordinates for n vertices.

Another similar function in the **quadrilateral mesh**, which generates a mesh of (n-1) by (m-1) quadrilaterals, given the coordinates for an n by m array of vertices. Figure shows 20 vertices forming a mesh of 12 quadrilaterals.



Curved Lines and Surfaces Displays of three dimensional curved lines and surface can be generated from an input set of mathematical functions defining the objects or from a set of user specified data points. When functions are specified, a package can project the defining equations for a curve to the display plane and plot pixel positions along the path of the projected function. For surfaces, a functional description in decorated to produce a polygon-mesh approximation to the surface.

Spline Representations A Spline is 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 drawn.

The **Spline curve** refers to any sections curve formed with polynomial sections satisfying specified continuity conditions at the boundary of the pieces.

A **Spline surface** can be described with two sets of orthogonal spline curves. Splines are used in graphics applications to design curve and surface shapes, to digitize drawings for computer storage, and to specify animation paths for the objects or the camera in the scene. CAD applications for splines include the design of automobiles bodies, aircraft and spacecraft surfaces, and ship hulls.

Interpolation and Approximation Splines Spline curve can be specified by a set of coordinate positions called **control points** which indicates the general shape of the curve. These control points are fitted with piecewise continuous parametric polynomial functions in one of the two ways. 1. When polynomial sections are fitted so that the curve passes through each control point the resulting curve is said to **interpolate** the set of control points.

AMSCE - 1101

A set of six control points interpolated with piecewise continuous polynomial sections

1. When the polynomials are fitted to the general control point path without necessarily passing through any control points, the resulting curve is said to **approximate** the set of control points.

A set of six control points approximated with piecewise continuous polynomial sections

Interpolation curves are used to digitize drawings or to specify animation paths. Approximation curves are used as design tools to structure object surfaces. A spline curve is designed , modified and manipulated with operations on the control points.The curve can be translated, rotated or scaled with transformation applied to the control points. The convex polygon boundary that encloses a set of control points is called the **convex hull**. The shape of the convex hull is to imagine a rubber band stretched around the position of the control points so that each control point is either on the perimeter of the hull or inside it. **Convex hull shapes (dashed lines) for two sets of control points**

AMSCE - 1101

Parametric Continuity Conditions

For a smooth transition from one section of a piecewise parametric curve to the next various **continuity conditions** are needed at the connection points.

If each section of a spline in described with a set of parametric coordinate functions or the form
x = x(u), y = y(u), z = z(u), u1<= u <= u2 -----(a)

We set **parametric continuity** by matching the parametric derivatives of adjoining curve sections at their common boundary.

Zero order parametric continuity referred to as C0 continuity, means that the curves meet. (i.e) the values of x,y, and z evaluated at u2 for the first curve section are equal. Respectively, to the value of x,y, and z evaluated at u1 for the next curve section.

First order parametric continuity referred to as C1 continuity means that the first parametric derivatives of the coordinate functions in equation (a) for two successive curve sections are equal at their joining point.

Second order parametric continuity, or C2 continuity means that both the first and second parametric derivatives of the two curve sections are equal at their intersection.

Higher order parametric continuity conditions are defined similarly.

Piecewise construction of a curve by joining two curve segments using different orders of continuity

a) Zero order continuity only

b) First order continuity only

c) Second order continuity only

Geometric Continuity Conditions

To specify conditions for geometric continuity is an alternate method for joining two successive curve sections.

The parametric derivatives of the two sections should be proportional to each other at their common boundary instead of equal to each other.

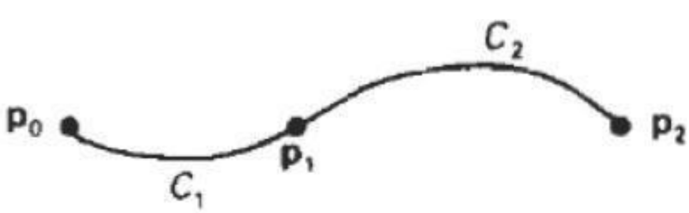
Zero order Geometric continuity referred as G0 continuity means that the two curves sections must have the same coordinate position at the boundary point.

First order Geometric Continuity referred as G1 continuity means that the parametric first derivatives are proportional at the interaction of two successive sections.

Second order Geometric continuity referred as G2 continuity means that both the first and second parametric derivatives of the two curve sections are proportional at their boundary. Here the curvatures of two sections will match at the joining position.

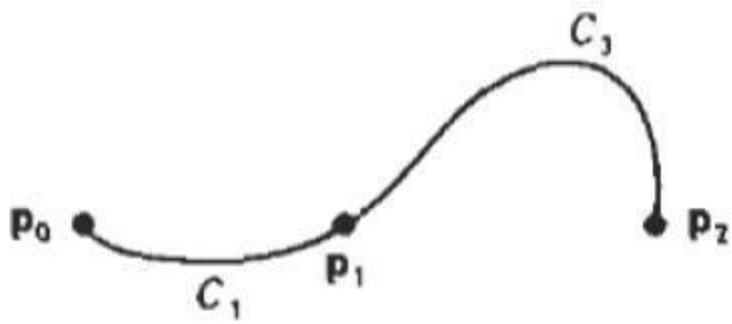
Three control points fitted with two curve sections joined with

a) parametric continuity



AMSCE - 1101

b)geometric continuity where the tangent vector of curve C3 at point p1 has a greater magnitude than the tangent vector of curve C1 at p1.



(b)

Spline specifications There are three methods to specify a spline representation:

1. We can state the set of boundary conditions that are imposed on the spline; (or)
2. We can state the matrix that characterizes the spline; (or)
3. We can state the set of **blending functions** that determine how specified geometric constraints on the curve are combined to calculate positions along the curve path.

To illustrate these three equivalent specifications, suppose we have the following parametric cubic polynomial representation for the x coordinate along the path of a spline section.

x(u)=axu3 + axu2 + cxu + dx 0<= u <=1 ----------(1) Boundary conditions for this curve might be set on the endpoint coordinates x(0) and x(1) and on the parametric first derivatives at the endpoints x'(0) and x'(1). These boundary conditions are sufficient to determine the values of the four coordinates ax, bx, cx and dx. From the boundary conditions we can obtain the matrix that characterizes this spline curve by first rewriting eq(1) as the matrix product

x(u) = [u3 u2 u1 1] ax

bx

cx -------(2

) dx

= U.C

where U is the row matrix of power of parameter u and C is the coefficient column matrix. Using equation (2) we can write the boundary conditions in matrix form and solve for the coefficient matrix C as

C = Mspline . Mgeom -----(3) Where Mgeom in a four element column matrix containing the geometric constraint values on the spline and Mspline in the 4 * 4 matrix that transforms the geometric constraint values to the polynomial coefficients and provides a characterization for the spline curve.

Matrix Mgeom contains control point coordinate values and other geometric constraints.
We can substitute the matrix representation for C into equation (2) to obtain.

x (u) = U . Mspline . Mgeom ------(4)

The matrix Mspline, characterizing a spline representation, called the **basis matriz** is useful for transforming from one spline representation to another.

Finally we can expand equation (4) to obtain a polynomial representation for coordinate x in

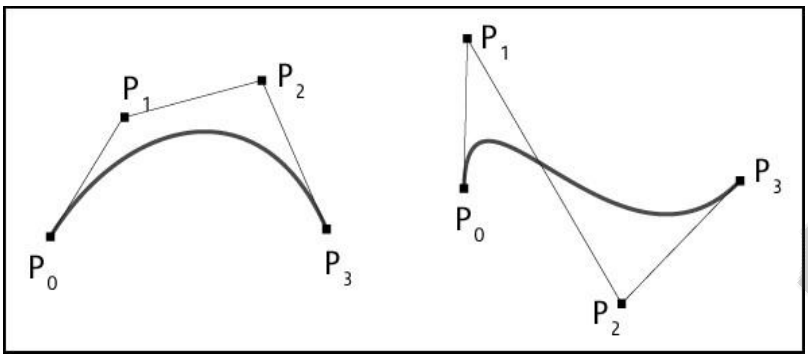
AMSCE - 1101

terms of the geometric constraint parameters.

x(u) = **Σ gk. BFk(u)** where gk are the constraint parameters, such as the control point coordinates and slope of the curve at the control points and BFk(u) are the polynomial blending functions.

Parametric equation for a Beizer curve:

Beizer curve can be fitted to any number of control points. The number of control points to be approximated and their relative positions determine the degree of the Beizer polynomial.



Cubic Bezier curves:

Cubic Beizer curves are generated with four control points. The four blending functions for cubic curves are

$$BEZ_{0,3}(u) = (1-u)^3$$

$$BEZ_{1,3}(u) = 3u(1-u)^2$$

$$BEZ_{2,3}(u) = 3u^2(1-u)$$

$$BEZ_{3,3}(u) = u^3$$

At the end positions of the cubic Bezier curve, the parametric first derivatives (slopes) are

$$P'(0) = 3(P_1 - P_0), \qquad P'(1) = 3(P_3 - P_2)$$

And the parametric second derivatives are

$$P''(0) = 6(P_0 - 2P_1 + P_2), \qquad P''(1) = 6(P_1 - 2P_2 + P_3)$$

By expanding the polynomial expression for blending functions

$$P(u) = [u^3\ u^2\ u\ 1].\, M_{Bez} . \begin{bmatrix} P_0 \\ P_1 \\ P_2 \\ P_3 \end{bmatrix}$$

Where the **Bezier matrix** is

AMSCE - 1101

$$M_B = \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

i) B- Spline Curves:

A general expression for the calculation of coordinate positions along a B-spline curve in a blending function formulation as

$$P(u) = \sum_{k=0}^{n} p_k B_{k,d}\ (u), u_{min} \leq u \leq u_{max}\ , 2 \leq d \leq n+1$$

Where p_k are an input set of n+ 1 control points and spline blending functions $B_{k,d}$ are polynomials of degree d-1, d can be chosen any integer from 2 up to n+1 control points. Blending functions for B-spline curves are define by Cox-deBoor recursion formulas:

$$B_{k,1}\ (u) = \begin{cases} 1, if\ u_{k \leq u \leq u_{K+1}} \\ 0, otherwise \end{cases}$$

$$B_{k,d}\ (u) = \frac{u - u_k}{u_{k+d-1} - u_k} B_{k,d-1}(u) + \frac{u_{k+d} - u}{u_{k+d} - u_{k+1}} B_{k+1,d-1}(u)$$

Properties:

- Each basis function has precisely one maximum value, except for k=1.
- The maximum order of the curve is equal to the number of vertices of defining polygon.

AMSCE - 1101

- The degree of B-spline polynomial is independent on the number of vertices of defining polygon.
- B-spline allows the local control over the curve surface because each vertex affects the shape of a curve only over a range of parameter values where its associated basis function is nonzero.
- The curve exhibits the variation diminishing property.
- The curve generally follows the shape of defining polygon.
- Any affine transformation can be applied to the curve by applying it to the vertices of defining polygon.
- The curve line within the convex hull of its defining polygon.

ii) Spline representation:

- Spline curve refers to any composite curve formed with polynomial sections satisfying specified continuity conditions at the boundary of the pieces.
- A spline surface is generated with two sets of orthogonal spline curves.

Interpolation and Approximation curves:

Spline curve is specified by set of coordinate positions called control points, which indicates the general shape of the curve.

- When polynomial sections are fitted so that the curve passes through each control point, the resulting curve is said to interpolate the set of control points.
- When the polynomials are fitted to the general control path without necessarily passing through any control point, the resulting curve is said to approximate the set of control points.
- A spline curve is defined, modified and manipulated with operations on control points

AMSCE - 1101

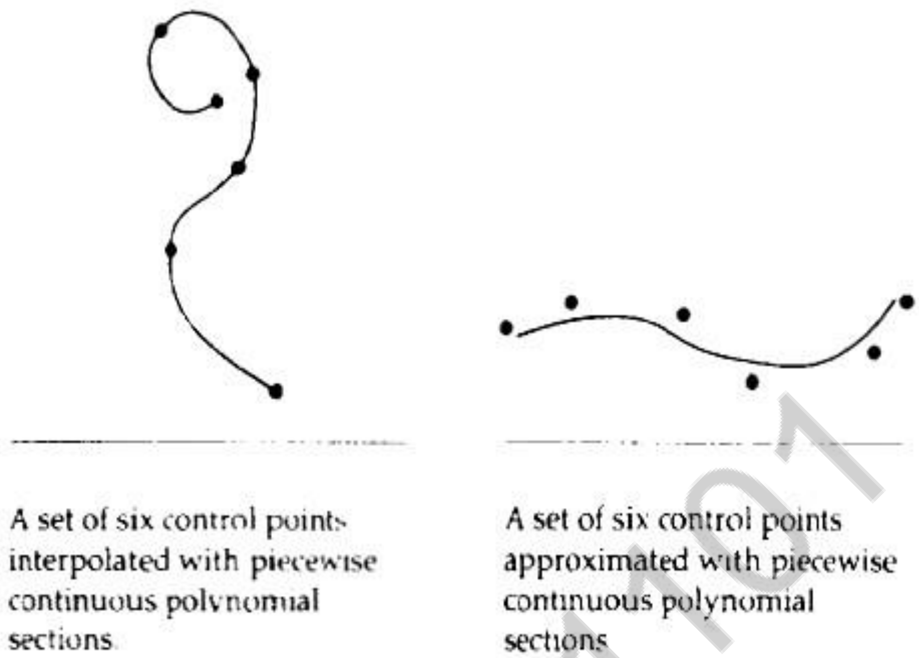


Figure 3.13 Spline through control points

AMSCE - 1101

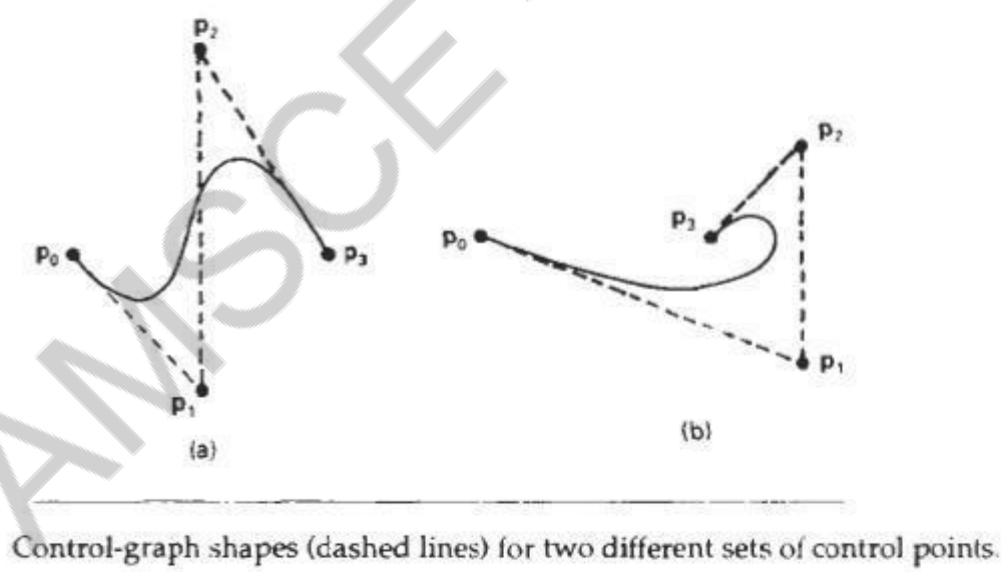


Figure 3.14 Convex hull

- The convex polygon boundary that encloses a set of control points is called convex hull. A way to envision the shape of convex hull is to imagine rubber band stretched around the positions of the control points

Parametric continuity conditions:

To ensure smooth transition continuity conditions at connection points can be imposed.

$x = x(u), y = y(u), z = z(u), u_1 \le u \le u_2$

Figure 3.15 Parametric continuity

- Zero order parametric continuity means simply that the curves meet
- First order parametric continuity referred as C^1 means that the first parametric derivatives are equal at the joining point.
- Second order parametric continuity referred as C^2 means that both first and second order parametric derivatives are same at intersection.

Spline specifications:

There are three equivalent methods for specifying spline representation

- The set of boundary conditions that are imposed on the spline are stated.
- The matrixes that characterize the spline are stated.
- The set of blending functions are stated.

AMSCE - 1101

Back Face detection:

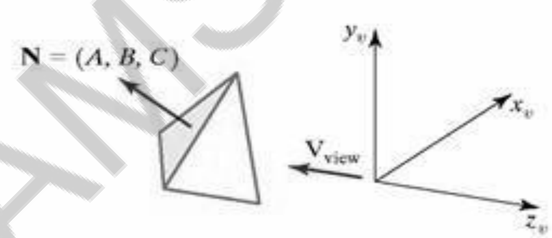
- A fast and simple object-space method for identifying the back faces of a polyhedron is based on the "inside-outside" tests.

- A point (x, y, z) is "inside" a polygon surface with plane parameters A, B, C, and D if When an inside point is along the line of sight to the surface, the polygon must be a back face
- This test can be done by considering the normal vector **N** to a polygon surface, which has Cartesian components (A, B, C).
- In general, if V is a vector in the viewing direction from the eye (or "camera") position, then this polygon is a back face if

  V.N > 0
- Furthermore, if object descriptions are converted to projection coordinates and your viewing direction is parallel to the viewing z-axis, then

  $V = (0, 0, V_z)$ and $V.N = V_Z C$
- In a right-handed viewing system with viewing direction along the negative Z_V axis, the polygon is a back face if $C < 0$.

  Also, any face cannot be seen whose normal has z component C = 0, since your viewing



A polygon surface with plane parameter $C < 0$ in a right-handed viewing coordinate system is identified as a back face when the viewing direction is along the negative z_v axis.

direction is towards that polygon.

Figure 3.6 polygon surface with plane parameters

AMSCE - 1101

Depth Buffer Method:

- It is an image-space approach. The basic idea is to test the Z-depth of each surface to determine the closest (visible) surface.
- In this method each surface is processed separately one pixel position at a time across the surface. The depth values for a pixel are compared and the closest (smallest y) surface determines the color to be displayed in the frame buffer.
- It is applied very efficiently on surfaces of polygon. Surfaces can be processed in any order. To override the closer polygons from the far ones, two buffers named frame buffer and depth buffer are used.
- **Depth buffer** is used to store depth values for (x, y) position, as surfaces are processed ($0 \le depth \le 1$).
- The **frame buffer** is used to store the intensity value of color value at each position (x, y).
- The z-coordinates are usually normalized to the range [0, 1]. The 0 value for z-coordinate indicates back clipping pane and 1 value for z-coordinates indicates front clipping pane.

AMSCE - 1101

Algorithm

1. Initialize the depth and refresh buffer

   Depthbuffer (x, y) = 0

   Framebuffer (x, y) = background color

2. For each position on each polygon surface, compare depth values to previously stored values in the depth buffer to determine visibility.

   For each projected (x, y) pixel position of a polygon, calculate depth z. If Z > depthbuffer (x, y)

   set depthbuffer (x, y) = z,

   framebuffer (x, y) = surfacecolor (x, y)

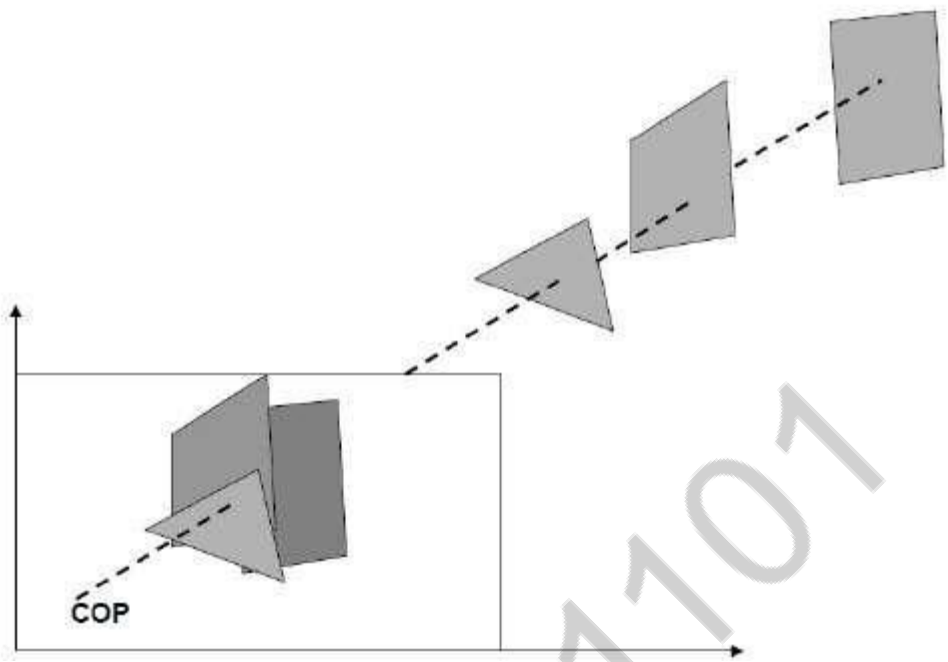


Figure 3.7 A view plane position

A buffer Method:

- The A-buffer method is an extension of the depth-buffer method.
- The A-buffer method is a visibility detection method developed at Lucas film Studios for the rendering system Renders Everything You Ever Saw (REYES).
- The A-buffer expands on the depth buffer method to allow transparencies. The key data structure in the A-buffer is the accumulation buffer.
- Each position in the A-buffer has two fields −
- **Depth field** − It stores a positive or negative real number
- **Intensity field** − It stores surface-intensity information or a pointer value

If depth >= 0, the number stored at that position is the depth of a single surface overlapping the corresponding pixel area. The intensity field then stores the RGB components of the surface color at that point and the percent of pixel coverage.

AMSCE - 1101

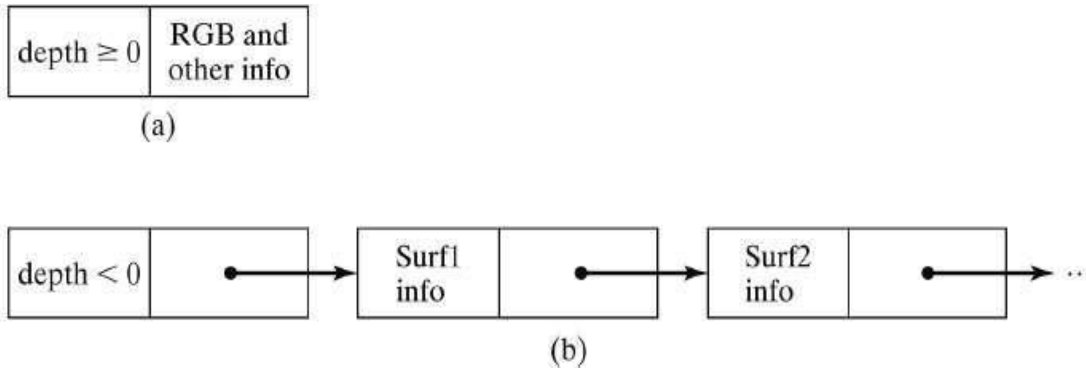


Fig.3.8 : Organization of an A buffer pixel position (a) Single-surface overlap (b) Multiple-surface overlap

- If depth < 0, it indicates multiple-surface contributions to the pixel intensity. The intensity field then stores a pointer to a linked list of surface data. The surface buffer in the A-buffer includes
  - RGB intensity components
  - Opacity Parameter
  - Depth
  - Percent of area coverage
  - Surface identifier

AMSCE - 1101

Scan line Method:

- It is an image-space method to identify visible surface.
- This method has depth information for only single scan-line. In order to require one scan-line of depth values, it is necessary to group and process all polygons intersecting a given scan-line at the same time before processing the next scan-line.
- Two important tables -edge table and polygon table**,** are maintained for this.
- **The Edge Table** − It contains coordinate endpoints of each line in the scene, the inverse slope of each line, and pointers into the polygon table to connect edges to surfaces.

- **The Polygon Table** − It contains the plane coefficients, surface material properties, other surface data, and may be pointers to the edge table.

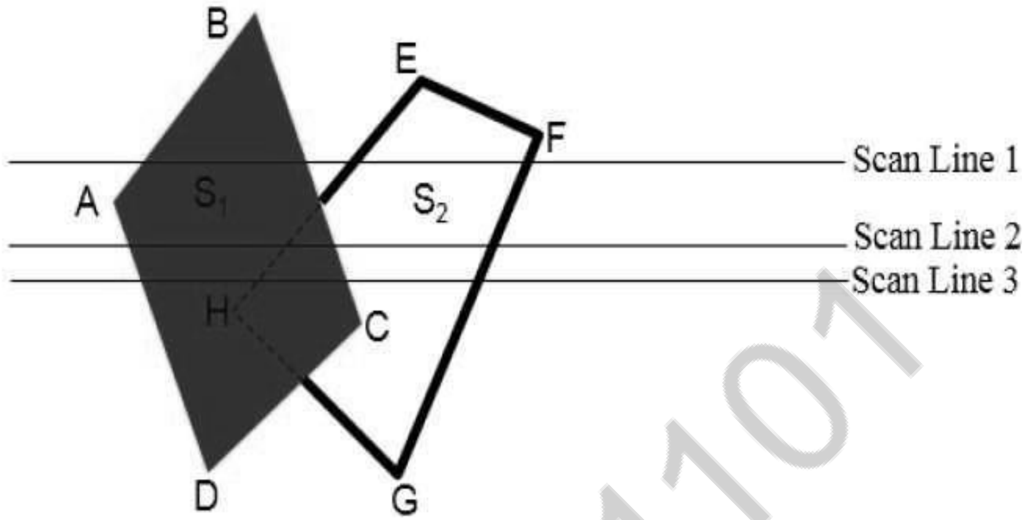


Fig.3.9 scan line intersecting a polygon surface

- To facilitate the search for surfaces crossing a given scan-line, an active list of edges is formed. The active list stores only those edges that cross the scan-line in order of increasing x.
- Also a flag is set for each surface to indicate whether a position along a scan-line is either inside or outside the surface.

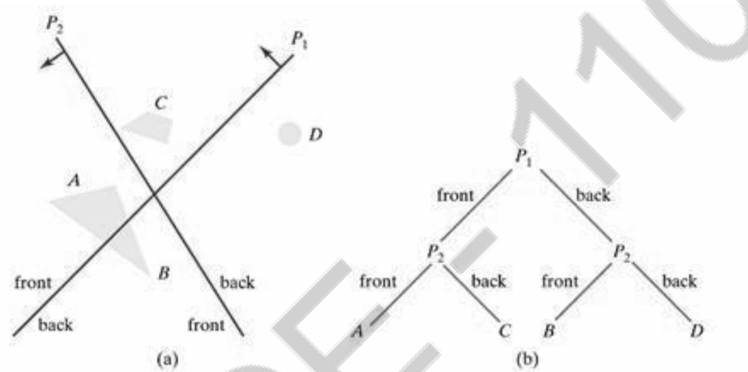
Depth Sorting Method:

- Depth sorting method uses both image space and object-space operations. The depth-sorting method performs two basic functions
- First, the surfaces are sorted in order of decreasing depth.
- Second, the surfaces are scan-converted in order, starting with the surface of greatest depth.
- The scan conversion of the polygon surfaces is performed in image space. This method for solving the hidden-surface problem is often referred to as the painter's algorithm.

AMSCE - 1101

BSP Tree Method:

- Binary space partitioning is used to calculate visibility.
- To build the BSP trees, one should start with polygons and label all the edges. Dealing with only one edge at a time, extend each edge so that it splits the plane in two. Place the first edge in the tree as root.
- Add subsequent edges based on whether they are inside or outside. Edges that span the extension of an edge that is already in the tree are split into two and both are added to the tree.



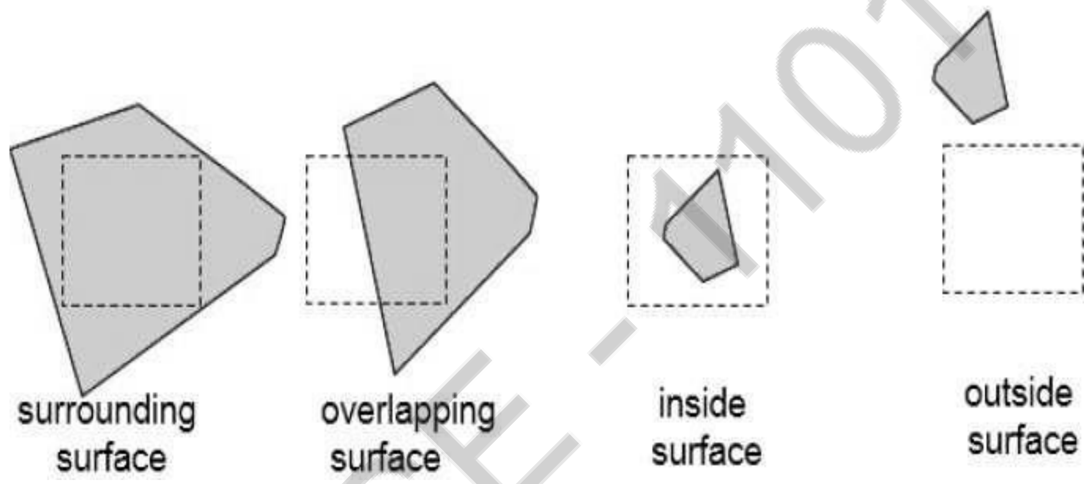
A region of space (a) is partitioned with two planes P_1 and P_2 to form the BSP tree representation shown in (b).

Area Subdivision Method:

- The area-subdivision method takes advantage by locating those view areas that represent part of a single surface.
- Divide the total viewing area into smaller and smaller rectangles until each small area is the projection of part of a single visible surface or no surface at all.
- Continue this process until the subdivisions are easily analyzed as belonging to a single surface or until they are reduced to the size of a single pixel.

AMSCE-1101

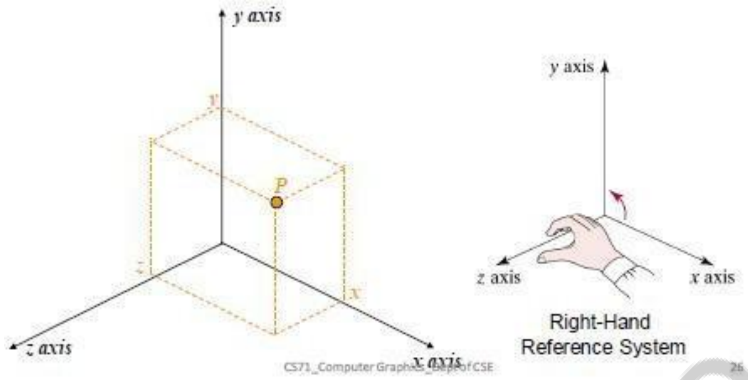
- An easy way to do this is to successively divide the area into four equal parts at each step. There are four possible relationships that a surface can have with a specified area boundary.
  - **Surrounding surface** − One that completely encloses the area.
  - **Overlapping surface** − One that is partly inside and partly outside the area.
  - **Inside surface** − One that is completely inside the area.
  - **Outside surface** − One that is completely outside the area.



AMSCE - M01

LP 2: 3D Transformation

3-D Coordinate Spaces

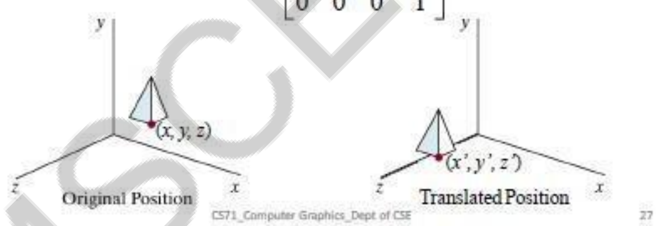


Translations In 3-D

To translate a point in three dimensions by dx, dy and dz .calculate the new points as follows:

$$x' = x + dx \quad y' = y + dy \quad z' = z + dz$$

$$T(d_x, d_y, d_z) = \begin{bmatrix} 1 & 0 & 0 & d_x \\ 0 & 1 & 0 & d_y \\ 0 & 0 & 1 & d_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



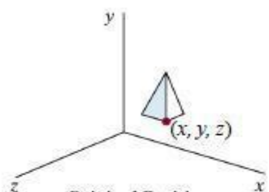
AMSCE - 1101

Scaling In 3-D

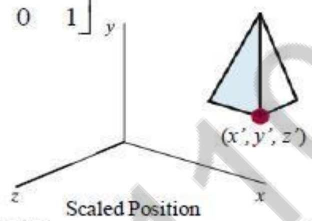
To sale a point in three dimensions by *sx*, *sy* and *sz* simply calculate the new points as follows:

$x' = sx*x \qquad y' = sy*y \qquad z' = sz*z$

$$S(s_x, s_y, s_z) = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



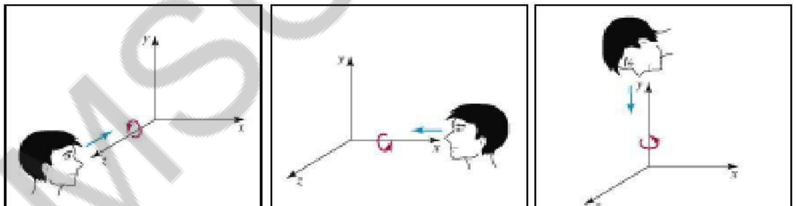
Original Position



Scaled Position

Rotations In 3-D (cont…)

The equations for the three kinds of rotations in 3-D are as follows: (Assume – X→Y→Z→X)

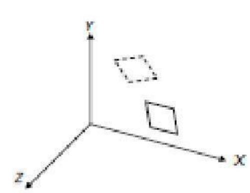


$x' = x\cdot\cos\theta - y\cdot\sin\theta$ $y' = x\cdot\sin\theta + y\cdot\cos\theta$ $z' = z$	$x' = x$ $y' = y\cdot\cos\theta - z\cdot\sin\theta$ $z' = y\cdot\sin\theta + z\cdot\cos\theta$	$x' = z\cdot\sin\theta + x\cdot\cos\theta$ $y' = y$ $z' = z\cdot\cos\theta - x\cdot\sin\theta$

3D Transformations (*cont.*)

The 2D rotation introduced previously is just a 3D rotation about *z* axis.

$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



similarly we have:

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad R_y(\theta) = \begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

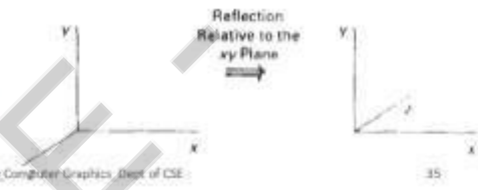
AMSCE - 1101

www.padeepz.net

General 3D Rotations • Rotation about an axis that is parallel to one of the coordinate axes : 1. Translate the object so that the rotation axis coincides with the parallel coordinate axis 2. Perform the specified rotation about the axis 3. Translate the object so that the rotation axis is moved back to its original position • Not parallel : 1. Translate the object so that the rotation axis passes 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 the axis 4. Apply inverse rotations to bring the rotation axis back to its original orientation 5. Apply the inverse translation to bring back the rotation axis to its original position

3 D Transformation functions • Functions are – translate3(translateVector, matrixTranslate) – rotateX(thetaX, xMatrixRotate) – rotateY(thetaY, yMatrixRotate) – rotateZ(thetaZ, zMatrixRotate) – scale3(scaleVector,matrixScale) • To apply transformation matrix to the specified points , – transformPoint3(inPoint, matrix,outPoint) • We can construct composite transformations with the following functions – composeMatrix3 – buildTransformationMatrix3 – composeTransformationMatrix3 CS71_Computer Graphics_Dept of CSE 34 Reflections In 3-D • Three Dimensional Reflections can be performed relative to a selected reflection axis or a selected reflection plane • Consider a reflection that converts coordinate specifications from a right handed system to left handed system. • This transformation changes the sign of Z coordinate leaving x and y coordinates

$$T(d_x,d_y,d_z)=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$$



Shears In 3-D

Shearing transformations are used to distortions in the shape of an object. In 2D, shearing is applied to x or y axes. In 3D it can applied to z axis also

The following transformation produces an Z axis shear

$$SHz = \begin{bmatrix}1 & 0 & a & 0\\ 0 & 1 & b & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

Parameters a and b can be assigned any real values

AMSCE - 1101

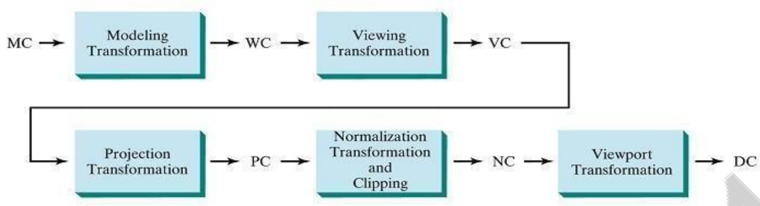
LP 3: 3D Viewing

- World coordinate system(where the objects are modeled and defined)

- Viewing coordinate system(viewing objects with respect to another user defined coordinate system)

- Projection coordinate system(coordinate positions to be on the projection plane)

- Device Coordinate System (pixel positions in the plane of the output device)



steps to establish a Viewing coordinate system or view reference coordinate system and the view plane

Transformation from world to viewing coordinates

- Translate the view reference point to the origin of the world coordinate system

- Apply rotations to align the axes

  – Three rotations of three axes

  – Composite transformation matrix (unit vectors u, v, n)

    - n= N/|N|

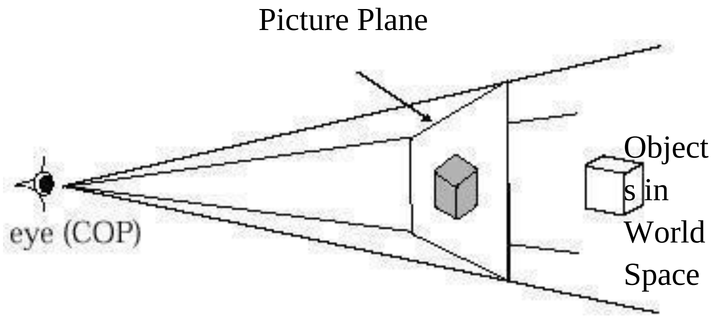
    - u = (V*N) / |V*N|

    - v = n * u

AMSCE - 1101

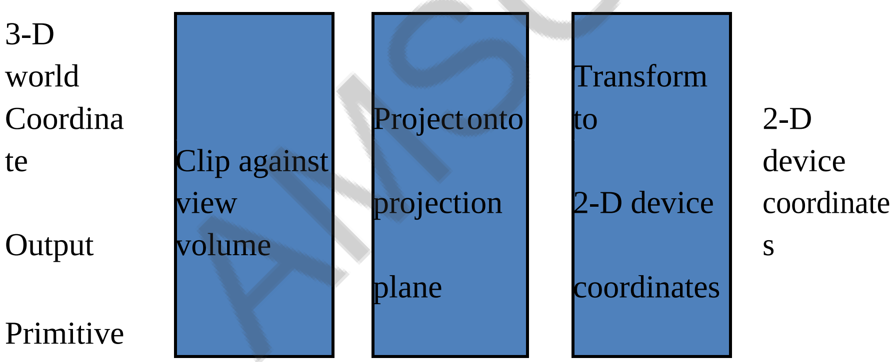
Projections

Our 3-D scenes are all specified in 3-D world coordinates. To display these we need to generate a 2-D image - project objects onto a picture plane



Converting From 3-D To 2-D

- Projection is just one part of the process of transforming 3D world coordinates to a 2-D projection plane

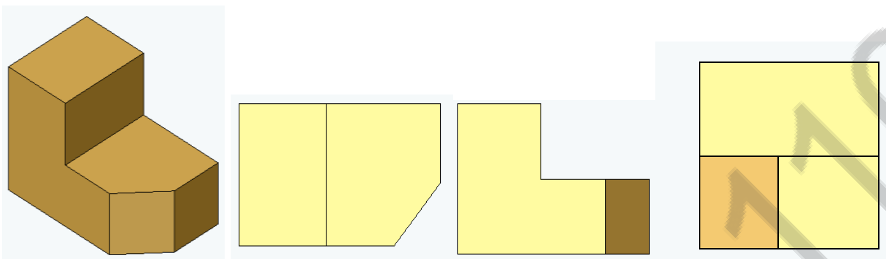


AMSCE - 1101

1. Three-Dimensional Display Methods

Parallel projection

- Project points on the object surface along parallel lines onto the display plane.
- Parallel lines are still parallel after projection.
- Used in engineering and architectural drawings.
- By selecting different viewing positions, we can project visible points on the object onto the display plane to obtain different two-dimensional views of the object.



Perspective projection

- Project points to the display plane along converging paths.
- This is the way that our eyes and a camera lens form images and so the displays are more realistic.

It has two major characteristics

- Smaller as their distance from the observer increases.
- Foreshortened: the size of an object‟s dimension along the line of sight are relatively shorter than dimensions across the line of sight.

Depth Cueing

- To easily identify the front and back of display objects.
- Depth information can be included using various methods.
- A simple method to vary the intensity of objects according to their distance from the viewing position. Eg: lines closest to the viewing position are

displayed with the highest intensities and lines farther away are displayed with decreasing intensities.

- Application is modeling the effect of the atmosphere on the pixel intensity of objects. More distant objects appear dimmer to us than nearer objects due to light scattering by dust particles, smoke etc.

Visible line and surface identification

- Highlight the visible lines or display them in different color
- Display non visible lines as dashed lines
- Remove the non visible lines

Surface rendering

- Set the surface intensity of objects according to
  - Lighting conditions in the scene
  - Assigned surface characteristics
- Lighting specifications include the intensity and positions of light sources and the general background illumination required for a scene.
- Surface properties include degree of transparency and how rough or smooth the surfaces are to be.
- Exploded and Cutaway Views
  - To maintain a hierarchical structures to include internal details.
  - These views show the internal structure and relationships of the object parts

Cutaway view

- Remove part of the visible surfaces to show internal structure.

AMSCE - 1101

UNIT IV MULTIMEDIA SYSTEMS DESIGN - PART 1

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

Multimedia Basics

Multimedia is a combination of text, graphic art, and sound, animation and video elements.

The IBM dictionary of computing describes multimedia as "comprehensive material, presented in a combination of text, graphics, video, animation and sound. Any system that is capable of presenting multimedia, is called a multimedia system".

A multimedia application accepts input from the user by means of a keyboard, voice or pointing device. Multimedia applications involve using multimedia teclmology for business, education and entertainment. Multimedia is now available on standard computer platforms. It is the best way to gain attention of users and is widely used in many fields as follows:

* Business - In any business enterprise, multimedia exists in the form of advertisements, presentations, video conferencing, voice mail, etc.

- Schools - Multimedia tools for learning are widely used these days. People of all ages learn easily and quickly when they are presented information with the visual treat.
- Home PCs equipped with CD-ROMs and game machines hooked up with TV screens have brought home entertainment to new levels. These multimedia titles viewed at home would probably be available on the multimedia highway soon.
- Public places - Interactive maps at public places like libraries, museums, airports and the stand-alone terminal
- Virtual Reality (VR) - This technology helps us feel a 'real life-like' experience. Games using virtual reality effect is very popular

MULTIMEDIA ELEMENTS

High-impact multimedia applications, such as presentations, training and messaging, require the use of moving images such as video and image animation, as well as sound (from the video images as well as overlaid sound by a narrator) intermixed with document images and graphical text displays. Multimedia applications require dynamic handling of data consisting of a mix of text, voice, audio components, video components, and image animation. Integrated multimedia applications allow the user to cut sections of all or any of these components and paste them in a new document or in another application such as an animated sequence of events, a. desktop publishing system, or a spreadsheet. The components that fall under our definition of multimedia are: '

Data elements for Multimedia Systems

Facsimile

Facsimile transmissions were the first practical means of transmitting document images over telephone lines. The basic technology, now widely used, has evolved to allow higher scanning density for better-quality fax

Document images

Document images are used for storing business documents that must be retained for long periods oftime or may need to be accessed by a large number of people. Providing multimedia access to such documents removes the need far making several copies ofthe original for storage or *distribution*

Photographic images

Photographic images are used for a wide range of applications . such as employee records for instant identification at a security desk, real estates systems with photographs of houses in the database containing the description of houses, medical case histories, and so on.

AMSCE - 1101

Geographic information systems map (GIS)

Map created in a GIS system are being used wildly for natural resources and wild life management as well as urban planning. These systems store the geographical information of the map along with a database containing information relating highlighted map elements with statistical or item information such as wild life statistics or details of the floors and rooms and workers in an office building

Voice commands and voice synthesis

Voice commands and voice synthesis are used for hands-free operations of a computer program. Voice synthbsis is used for presenting the results of an action to the user in a synthesized voice. Applications such as a patient monitoring system in a surgical theatre will be prime beneficiaries of these capabilities. Voice commands allow the user to direct computer operation by spoken commands

Audio message

Annotated voice mail already uses audio or voice message as attachments to memos and documents such as maintenance manuals.

Video messages

Video messages are being used in a manner similar to annotated voice mail.

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.

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.

MULTIMEDIA APPLICATIONS

The first widely used application of multimedia is document image management. It is primarily intended for scanning documents and retaining their images.

Another application is image processing. It is also known as Image recognition. It is intended for recognizing objects by analyzing their raster images. Applications that present a view of generic multimedia applications are:

1. Document Imaging

The fundamental concepts of storage, compression and decompression, and display technologies used for multimedia systems were developed for document image management. Organizations such as insurance agencies law offices,country and state governments, and the federal government manage large volumes of documents.

Document image technology is adopted by Department of Defence for applications ranging from military personnel records to maintenance manuals and high-speed printing systems. Almost all document image system use workflows that are customized for the purpose for which they are being used. The workflow defines the sequence for scanning images, performing data *entry* based on the contents of the lmages, indexing them and storing them on optical media.

Document Image Hardware requirements:

Realtime 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 nano-second. For high-resolution images, processing of the order of 10 pixels/ ns is enough for monochrome still images. Gray scale images consist of pixels that have shades of gray ranging from 16 to 256. Color images feature color hues instead of shades of gray. Most high-resolution monitors support 16 to 256 colors display capability. The number of colors that can be depicted depends on the number of bits used to define the palette.

AMSCE - 1101

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 processing systems employ the compression and decompression techniques, a wide range of algorithm for object recognition, comparing images of objects with pre-defined objects, extrapolating finer details to view edges more clearly, gray-scale balancing and gray-scale and color adjustments.

Let us briefly review the various aspects of image processing and recognition.

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

Capabilities built in the compression boards might include the following

* **Image calibration**: The overall image density is calibrated, and the image pixels are adjusted to a predefined level. *** Real time alignment**: The image is aligned in real-time for skewing caused by improper feeding of paper. * **Gray-Scale normalization**: The overall gray level of an image or picture is evaluated to determine if it is skewed in one direction and if it needs cOlTection. * **RGB hue intensity adjustment**: Too much color makes picture garish and fuzzy. Automatic hue intensity adjustment brings the hue intensity within pre-defined ranges. * **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. * **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 Animation

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

The basic concept of displaying successive images at short intervals to give the perception of motion is being used successfully in designing moving parts such as automobile engines.

Image annotation

Image annotation 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 overlayed over the original image for display purposes. It requires tracking multiple image components associated with a single page, decompressing all of them, and ensuring correct spatial alignment they are overlayed.

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.

Initially, people used dedicated OCR scanners. Now, OCR Technology is available in software. OCR technology, used as a means of data entry, may be used for capturing entire paragraphs of text. The capturing text is almost a\ways entered as a field in a database or in an editable document

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. Handwriting recognition engines use complex algorithms designed to capture data in real time as it is being input or from an image displayed in a window, depending on the application. Two factors are important for handwriting recognition. They are the strokes or shapes being entered, and the velocity of input or the vectoring that is taking place. The strokes are parsed and processed by a shape recognizer that tries to determine the geometry and topology of the strokes. It attempts to compared it to existing shapes, such as predefined characters. The stroke is compare with the prototype character

AMSCE - 1101

set until a match is found or all pre-defined prototypes have been checked without a match.
Multimedia system will use handwriting recognition as another means of user input.

Non-Textual Image Recognition

Image recognition is a major technology component in designing, medical and manufacturing fields. Let us review the basic concepts of image recognition architecture.

For example, a general Image recognition system,- the Image Understanding Architecture has the design which calls for three processing layers.

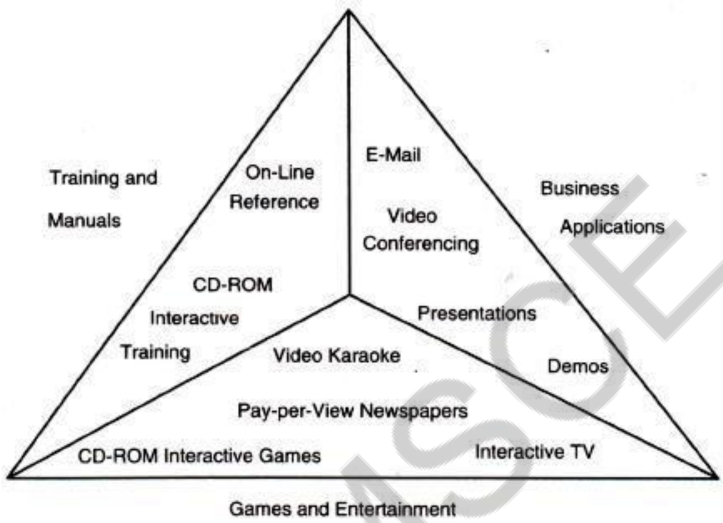
(i) 512 x 512 array of custom pixel processors that extract basic features such as lines and object boundaries. (ii) The features of an object extracted by the first layer are tracked by the DSP array, and that information is fed into 5l2-M byte RAM. (iii) At the highest level, sophisticated AI algorithms perform the difficult task of object and scene recognition.

3.Full motion Digital video Applications

Full motion video has applications in the games industry and training, as well as the business world. Full motion video is the most complex and most demanding component of multimedia applications. For business applications, some core requirements are needed.

(i) Full-motion video clips should be sharable but should have only one sharable copy.

(ii) It should be possible to attach full-motion video clips to other documents such as memos, chapter text, presentation, and so on.



The following features should be available:

(a) Features, of a VCR metaphor, such as, rewind, fast-forward, play, and search.
(b) Ability to move and resize the window displaying the video clip.
(c) Ability to view the same clip on a variety of display terminal types with varying resolution capabilities without the need for storing multiple copies in different form
(d) Ability to adjust the contrast and brightness of the video clip.
(e) Ability to adjust the volume of the associated sound.
(f) It should enable the users to place their own indexing marks to locate segments in video clip.

4.Electronic Messaging

The first generation mail system provided a basic text link between users and provided a valuble communications medium for users within a department or enterprise. These systems were the first alternative to paper based inter-office memos. The second generation of electronic mail system expands this capability tremendously by providing cross-platform and cross-network electronic mail with a bapability to attach other files ranging from editable text files to bit mapped graphics and program executables.

A multimedia enabled electronic messaging system requires a sophisticated infrastructure consisting of the following to support it:

AMSCE - 1101

- Message storage and forward facility.
- Message transfer agents to route messages to their final destinations across various nodes in a multilevel network.
- Message repositories to store documents in a filing cabinents.
- Repositories for dense multimedia components such as images, videos, frames, audio messages, and full-motion video clips.
- Ability for multiple electronic hypermedia messages to share the same multimedia components residing in various repositories on the enterprise network.
- Local and global Directories to locate users and servers across an enterprise network.
- Automatic database synchronization of dynamic electronics messaging databases.
- Automatic protocol conversions and data formats conversions.
- Administrative tools to manage enterprise-wide networks.

5. A Universal Multimedia Application

It is an application that works on universal data type. This means that the application manipulates datatypes that can be combined in a document, displayed 'on a screen, or printed, with no special manipulations that the user needs to perform. A document of this type may be a phonebook,a color brochure with pictures and drawings, a memo,a phone message, a video phone message,or live teleconferencing. The application is truly distributed in nature. An important consideration for such a universal application is the methodology for dissemination of the information on a network.

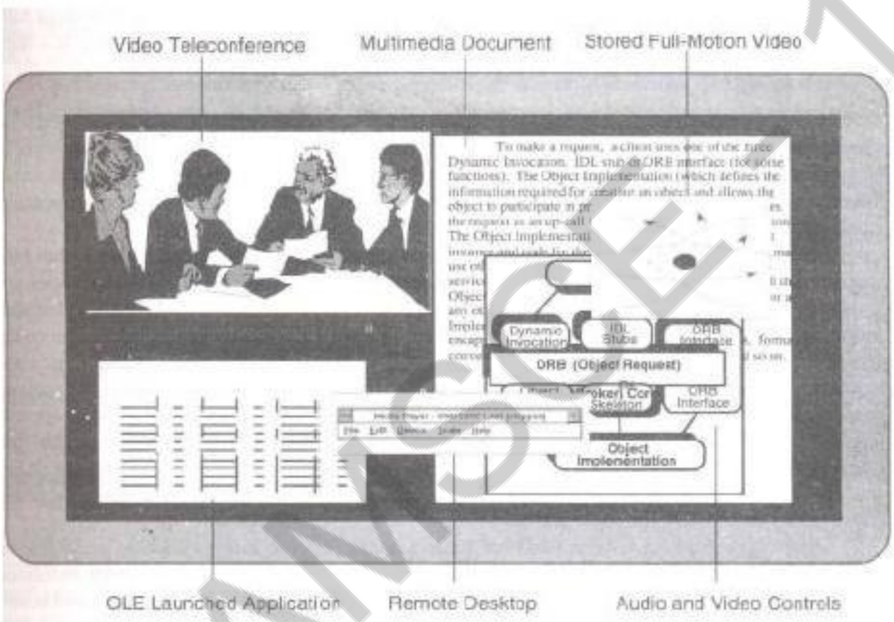


Figure describes the user screen for a universal multimedia application. In this screen, mix of windows for displaying still video and document images, a video conference window with a live session in progress, a remote live desk top, and a couple of other windows for applications such as electronic mail and desk top publishing.

To maintain all of these windows requires a substantial amount of CPU power. Digital Signal Processing assistance is needed to manage the multiple simultaneous decompressions for JPEG, MPEG and windows applications.

Full-Motion Video Messages

In addition to textual messages, electronic mail capability allows embedding of voice messages and video messages. Video messages may consist of video snapshots or live video with full-motion picture and sound.

Two technological concepts at play in the implementation of full motion video messages:

i) The storage and transmitted of a very large volume of data at a high rate,

(ii) Decompression of that data to present a continuous play back .

Viewer interactive Live video:

- The key difference between full motion video and viewer interactive video is that full motion video is a play back of stored video clips while viewer interactive video is a live. It may be possible to manage decompression and display of stored video clips more easily.
- Interactive Live video are the interesting applications used for direct interaction,medical application, manufacturing application and various process control application.
- Whereas full motion video are useful for for messagins, information dissemination.

Audio and Video Indexing.

Indexing is an important and complex subject for multimedia design. Marking a position is called Indexing. Audio and video indexing are used in full-motion video in a manner similar to any video sequence, i.e.,. just as it would in a home movie, taped performance and so on.

The needs of the application must be a strong consideration for the type of indexing provideq with the system.

Key points for indexing of stored video clips:

* Indexing is useful only if the video is stored, indexing information is lost.

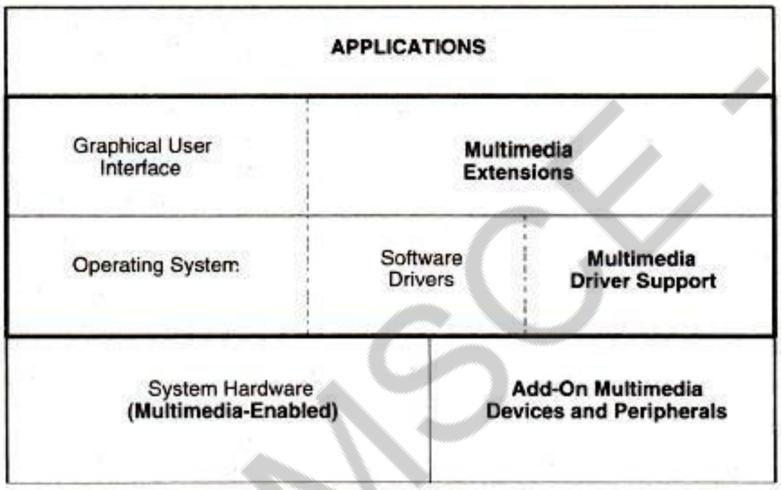
* When sound and video are decompressed and managed separately, synchronization is very important.

* Depending on the application, indexing information must be maintained separately for sound and video components of a video clip.

3.3 MULTIMEDIA SYSTEMS ARCHITECTURE

Multimedia encompasses a large variety of technologies and integration of multiple architectures interacting in real time. All of these multimedia capabilities must integrate with the standard user interfaces such as Microsoft Windows.

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

The various graphics standards such as MCA, GGA and XGA have demonstrated the increasing demands for higher resolutions for GUls.

Combined graphics and imaging applications require functionality at three levels. They are provided by three classes of single-monitor architecture.

(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.

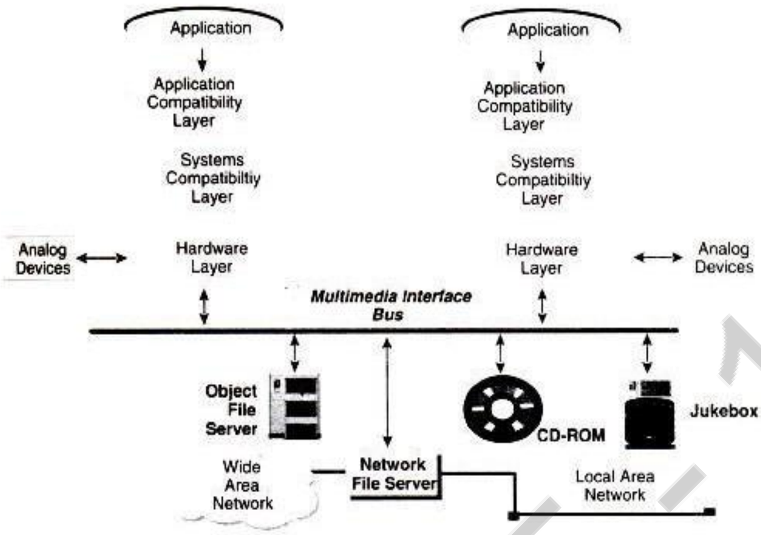
(iii) Dual-buffered VGA/Mixing/Scaling: Double buffer schemes maintain the original images in a decompression buffer and the resized image in a display buffer.

AMSCE - 1101

The IMA Architectural Framework

The Interactive Multimedia Association has a task group to define the architectural framework for multimedia to provide interoperability. The task group has C0ncentrated on the desktops and the servers. Desktop focus is to define the interchange formats. This format allows multimedia objects to be displayed on any work station.

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 service"s, including filters and translators **Figure 3.4** describes the generalized architectural approach



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 W ANs.

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.

High speed server to server Links

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. **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.

Networking Standards:

The two well-known networking standards are

1. Ethernet
2. token ring.

ATM and FDDI are the two technologies which are going to be discussed in detail.

AMSCE - 1101

ATM:

- ATM is a acronym for Asynchronous Transfer Mode. It's topology was originally designed for broadband applications in public networks.
- ATM is a method of multiplexing and relaying (cell-switching) 53 byte cells. (48 bytes of user information and 5 bits of header information).
- It has been increasingly used for transferring real time multimedia data in local network at a speed higher than 100Mbits/sec. ANSI has adopted ATM as the cell switching standard.
- **Cell Switching:** It is a form of fast packet switching based on the us of cells. **Cells:** Short, fixed length packets are called cells.
- The ANSI standard for FDDI allows large-distance networking. It can be used as high-performance backbone networks to complement and extend current LANs.
- ATM provides high capacity, low-latency switching fabric for data. It is independent of protocol and distances. ATM effectively manage 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. 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 (American National Standard Institute) standard for FDDI allows for single-mode fiber supporting up to 40 km between stations.
- It extends the current LAN speed from 100 Mbits/sec to several Gigabits per seconds, and large-distance networking.

Difference between ATM and FDDI

ATM	FDDI II
ATM pushes network speed as high as 622Mbits/sec	FDDI II pushes network speed as high as 100 Mbits /sec
ATM is capable of lower speeds at the workstations.It reduces number of devices protocol translation require for communication between local and wide area network	FDDI II does not allow a user to connect to the network at the spped required by the user, rather it requires the user to be capable of supporting the network speed.

Benefits of Shared media Networks:

- Ease of installation
- Lack of common equipment
- Connectionless operation
-

Difficulties of Shared Media Networks:

- Wiring existing buildings
- Fault isolation

3.4 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.

1.Hypermedia documents

Hypermedia documents are documents which have text, embedded or linked multimedia objects such as image, audio, hologram, or full-motion video.The network speed and computing efficiency with which these hypermedia documents can manipulated has special implications for multimedia application such as messaging. Hypermedia has its roots in hypertext.

Hypertext

Hypertext systems allow authors to link information together, create information paths through a large volume of related text in documents.

It also allows to annotate existing text, and append notes.

It allows fast and easy searching and reading of selected excerpts.

Hypermedia

It is an extension of hypertext.

In that, we can include texts, any kind of information that can be stored in electronic storage, such as audio, animated video, graphics 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.

Hyper Speech

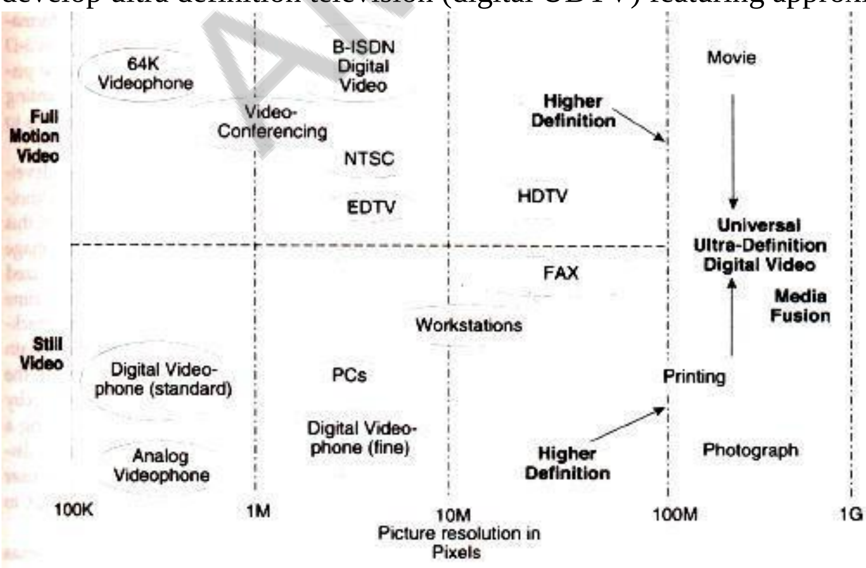
Multimedia stimulated the development of general-purpose speech interfaces. Speech synthesis and speech recognition are fundamental requirement for hyperspeech systems. Speech recognition is nothing but converting the analog speech into a computer action and into ASCII text. Speech-recognition systems 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.
Speech synthesis and speech recognition requires substantial processing power. High performance microprocessors, such as main CPU in workstations,and Digital Signal Processingand codecs supporting encoding and decoding of sound based on emerging standards can handle speech recognition and speech synthesis.

2.HDTV AND UDTV

HDTV is an acronym of High-Definition Television.

The broadcasting standards such as NTSC, PAL, SECAM, NHK have an idea of bringing the world together on a single high-definition Television broadcasting standard.

The japanese broadcasting services developed a 1125-line, along MUSE system. A competing standard in the U.S. changed direction from analog to digital technology:A 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



AMSCE - 101

3.3D 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.

The omniview three dimentional volumetric display device, developed by texas Instruments,Inc., uses lasers of different colors to project images on a moving surface sweeping a 3-D cylindrical display volume.

4.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.

Like Digital Signal Processing(DSP) the Fuzzy Logic Signal Processing(FLSPs) provide interesting applications for multimedia systems. Fuzzy logic is a integral part of multimedia hardware and it is the key for emerging graphical interface.

The benefits of FLSPs are

5. Digital Signal Processing

Digital Signal Processing are 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:

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

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

6. **Intertask Synchronization And Communication**

Mechanisms for intertask 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.

3.5 DEFINING OBJECTS FOR MULTIMEDIA SYSTEMS

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

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

HYPERTEXT

AMSCE - 1101

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

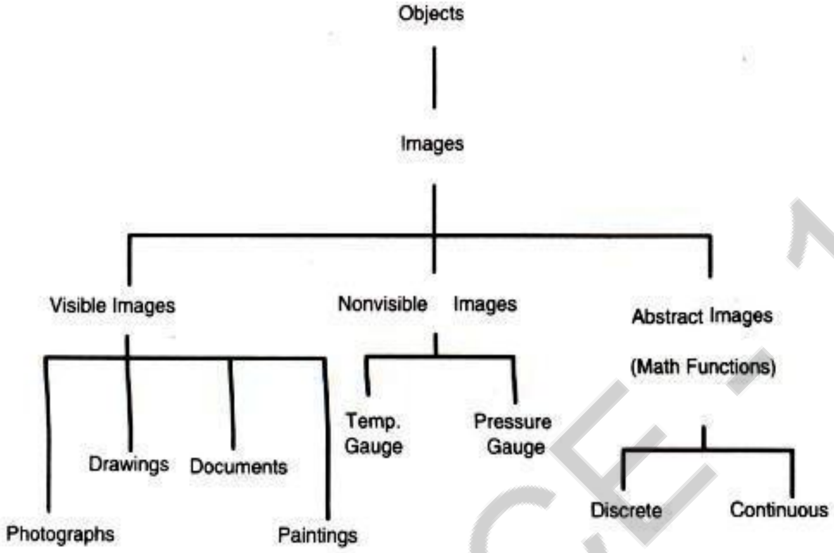
IMAGES

Image object is an object that is represented in graphics or encoded form. Image object is a subobject 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 do not have a temporal property associated with them.

Thc data types such as document images, facsimile systems, fractals, bitmaps, meta files, and still pictures or still video frames are grouped together.

Figure 3.6 describes a hierarchy of the object classes



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

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.

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 thc sound clip.

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

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 need 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.

AMSCE - 1101

MULTIMEDIA DATA INTERFACE STANDARDS

File Formats for Multimedia Systems:

(i) Device-independent Bitrnap (DIB): This file format contains bit map, color, and color paJlette information.

(ii) RIFF device Independent Bitrnap (RDIB): Resource Interchange File Frmat **(RIFF)** is the standard file format defined for Microsoft Windows and OS/2. It allows a more complex set of bit maps than can be handled by DIB.

(iii) Musical Instrument Digital interface (MIDI): This is the interface standard for file transfer between a computer and a musical instrument such as a digital piano. It is also, used for full-motion video and voice-mail messaging systems. It has the advantage of ready availability of MIDI device controller boards for personal computers.

RIFF Musical Instrument Digital Interface

A MIDI format within a RIFF envelope provides a more complex interface.

Palette File Format (PAL) An interface that allows defining a palette of 1 to 256 colours in a representation as RGB values.

Rich Text Format (RTF) This file format allows embedding graphics and other file formats within a document. This format is used by products such as Lotus Notus. This format is also the basis for the use of OLE.

Waveform Audio File Format (WAVE) A digital file representation of digital audio.

Windows Metafile Format (WMF) This is a vector graphic format used by Microsoft Windows as an interchange format.

Multimedia Movie Format (MMM) This is a format used for digital video animation.

Apple's Movie Format This format was defined as the standard for file exchange by Quick Time enabled systems.

Digital Video Command Set (DVCS) This is the set of digital video commands simulating VCR controls.

Digital Video Media Control Interface Microsoft's high level control interface for VCR controls, including play, rewind, record and so on.

Vendor - Independent Messaging (VIM) Developed by a consortium of Vendors providing a standardized format for cross-product messages.

Apple's Audio Interchange File Format Apple's standard file format for compressed audio and voice data.

SDTS GIS Standard The Spatial Data Transfer Standard (SDTS) is designed to provide a common storage format for geographic and cartographic data.

VIDEO PROCESSING STANDARDS

INTELS DVI

DVI is an achronym of Digital Video Interface.

DVI standard is to provide a processor independent specification for a video interface. That video interface should accomodate most compression algorithms for fast multimedia displays. An example of custom-designed chip which supports DVI is Intel's i750 B. This chip is designed for enhancing low-end, software based PC video.

Advantages of the DVI Chip

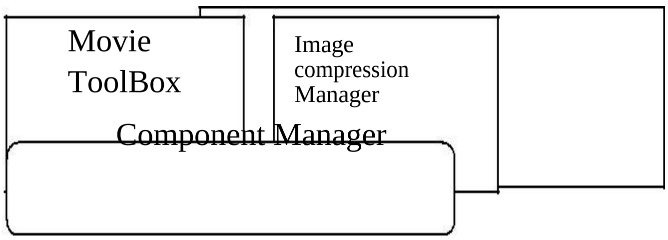
(i) It can operate software video processing real time. (ii) It can share the processing with the host CPU. (iii) It can handle additional vector-quantization-type algorithms in conjunction with host processing. DVI silicon chip relies on a programmable video processor. It gives potential to DVI chips to run a range of compression algorithms.

APPLE QUICK TIME

Quick Time standard is developed by Apple Computer. It is designed to Support multimedia applications. It is integrated with the operating system. Quick time refers to both the extensions to the Mac Operating system and to the compression/decompression functionality Of the environment. Quick Time is designed to be the graphics standard for timebased graphic data types.

AMSCE - 1101

Quick Time's definition has been extended to include (i) System Software, (ii) File Formats, (Hi) Compression! decompression algorithms, (iv) Human Interface Standards.
Figure Shows the components in the Quick Time Architecture.



Quick Time adjust automatically to the hardware being used by the user. MPEG is another competing standard which is comparitively higher-end, hardware-assisted standard. It can produce better resolutions at faster rates.

MICROSOFT AVI

A VI is an achronym for Audio Video Interleave Standard. It is similar to Apple's Quick Time. It offers low-cost, low-resolution video processing for the average desktop user. It is a layered product. A VI is scalable. It allows users to set parameter such as window size, frame rate, quality and compression algorithm through a number of dialog boxes. AVI-compatible hardware allows enhancing performance through hardware-accelerated compression algorithms such as DVI and MPEG. A VI supports several compression algorithms

Multimedia Databases

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

Multimedia storage and retrieval 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

Massive Data Volumes

A single multimedia document may be a combination of different media Hence indexing of documents, fi lms 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 eff0l1 to ensure that they are returned in proper sequence to their original storage location.

storage technologies

There are two major mass storage technologies used currently for storage of multimedia documents.

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

Advantages of Optical disk storage systems:

(i) Managing a few optical disk platters in a juke box is much simpler than man;lging a large magnetic disk farm. (ii) Optical disk storage is excellent storage system for off line archival of old and infrequently referenced documents for significant periods of time

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 t6 various components of hypermedia database record. Optical media providcs 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.

AMSCE - 1101

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.

Database Management Systems for Multimedia Systems

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 storge choices are available.
The choices available are:

* Extending the existing relational database management systems, (RDBMSs) to support the various objects for multimedia as binary objects.
* Extending RDBMSs beyond basis binary objects to the concepts of inheritance and classes. RDBMSs supporting these . features provide extensions for object-programming front ends and/or C++ support.
* Converting to a full fledged object oriented database that supports the standard SQL language.
* Converting the database and the application to an objectoriented database and using an object-oriented language, or an object-enabled SQL for development.

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 dominent 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).

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
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.
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.
Object-oriented software technology has three important concepts. They are:
Encapsulation: It is the ability to deal with software entities as units that interact in pre-defined and controllable manner, and where the control routines are integral with entity.
Association: It is the ability to define a software entity in terms of its di fferences 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. It also provides autonomy; **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.
Database Organization for Multimedia Applications
Data organization for multimedia systems has some key issues. They are:

(l) Data independence (2) Common distributed database architecture

(3) Distributed database servers· (4) Multimedia object management.

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.

Distributed Data 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.

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.

Transaction management for Multimedia Systems

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 .

In most simple applications based on text and textual or numeric data a transaction managed generally by the server that provides the storace for the data.

Even these transactions becomes more complex when data has to be retrieved from multiple data servers that can be accessed simultaneously by a large number of users.

Conflicts arise when two users attempts to read from and write to same data record.

A multi-phase commit methodologies are used to address conflicts in relational database.

Multimedia Database	Conventional Database
A **Multimedia database** (**MMDB**) is a collection of related multimedia data.	A **relational database management system** (**RDBMS**) is a database management system (DBMS) that is based on the relational model
A Multimedia Database (MMDB) hosts one or more multimedia data types[3] (i.e. text, images, graphic objects, audio, video, animation sequences.	Each database is a collection of tables, which are called relations, hence the name "relational database
Multimedia data consists of a variety of media formats or file representations including TIFF,BMP, PPT, IVUE, FPX, JPEG, MPEG, AVI, MID, WAV, DOC, GIF, EPS, PNG, etc.	relational model contains the following components: • Collection of objects or relations • Set of operations to act on the relations • Data integrity for accuracy and consistency
• Multimedia database consume a lot of processing time, as well as bandwidth.	Extremely fast retrieval times for multi-user, transactional environment. • Ease the use compared to other database systems
Examples of multimedia database application areas: • Digital Libraries • News-on-Demand	examples of relational database system • process control • internet service management • spacecraft control system

• Video-on-Demand • Music database • Geographic Information Systems (GIS) • Telemedicine	• network management system

AMSCE - 1101

UNIT IV –PART II

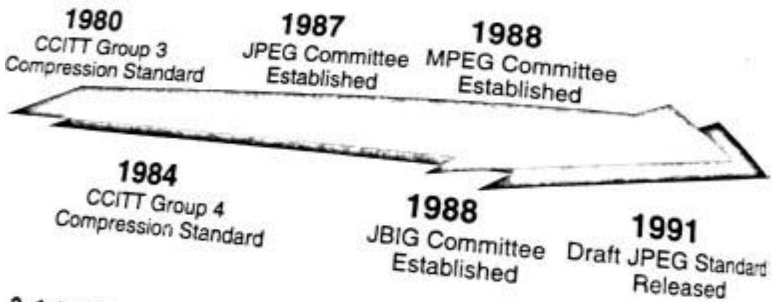
Compression and decompression − Data and file format standards − Multimedia I/Otechnologies − Digital voice and audio − Video image and animation − Full motionvideo − Storage and retrieval technologies.

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. The figure shows the evolving path for compression standards.



2–1 Multimedia Standards Evolution.

Types Of Compression

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. .

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 perceptly 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 lossy compression. Audio compression, on the other hand, is not lossy. It is called lossless compression.

Lossless Compression.

In lossless compression, data is not altered or lost in the process of compression or decompression. Decompression generates an exact replica ofthe 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:

- **Packpits encoding (Run-length encoding)**
- **CCITT Group 3 I D**
- **CCITT Group 3 2D**
- **CCITT Group 4**
- **Lempe l-Ziv and Welch algorithm LZW.**

Lossy compression is that some loss would occur while compressing information objects.

Lossy compression is used for compressing audio, gray-scale or color images, and video objects in which

AMSCE - 1101

absolute data accuracy is not necessary.

The idea behind the lossy 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 tele-conferencing, and multimedia electronic messaging systems.

Lossy compressions techniques can be used alone o~' in combination 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 lossy compression mechanisms:

- Joint Photographic Experts Group (JPEG)
- Moving Picture Experts Group (MPEG)
- Intel DVI
- CCITT H.261 (P * 24) Video Coding Algorithm
- Fractals.

4.1.2Binary 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 used primarily for documents that do not contain any continuous-tone information or where the continuous-tone information can be captured in a black and white mode to serve the desired purpose.

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.

1 .Packpits 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 need not to have a standard. 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 seven bits to represents the run length.
Eg for packbit encoding:

00000000000000000000111111100000000000000000001111

is represented as:

```
Byte1, Byte2, Byte3, Byte4, Byte5, Byte6, Byte7, Byte8...ByteN-1, ByteN.
0x14, 0x00, 0x07, 0x01, 0x11, 0x00, 0x04, 0x01.....
```

2. CCITT Group 3 1-D Compression

This scheme is based on run-length encoding and assumes that a typical scanline 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

AMSCE - 1101

A modified version of run-length encoding is Huffman encoding.

It is used for many software based document imaging systems. It is used for encoding the pixel run length in CCITT Group 3 1-dGroup 4.

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:

Huffman encoding scheme is based on a coding tree.

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

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

AMSCE - 1101

White Run Length	Code Word	Black Run Length	Code Word
0	00110101	0	0000110111
1	000111	1	010
2	0111	2	11
3	1000	3	10
4	1011	4	011
5	1100	5	0011
6	1110	6	0010
7	1111	7	00011
8	10011	8	000101
9	10100	9	000100
10	00111	10	0000100
11	01000	10	0000100
11	01000	11	0000101
12	001000	12	0000111
13	000011	13	00000100
14	110100	14	00000111
15	110101	15	000011000
16	101010	16	0000010111
17	101011	17	0000011000
18	0100111	18	0000001000
19	0001100	19	0000 11 00 III
20	0001000	20	00001101000
21	0010111	21	00001101100
22	0000011	22	00000110111
23	0000100	23	00000101000
24	0101000	24	00000010111
25	0101011	25	00000011000
26	0010011	26	000011001010
27	0100100	27	000011001011
28	0011000	28	000011 00 11 00
29	00000010	29	000011001101

AMSCE - 1101

For example, from **Table 2,** the run-length code of 16 white pixels is 101010, and of 16 black pixels 0000010111. Statistically, the occurrence of 16 white pixels is more frequent than the occurrence of 16 black pixels. Hence, the code generated for 16 white pixels is much shorter. This allows for quicker decoding. For this example, the tree structure could be constructed.

36	00010101	36	000011010100
37	00010110	37	000011010101
38	000101 II	38	000011010110
9	00101000	39	000011 0 1 0 1 1 1
40	00101001	40	000001I01100
41	00101010	41	000001101101
42	00101011	42	000011011010
43	00101100·	43	0000 11 0 1 1011
44	00101101	44	000001010100
45	00000100	45	000001010101
46	00000101	46	000001010110
47	00001010	47	000001010111
48	00001011	48	000001100100
49	01010010	49	000001100101
50	010100II	50	000001010010
51	01010100	51	000001010011
52	01010101	52	000000100100
53	00100100	53	000000110111

The codes greater than a string of 1792 pixels are identical for black and white pixels. A new code indicates reversal of color, that is, the pixel Color code is relative to the color of the previous pixel sequence.
Table 3 shows the codes for pixel sequences larger than 1792 pixels.

AMSCE - 1101

30	00000011	30	000001101000
31	00011010	31	000001101001
32	00011011	32	000001101010
33	00010010	33	000001101011
34	00010011	34	000011010010
35	00010100	35	000011 0 10011

AMSCE - 1101

Run Length (Black and White)	**Make-up Code**
1792	00000001000
1856	00000001100
1920	00000001101
1984	000000010010
2048	000000010011
2112	000000010100
2176	000000010101
2240	000000010110
2304	000000010111
2368	000000011100

2432	000000011101
2496	000000011110
2560	000000011111

CCITT Group 3 compression utilizes Huffman coding to generate a set of make-up codes and a set of terminating codes for a given bit stream. Make-up codes are used to represent run length in multiples of 64 pixels. Terminating codes are used to represent run lengths of less than 64 pixels.

As shown in **Table** 2; run-length codes for black pixels are different from the run-length codes for white pixels. For example, the run-length code for 64 white pixels is 11011. The run length code for 64 black pixels is 0000001111. Consequently, the run length of 132 white pixels is encoded by the following two codes:

Makeup code for 128 white pixels - 10010
Terminating code for 4 white pixels - 1011

The compressed bit stream for 132 white pixels is 100101011, a total of nine bits. Therefore the compression ratio is 14, the ratio between the total number of bits (132) divided by the number of bits used to code them (9).
CCITT Group 3 uses a very simple data format. This consists of sequential blocks of data for each scanline, as shown in **Table 4.**
Coding tree for 16 white pixels

Coding tree for 16 black pixels

AMSCE - 1101

Note that the file is terminated by a number of EOLs (End of. Line) if there is no change in the line [rom the previous line (for example, white space).
TABLE 4: CCITT Group 3 1D File Format

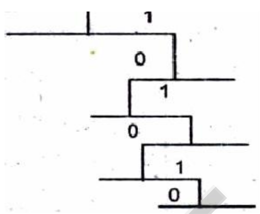
Advantages of CCITT Group 3 ID

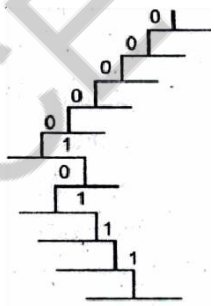
CCITT Group 3 compression has been used extensively due to the following two advantages: It
is simple to implement in both hardware and software .

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

CCITT group 3 compressions utilizes Huffman coding to generate a set of make-up codes and a set of terminating codes for a give bit stream.

CCITT Group 3 uses a very simply data format. This consists of sequential blocks of data for each scanline.





EOL DATA FILL EOL DATA FILL		EOL… DATA FILL EOL EOL EOL
LINE	LINE	LINE
1	2	n

AMSCE - 1101

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 scanline is redundant.

If black and white transition occurs on a given scanline, chances are the same transition will occur within + or - 3 pixels in the next scanline.

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.

Data formatting for CClIT 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.

The steps for pseudocode to code the code line are:

(i) Parse the coding line and look for the change in the pixel value. (Change is found at al location). **(ii)** Parse the reference line and look for the change in the pixel value. (Change is found at bl location). **(iii)** . Find the difference in location between bland a 1: delta = b1- al

Advantage of CClIT Group 3 2D

The implementation of the k factor allows error-free transmission . Compression ratio achieved is better than CClTT Group 3 1 D .

It is accepted for document imaging applications.

DisadvantageIt doesn’t provide dense compression

CCITT Group 4 2D compression

CClTT Group 4 compression is the two dimensional coding scheme without the k-factor.

In this method, the first reference line is an imaginary all-white line above the top of the image.The first group of pixels (scanline) is encoded utilizing the imaginary white line as the reference line.

The new coded line becomes the references line for the next scan line. The k-factor in this case is the entire page of line. In this method, there are no end-of-line (EOL) markers before the start of the compressed data

4.1.3 COLOR, GRAY SCALE AND STILL-VIDEO IMAGE COMPRESSION

Color:

Color is a part of life we take for granted. Color adds another dimension to objects. It helps in making things standout.

Color adds depth to images, enhance images, and helps set objects apart from -background.

Let us review the physics of color. Visible light is a form of electromagnetic radiation or radiant energy, as are radio frequencies or x-rays. The radiant energy spectrum contains audio frequencies, radio frequencies, infrared, visible light, ultraviolet rays, x-rays and gamma rays.

Radian energy is measured in terms of frequency or wavelength. The relationship between the two is

AMSCE - 1101

AMSCE - 1101

$$\lambda = \frac{c}{f} \text{ meters}$$

where λ is the wavelength in meters
c is the velocity of light in meters per second.
f is frequency of the radiation in hertz.

Since all electromagnetic waves travel through space at the velocity of light i.e., 3×10^8 meters/second - the wavelength is calculated by

$$\lambda = \frac{3 \times 10^8}{f} \text{ meters}$$

Color Characteristics
We typically define color by its brightness, the hue and depth of color.
Luminance or Brightness
This is the measure of the brightness of the light emitted or reflected by an object; it depends on the radiant, energy of the color band.

Hue This is the color sensation produced in an observer due to the presence of certain wavelengths of color. Each wavelength represents a different hue.

Saturation This is a measure of color intensity, for example, the difference between red and pink. **Color Models** Several calm' models have been developed to represent color mathematically. **Chromacity Model** It is a three-dimensional model with two dimensions, x and y, defining the color, and the third dimension defining the luminance. It is an additive model since x and yare added to generate different colors.

RGBModel RGB means Red Green Blue. This model implements additive theory in that different intensities of red, green and blue are added to generate various colors.

HSI Model The Hue Saturation and Intensity (HSI) model represents an artist's impression of tint, shade and tone. This model has proved suitable for image processing for filtering and smoothing images. **CMYK Model** The Cyan, Magenta, Yellow and Black color model is used in desktop publishing printing devices. It is a color-subtractive model and is best used in color printing devices only.

YUV Representation The NTSC developed the YUV three-dimensional color model. y - Luminance Component UV -Chrominance Components.

Luminance component contains the black and white or grayscale information. The chrominance component contains color information where U is red minus cyan and V is megenta minus green.
YUV Model for JPEG
The *JPEG* compression scheme uses several stages.

The first stage converts the signal from the spatial RGB domain to the YUV frequency domain by performing discrete cosine transform. This process allows separating luminance or gray-scale components from the chrominance components of the image.

4.1.3.1JOINT 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.
JPEG has been released as an ISO standard in two parts

Part I specifies the modes of operation, the interchange formats, and the encoder/decoder specifies for these modes along with substantial implementation guide lines .

Part 2 describes compliance tests which determine whether the implementation of an encoder or decoder conforms to the standard specification of part I to ensure interoperability of systems compliant with JPEG standards

Requirements addressed by JPEG

The design should address image quality .

The compression standard should be applicable to practically any kind of continuous-tone digital source image .

It should be scalable from completefy lossless to lossy ranges to adapt it. It should provide sequential encoding .

It should provide for progressive encoding .

It should also provide for hierarchical encoding .

The compression standard should provide the option of lossless encoding so that images can be guaranteed to provide full detail at the selected resolution when decompressed.

Definitions in the JPEG Standard
The JPEG Standards have three levels of definition as follows:

AMSCE - 1101

AMSCE - 1101

* Base line system
* Extended system
* Special lossless function.

The base line system must reasonably decompress color images, maintain a high compression ratio, and handle from 4 bits/pixel to 16 bits/pixel.

The extended system covers the various encoding aspects such as variable-length encoding, progressive encoding, and the hierarchical mode of encoding.

The special lossless function is also known as predictive lossless coding. It ensures that at the resolution at which the image is no loss of any detail that was there in the original source image.

Overview of JPEG Components JPEG Standard components are:

(i) Baseline Sequential Codec
(ii) OCT Progressive Mode
(Hi) Predictive Lossless Encoding
(iv) Hierarchical Mode.

These four components describe four four different levels of JPEG compression.

The baseline sequential code defines a rich compression scheme the other three modes describe enhancements to this baseline scheme for achieving different results.

Some of the terms used in JPEG methodologies are:

Discrete Cosine Transform (DCT)

DCT is closely related to Fourier transforms. Fourier transforms are used to represent a two dimensional sound signal.DCT uses a similar concept to reduce the gray-scale level or color signal amplitudes to equations that require very few points to locate the amplitude in Y-axis X-axis is for locating frequency.

DCT Coefficients

The output amplitudes of the set of 64 orthogonal basis signals are called OCT Co-efficients. **Quantization** This is a process that attempts to determine what information can be safely discarded without a significant loss in visual fidelity. It uses OCT co-efficient and provides many-to-one mapping. The quantization process is fundamentally lossy due to its many-to-one mapping.

De Quantization This process is the reverse of quantization. Note that since quantization used a many-to-one mapping, the information lost in that mapping cannot be fully recovered

Entropy Encoder / Decoder Entropy is defined as a measure of randomness, disorder, or chaos, as well as a measure of a system's ability to undergo spontaneous change. The entropy encoder compresses quantized DCT co-efficients more compactly based on their spatial characteristics. The baseline sequential. codec uses Huffman coding. Arithmetic coding is another type of entropy encoding **Huffman Coding** Huffman coding requires that one or more sets of huff man code tables be specified by the application for encoding as well as decoding. The Huffman tables may be pre-defined and used within an application as defaults, or computed specifically for a given image.

Baseline Sequential codec

It consists of three steps: Formation of DCT co-efficients quantization, and entropy encoding. It is a rich compression scheme.

DCT Progressive Mode

The key steps of formation of DCT co-efficients and quantization are the same as for the baseline sequential codec. The key difference is that each image component is coded in multiple scans instead ofa single scan.

Predictive Lossless Encoding

It is to define a means of approaching lossless continuous-tone compression. A predictor combines sample areas and predicts neighboring areas on the basis of the sample areas. The predicted areas are checked against the fully loss less sample for each area.

The difference is encoded losslessly using huffman on arithmetic entropy encoding .

Hierarchical Mode

The hierarchical mode provides a means of carrying multiple resolutions. Each successive encoding of the image is reduced by a factor of two, in either the horizontal or vertical dimension.

JPEG Methodology

The JPEG compression scheme is lossy, and utilizes forward discrete cosine transform (or forward DCT mathematical function), a uniform quantizer, and entropy encoding. The DCT function removes data

AMSCE - 1101

redundancy by transforming data from a spatial domain to a frequency domain; the quantizer quantizes DCT co-efficients with weighting functions to generate quantized DCT co-efficients optimized for the human eye; and the entropy encoder minimizes the entropy of quantized DCT co-efficients.

The JPEG method is a symmetric algorithm. Here, decompression is the exact reverse process of compression.

Figure below describes a typical DCT based encoder and decoder. Symmetric Operation of DCT based Codec

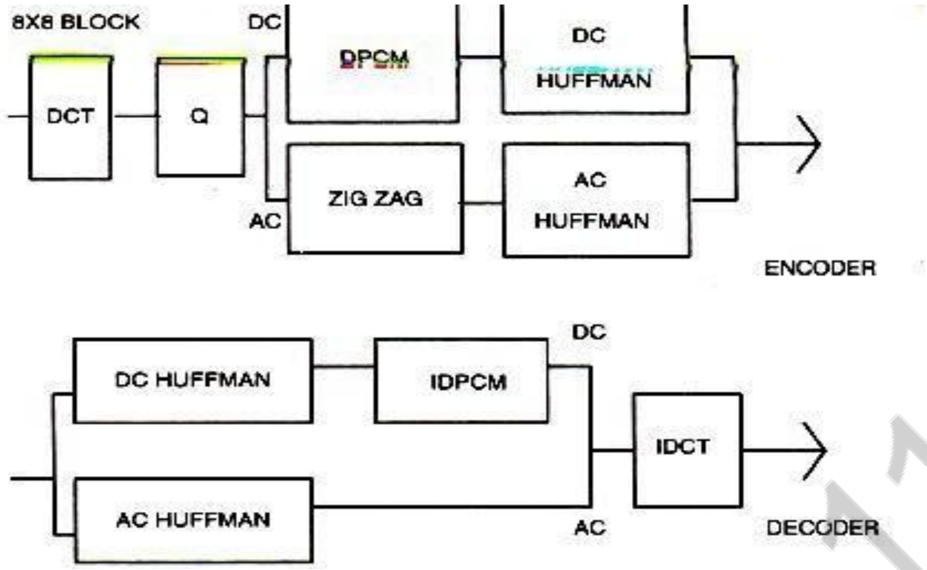


Figure below shows the components and sequence of quantization 5 * 8 Image blocks

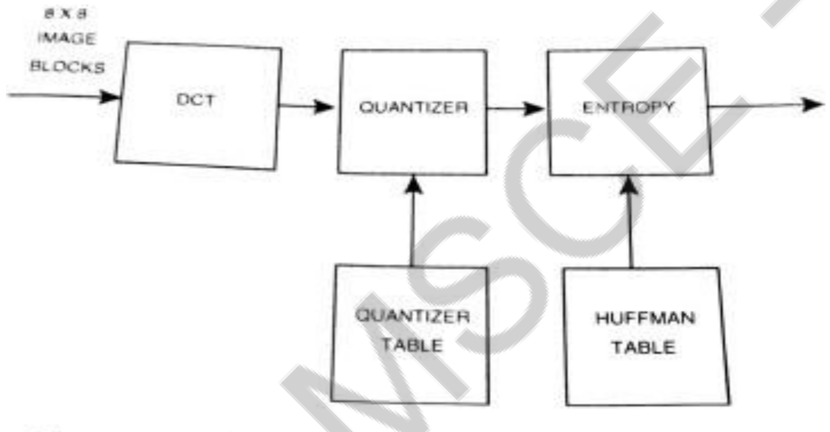


Fig. 2–4 Codec Components

AMSCE - 1101

The Discrete Cosine Transform(DCT)

DCT is closely related to Fourier transforms. Fourier transforms are used to represent a two dimensional sound signal.DCT uses a similar concept to reduce the gray-scale level or color signal amplitudes to equations that require very few points to locate the amplitude in Y-axis X-axis is for locating frequency.

The benefits provided by DCT transformations are as follows

- DCT is proven to be the optimal transform for large classes of images.
- DCT is an orthogonal transform; it allows converting the spatial representation of an 8x8 image to the frequency domain where only a few data points are required to represent the image.
- DCT generatescoefficients that are easily quantized to achieve good compression of the block.
- The DCT algorithm is well-behaved and can be computed efficiently by making it easy to implement in both hardware and software.
- The DCT algorithm is symmetrical, and an inverse DCT algorithm can be used to decompress an image.

DCT Calculations The formula for discrete cosine transform (creating DCT coefficients) is as follows:

$$DCT(i,j) = \frac{1}{\sqrt{2N}} C(i)C(j) \sum \sum pixel(x,y) \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)j\pi}{2N}\right]$$

The formula for inverse discrete cosine transform (restoring original pixel information from a DCT coefficient) is:

$$pixel(x,y) = \frac{1}{\sqrt{2N}} \sum \sum C(i)C(j) DCT(i,j) \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)j\pi}{2N}\right]$$

Let us consider the gray-scale image first. The image is first divided into an 8-pixel-by-8-pixel block. The 8×8 block is represented by an 8×8 matrix of gray values for the block. DCT coefficients are generated by applying the discrete cosine transform on the 8×8 block. Table 2-9 shows the input matrix of an 8×8 block of an image with its pixel values.

Table 2–9 Input Matrix of DCT Coefficients

132	136	138	140	144	145	147	155
136	140	140	147	140	148	155	156
140	143	144	148	150	152	154	155
144	144	146	145	149	150	153	160
150	152	155	156	150	145	144	140
144	145	146	148	143	158	150	140
150	156	157	156	140	146	156	145
148	145	146	148	156	160	140	145

Table 2–10 Output Matrix Showing DCT Coefficients

172	−18	15	−8	23	−9	−14	19
21	−34	24	−8	−10	11	14	7
−9	−8	−4	6	−5	4	3	−1
−10	6	−5	4	−4	4	2	1
−8	−2	−3	5	−3	3	4	6
4	−2	−4	6	−4	4	2	−1
4	−3	−4	5	6	3	1	1
0	−8	−4	3	2	1	4	0

AMSCE - 1101

Quantization

Quantization is a process of reducing the precision of an integer, thereby reducing the number of bits required to store the integer, thereby reducing the number of bits required to store the integer.

The baseline JPEG algorithm supports four color quantization tables and two huffman tables for both DC and AC DCT co-efficients. The quantized co-efficient is described by the following equation:

$$\text{Quantized Co-efficient } (i, j) = \frac{DCT(i,j)}{Quantum(i,j)}$$

ZigZag Sequence

Run-length encoding generates a code to represent the Count of zero-value OCT co-efficients. This process of run-length encoding gives an excellent compression of the block consisting mostly of zero values.

Further empirical work proved that the length of zero values in a run can be increased to give a further increase in compression by reordering the runs. JPEG came up with ordering the quantized OCT co-efficients in a ZigZag sequence. ZigZag sequence the sequence in which the cells are encoded.

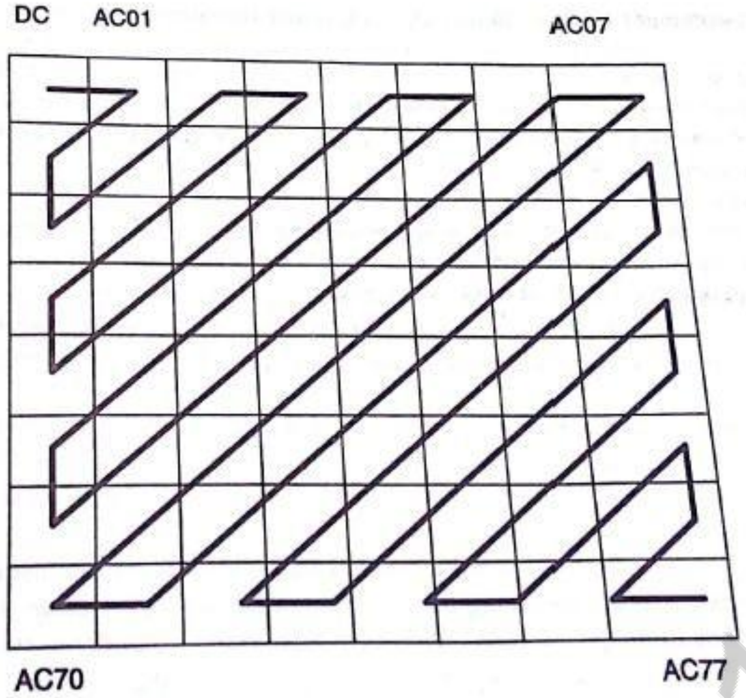


Fig. 2–5 Use of Zigzag Sequence for DCT Coefficients.

Entropy Encoding

Entropy is a term used in thermodynamics for the study of heat and work. Entropy, as used in data compression, is the measure of the information content of a message in number of bits. It is represented as

Entropy in number of bits = log2 (probability of Object)

Huffman versus Arithmetic coding

Huffman coding requires that one or more sets of Huffman code tables be specified by the application for coding as well as decoding. For arithmetic coding JPEG does not require coding tables.It able to adapt to the image statistics as it encodes the image.

DC coefficient coding

Before DC coefficients are compressed the DC prediction is processed first.In DC prediction the DC coefficient of the previous 8x8 block is subtracted from the current 8x8 block.

Two 8x8 blocks of a quantized matrix are shown in figure2.6. The Differential DC coefficient is delta $D=DC_x - DC_{x-1}$.

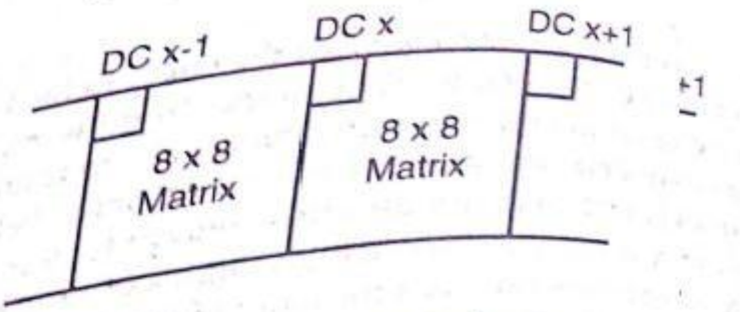


Fig. 2–6 Successive Blocks of Quantized Matrices.

AC coefficient coding

Each AC coefficient is encoded by utilizing two symbols symbol-1 and symbol-2. Symbol-1 represents two

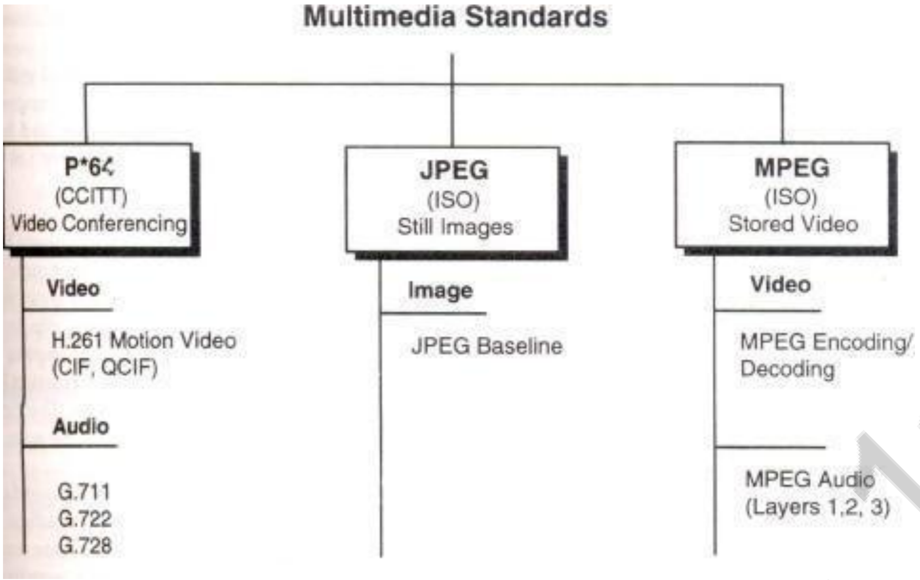
AMSCE - 1101

piece of information called “run length” and “size”. Symbol-2 represents the amplitude of the AC coefficient.

4.1.4 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

Multimedia standards for Video:



Requirements for full-motion Video 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 standard has identified a number of such issues that have been addressed by the standards activity. Let us review these Issues.

Random Access

The expectations generated for multimedia systems are the ability to playa sound or video clip from any frame with that clip, irrespective of on what kind-of media the information is stored

VCR paradigm

The VCR paradigm consists of the control functions typically found on a VCR such as play, fast forward, rewind, search forward and rewind search.

Multiplexing Multiple Compressed Audio and Video Bit Streams

It is a special requirement retrieved from different storage centers on a network. It may be received from different storage centers on a network. It may have to be achieved in a smooth manner to avoid the appearance of a jumpy screen.

Editability

Playback Device Flexibility

CCITT H.261 Video Coding Algorithms (P x 64)

The linear quantizer uses a step algorithm that can be adjusted based on picture quality and coding efficiency. The H.261 is a standard that uses a hybrid of OCT and OPCM (differential pulse Code Modulation) schemes with motion estimation.

It also defines the data format. Each MB contains the OCT coefficients (TCOEFF) ofa block followed by an EOB (a fixed length end-of-block marker). Each MB consists of block data and an MB header. A GOB (Group of Blocks) consists of a GOB header. The picture layer consists of a picture header.

AMSCE - 1101

The H.261 is designed for dynamic use and provides a fully contained organization and a high level of interactive control.

Moving Picture Experts Group Compression

The MPEG standards consist of a number of different standards.

The MPEG 2 suite of standards consist of standards for MPEG2 Video, MPEG - 2 Audio and MPEG - 2 systems. It is also defined at different levels, called profiles.

The main profile is designed to cover the largest number of applications. It supports digital video compression in the range of2 to 15 M bits/sec. It also provides a generic solution for television worldwide, including cable, direct broadcast satellite, fibre 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 interframe coding. If we access information randomly by frame requires coding confined to a specific frame, then it is known as intraframe coding.

The MPEG standard addresses these two requirements by providing a balance between interframe 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 YUVcolor 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 noncausal coding to further reduce spatial redundancy.

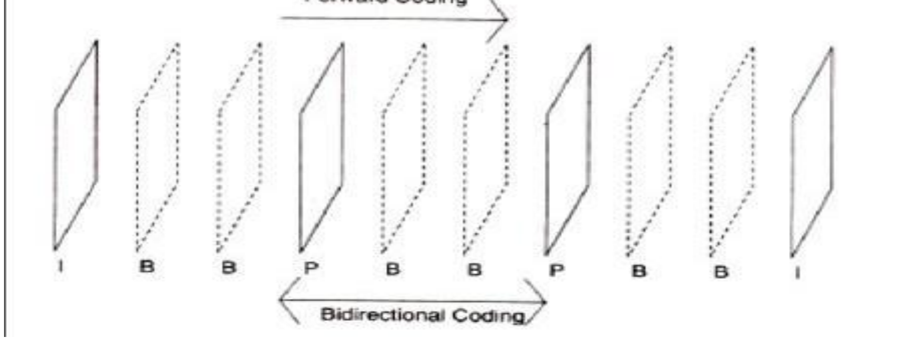
The MPEG algorithm for spatial reduction is lossy and is defined as a hybrid which employs motion compensation, forward discrete cosine transform (DCF), a uniform quantizer, and Huffman coding. Block-based motion compensation is *utilized* for *reducing temporal* redundancy (i.e. to reduce the amount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

Moving Picture Types

Moving pictures consist of sequences of video pictures or t1'ame'S 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)
Undirectionally predicted pictures (U)
Bidirectiomflly predicted pictures (B)

A Gap consists of consecutive pictures that begin with an intrapicture. The intrapicture is coded without any reference to any other picture in the group.



Let us review the concept of Macroblocks and understand the role they play in compression

MACRO BLOCKS

For the video coding algorithm recommended by CCITT, CIF and QCIF are diviqed into a hierarchical block structure consisting of pictures, groups of blocks (GOBs), Macro Blocks(MBs), and blocks. Each picture frame is divided into 16 x 16 blocks. Each Macroblock is composed of four 8 x 8 (Y) luminance blocks and two 8 x 8 (C_b and Cn) chrominance blocks. This set of six blocks, called a macroblock; 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 reference, frames are averaged.

Picture Coding Method

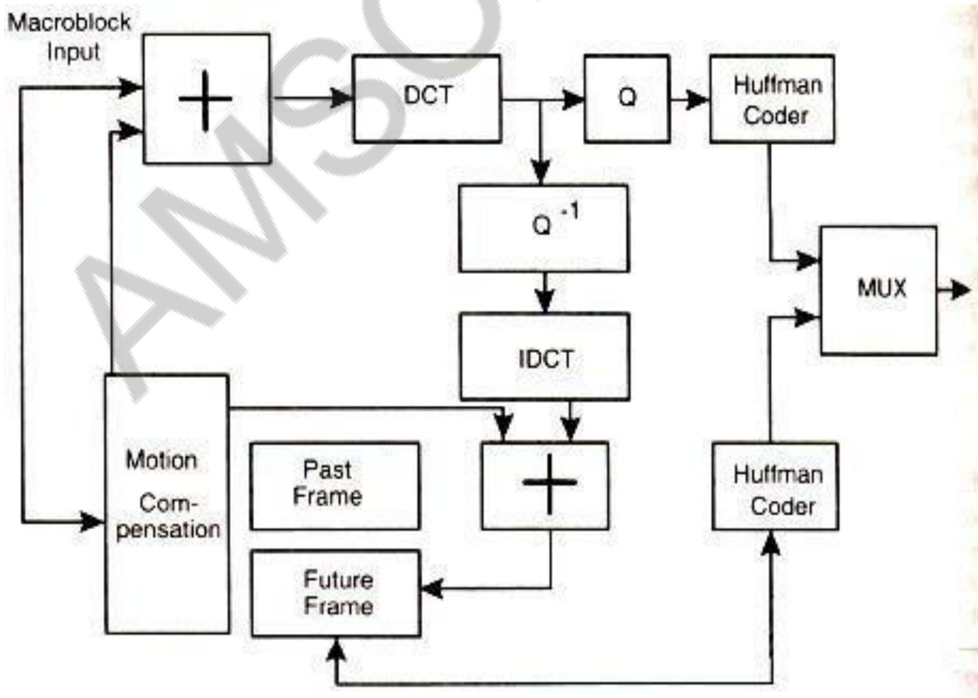
In this coding method, motion compensation is applied bidirectionally. In MPEG terminology, the motion-compensated units are called macro blocks (MBs).

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 quantizer, Huffman coder and Motion compensation. These represent the key modules in the encoder.

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

AMSCE - 1101

encoding. For full-motion video, several more levels of motion compensation compression and coding are applied.

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 fro backward compatibility.
   * Layer-2 audio definitions for MPEG-2 and stereo sound.
   * Multichannel sound.
3. Multiplexing: MPEG-2 definitions

MPEG-2, "The Grand Alliance"

It consists of following companies AT&T, MIT, Philips, Sarnoff Labs, GI Thomson, and Zenith.

The MPEG-2committee and FCC formed this alliance. These companies together have defined the advanced digital television system that include the US and European HDTV systems. The outline of the advanced digital television system is as follows:

1. Format: 1080/2: 1160 or 720/1.1160
2. Video coding: MPEG-2 main profile and high level
3. Audio coding: Dolby AC3
4. Multiplexor: As defined in MPEG-2

Modulation: 8- VSB for terrestrial and 64-QAM for cable.

Vector Quantization

Vector quantization provides a multidimensional representation of information stored in look-up tables, vector quantization is an efficient pattern-matching algorithm in which an image is decomposed into two or more vectors, each representing particular features of the image that are matched to a code book of vectors. These are coded to indicate the best fit.

In image compression, source samples such as pixels are blocked into vectors so that each vector describes a small segment or sub block of the original image.

The image is then encoded by quantizing each vector separately

Intel's Indeo Technology

It is developed by Intel Architecture Labs Indeo Video is a software technology that reduces the size of uncompressed digital video files from five to ten times.

Indeo technology uses multiple types of 'lossy' and 'lossless' compression techniques.

4.1.5 AUDIO COMPRESSION

Audio consists of analog signals of varying frequencies. The audio signals are converted to digital form and then processed, stored and transmitted. Schemes such as linear predictive coding and adaptive differential. pulse code modulation (ADPCM) are utilized for compression to achieve 40-80% compression.

FRACTAL COMPRESSION

A fractal is a multi-dimensional object with an irregular shape or body that has approximately the same shape or body irrespective of size. For example, if you consider 'stick' as your object, the fractal is defined, mathematically as

$$D = lim \frac{\ln[N(L)]}{\ln[1/L]}$$

where L - approaches 0,

N(L) ~ number of stick L, and L is the length of the stick.

4.2 DATA AND FILE FORMATS STANDARDS

There are large number of formats and standards available for multimedia system. Let us discuss about the following file formats:

.:. Rich-Text Format (RTF)

AMSCE-1101

.:. Tagged Image file Format (TIFF)
.:. Resource Image File Format (RIFF)
.:. Musical Instrument Digital Interface (MIDI)
.:. Joint Photographic Experts Group (JPEG)
.:. Audio Video Interleaved (AVI) Indeo file format
.:. TWAIN.

Rich Text Format

This format extends the range of information from one word processor application or DTP system to another.

The key format information carried across in RTF documents are given below: Character Set: It determines the characters that supports in a particular implementation.

Font Table: This lists all fonts used. Then, they are mapped to the fonts available in receiving application for displaying text.

Color Table: It lists the colors used in the documents. The color table then mapped for display by receiving application to the nearer set of colors available to that applications.
Document Formatting: Document margins and paragraph indents are specified here.

Section Formatting: Section breaks are specified to define separation of groups of paragraphs.
Paragraph Formatting: It specifies style sheds. It specifies control characters for specifying paragraph justification, tab positions, left, right and first indents relative to document margins, and the spacing between paragraphs.
General Formatting: It includes footnotes, annotations, bookmarks and pictures.

Character Formatting: It includes bold, italic, underline (continuous, dotted or word), strike through, shadow text, outline text, and hidden text.
Special Characters: It includes hyphens, spaces, backslashes, underscore and so on

TIFF File Format

TIFF is an industry-standard file format designed to represent raster image data generated by scanners, frame grabbers, and paint/ photo retouching applications.
TIFF Version 6.0 .
It offers the following formats:

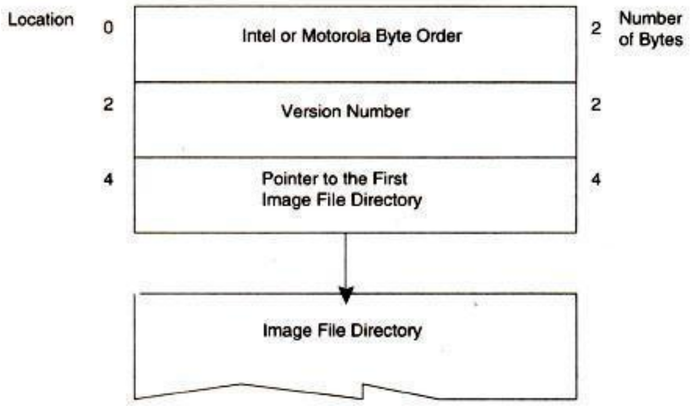
(i) Grayscale, palette color, RGB full-color images and black and white.
(ii) Run-length encoding, uncompressed images and modified Huffman data compression schemes.

The additional formats are:

(i) Tiled images, compression schemes, images using CMYK, $YCbC_r$ color models.

TIFF Structure

TIFF files consists of a header. The header consists of byte ordering flag, TIFF file format version number, and a pointer to a table. The pointer points image file directory. This directory contains table of entries of various tags and their information.
TIFF file format Header:



The next figure shows the IFD (Image File Directory) as its content. The IFD is avariable –length table containing directory entries. The length of table depends on the number of directory entries in the table. The first two bytes contain the total number of entries in the table followed by directory entrie. Each directory entry consists of twelve bytes.The Last item in the IFD is a four byte pointer that points to the next IFD. The byte content of each directory entry is as follows:

- The first two byte contains tag number-tag ID.
- The second two byte represent the type of data as shown in table3-1 below.
- The next four bytes contains the length for the data type.
- The final four bytes contain data or a pointer.

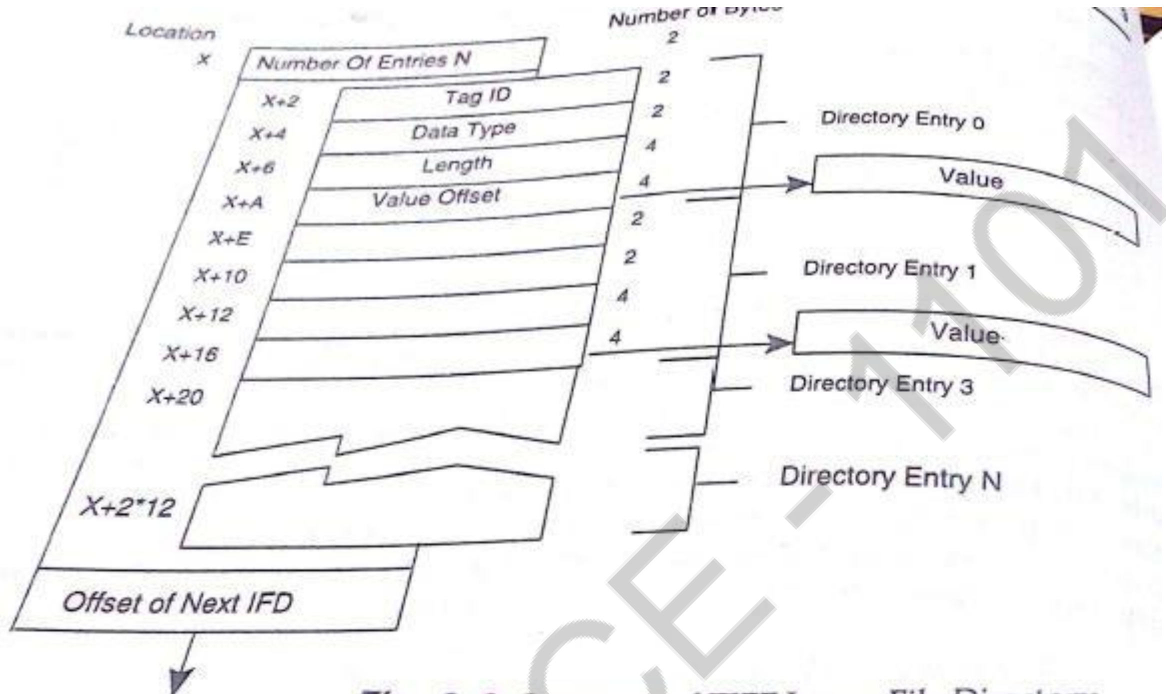


Fig. 3–3 Structure of TIFF Image File Directory

Table 3-1 Bytes 3 and 4 of Directory Entry Representing Data Type

Value	Data Type	Description	Length
1	Byte	8 bit unsigned byte	1
2	ASCII	8 bit bytes to represent ASCII codes; the last byte of the ASCII data is null for null termination	1
3	SHORT	16 bit (2 bytes) unsigned integer	2
4	LONG	32 bit (4 bytes) unsigned word	4
5	RATIONAL	Two LONGS to represent the numerator and the denominator of a fraction	8

AMSCE - 1101

TIFF Tags

The first two bytes of each directory entry contain a field called the Tag ID.

Tag IDs arc grouped into several categories. They are Basic, Informational, Facsimile, Document storage and Retrieval.

TIFF Classes: (Version 5.0)It has five classes

1. Class B for binary images
2. Class F for Fax
3. Class G for gray-scale images
4. Class P for palette color images

5. Class R for RGB full-color images.

4.2.3 Resource Interchange File Format (RIFF)

RIFF provides a framework or an envelope for multimedia file formats for Microsoft Windows based application and it can be used to convert a custom file format to a RIFF file format by wrapping a RIFF structure in the form of riff chunks to a MIDI file.

The **RIFF** file formats consist' of blocks of data called chunks. They are

RIFF Chunk - defines the content of the RIFF file.

List Chunk - allows to embed archival location copy right information and creating date.

Subchunk - allow additional information to a primary chunk

The first chunk in a RIFF file must be a RIFF chunk and it may contain one or more sub chunk .The first four bytes of the RIFF chunk data field are allocated for the form type field containing four characters to identify the format of the data stored in the file: AVI, WAV, RMI, PAL and so on.

The following figure shows the organization of RIFF chunk

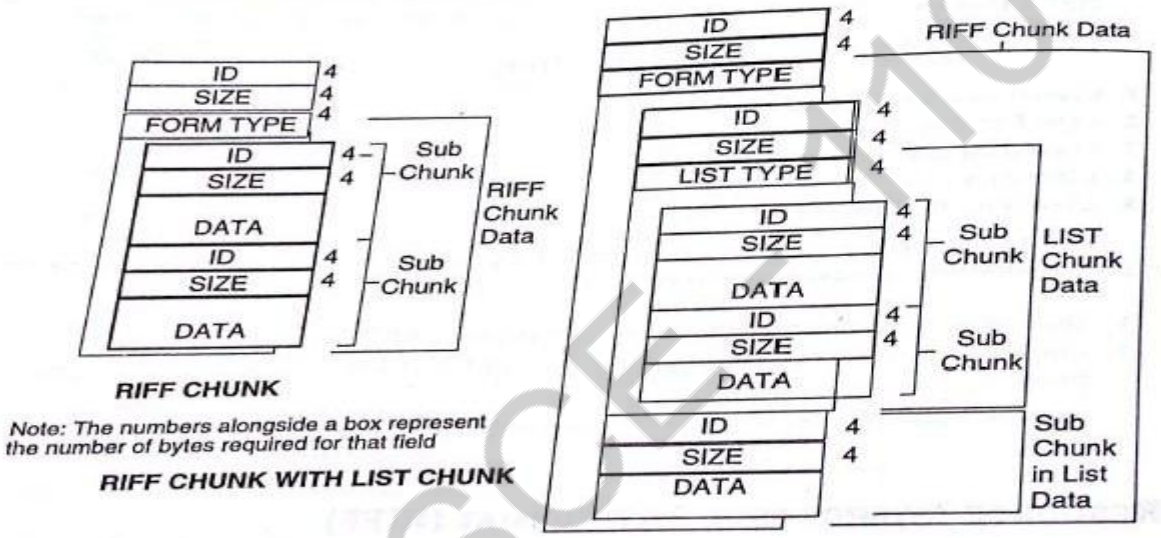


Fig. 3–4 Organization of RIFF Chunks

AMSCE - 101

File name extension for RIFF:

The sub chunk contains a four-character ASCII string 10 to identify the type of data.

Four bytes of size contains the count of data values, and the data. The data structure of a chunk is same as all other chunks.

RIFF chunk with two sub chunk:

The first 4 characters of the RlFF chunk are reserved for the "RIFF" ASCII string. The next four bytes define the total data size.

File type	Form typl	File extension
Waveform Audio File	WAVE	.WAV
Audio Video Interleaved file	AVI	.AVI
MIDI File	RMID	.RMI
Device Independent Bitmap file	RDIB	.RDI
Pallette File	PAL	.PAL

AMSCE - 1101

The first four characters of the data field are reserved for form tyPe. The rest of the data field contains two subchunk:

(i) fmt ~ defines the recording characteristics of the waveform.

(ii) data ~ contains the data for the waveform.

```
RIFF('WAVE'
    { SubChunk1 'fmt' < Describes recording characteristics of the waveform>
      SubChunk2 'data'< Waveform data>
    }
  )
```

LIST Chunk

RlFF chunk may contains one or more list chunks.

List chunks allow embedding additional file information such as archival location, copyright information, creating date, description of the content of the file.

***RlFF* MIDI FILE FORMAT**

RlFF MIDI contains a *RlFF* chunk with the form type "RMID"and a subchunk called "data" for MIDI data.

The 4 bytes are for ID of the *RlFF* chunk. 4 bytes are for size 4 bytes are for form type 4 bytes are for ID of the subchunk data and 4 bytes are for the size of MIDI data.

RIFF DIBS (Device-Independent Bit Maps) .

DIB is a Microsoft windows standard format. It defines bit maps and color attributes for bit maps independent of devices. DIEs are normally embedded in .BMP files, .WMF meta data files, and .CLP files.

DIB Structure

BIT MAP INFOHEADER is the bit map information header.
RGBEQUAD is the color table structure.
PIXELs are the array of bytes for the pixel bit map.
The following shows the DIE file format

A RIFF DIB file format contains a RIFF chunk with the Form Type "RDIB" and a subchunk called "data" for DIB data.

4 bytes denote ID of the RIFF chunk
4 bytes refer size ofXYZ.RDI 4 bytes define Forum Type
4 bytes describe ID of the sub chunk data 4 bytes define size of DIB data.

RIFF PALETTE File format

The RIFF Palette file format contains a RIFF chunk with the Form Type "RP AL" and a subchunk called "data" for palette data. The Microsoft Windows logical palette structure is enveloped in the RIFF data subchunk. The palette structure contains the palette version number, number of palette entries, the intensity of red, green and blue colours, and flags for the palette usage. The palette structure is described by the following code segment:

typedef struct tagLOGP ALETTE {
WORD palVersion; *//Windows* version number for the structure
WORD palNumEntries;
PALETIEENTRY palpalEntry []; *//array* of PALEN TRY data
} LOGPALETTE;

AMSCE - 1101

BITMAPINFOHEADER	RGBQUAD	PIXELS

BITMAPINFOHEADER	BITMAPINFO=BITMAPINFOHEADER+ RGBQUAD	PIXELS

AMSCE - 1101

RIFF Audio Video Interleaved(AVI) file format:
AVI files can be envelope within the RIFF format to create the RIFF AVI file . A RIFF AVI file contains a RIFF chunk withi the form type “AVI” and two mandatory list chunks "hdr 1" and "movi". The "hdr 1" defines the format of the data "Movi" contains the data for the audio-video streams. The third list chunk called "id xl", is an optional index chunk.

Boundary condition Handling for AVI files

Each audio and video stream is grouped together to form a ree chunk. If the size of a rec chunk is not a multiple of2048 bytes, then the rec chunk is padded to make the size of each rec chunk a multiple of 2048 bytes. To align data on a 2048 byte boundary, dummy data is added by a "JUNK" data chunk. The JUNK chunk is a standard RIFF chunk with a 4 character identifier, "JUNK," followed by the dummy data.

RIFF Waveform Audio File Format with INFO List Chunk

The pseudo code of the RIFF waveform file format with INFO list chunks is as follows:

```
RIFF(' WAVE '
  LIST ('INFO'
    { SubChunk   'INAM'   < Name of the company >
      SubChunk   'ICOP'   < Copyright notice >
      SubChunk   'ICRD'   < Creation data >
    }
  )
  { SubChunk   'fmt'    < WAVEFORMAT structure describes
                          recording characteristics of
                          the waveform>
    SubChunk   'data'   < Waveform data>
  }
)
```

The subchunk '"fmt" contains the following data structure:

```
typedef struct tagwaveformat {
  WORD   wFormatTag;       // Waveform format type
  WORD   nChannels;        // Number of channels
  DWORD  nSamplesPerSec;   // Sampling rate
  DWORD  nAvgBytesPerSec;  // Average transfer rate for buffering
  WORD   nBlockAlign;      // Block alignments
  UINT   nBitsPerSample;   // Number of bits per sample
}WAVEFORMAT;
```

AMSCE - 1101

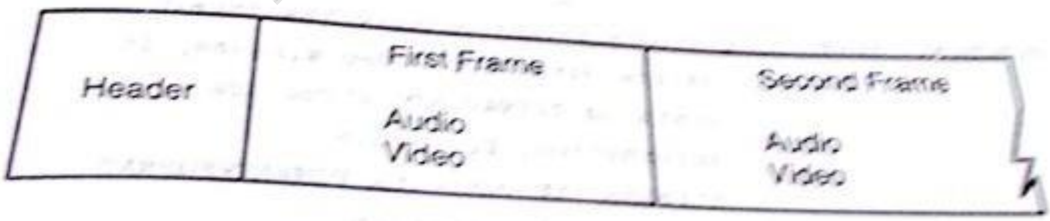


Fig. 3–5 Interleaved Audio and Video for AVI Files

4.2.4 MIDI File Format

The MIDI file format follows music recording metaphor to provide the means of storing separate tracks of music for each instrument so that they can be read and syn~hronized when they are played.

The MIDI file format also contains chunks (i.e., blocks) of data. There are two types of chunks: (i) header chunks (ii) track chunks.

Header Chunk

It is made up of 14 bytes .

The first four-character string is the identifier string, "MThd" .

The second four bytes contain the data size for the header chunk. It is set to a fixed value of six bytes.

The last six bytes contain data for header chunk.

Track chunk

The Track chunk is organized as follows:

- .:. The first 4-character string is the identifier.
- .:. The second 4 bytes contain track length.

MIDI Communication Protocol

This protocol uses 2 or more bytes messages.

The number of bytes depends on the types of message. There are two types of messages:

(i) Channel messages and (ii) System messages.

Channel Messages

A channel message can have up to three bytes in a message. The first byte is called a status byte, and other two bytes are called data bytes. The channel number, which addresses one of the 16 channels, is encoded by the lower nibble of the status byte. Each MIDI voice has a channel number; and messages are sent to the channel whose channel number matches the channel number encoded in the lower nibble of the status byte. There are two types of channel messages: voice messages and the mode messages.

Voice messages

Voice messages are used to control the voice of the instrument (or device); that is, switch the notes on or off and sent key pressure messages indicating that the key is depressed, and send control messages to control effects like vibrato, sustain, and tremolo. Pitch wheel messages are used to change the pitch of all notes

.**Mode messages**

Mode messages are used for assigning voice relationships for up to 16 channels; that *is,* to set the device to MOWO mode or POLY mode. Omni Mode on enables the device to receive voice messages on all channels.

System Messages

System messages apply to the complete system rather than specific channels and do not contain any channel numbers. There are three types of system messages: common messages, real-time messages, and exclusive messages. In the following, we will see how these messages are used.

Common Messages These messages are common to the complete system. These messages provide for functions such as select a song, setting the song position pointer with number of beats, and sending a tune request to an analog synthesizer.

System Real Time Messages

These messages are used for setting the system's real-time parameters. These parameters include the timing clock, starting and stopping the sequencer, ressuming the sequencer from a stopped position, and resetting the system.

System Exclusive messages

These messages contain manufacturer-specific data such as identification, serial number, model number, and

AMSCE - 1101

Table 3–26 Example of Header Chunk for MIDI file

Header Field	Byte #	Value
Identifier String	1 - 4	4D 54 68 64
Data Size	5 - 8	00 00 00 06
Data	9 - 14	00 00 00 01 01 E0

AMSCE - 1101

other information. Here, a standard file format is generated which can be moved across platforms and applications.

JPEG Motion Image:

JPEG Motion image will be embedded in A VI RIFF file format.
There are two standards available:
(i) MPEG ~ In this, patent and copyright issues are there.
(ii) MPEG 2 ~ It provide better resolution and picture quality.

TWAIN

A standard interface was designed to allow application to interface with different types of input devices such as scanners, digital still cameras, and so on, using a generic TWAIN interface without creating device- specific driver. The benefits of this approach areas follows

I. Application developers can code to a single TWAIN specification that allows application to interface to all TWAIN complaint input devices.
2. Device manufactures can write device drivers for their proprietary devices and, by complying to the TWAIN specification , allow the devices to be used by all TWAIN-compliant applications

TWAIN Specification Objectives

The TWAIN specification was started with a number of objectives:

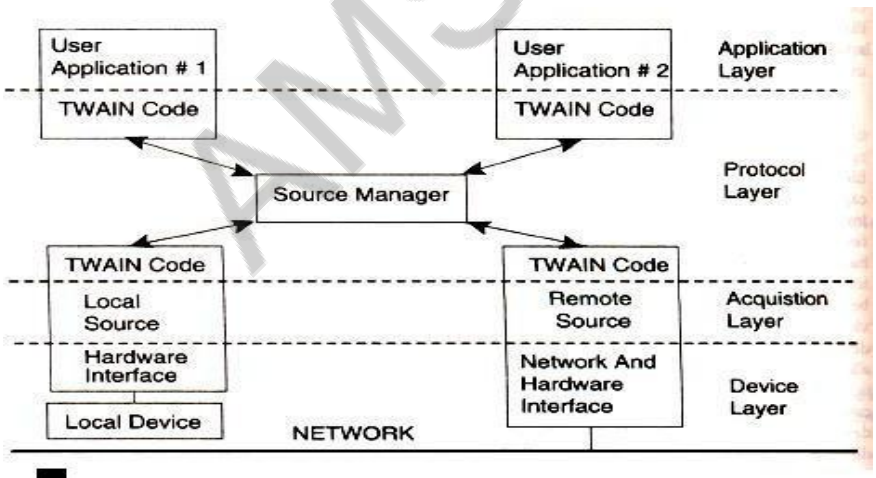
- Supports multiple platforms: including Microsoft Windows, Apple Macintosh OSSystem6.xor7.x, UNIX, andIBMOSl2.
- Supports multiple devices: including scanners, digital camera, frame grabbers etc.
- Standard extendibility and backward compatibility: The TWAIN architecture is extensible for new types of devices and new device functionality. New versions of the specification are backward compatible.
- Easy to use: The standard is well documented and easy to use.

The TWAIN architecture defines a set of application programming interfaces (APls) and a protocol to acquire data from input devices. It is a layered architecture consisting of a protocol layer and an acquisition layer sandwiched between the application and device layers. The protocol layer is responsible for communication between the application and acquisition layers. The acquisition layer containsthevirtualdevicedrivertocontrolthedevice.Thisvirtuallayerisalsocalledthesource.

TWAIN ARCHITECHTURE:

The Twain architecture defines a set of application programming interfaces (APls) and a protocol to acquire data from input devices.
It is a layered architecture.
It has application layer, the protocol layer, the acquisition layer and device layer.

AMSCE - 1101


User Application # 1
TWAIN Code
User Application # 2
TWAIN Code
Application Layer
Source Manager
Protocol Layer
TWAIN Code
Local Source
Hardware Interface
Local Device
TWAIN Code
Remote Source
Network And Hardware Interface
Acquistion Layer
Device Layer
NETWORK

AMSCE - 1101

Application Layer: A TWAIN application sets up a logical connection with a device. TWAIN does not impose any rules on the design of an application. However, it set guidelines for the user interface to select sources (logical device) from a given list of logical devices and also specifies user interface guidelines to acquire data from the selected sources.

The Protocol Layer: The application layer interfaces with the protocol layer. The protocol layer is responsible for communications between the application and acquisition layers. The protocol layer does not specify the method of implementation of sources, physical connection to devices, control of devices , and other device-related functionality. This clearly highlights that applications are independent of sources. The heart of the protocol layer, as shown in Figure is the Source Manager. It manages all sessions between an application and the sources, and monitors data acquisition transactions.

The functionality of the Source Manager is as follows:

- Provide a standard API for all TWAIN compliant sources
- Provides election of sources for a user from within an application
- Establish logical sessions between applications and sources, and also manages essions between multiple applications and multiple sources
- Act as a traffic cop to make sure that transactions and communication are routed to appropriate sources, and also validate all transactions
- Keep track of sessions and unique session identities
- Load or unload sources as demanded by an application
- Pass all return code from the source to the application
- Maintain a default source

The Acquisition Layer: The acquisition layer contains the virtual device driver, it interacts directly with the device driver. This virtual layerisalsocalledthesource.Thesourcecanbelocalandlogicallyconnected to a local device, or remote and logically connected to a remote device(i.e.,a device ove rthe network). The source performs the following functions:

~ Control of the device.
~ Acquisition of data from the device.
~ Transfer of data in agreed (negotiated) format. This can be transferred in native format or another filtered format.
~ Provision of a user interface to control the device.

The Device Layer: The purpose of the device driver is to receive software commands and control the device hardware accordingly. This is generally developed by the device manufacturer and shipped with the device.

NEW WAVE RIFF File Format: This format contains two subchunks:
(i) Fmt **(ii)** Data.
It may contain optional subchunks:

(i) Fact
(ii) Cue points
(iii) Play list
(iv) Associated datalist.

Fact Chunk: It stores file-dependent information about the contents of the WAVE file.
Cue Points Chunk: It identifies a series of positions in the waveform data stream.
Playlist Chunk: It specifies a play order for series of cue points.
Associated Data Chunk: It provides the ability to attach information, such as labels, to sections of the waveform data stream.
Inst Chunk: The file format stores sampled sound synthesizer's samples.

4.3 MULTIMEDIA INPUT/OUTPUT TECHNOLOGIES

Multimedia Input and Output Devices

Wide ranges of Input and output devices are available for multimedia.

Image Scanners: Image scanners are the scanners by which documents or a manufactured part are scanned.

AMSCE - 1101

The scanner acts as the camera eye and take a photograph of the document, creating an unaltered electronic pixel representation of the original.

Sound and Voice: When voice or music is captured by a microphone, it generates an electrical signal. This electrical signal has analog sinusoidal waveforms. To digitize, this signal is converted into digital voice using an analog-to-digital converter.

Full-Motion Video: It is the most important and most complex component of Multimedia System. Video Cameras are the primary source of input for full-motion video.

. **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-ordinating windows pen applications with the pen. It works with Recognizer, dictionary, and display driver to recognize and display pen drawn objects.

Recognizor: 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.

Video and Image Display Systems Display System Technologies

There are variety of display system technologies employed for decoding compressed data for displaying.
Mixing and scaling technology: For VGA screen, these technologies are used.

VGA mixing: Images from multiple sources are mixed in the image acquisition memory.

VGA mixing with scaling: Scalar ICs are used to sizing and positioning of images in predefined windows.

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.

Visual Display Technology Standards

MDA: Monochrome Display Adapter.

It was introduced by IBM . displays 80 x 25 rows and columns .

:. It could not display bitmap graphics .

:. It was introduced in 1981.

CGA: Color Graphics Adapter .

:. It was introduced in 1981.

.:. It was designed to display both text and bitmap graphicsi

it supported RGB color display,

.:. It could display text at a resolution of 640 x 200 pixels .

:. It displays both 40 x 25 and 80 x 25 row!' and columns of text characters.

MGA: Monochrome Gr.aphics Adapter .

:. It was introduced in 1982 .

:. It could display both text and graphics .

:. It could display at a resolution 720 x 350 for text and
720 x 338 for Graphics . MDA is compatible mode for this standard.

EGA: Enhanced Graphics Adapter .

:. It was introduced in 1984 .

:. It emulated both *MDt.* and CGA standards .

:. It allowed the display of both text and graphics in 16 colors at a resolution of 640 x· 350 pixels.

PGA: Professional Graphics Adapter.

.:. It was introduced in 1985 .

:. It could display bit map graphics at 640 x 480 resolution and 256 colors .

• :. Compatible mode of this standard is CGA.

VGA: Video Graphics Array . :. It was introduced by IBM in 1988 .

:. It offers CGA and EGA compatibility .

:. It display both text and graphics .

AMSCE - 1101

AMSCE - 1101

:. It generates analog RGB signals to display 256 colors .

:. It remains the basic standard for most video display systems.

SVGA: Super Video Graphics Adapter. It is developed by VESA (Video Electronics Standard Association) . It's goal is to display with higher resolution than the VGA with higher refresh rates with minimize flicker.

XGA: Extended Graphics Array

It is developed by IBM . It offers VGA compatible mode . Resolution of 1024 x 768 pixels in 256 colors is offered by it. XGA utilizes an interlace scheme for refresh rates.

Flat Panel Display system

Flat panel displays use a fluorescent tube for backlighting to give the display a sufficient level of brightness. The four basic technologies used for flat panel display are:

1. Passive-matrix monochrome
2. Active-matrix monochrome
3. Passive-matrix color
4. Active-matrix color.

LCD (Liquid Crystal Display)

Construction: Two glass plates each containing a light polarizer at right angles to the other plate, sandwich the nematic (thread like) liquid crystal material.

Liquid crystal is the compounds having a crystalline arrangement of molecules. But it flow like a liquId. Nematic liquid crystal compounds are tend to keep the long axes of rod-shaped molecules aligned. Rows of horizontal transparent conductors are built into one glass plate, and columns of vertical conductors are put into the other plate. The intersection of two conductors defines a pixel position.

Passive Matrix LCD

Working: Normally, the molecules are aligned in the 'ON' state.

Polarized light passing through the materials is twisted so that it will pass through the opposite polarizer. The light is then reflected back to the viewer. To turn off the pixel, we have to apply a voltage to the two intersecting conductors to align molecules so that the light is not twisted.

ACTIVE Matrix LCD

In this device, a transistor is placed at each pixel position, using thin-film transisor technology.

The transistors are used to control the voltage at pixel locations and to prevent charge from gradually leaking out of the liquid crystal cells.

PRINT OUTPUT TECHNOLOGIES

There are various printing technologies available namely Dot matrix, inkjet, laser print server 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 software to access the features ofthe printer.

IMAGE SCANNERS

In a document imaging system, documents are scanned using a scanner. \The document being scanned is placed on the scanner bed or fed into the sheet feeder of the scanner .The scanner acts as the camera eye and takes a photograph of the document, creating an image of the original. The pixel representation (image) is recreated by the display software to render the image of the original document on screen or to print a copy

AMSCE - 1101

of it.

Types of Scanners

A and B size Scanners, large form factor scanners, flat bed scanners, Rotory drum scanners and hand held scanners are the examples of scanners.

Charge-Coupled Devices All scanners use charge-coupled devices as their photosensors. CCDs consists 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.

Image Enhancement Techniques

HalfTones In a half-tone process, patterns of dots used to build .scanned or printed image create the illusion of continuous shades of gray or continuous shades of color. Hence only limited number of shades are created. This process is implemented in news paper printers.
But in black and white photograph or color photograph, almost infinite levels of tones are used.

Dithering

Dithering is a process in which group of pixels in different patterns are used to approximate halftone patterns by the scanners. It is used in scanning original black and white photographs.

Image enhancement techniques includes controls of brightness, deskew (Automatically corrects page alignment), contrast, sharpening, emphasis and cleaning up blacknoise dots by software.

Image Manipulation

It includes scaling, cropping and rotation.

Scaling: Scaling can be up or down, the scaling software is available to reduce or enlarge. This software uses algorithms.

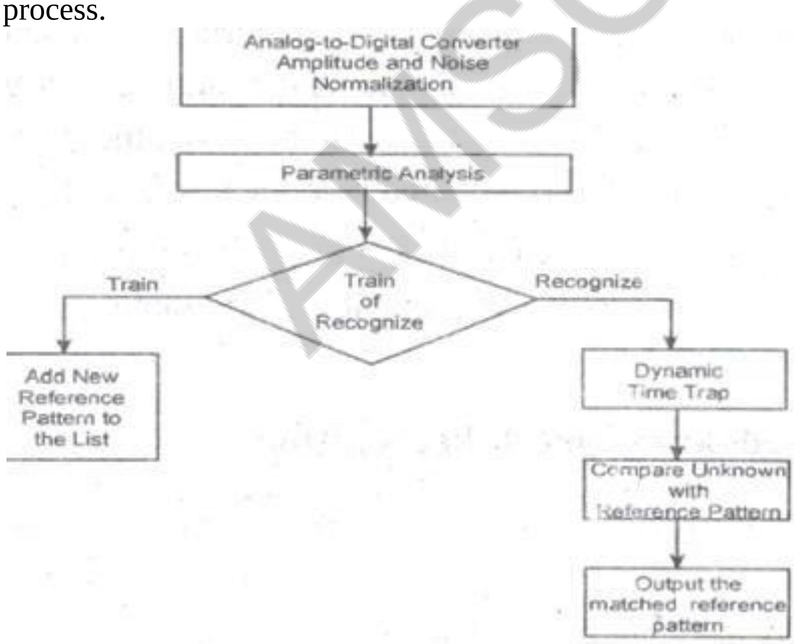
Cropping: To remove some parts of the image and to put the rest of the image as the subset of the old image.

Rotation: Image could be rotated at any degree for displaying it in different angles.

4.4 DIGITAL VOICE AND AUDIO

4.4.1Digital 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 fornlat by analog-to-digital converter (ADC) using sampling process.



Sampling process

Sampling is a process where the analog signal is sampled over time at regular intervals to obtain the amplitude of the analog signal at the sampling time.

Sampling rate

The regular interval at which the sampling occurs is called the sampling rate.

Digital Voice

Speech is analog in nature and is cOl1veli 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.

The four important factors governing the

ADC process are **sampling rate resolution linearity and conversion speed.**

Sampling Rate: The rate at which the ADC takes a sample of an analog signal. **Resolution:** The number of bits utilized for conversion determines the resolution of ADC.

Linearity: Linearity implies that the sampling is linear at all frequencies and that the amplitude tmly represents the signal.

Conversion Speed: It is a speed of ADC to convert the analog signal into Digital signals. It must be fast enough.

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 relevent 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 concatenatic basic phonemes) time duration. The unknown pattern is then compared with the reference patterns.

A speaker independent isolated word recognizer can be achieved by groupi.ng a large number of samples corresponding to a word into a single cluster.

2Connected-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.

Continuous Speech Recognition

This sytem can be divided into three sections:

(i) A section consisting of digitization, amplitude normalization, time nonnalization and parametric representation.

(ii) Second section consisting of segmentation and labeling of the speech segment into a symbolic string based on a knowledgebased or rule-based systems.

(iii) The final section is to match speech segments to recognize word sequences.

Voice Recognition performance

It is categorized into two measures: Voice recognition performance and system performance. The

AMSCE - 1101

following four measures are used to determine voice recognition performance.

1. **Voice Recognition Accuracy**

Voice Recognition Accuracy = *Number of correctly recognized words* x l00
Number of test words

2. **Substitution Error**

Substitution error = *Number of substituted words* x l00 *Number of test words*

3. **No Response Error**

Number of no responses x 100
Number of test words

4. **Insertion Error**

Insertion error =*Number of insertion error* x 100
Number of test words

Voice Recognition Applications

Voice mail integration: The voice-mail message can be integrated with e-mail messages to create an integrated message.

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.

- 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 data base to take follow-up action ifhe 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.

4.4.2 Musical Instrument Digital Interface (MIDI)

MIDI interface is developed by Daver Smith of sequential circuits, inc in 1982. It is an universal synthesizer interface

MIDI Specification 1.0MIDI is a system specification consisting of both hardware and software ~omponents which define inter-coimectivity and a communication protocol for electronic sysnthesizers, sequences, rythm machines, personal computers, and other electronic musical instruments. The inter-connectivity defines the standard cabling scheme, connector type and input/output circuitry which enable these different MIDI instruments to be interconnected. The communication protocol defines standard multibyte messages that allow controlling the instrument"s voice and messages including to send response, to send status and to send exclusive.

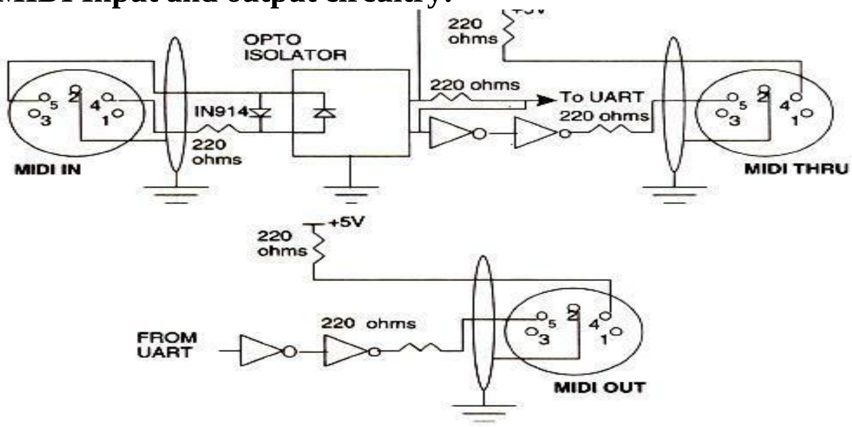
MIDI Hardware Specification

The MIDI. hardware specification require five pin panel mount requires five pin panel mount receptacle DIN connectors for MIDI IN, MIDI OUT and MIDI THRU signals. The MIDI IN connector is for input signals The MIDI OUT is for output signals MIDI THRU connector is for daisy-chaining multiple MIDI instruments.

AMSCE - 1101

AMSCE - 1101

MIDI Input and output circuitry:



MIDI Interconnections

The MIDI IN port of an instrument receives MIDI ncssages to play the instrument's internal synthesizer. The MIDI OUT port sends MIDI messages to play these messages to an external synthesizer. The MIDI THRU port outputs MIDI messages received by the MIDI IN port for daisy-chaining external synthesizers.

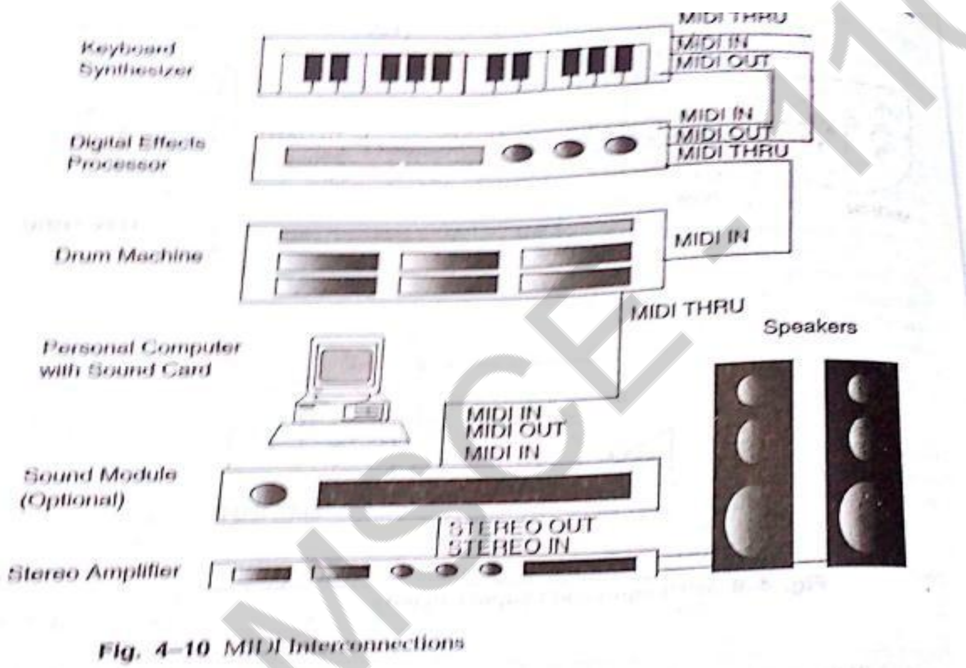


Fig. 4–10 MIDI Interconnections

AMSCE - 110

MIDI Communication Protocol

This protocol uses 2 or more bytes messages.

The number of bytes depends on the types of message. There are two types of messages:

(i) Channel messages and (ii) System messages.

Channel Messages

A channel message can have up to three bytes in a message. The first byte is called a status byte, and other two bytes are called data bytes. The channel number, which addresses one of the 16 channels, is encoded by the lower nibble of the status byte. Each MIDI voice has a channel number; and messages are sent to the channel whose channel number matches the channel number encoded in the lower nibble of the status byte. There are two types of channel messages: voice messages and the mode messages.

Voice messages

Voice messages are used to control the voice of the instrument (or device); that is, switch the notes on or off and sent key pressure messages indicating that the key is depressed, and send control messages to control effects like vibrato, sustain, and tremolo. Pitch wheel messages are used to change the pitch of all notes

Mode messages

Mode messages are used for assigning voice relationships for up to 16 channels; that *is,* to set the device to MOWO mode or POLY mode. Omny Mode on enables the device to receive voice messages on all channels.

System Messages

System messages apply to the complete system rather than specific channels and do not contain any channel numbers. There are three types of system messages: common messages, real-time messages, and exclusive messages. In the following, we will see how these messages are used.

Common Messages These messages are common to the complete system. These messages provide for functions such as select a song, setting the song position pointer with number of beats, and sending a tune request to an analog synthesizer.

System Real Time Messages

These messages are used for setting the system's real-time parameters. These parameters include the timing clock, starting and stopping the sequencer, ressuming the sequencer from a stopped position, and resetting the system.

System Exclusive messages

These messages contain manufacturer-specific data such as identification, serial number, model number, and other information. Here, a standard file format is generated which can be moved across platforms and applications.

4.4.3 Sound Board Architecture

A sound card consist 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 WAVfiles, a decompression and compression chip to compress and decompress audio files, a speech synthesizer to synthesize speech output, a speech

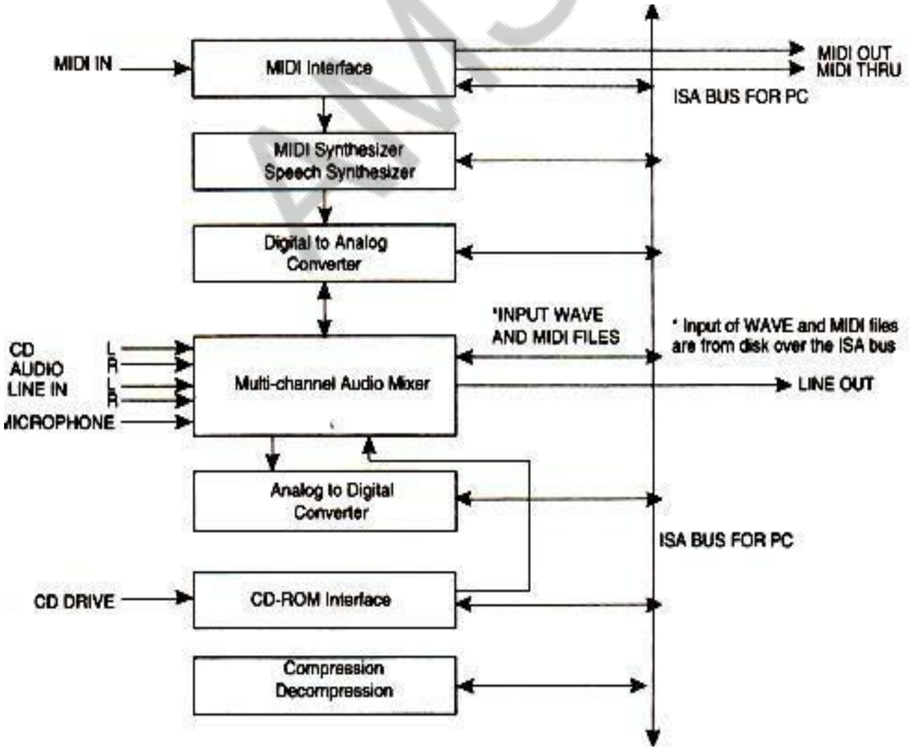
recognition circuitry to recognize speech input and output circuitry to output stereo audio OUT or LINEOUT.

AUDIO MIXER

The audio mixer c:omponent 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-digitaf conversion in conjunction with PCM or ADPCM to generate digitized samples.

SOUND BOARD ARCHITECTURE:



AMSCE - 1101

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 'foml 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 u CD ROM drive.to the sound board.

4.5 VIDEO IMAGES AND ANIMATION

4.5.1video Frame Grabber Architecture

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.

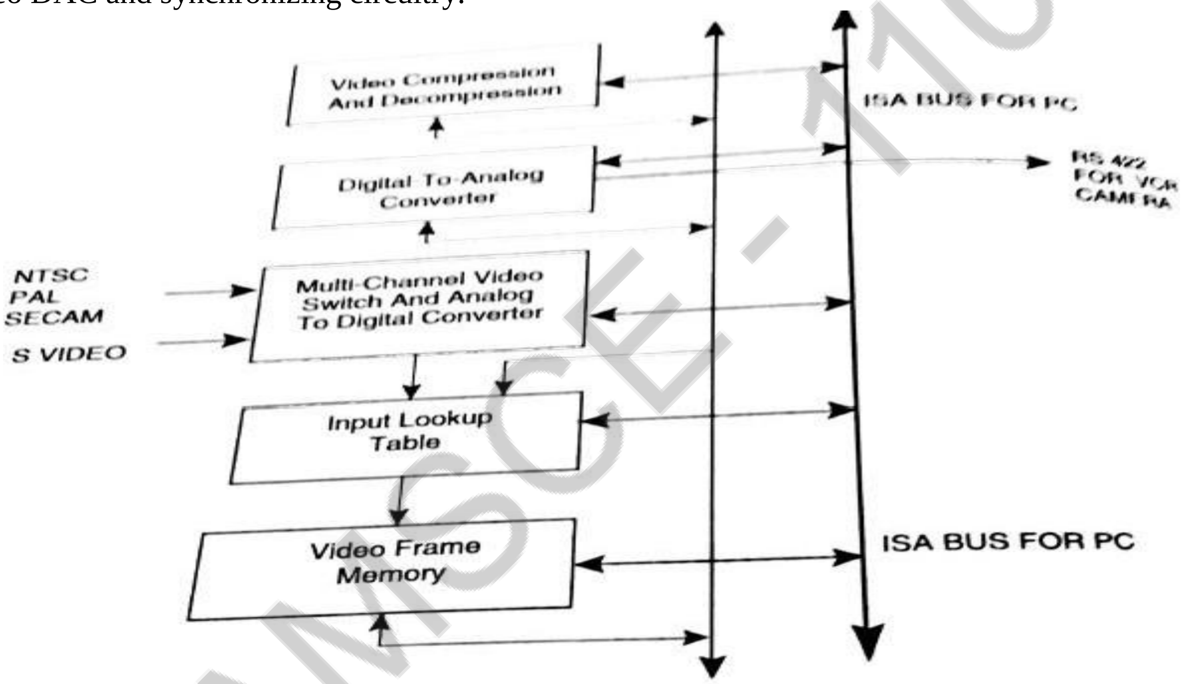


Fig. 4–13 Video Frame Grabber

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 aTV 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 colour.

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 a l024 x 1024 x 24 storage buffer to store image for image processing and display.

Video Compression-Decompression: The video compressiondecompression 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 is 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. The figure shows the histogram of low contrast image.

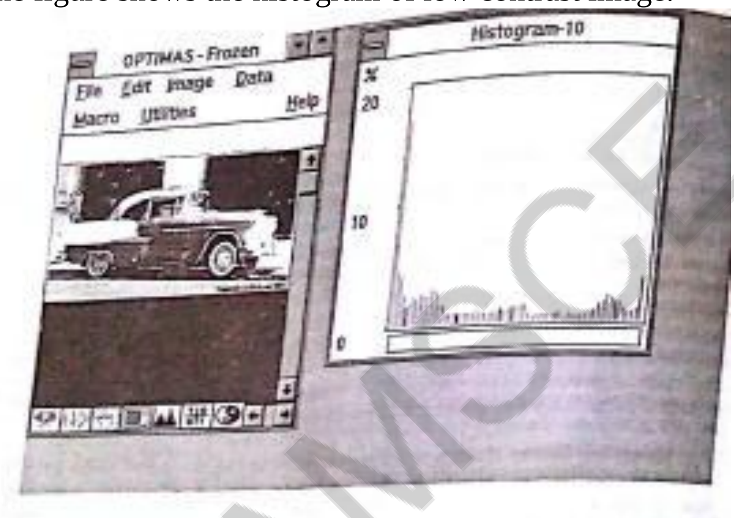


Fig. 4-14 Histogram of Picture with Low Contrast

The next figure shows the histogram of high contrast image

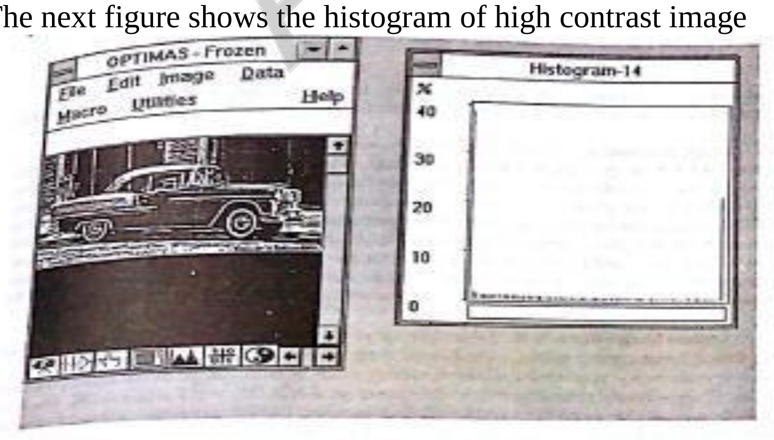


Fig. 4-15 Histogram of Picture with High Contrast

AMSCE - 1101

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. The following figure shows the one pixel calculation using a pixel map.

Table 4–5 Pixel Calculations

	p1	p2	p3	
	p4	p5	p6	
	p7	p8	p9	

Table 4–6 Convolution Mask

a11	a12	a13
a21	a22	a23
a31	a32	a33

Low Pass Filter: A low pass filter causes blurring of the image and appears to cause a reduction in noise.

Table 4–7 Low-Pass Filter Convolution Mask

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

High Pass Filter: 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.

Table 4–8 High-Pass Filter Convolution Mask

-1	-1	-1
-1	-9	-1
-1	-1	-1

Laplacian Filter: This filter sharply attenuates low-spatial-frequency components without affecting and

AMSCE - 1101

high-spatial frequency components, thereby enhancing edges sharply.

Table 4–9 Filter Convolution Mask for a Laplacian Filter

-1	-1	-1
-1	8	-1
-1	-1	-1

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 Q + 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.

4.5.2 Image Animation Techniques

Animation: Animation is an illusion of movement created by sequentially playing still image frames at the rate of 15-20 frames per second.

Toggling between image frames: We can create simple animation by changing images at display time. The simplest way is to toggle between two different images. This approach is good to indicate a "Yes" or "No" type situation.

Rotating through several image frames: The animation contains several frames displayed in a loop. Since the animation consists of individual frames, the playback can be paused and resumed at any time.

4.6 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

Video Capture Board Architecture: 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.

AMSCE - 1101

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.

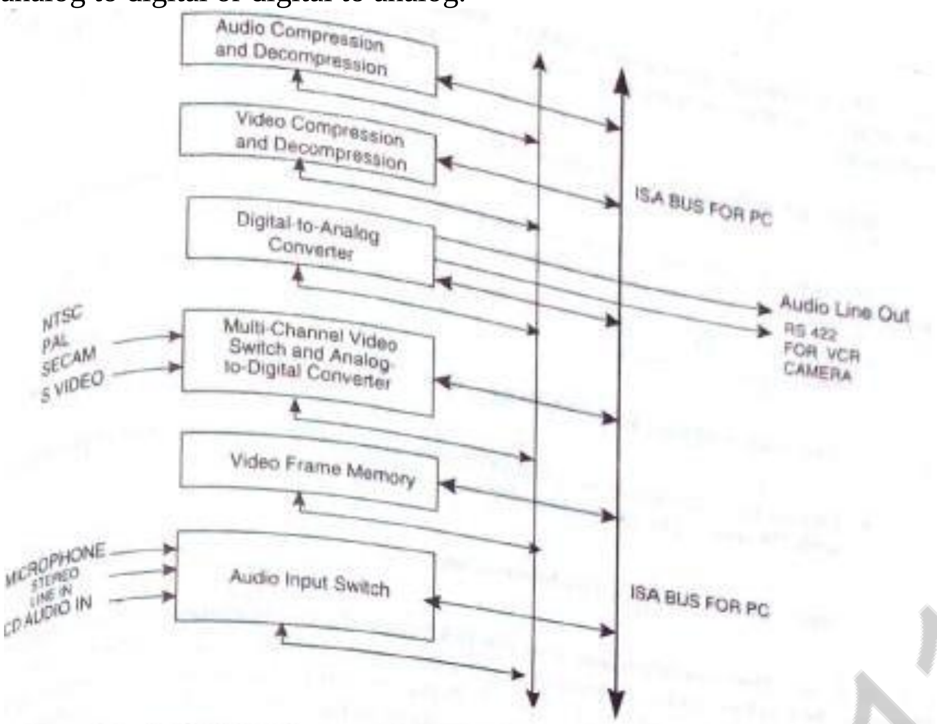


Fig. 4–17 Video Capture Board Architecture

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, ZigZag 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.

Performance issues for full motion video:

During the capture the video hardware and software must be able to keep up with the output of the camera to prevent loss of information. The requirements for playback are equally intense although there is no risk of permanent loss of information. Consider the eg below,

AMSCE - 1101

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.

Bus bandwidth The bandwidth required for display of full-motion video is:

(Resolution × acquisition frames per second × pixel per bits)/8 MBytes/sec

(640 × 480 × 30 × 24) / 8 = 27.648 MBytes/sec

AMSCE - 1101

4.7.1 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.

HARD DISK TECHNOLOGY

Magnetic hard disk storage remains a much faster mass storage to play an important rol~ in multimedia systems.

It remains a much faster mass storage medium than any other mass storage medium.

ST506 and MFM Hard drives: ST506 is an interface that defines the signals and the operation of signals between a hard disk controller and the hard disk. It is developed by seagate. It is used to control platter speed and the movement of heads for a drive. Parallel data is converted to a series of encoded pulses by using a scheme called MFM (modified frequency modulation). The MFM encoding scheme offers greater packing of bits and accuracy than the FM encoding scheme. Other encoding scheme is Run-Length-Limited. Its drive capacity varies from 20 M Bytes to 200 M Bytes.

ESDI Hard Drive: ESDI (Enhanced Small Device Interface) was developed by a consortium of several manufacturers. It converts the data into serial bit streams.

It uses the Run-Length-Limited Scheme for encoding. The drive has data separator circuitry Drive capacity varies from 80 M Bytes to 2 GB. ESDI interface has two ribbon cables: (i) 36 pin cable for control signals. (ii) 20 pin cable for data signals.

IDE: Integrated Device Electronics (IDE) contains a,n integrated controller with drive.

The interface is 16 bit parallel data interface. The IDE interface supports two IDE drives. One is master drive and other is slave drive. Here, Jumper setting is required. The transfer rate is 8 MHz at bus speed.

New Enhanced IDE Interface

This new interface has a transfer rate of 9-13 M Bytes/See with maximum capacity around 8 GB. It supports upto four drives CD ROM and tape drives.

SCSI (Small Computer System Interface)

It is an ANSI X3T9.2 standard which supports SCSI and SCSI2 Standards. The Standard defines both software and hardware.

SCSI-I:It defines an 8-bit parallel data path between host adapter and device.

Here, host adapter is known as initiator and the device is known as target. There are one initiator and seven targets.

Nine control signals define the activity phases of the SCSI bus during a transaction between an initiator and a target. The phases are:

(i) arbitration phase (ii) selection phase (iii) command phase (iv) data phase (v) status phase
(vi) message phase (vii) bus free phase.

Arbitrary Phase: In this phase an initiator starts arbitration and tries to acquire the bus.

Selection Phase: In this phase, an initiator has acquired the bus and selects the target to which it needs to communicate.

Command Phase: The target now enters into this phase. It requests a command from the initiator. Initiator places a command on the bus. It is accepted by the target.

Data Phase: The target now enters in this phase. It requests data transfer with the initiator. The data is placed on the bus by the target and is then accepted by the initiator.

Status Phase: Now, the target enters in status phase. It indicates the end of data transfer to the initiator. **Message Phase**: This is the last phase. It is to interrupt the initiator signaling completion of the read message. The bus free phase is a phase without any activity on the bus so that the bus can settle down before the next transaction. SCSI-l transfers data in 8-bit parallel form, and the transfer rate vades rom I M Bytes/See to 5 M Bytes/Sec. SCSI-I'drive capacity varies from 20 M bytes to 2 GB. SCSI-1 has over 64 commands specified to carry out transactions.

Commands include read, write, seek, enquiry, copy, verify, copy and verify, compare and so on.

SCSI-2

It has the same aspects of SCSI -1, But with faster data transfer . rates, and wider data width.

AMSCE - 1101

It includes few more new commands, and vender-unique command sets for optical drives, tape drives, scanners and so on. To make the bus wider, a system designer uses a second 68-pin connector in addition to the standard 50 pin connector.

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 seccors. It determines the organization of sectors.

Transfer Rate and I/O per Second: I/O transfer nite varies from 1.2 M bytes/Sec. to 40 M bytes/Sec. Transfer rate is defined as the rate at which data is transferred from the drive buffer to the host adapter memory.

Data Management: It includes Command queueing and Scattergather. Command queueing allows execution of multiple sequential commands with system CPU intervention.Scatter is a process of setting the data for best fit in available block of memory or disk. Gather is a process which reassembles data into contiguous blocks on memory or disk ..

Figure below shows the relationship between seek latency, Rotational latency and Data transfer.

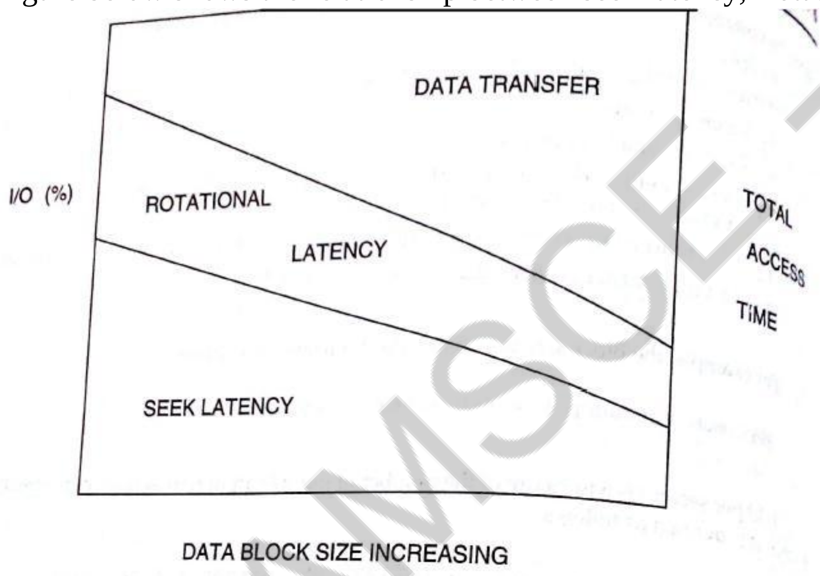


Fig. 5–1 Seek Latency. Rotational Latency, and Data Transfer Relationships.

It is a method of attaching multiple drives to a single host adapter. The data is written to the first drive first, then after filling it, the controller, allow the data to write in second drive, and so on. Meantime Between Failure (MTBF) = MTBF of single/drivel Total no. of drives.

RAID (Redundant Array of Inexpensive Disks)

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.

If we use RAID as our hot backups, it will be economy. A number of RAID schemes havebeen developed:

1. Hot backup of disk systems
2. Large volume storage at lowercost

3.Higher performance at lower cost
4.Ease of data recovery
5.High MTBF.
There are six levels of RAID available.

(i) RAID Level 0 Disk Striping

It spreads data across drives. Data is striped to spread segments of data across multiple drives. Data striping provides high transfer rate. Mainly, it is used for database applications.

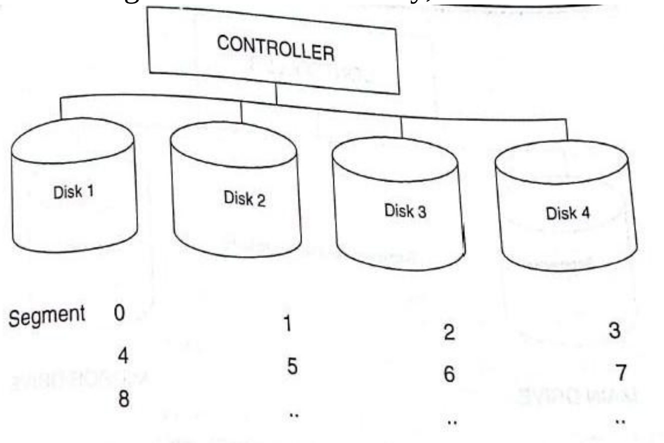


Fig. 5–3 Disk Striping for RAID Level 0

RAID level 0 provides performance improvement. It is achieved by overlapping disk reads and writes. Overlapping here means, while segment I is being written to drive 1, segment 2 writes can be initiated for drive 2.
The actual performance achieved depends on the design of the controller and how it manages disk reads and writes.

2.RAID Level 1 Disk Mirroring

The Disk mirroring causes two copies of every file to be written on two separate drives. (Data redundancy is achieved).

These drives are connected to a single disk controller. It is useful in mainframe and networking systems. Apart from that, if one drive fails, the other drive which has its copy can be used.

Performance: Writing
is slow.
Reading can be speeded up by overlapping seeks.

Read transfer rate and number ofI/O per second is better than a single drive. I/O transfer rate (Bandwidth) = No. of drives x drive I/O transfer rate

No of I/O's Per second = $\frac{I / Otransferr\ ate}{Averagesizeoftransfe\ r}$

controller

Segment 0

Segment 0

Disk controller arrangement for RAID Level1

Uses:

Provide backup in the event of disk failures in file servers.
Another form of disk mirroring is Duplexing uses two separate controllers,this sectioned controller enhances both fault tolerance and performance.

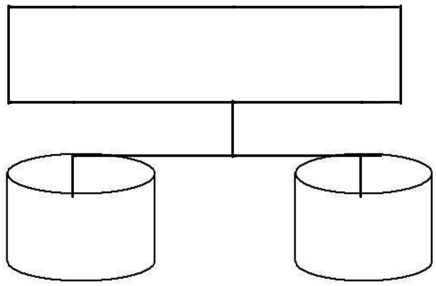
3.RAID Level 2, - Bit interleaving of Data:

It contains arrays of multiple drives connected-to a disk array controller.

Data (written one bit at a time) is bit interleaved across multiple drives. Multiple check disks are used to

AMSCE - 1101

AMSCE - 1101



detect and correct errors

Host Adapter

.

organization of bit interleaving for RAID level2

It provides the ability to handle very large files, and a high level of integrity and reliability. It is good for multimedia system. RAID Level 2 utilizes a hamming error correcting code to correct single-bit errors and doublebit errors.

Drawbacks:

(i) It requires multiple drives for error correction (ii) It is an expensive approach to data redundancy. (iii) It is slow.

Uses: It is used in multimedia system. Because we can store bulk of video and audio data.

4.RAID Level-3 Parallel Disk Array:

RAID 3 subsystem contains an array of multiple data drives and one parity drive, connected to a disk array controller.

The difference between RAID 2 and RAID 3 is that RAID 3 employs only parity checking instead of the full hamming code error detection and correction. It has the advantages of high transfer rate, cost effective than RAID 2, and data integrity.

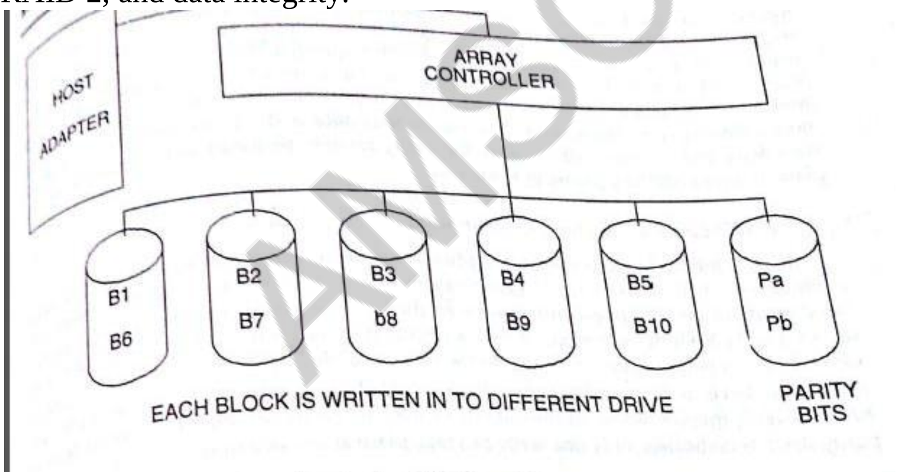


Fig. 5–7 Bit Interleaving for RAID Level 3

Performance and Uses:

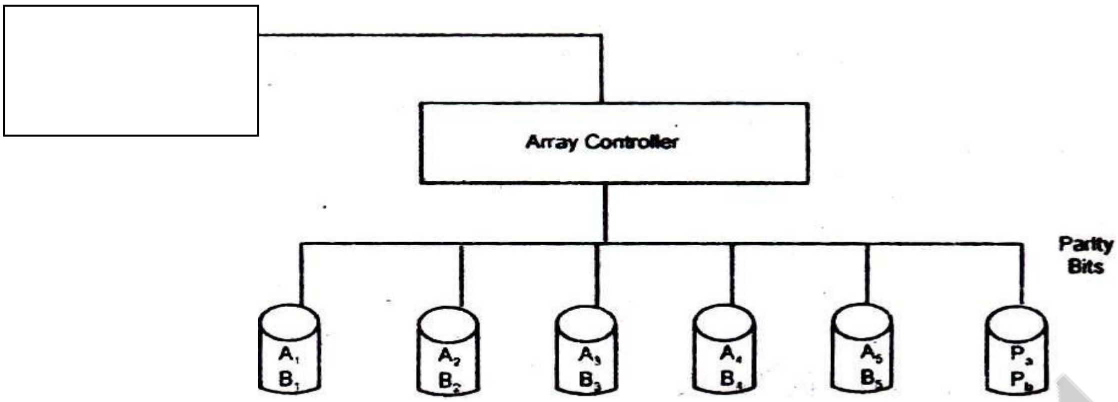
RAID 3 is not suitable for small file transfers because the data is distributed and block-interleaved over multiple drives.

It is cost effective, since it requires one drive for parity checking.

5.RAID Level-4 Sector Interleaving: Sector interleaving means writing successive sectors of data on different drives.

As in RAID 3, RAID 4 employs multiple data drives and typically a single dedicated parity drive. Unlike RAID 3, where bits of data are Written to successive disk drives, an Ri\ID 4, the first sector of a block of

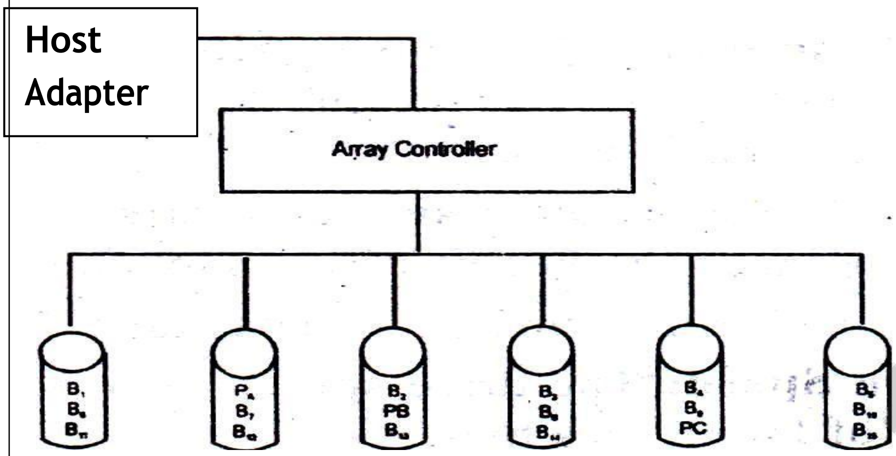
AMSCE - 1101



AMSCE - 1101

data is written to the first drive, the second sector of data is written to the secohd drive, and so on. The data is interleaved at the data level.
RAID *Level1-4* offers *cost-effective* improvement in performance with data.

RAID Level-5 Block Interleaving: In RAID LevelS, as in all the other RAID systems, multiple drives are connected to a disk array controller.
The disk array controller contains multiple SCSI channels.
A RAID 5 system can be designed with a single SCSI host adapter with multiple drives connected to the single SCSI channel.
Unlike RAID Level-4, where the data is sector-interleaved, in RAID Level-5 the data is block-interleaved.



RAID level 5 disk arrays

Table 5–2 Comparison of RAID Levels

RAID Type	Description	Request Rate # of I/Os/sec (Read/Write)	Data Rate per I/O (Read/Write)	Types of Applications
Level 0	Striping, no parity, no redundancy	Excellent for small chunks	Excellent for small chunks	High performance for noncritical data
Level 1	Disk mirroring (degrades write performance)	Good for read, average for write	Good for read, average for write	System drives and critical files. 100% overhead.
Level 2	Striping, as bit inter-leave, hamming code error protection	Average	Low due to hamming code	Not used due to high error-correct-ing overhead
Level 3	Striping as bit or byte interleave with dedicated parity drive, synchronized spindles	Good for large read/write transfers	Excellent for all sizes	Large I/O requests, e.g., CAD, image, video, poor efficiency on small blocks
Level 4	Striping at sector level Reads and writes on independent drives; dedicated parity drive	Good for small block sizes	Good for all sizes	Same as for Level 5, parity drive bottleneck is an issue
Level 5	Striping at block level, and parity (on all drives)	High I/O rate for small block sizes	Excellent for all sizes, degrada-tion during recovery and reconstruction	High request rate, read-intensive data lookups, transparent to system software

AMSCE - 1101

4.7.2 Optical Media

CD ROM, WORM (Write once, Read many) and rewritable optical systems are optical drives.

· CD-ROMs have become the primary media of choice for music due to the quality of ,sound. WORMs and erasable opticel drives both use lasers to pack information densely on a removable disk.

Optical Media can be classified by technology as follows:

- **CD-ROM - Compact Disc Read Only Memory**
- **WORM - Write Once Read Many Rewritable – Erasable**
- **Multifunction - WORM and Erasable.**

1. **CD-ROM**

Physical Construction of CD ROMs:

It consists of a polycarbonate disk. It has 15 mm spindle hole in the center. The polycarbonate substrate contains lands and pits.

The space between two adjacent pits is called a land. Pits, represent binary zero, and the transition from land to pits and from pits to land is represented by binary one.

The polycarbonate substrate is covered by reflective aluminium or aluminium alloy or gold to increase the reflectivity of the recorded surface. The reflective surface is protected by a coat oflacquer to prevent oxidation. A CD-ROM consists of a single track which starts at the center from inside and spirals outwards. The data is encoded on this track in the form of lands and pits. A single track is divided into equal length sectors and blocks**.** Each sector or block consists of2352 bytes, also called a frame. For Audio CD, the data is indexed on addressed by hours, rninutes, seconds and frames. There are 75 frames in a second.

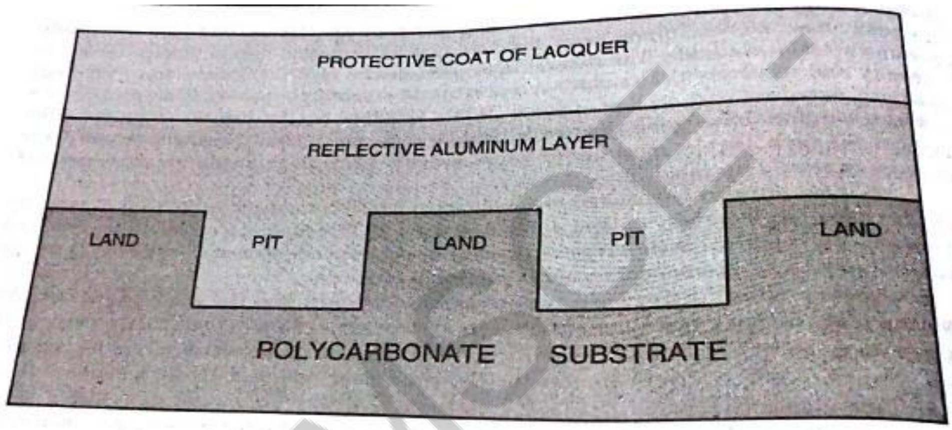


Fig. 5–9 CD-ROM Physical Layers

Physical Layer of Recordable CD-ROM:

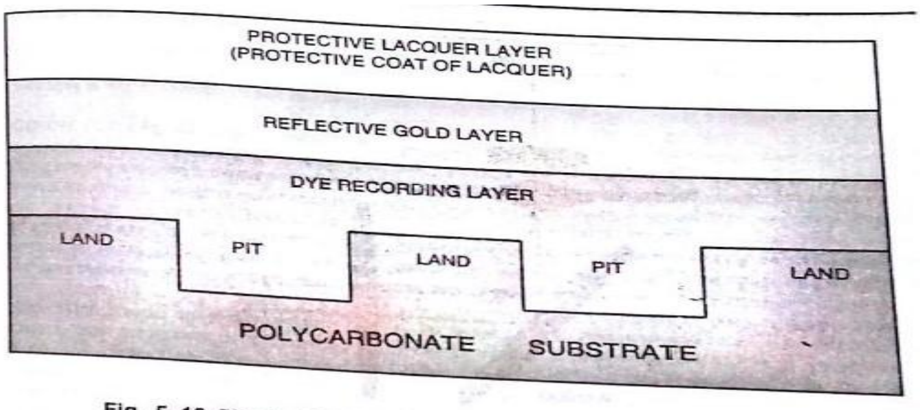
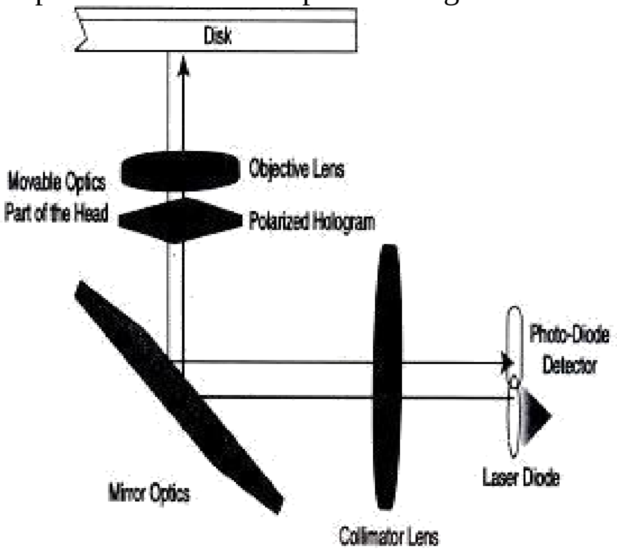


Fig. 5–10 Physical Layers of a Recordable CD-ROM

AMSCE - 1101

Magnetic Disk Organization: Magnetic disks are organized by CYlinder, track and sector. Magnetic hard disks contain concentric circular tracks. They are divided into sector.

Component of rewritable phase change cd-rom



Organization of magnetic media

CD-ROM Standards : A number of recording standards have emerged for CD-ROMs.
They are:

CD-DA (DD-Digital Audio) Red Book: CD-ROM is developed by philips and sony to store audio information. CD-DA is the basic medium for the music industry.

The standard specifies multiple tracks, with one song per track. One track contains one frame worth of data: 2352 bytes. There are 75 frames in a second. Bandwidth = 176 *KB/s.*

CD-ROM Mode 1 Yellow Book: The Mode 1 Yellow Book Stnadard was developed for error correction. The Yellow Book Standard dedicates 288 bytes for error detection codes (EDCs) and error correction codes (ECCs).

CD-ROM Mode 2 Yellow Book

The Mode 2 Yellow Book standard was developed for compressed audio and video applications where, due to lossy compression, data integrity is not quite as important. This standard maintains the frame stmcture but it does not contain the *ECC/EDC* bytes. Removing the *ECC/EDC* bytes allows a frame to contain an additional 288 bytes of data, resulting in an increase of 14% more data. The frame stmcture is shown in the Table below:

Synchronization	Header	Data
12 Bytes	4 Bytes	2336 Bytes
0-11	13-15	16-2351

CD-ROMXA

XA stands for Extended Architecture. The standard was created for extending the present CD-ROM format.

CD-ROM XA contains multiple tracks. Each track's content is desclibed by mo~e. CD-ROM XA also allows interleaving audio and video objects with data for synchroni~ed playback. It does not support video compression. It supports audio compression. It uses Adaptive differential pulse Code Modulation algorithms.

CD-MO Orange Book Part 1

This standard defines an optional pre-mastered area conforming to the Red, Yellow or Green book standards for read-only, and a recordable area. It utilizes a read/write head similar to that found in magnetooptical drives. We can combine the pre-master multimedia objects as the base and develop their own versions.

CD-R Orange Book Part 2

This standard allows writing data once to a writeable disk. Here, the CD contains a polycarbonate substrate with pits and lands.

AMSCE-1101

AMSCE - 1101

The polycarbonate layer is covered with an organic dye recording layer.

As in CD-ROM construction, the track starts from the center and spirals outwards. CD-R uses a high powered laser beam. The laser beam alters the state of the organic dye such that when the data is read, the altered state of dye disperses light instead of reflecting it. The reflected beam is measured for reading the state of each bit on the disk.

2. **Mini-Disk**

Mini-Disk for Data is known as MD-Data. It was developed by Sony Corporation. It is the data version of the new rewritable storage format. It can be used in three formats to support all users.

A premastered optical disk.

A recordable magneto-optical disk.

A hybrid of mastered and recorded.

Its size is 2.5 inch. It provides large capacity. It is low cost. It is used in multimedia applications.A MD demands as a replacement for audio cassette. A 2-1/2 inch MD-Data disk stores 140Mbytes of data and transfer data at 150Kbytes/sec. the figure shows the format for MD-Data standard.

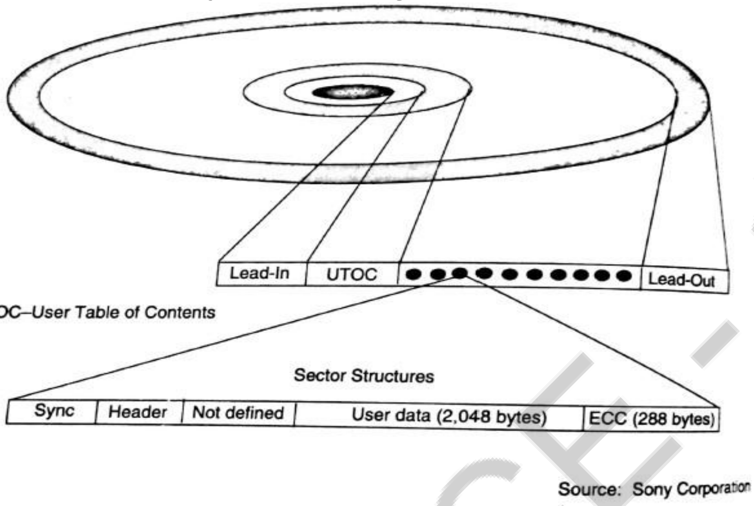


Fig. 5–14 Block Format for MD-Data Standard

3. **WORM Optical Drives**

It records data using a high power laser to create a permanent burnt-in record of data. The laser beam makes permanent impressions on the surface of the disk.

It creates pits. Information is written once. It cannot be written over and cannot be erased. i.e., Here data cannot be edited.

Layers of WORM Drive:

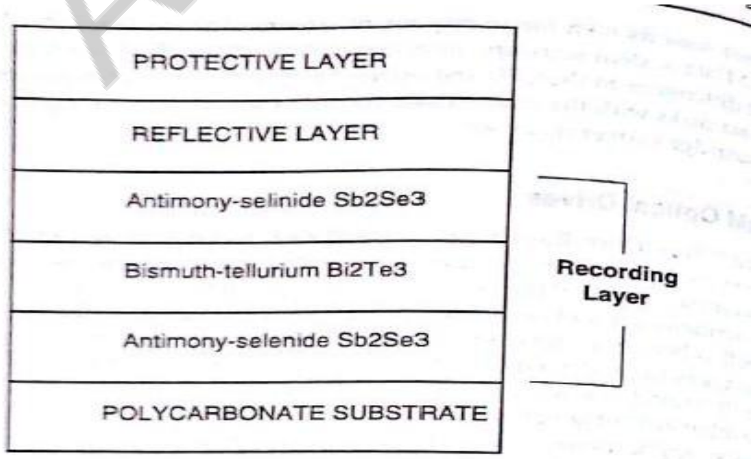


Fig. 5–15 Layers on WORM Drives

AMSCE - 1101

AMSCE - 1101

- The optical disk of WORM consists of six layers, the first layer is a poly carbonate substrate.
- The next three layer are multiple recording layers made from antimony-selenide(Sb_2Se_3)and bismuth-tellurium(Bi_2Te_3).
- The Bismuth-tellurium is sandwiched between antimony-selenide as shown in the figure.
- The recording layer are covered by aluminium alloy or gold to increasethe reflectivity of recorded surface.
- The reflective surface is protected by a coat of lacquer to prevent oxidation.

Recording(writing) of information: During recording, the input Signal is fed to a laser diode. The laser beam from the laser diode is modulated by the input signal. It switches the laser beam on and off. if the beam is on, it strikes the three recording layers.

The beam is absorbed by the bismuth-tellurium layer. Heat is generated within the layer. This heat diffuses the atoms in the three recording layers. It forms four-element alloy layer. Now, the layer becomes recorded layers.

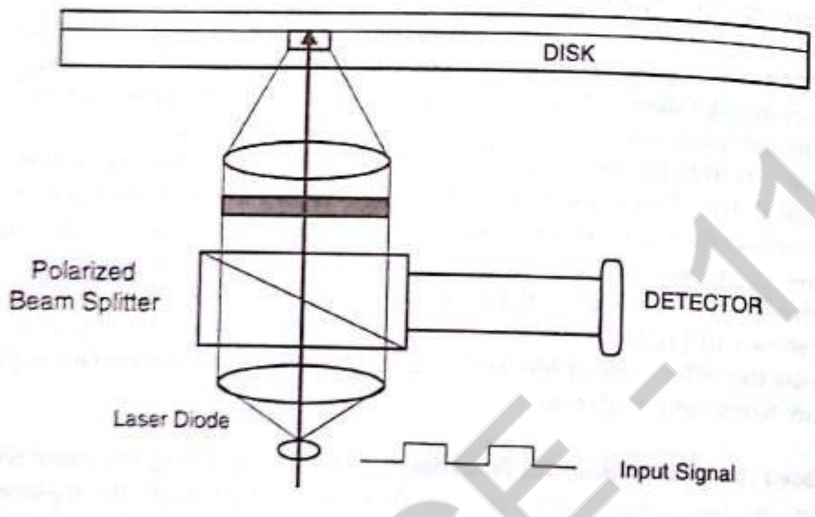


Fig. 5–16 Beam Splitting on WORMs for Disk Writes

Reading Information from disk:

During disk read, a weaker, laser beam is focused on to the disk. It is reflectted back. The beam splitter mirror and lens arrangement sends the reflected beam to the photo detector. The photo sensor detects the beam and converts it into an electrical signal.

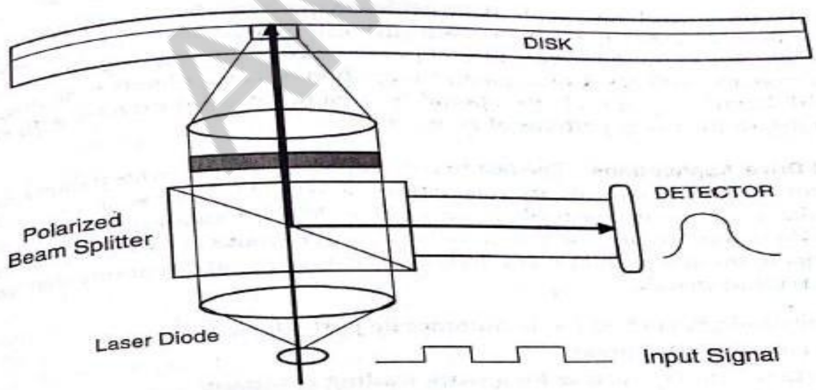


Fig. 5–17 Beam-Splitting on WORMs for Disk Reads

WORM Format Standards:

While WORM drives originated in 14" and 12" form factors, by and large, the 5-1/4" form factor has become standard in industry. The smaller size of the optical disk library is a major factor in this move. There

AMSCE-1101

is no standard for logical format for WORM drives.

WORM performance:

A WORM drive is not known for performance.

Average seek time is between 70-120ms as compared to average seek times 10-25ms for PC-class magnetic drives.

They are typically resident in an optical disk library.

It provides a cost effective solution for large volume of storage.

WORM DRIVE Applications

On-line catalogs – In online catalogs the overall size of the data is very high.

Large-volume distribution-

Transaction logging – every transaction and conversation with the client is logged and stored on optical media.

Multimedia archival- The optical disk libraries have become the storage of choice for archiving images in document image management system

4. Rewritable Optical Disk Technologies

In contrast to WORM technology this technology allows erasing old data and rewriting new data over old data. There are two types of rewritable technology: (i) Magneto-optical ii)Phase change.

Magneto-Optical Technology

It uses a combination of magnetic and laser technology to achieve read/write capability. The disk recording layer uses a weak magnetic field to record data under high temperature. High temperature is achieved by laser beam.

When the beam is on, it heats the spot on the magneto optical disk to its curie temperature. The rise in temperature makes the spot extra sensitive to the magnetic field of bias field.

Magneto-optical drives require two passes to write data; in the first pass, the magneto optical head goes through an erase cycle, and in the second pass, it writes the data.

During the erase cycle, the laser beam is turned on and the bias field is modulated to change the polarity of spots to be erased. During the write cycle, the bias field is turned on and the laser beam is modulated to change the polarity of some spots to 1 according to the bit value.

Magneto optical Construction:

- The optics for magneto optical drive is divided into two sections fixed optics and movable optics.
- There is fixed set of lens and mirror in an optical arrangement which consists of laser diode, a photodetector diode lens and mirrors. This parts is called fixed.
- The movable optics is a part of head and moves during seek , read and write operations.

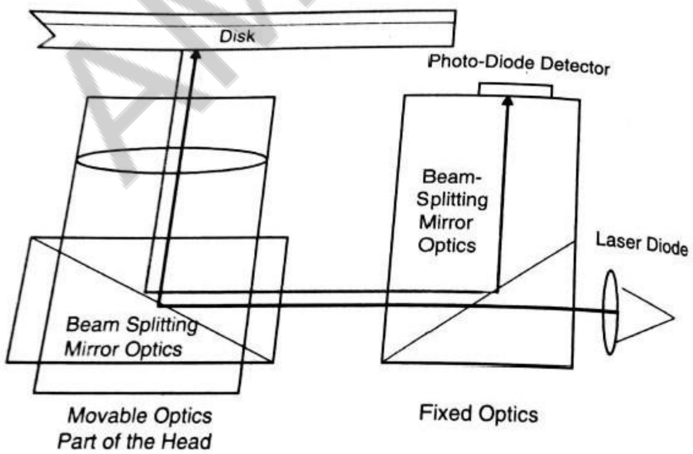


Fig. 5–18 Optics for Magneto-Optical Drives

Reading Magneto-optical Disks:

- During disk reads a low-power laser beam is transmitted to the surface of the disk.

AMSCE - 1101

- The laser beam gets reflected off the surface of the disk.
- The weak magnetic field makes the laser beam polarized and the plane of the beeam is rotated clockwise or counter clockiwise this phenomenon is called **Kerr Effect.**
- The direction of the rotation for the beam dependson the polarity of the magnetic field.

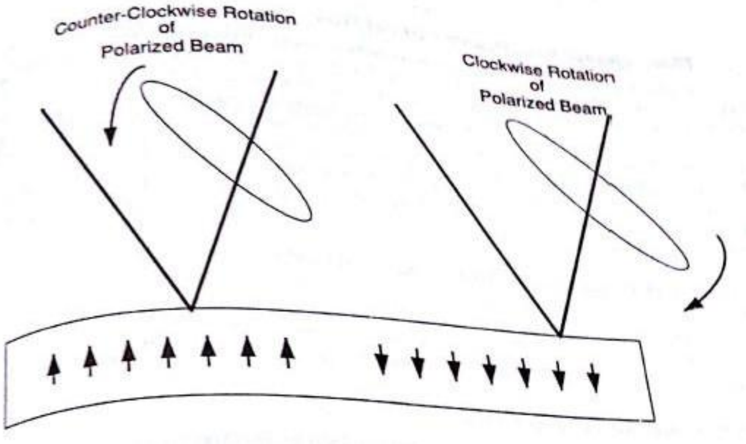


Fig. 5–19 Polarized Beam from Kerr Effect

Uses of Magneto-optical Disk Drives:
Provides very large volume storage.
They exibit performance characteristics same as WORM.
It serves as large online caches for multimedia objects.
Probides an excellent intermediate caching medium.

Standards for Magneto-optical Disk Drives:
ISO and ANSI standard have defined both physical and logical formats for 5-1/4” magneto optical disk.
ISO and ANSI also have settled on physical and logical standard for 3-1/2” magneto optical disk.
Magneto optical drives range in size from 128 to over 500 Mbytes.

Phase change Rewritable optical Disk

In phase change technology the recording layer changes the physical characteristics from crystalline to amorphous and back under the influence of heat from a laser beam.

To read the data, a low power laser beam is transmitted to the disk. The reflected beam is different for a crystalline state than for an amorphous state. The difference in reflectivity determines the polarity of the spot.

Benefits: it requires only one pass to write.

Dye Polymer Rewritable Disk

There is no need of magnetic technology here.

This technology consists of giant molecules formed from smaller molecules of the same kind with light-sensitive dye. This technology is also used in WORM drives.

4.7.3 HIERARCHICAL STORAGE MANAGEMENT

multi-function drive is a single drive unit. It is capable of reading and writing a variety of disk media. Three types of technologies are used for multi-function drives. They are:

(i) Magneto-optical disk for both rewritable and WORM capability.
(ii) Magneto-optical disk for rewritable and dye polymer disk for WORM capability.
(iii) Phase change technology for both rewritable and WORM capability.

The storage hierarchies described in thE pyramid consist of random access memory (RAM), on-line fast magnetic hard disks, optical disks and juke boxes, diskettes, and tapes.

AMSCE - 1101

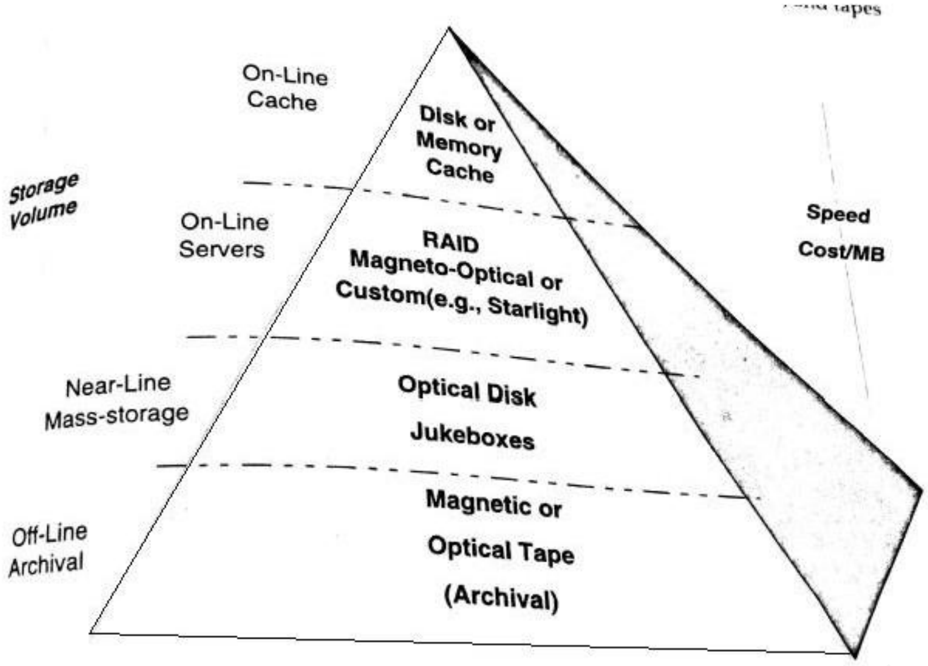


Fig. 5–20 Hierarchical Storage Pyramid

Permanent Vs. Transient Storage issues

The process of moving an object from one level in the storage hierarchy to another level in that hierarchy is called migration. Migration to objects to off-line media and removal of these objects from on-line media is called archiving. Migration Can be set up to be manual or automatic.

Manual migration requires the user or the system administrator to move objects from one level of storage to another level. Systems with automatic migration perform this task automatically. In document-imaging systems, compressed image files are created in magnetic cache areas on fast storage devices when documents are scanned.

Optical Disk Library (Juke box)

An optical juke box stacks disk platters to be played. In the optical disk library, the platters are optical and contain objects such as data, audio, video, and images.

An optical disk library has one or more optical drives. An optical disk library uses a very-high-speed and accurate server-controlled electromechanical robotics elevator mechanism for moving the optical platters between their slots on a disk stacks and the drives. The robotics mechanism removes disk platter from a drive and returns it to its slots on the stack after the disk has finished playing (usually when the drive is required for another disk). The robotics device operates and manages multiple drives under program control.

.

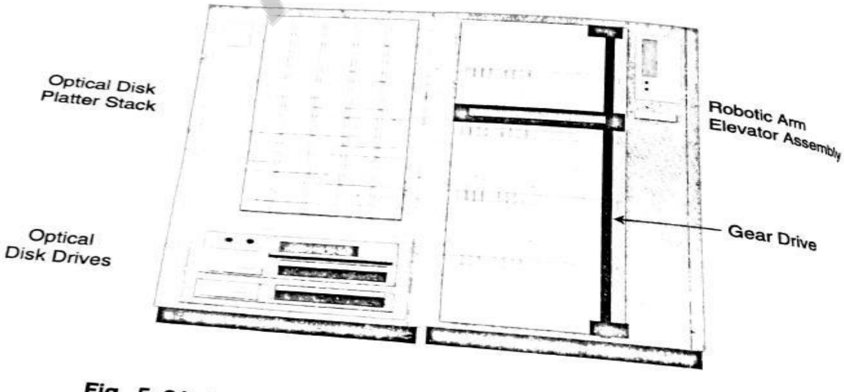


Fig. 5–21 Organization of Optical Disk Libraries

AMSCE - 1101

A juke box may contain drives of different types, including WORM, rewritable, or multifunction. Juke boxes contain one or more drives. A juke box is used for storing large volumes of multimedia information in one cost effective store.

Juke box-based optical disk libraries can be networked so that multiple users can access the information. Optical disk libraries serve as near-line storage for infrequently used data.

Hierarchical Storage Applications: Banks, insurance companies, hosptials, state and federal governments, manufacturing companies and a variety of other business and service organizations need to permanently store large volumes of their records, from simple documents to video information, for audit trail use.any Cache designs use a high-water mark and a low-water mark to trigger cache management operations. When the cache storage fills up to the high-water mark, the cache manager starts creating more space in cache storage. Space is created by discarding objects.

The cache manager maintains a data base of objects in the cache. Cache areas containing updated objects are frequently called dirty cache.

Objects in dirty cache are written back at predetermined time intervals or before discarding an object.

AMSCE - 1101

AMSCE - 1101

UNIT V HYPERMEDIA 9

Multimedia authoring and user interface − Hypermedia messaging − Mobilemessaging − Hypermedia message component − Creating hypermedia message −Integrated multimedia message standards − Integrated document management −Distributed multimedia systems.

5.1 Multimedia authoring and User Interface

5.1.1Multimedia Authoring Systems

Multimedia authoring systems are designed with two primary target users: They are

(i) Professionals who prepare documents, audio or sound tracks, and full motion video clips for wide distribution.

(il) Average business users preparing documents, audio recordings, or full motion video clips for stored messages' or presentations.

The authoring system covers user interface. The authoring system spans issues such as data access, storage structures for individual components embedded in a document, the user's ability to browse through stored objects, and so on.

Most authoring systems are managed by a control application.

Design Issues for Multimedia Authoring

Enterprise wide standards should be set up to ensure that the user requirements are fulfilled with good quality and made the objects transferable from one system to another.

So standards must be set for a number of design issues

1. Display resolution
2. Data formula for capturing data
3. Compression algorithms
4. Network interfaces
5. Storage formats.

Display resolution

A number of design issues must be considered for handling different display outputs. They are:

(a) Level of standardization on display resolutions.

(b) Display protocol standardization.

(c) Corporate norms for service degradations

(d) Corporate norms for network traffic degradations as they relate to resolution issues Setting norms will be easy if the number of different work station types, window managers, and monitor resolutions are limited in number. But if they are more in number, setting norms will be difficult. Another consideration is selecting protocols to use. Because a number of protocols have emerged, including AVI, Indeo, Quick Time and so on.So, there should be some level of convergence that allows these three display protocols to exchange data and allow viewing files in other formats.

File Format and Data Compression Issues

There are variety of data formats available for image, audio, and full motion video objects.

Since the varieties are so large, controlling them becomes difficult. So we should not standardize on a single format. Instead, we should select a set for which reliable conversion application tools are available.

Another key design Issue is to standardize on one or two compression formula for each type of data object. For example for facsimile machines, CCITT Group 3 and 4 should be included in the selected standard. Similarly, for full motion video, the selected standard should include MPEG and its derivatives such as MPEG 2.

While doing storage, it is useful to have some information (attribute information) about the object itself available outside the object to allow a user to decide if they need to access the object data. one of such attribute information are:

(i) Compression type (ii) Size of the object

(iii) Object orientation (iv)Data and time of creation

AMSCE - 1101

(v) Source file name (vi)Version number (if any)
(vii) Required software application to display or playback the object.

Service degradation policies: Setting up Corporate norms for network traffic degradation is difficult as they relate to resolution Issues:
To address these design issues, several policies are possible. They are:

1. Decline further requests with a message to try later.
2. Provide the playback server but at a lower resolution.
3. Provide the playback service at full resolution but, in the case of sound and full motion video, drop intermediate frames.

Design Approach to Authoring

Designing an authoring system spans a number of design issues. They include:

Hypermedia application design specifics, User Interface aspects, Embedding/Linking streams of objects to a main document or presentation, Storage of and access to multimedia objects. Playing back combined streams in a synchronized manner.
A good user interface design is more important to the success of hypermedia applications.

Types of Multimedia Authoring Systems

There are varying degrees of complexity among the authoring systems. For example, dedicated authoring systems that handle only one kind of an object for a single user is simple, where as programmable systems are most complex.

Dedicated Authority Systems

Dedicated authoring systems are designed for a single user and generally for single streams.

Designing this type of authoring system is simple, but if it should be capable of combining even two object streams, it becomes complex. The authoring is performed on objects captured by the local video camera and image scanner or an objects stored in some form of multimedia object library. In the case of dedicated authoring system, users need not to be experts in multimedia or a professional artist. But the dedicated systems should be designed in such a way that. It has to provide user interfaces that are extremely intuitive and follow real-world metaphors.
A structured design approach will be useful in isolating the visual and procedural design components.

TimeLine –based authoring

In a timeline based authoring system, objects are placed along a timeline. The timeline can be drawn on the screen in a window in a graphic manner, or it created using a script in a mann.er similar to a project plan. But, the user must specify a resource object and position it in the timeline.

On playback, the object starts playing at that point in the time Scale.

In most timeline based approaches, once the multimedia object has been captured in a timeline, .it is fixed in location and cannot be manipulated easily, So, a single timeline causes loss of information about the relative time lines for each individual object.

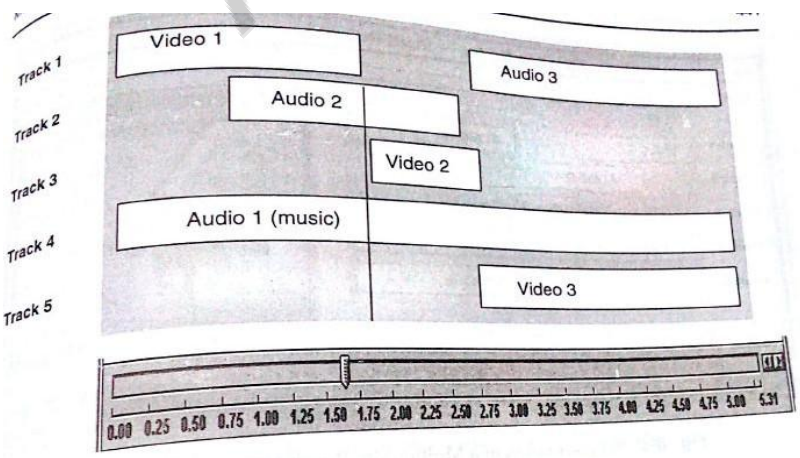


Fig. 8–1 Model of a Timeline-Based Authoring System

AMSCE - 1101

Structured Multimedia Authoring

A structured multimedia authoring approach was presented by Hardman. It is an evolutionary approach based on structured object-level construction of complex presentations. This approach consists of two stages:

(i) The construction of the structure of a presentation.

(ii) Assignment of detailed timing constraints.

A successful structured authoring system must provide the following capabilities for navigating through the structure of presentation.

1.Ability to view the complete structure.

2.Maintain a hierarchy of objects.

3.Capability to zoom down to any specific component. 4.View specific components in part or from start to finish.

5.Provide a running status of percentage full of the designated length of the presentation. 6.Clearly show the timing relations between the various components.

7.Ability to address all multimedia types including text, image, audio, video and frame based digital images.

The author must ensure that there is a good fit within each object hierarchy level. The navigation design of authoring system should allow the author to view the overall structure while examining a specific object segment more closely.

Programmable Authoring Systems

Early structured authoring tools were not able to allow the authors to express automatic function for handling certain routine tasks. But, programmable authoring system bas improved in providing powerful functions based on image processing and analysis and embedding program interpreters to use image-processing functious.

The capability of this authoring system is enhanced by Building user programmability in the authoring tool to perform the analysis and to manipulate the stream based on the analysis results and also manipulate the stream based on the analysis results. The programmability allows the following tasks through the program interpreter rather than manually. Return the time stamp of the next frame. Delete a specified movie segment. Copy or cut a specified movie segment to the clip board . Replace the current segment with clip board contents.

Multisource Multi-user Authoring Systems

We can have an object hierarchy in a geographic plane; that is, some objects may be linked to other objects by position, while others may be independent and fixed in position".

We need object data, and information on composing it. Composing means locating it in reference to other objects in time as Well as space.

Once the object is rendered (display of multimedia object on the screen) the author can manipulate it and change its rendering information must be available at the same time for display.If there are no limits on network bandwidth and server performance, it would be possible to assemble required components on cue at the right time to be rendered.

In addition to the multi-user compositing function A multi user authoring system must provide resource allocation and scheduling of multimedia objects.

Telephone Authoring systems

There is an application where the phone is linking into multimedia electronic mail application

1.Tele phone can be used as a reading device by providing fill text to-speech synthesis capability so that a user on the road can have electronic mail messages read out on the telephone.

2. The phone can be used for voice command input for setting up and managing voice mail messages. Digitized voice clips are captured via the phone and embedded in electronic mail messages.

3. As the capability to recognize continuous speech is deploy phones can be used to create electronic mail messages where the voice is converted to ASCII text on the fly by high-performance voice recognition engines.

AMSCE - 1101

Phones provide a means of using voice where the alternative of text on a screen is not available. A phone can be used to provide interactive access to electronic mail, calendar information databases, public information databass and news reports, electronic news papers and a variety of other applications. !ntegrating of all these applications in a common authoring tool requires great skill in planning.

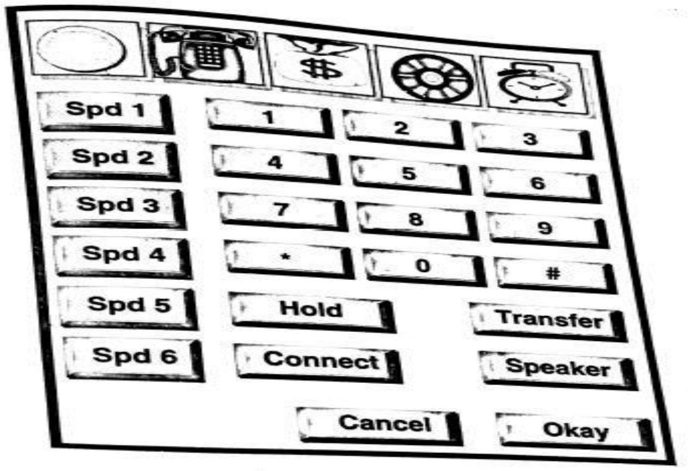


Fig. 8–4 Typical UI for Telephone Control from a Workstation

The telephone authoring systems support different kinds of applications. Some of them are:

1.Workstation controls for phone mail.
2.Voice command controls for phone mail.
3.Embedding of phone mail in electric mail.
4.Integration of phone mail and voice messages with electronic mail.
5.Voice Synthesis in integrated voice mail and electronic mail.
6.Local /remote continuous speech recognition.

5.1.2Hypermedia Application Design Consideration

The user interface must be highly intuitive to allow the user to learn the tools quickly and be able to use them effectively. In addition, the user interface should be designed to cater to the needs of both experienced and inexperienced user.

In addition to control of their desktop environments, user also need control of their system environment. This control should include some of the following:

- The ability to specify a primary server for each object class within a domain specified by the system administrative. A domain can be viewed as a list of servers to which they have unrestricted access.
- The ability to specify whether all multimedia -objects or only references should be replicated. The ability to specify that the multimedia object should be retrieved immediately for display versus waiting for a signal to "play" the object. This is more significant if the object must be retrieved from a remote server.
- Display resolution defaults for each type of graphics or video object.

Essential for good hypermedia design:

1.Determining the type of hypermedia application.
2.Structuring the information.
3.Determining the navigation throughout the application.
4.Methodologies for accessing the information.
5.Designing the user interface.

Integration of Applications

The computer may be called upon to run a diverse set of applications, including some combination of the following:

1.Electronic mail.
2.Word processing or technical publishing.

AMSCE - 1101

3.Graphics and formal presentation preparation software.
4.Spreadsheet or some other decision support software.
5.Access to a relational on object-oriented database.
6.Customized applications directly related to job function:

* Billing
* Portfolio management
* Others.

Integration of these applications consists of two major themes: the appearance of the applications and the ability of the applications to exchange of data.

Common UI and Application Integration

Microsoft Windows has standardized the user interface for a large number of applications by providing standardization at the following levels:

- **Overall visual look and feel of the application windows**
- **Menus**
- **Dialog Boxes**
- **Buttons**
- **Help Features**
- **Scroll Bars**
- **ool Bars**
- **File open and save** etc

This standardization level makes it easier for the user to interact with applications designed for the Microsoft Windows operational environment. Standardization is being provided for Object Linking and Embedding (OLE), Dynamic Data Exchange (DOE), and the Remote Procedure Call (RPC).

Data Exchange

The Microsoft Windows Clipboard allows exchanging data in any format. It can be used to exchange multimedia objects also. We can cut and copy a multimedia objects in one document and pasting in another. These documents can be opened under different applications.The windows clipboard allows the following formats to be stored:

- Text
- Bitrnap
- Image
- Sound
- Video (AVI format).

Distributed Data Access

If all applications required for a compound object can access the subobjects that they manipulate, then only application integration succeeds.

Fully distributed data access implies that any application at any client workstation in the enterprise-wide WAN must be able to access any data object as if it were local. The underlying data management software should provide transport mechanisms to achieve transparence for the application.

Hypermedia Application Design

Hypermedia applicati'ons are applications consisting of compound objects that include the multimedia objects. An authoring applicationn may use existing multimedIa objects or call upon a media editor to CD create new object.

Structuring the Information

A good information structure should consist the following modeling primitives:

.:. Object types and object hierarchies.
.:. Object representations.
.:. Object connections.
.:. Derived connections and representations.

AMSCE - 1101

The goal of information Structuring is to identify the information objects and to develop an information model to define the relationships among these objects.

Object Types and Object Hierarchies

Object types are related with various attributes and representations of the objects. The nature of the information structure determines the functions that can be performed on that information set. The object hierarchy defines a contained-in relationship between objects. The manner in which this hierarchy is approached depends on whether the document is being created or played back.Users need the ability to search for an object knowing very little about the object. Hypermedia application design should allow for such searches.

The user interface with the application depends on the design of the application, particularly the navigation options provided for the user.

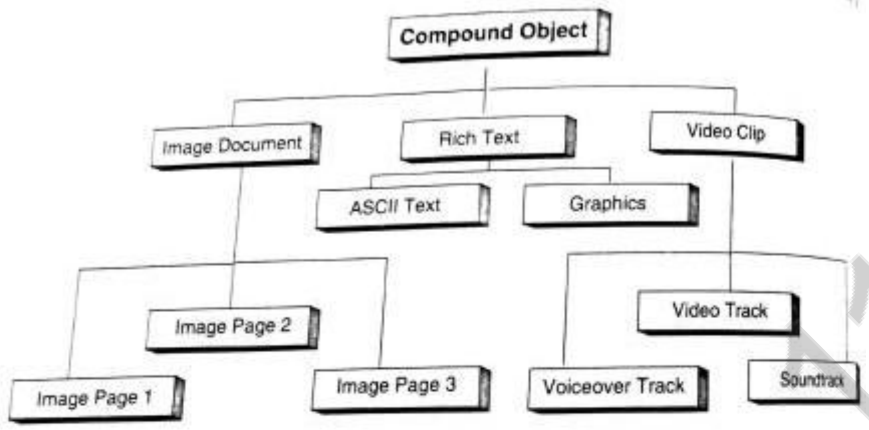


Fig. 8–6 Example of an Object Hierarchy

AMSCE - A101

Object representations

Multimedia objects have a variety of different object representations. A hypermedia object is a compound object, consists of s~ information elements, including data, text, image, and video

Since each of these multimedia objects may have its own sub objects, the design must consider the representation of objects.

An object representation may require controls that allow the user to alter the rendering of the object dynamically. The controls required for each object representation must be specified with the object.

Object connection

In the relational model, the connections are achieved through joins, and in the object oriented models, through pointers hidden inside objects. Some means of describing explicit connections is required for hypermedia design to define the relationships among objects more clearly and to help in establishing the navigation.

Derived Connections and Representations

Modeling of a hypermedia system should attempt to take derived objects into consideration for establishing connection guidelines.

User Interface Design

Multimedia applications contain user interface design. There are four kinds of user interface development tools. They are

1. Media editors
2. An authoring application
3. Hypermedia object creation
4. Multimedia object locator and browser

A media editor is an application responsible of the creation and editing of a specific multimedia object such as an image, voice, or Video object. Any application that allows the user to edit a multimedia object contains a media editor. Whether the object is text, ~voice, or full-motion video, the basic functions provided by the editor are the same: create, delete, cut, copy, paste, move, and merge.

Navigation through the application

Navigation refers to the sequence in which the application progresses and objects are created, searched and used.

Naviation can be of three modes:

(i) Direct: It is completely predefined. In this case, the user needs to know what to expect with successive navigation actions.

Free-form mode: In this mode~ the user determines the next sequence of actions.

Browse mode: In this mode, the user does not know the precise question and wnats to get general information about a particular topic. It is a very common mode in application based on large volumes of non-symbolic data. This mode allows a user to explore the databases to support the hypothesis.

5.1.3.1 Designing user Interfaces

User Interface should be designed by structured following design guidelines as follows: 1.Planning the overall structure of the application

2.Planning the content of the application

3.Planning the interactive behavior

4.Planning the look and feel of the application

A good user interface must be efficient and intuitive by most users.

The interactive behaviour of the application determines how the User interacts with the application. A number of issues are determined at this level.

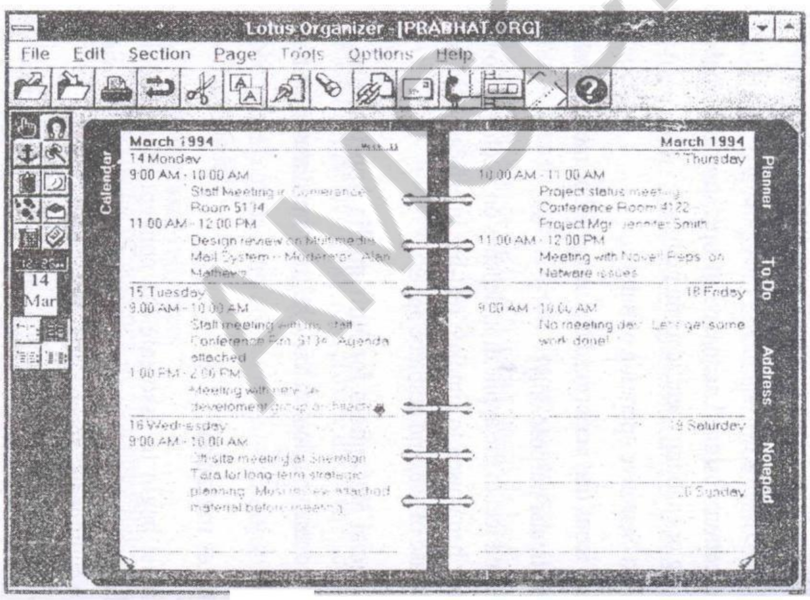
They are

Data entry dialog boxes

Application-designed sequence of operations depicted by graying or enabling specific menu items.

Context-sensitive operation of buttons

Active icons that perform ad hoc tasks(adhoc means created for particular purpose only)



A look and feel of the application depends on a combination of the metaphor being used to simulate real-life interfaces, Windows guidelines, ease of use, and aesthetic appeal.

Special Metaphors for Multimedia Applications

In this section let us look at a few key multimedia user interface metaphors.

The organizer metaphor

One must begin to associate the concept of embedding multimedia object in the appointment diary or notepad to get obvious view of the multimedia aspects of the organizer.

Other use of multimedia object in an organizer is to associate maps or voice mail directions with addresses in address books.

The lotus organizer was the first to use a screen representation of the office diary type organizer

Telephone Metaphor: The role of the telephone was changed b the advent of voice mail system.

AMSCE - 1101

Voice mail servers convert the analog voice and store it in digital form. With the standards for voice ~ail file formats and digital storage of sound for computer. Now, computer system is used to manage the phone system. The two essential components of a phone system are speakers and microphones. They are included in most personal computers.
Figure 8.8 shows how a telephone can be created on a screen to make it a good user interface

The telephone keypad on the screen allows using the interface just as a telephone keypad is used. Push buttons in dialog boxes and function selections in memos duplicate the function provided by the keypad. Push buttons, radio buttons, list boxes, and data entry fields and menu selections allow a range of functionality than can be achieved by the telephone.

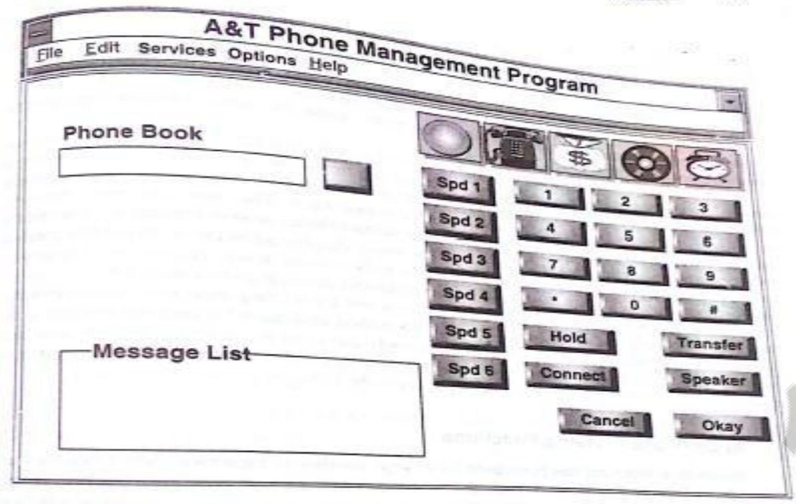


Fig. 8–8 Example of the Telephone Metaphor on a Computer Screen

Aural User Interface: A Aural user interface allows computer systems to accept speech as direct input and provide an oral response to the user actions. Speech enabling is an important feature in this UI. To design AUI system first, we have to create an aural desk top which substitutes voice and ear for the keyboard and display and be able to mix and match them Aural cues should be able to represent icons, voice, menus and the windows of graphical user interface.

AUl design involves human perception, cagnitive science and psycho-acoutic theory. AUI systems learn systems to perform routine functions without user's feedback. An AUI must be temporal and use time based metaphors.
AUI has to address the following issues

1. Recent user memory
2. Attention span
3. Rhythms
4. Quick return to missed oral cues

The VCR metaphor: The User interface metaphor for VCR is to draw a TV on screen and provide live buttons on it for selecting channels, increasing sound volume and changing channel.User interface for functions suchas video capture, channel play, and stored video playback is to emulate the camera, television and VCR on screen Fi5.6 shows all functions of typical video camera when it is in a video capture mode.

Audio/Video Indexing Functions

Index marking allowed users to mark the location on tape in the case of both audio and video to which they may wish to fast forward are rewind.

Other form of index marking is time based. In his form the tape counter shows playtime in hours, minutes, and secondsfrom the time the counter was reset.
Three paradigms for indexing audio and video tapes are

Counter identify tape locations, and the user maintains index listingSpecial events are used as index markersUsers can specify locations for index markings and the system maintains the index.Indexing is useful only if the video is stored. Unless live video is stored, indexing information is lost since the video cannot be repeated.In most systems where video is stored, the sound and video streams are decompressed and managed separately, so synchronization for playback is important. The

AMSCE-1101

indexing information n\must be stored on apermanent basis.

Information Access:

Access structure defines the way objects can be accessed and how navigation takes place through the information objects.
The common forms of navigations for information access are:

Direct: Direct information accessis completely predefined. User must have knowledge about the object that need to be accessed. That information includes object representations in a compound object. **Indexed:** Index access abstracts the real object from the access to the object. If the object ID of the object is an index entry that resolves to a filename on a specific server and disk partition, then the information access mechanism is an indexed mechanism. \

Random Selection: In this fonn, the user can pick one of several possible items. The items need not arranged in any logical sequence; and they need not to be displayed sequentially. The user need not have much knowledge about the infonnation. They must browse through the infornlation.

Path selection or Guided tour: In guided tour, the application guides the user through a predefined path acrosS a number of objects and operations. The user may pause to examine the objects at any stage, but the overall access is controlled by the application. Guided tours can also be used for operations such as controlling the timing for discrete media, such as slide show. It can be used for control a sound track or a video clip.

Browsing: It is useful when the user does not have much knowledge about the object to access it directly.

Object Display Playback Issues:

User expects some common features apart from basic functions for authoring systems. And to provide users with same special control on the display/ playback of these objects, designer have to address some of these issues for image, audio and video objects.

Image Display Issues Scaling: Image scaling is performed on the fly after decompressio The image is scaled to fit in an application defined window at t:' full pixel rate for the window.The image may be scaled by using factors. For eg: for the window 3600 x 4400 pixels can be scaled by a factor of 6 x 10 ie.60 x 440 (60 times).
Zooming: Zooming allows the user to see more detail for a specific area of the image. Users can zoom by defining a zoom factor (eg: 2: 1,5: 1 or 10: 1). These are setup as preselected zoom values.

Rubber banding: This is another form of zooming. In this case, the user uses a mouse to define two comers of the rectangle. The selected area can be copied to the clipboard, cut, moved or zoomed.
Panning: If the image window is unable to display the full image at the ·selected resolution for display. The image can be panned left to right or right to left as wellas top to bottom or bottom to top. Panning is useful for finding detail that is not visible in the full image.

Audio Quality: Audio files are stored in one of a number of formats, including WAVE and A VI. Playing back audio requires that the audio file server be capable of playing back data at the rate of 480 kbytes/min uncompressed or 48 kbytes/min for compressed 8 bit sound or 96 kbytes/min for 16 bit sound.

The calculation is based on an 8 MHz sampling rate and ADCPM compression with an estimated compression ratio. 32 bit audio will need to be supported to get concert hall quality in stored audio. Audio files can be very long. A 20 minute audio clip is over 1 MB long. When played back from the server, it must be transferred completely in one burst or in a controlled manner.

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

AMSCE - 1101

complex.To achieve isochronous playback, most video storage systems
use frame interleaving concepts. Video Frame Interleaving: Frame interleaving defines the structure o;the video file in terms of the layout of sound and video components.

Programmed Degradation: When the client workstation is unable to keep up with the incoming data, programmed degradation occurs. Most video servers are designed to transfer data from storage to the client at constant rates. The video server reads the file from storage, separate the sound and video components, and feeds them as a seperate streams over the network to the client workstations. Unless specified by the user, the video server defaults to favoring sound and degrades video playback by dropping frames. So, sound can be heard on a constant basis. But the video loses its smooth motion and starts looking shaky. Because intermediate frames are not seen.

The user can force the ratio of sound to video degradation by changing the interleaving factor for playback; ie the video server holds back sound until the required video frames are transferred. This problem becomes more complex when multiple streams of video and audio are being played back from multiple source servers. .

Scene change Frame Detection: The scene we see changes every few seconds or minutes and it replaced by a new image. Even within the same scene, there may be a constant motion of some objects in a scene.

Reason for scene change detection: Automating scene change detection is very useful for browsing through very large video clips to find the exact frame sequence of interest. Spontaneous scene change detection provides an automatic indexing mechanism that can be very useful in browsing. A user can scan a complete video clip very rapidly if the key frame for each new scene is displayed in an iconic (poster frame) form in a slide sorter type display. The user can then click on a specific icon to see a particular scene. This saves the user a significant amount of time and effort and reduces resource load by decompressing and displaying only the specific scene of interest rather than the entire video.

Scene change detection is of real advantage if it can be performed without decompressing the video object. Let us take a closer-look at potential techniques that can be employed for this purpose. Techniques:

(i) Histogram Generation: Within a scene, the histogram changes as the subject of the scene mover. For example, if a person is running and the camera pans the scene, a large part of the scene is duplicated with a little shift. But if the scene changes from a field to a room, the histogram changes quite substantially. That is, when a scene cuts over to a new scene, the histogram changes rapidly. Normal histograms require decompressing the video for the successive scenes to allow the optical flow of pixels to be plotted on a histogram. The foot that the video has to be decompressed does help in that the user can jump from one scene to the next. However, to show a slide sorter view requires the entire video to be decompressed. So this solution does not really of the job.

Since MPEG and JPEG encoded video uses DCT coefficients, DCT quantization analysis on uncompressed video or Audio provides the best alternatives for scene change detection without decompressing video

The efficiency can be managed by determining the frame interval for checks and by deciding on the regions within the frame that are being checked. A new cut in a scene or a scene change can be detected by concentrating on a very small portion of the frame

The scene change detection technology as is the case with video compression devices as well as devices that can process compressed video, the implementations of scene change detection can be significantly enhanced.

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 incombination 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. Therfore 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.

Three Dimensional Object Display and VR(Virtual Reality)

Number of 3D effects are used in home entertainment a advanced systems used for specialized applications to achieve find Ine results.

Let us review the approaches in use to determine the impact 0 multimedia display system design due to these advanced systems.

Planar Imaging Technique: The planar imaging technique, used in computer-aided tomography (CAT Scan) systems, displays a twodimensional [2D] cut of X-ray images through multidimensional data specialized display techniques try to project a 3D image constructed from the 2D data. An important design issue is the volume of data being displayed (based on the image resolution and sampling rate) and the rate at which 3D renderings need to be constructed to ensure a proper time sequence for the changes in the data.

Computed tomography has a high range of pixel density and can be used for a variety of applications. Magnetic resonance imaging, on the other hand, is not as fast, nor does it provide as high a pixel density as CT. Ultrasound is the third technique used for 3D imaging in the medical and other fields. .

5.2 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 nuch more complex.

We see how these components are added to messages.

5.2.1 Mobile Messaging

Mobile messaging represents a major new dimension in the users 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.

Handheld and desktop devices play an important growth area for messaging, require complementary back-end services to effectively manage communications for large organizations.

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.

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

AMSCE - 1101

messaging system. The messaging system must prompt the user to enter the name of the addressee forthe message.

The message system looks up the name in an online directory and convert it to an 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 facilitites 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 embed 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.

Typical Electronic mail message

Other capabilities of messaging systems includ~ 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 m 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 richtext 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.

Audio' (Music)

The Musical Instrument Digital interface (MIDI) was developed initially by the music industry to allow computer control of and music recordings from musical instruments such as digital pianos and electric keyboards. MIDI interfaces are now being used for a variety of peripherals, including digital pianos, digital organs, video games with high-fidelity sound output, and business presentations.

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 subset 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. .

AMSCE - 1101

2. Compression and decompression Interfaces for compressing the captured video as it is being captured.
3. A video editor with the ability to decompress, combine, edit, and compress digital video clips.
4. Video indexing and annotating software for marking sections of a videoclip and recording annotations.

Identifying and indexing video clips for storage.

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 facilites 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 perfonnance on playback.

The three main technologies for playing full motion video are microsoft's video for windows: Apple's Quicktime, and Intel's Indeo.

Video for Windows (VFW): It is the most common environment for multimedia messaging.

VFW provides capture, edit, and playback tools for full-motion video. The tools provided by VFW are: The VidCap tool, designed for fast digital video capture.

The VidEdit tool designed for decompression, edition, and compressing full-motion digital video. The VFW playback tool.

The VFW architecture uses OLE. With the development of DDE and OLE, Microsoft introduced in windows the capability to link or multimedia objects in a standardized manner. Hence variety :;windows based applications can interact with them. We can add full-motion video to any windows-based application with the help of VFW. The VFW playback tool is designed to use a number of codecs (software encoder/decoders) for decompressing and playing video files. The default is for A VI files.

Apple's QuickTime

An Apple QuickTime product is also an integrated system for playing back video files. The QuickTime product supports four compression methodologies.

Intel's Indeo

Indeo is a digital video recording format. It is a software technology that reduces the size of un compressed video files through successive compression methodologies, including YUV sub sampling, vector quantization, Huffman's run-length encoding, and variable content encoding. Indeo technology is designed to be scalable for playing back video; It determines the hardware available and optimizes playback for the hardware by controlling the frame rate. The compressed file must be decompressed for playback. The Indeo technology decompresses the video file dynamically in real time for playback. Number of operating systems provide Indeo technology as standard feature and with other software products (eg. VFW).

Hypermedia Linking and Embedding

Linking and embedding are two methods for associating multimedia objects with documents.

This topic deals with

- Linking as in hypertext applications. Hypertext system associate keywordsin a document with other document.
- Linking multimedia objects stored separately from the document and the link provides a pointer to its storage.

AMSCE - 1101

❖ Linking and embedding in a context specific to Microsoft Object linking and Embedding.

Linking in hypertext documents

Hypertext documents are indexed to locate keywords within the text component of the hypermedia document. An extension of this capability is to locate information within the linked component.

There are two types of links passive links and active links

Active Links- performs functions on theoir own based on readers customization. Active links are more intelligent and may use artificial intelligence technologies to monitor the nature of tasks performed by the user.

Passive Links- allow associating one document with another in a number of ways including the author to name the subject of a link and access it based on the content.

Linking and Embedding:

Linking and embedding are two ways of associating multimedia objectswith a hypermedia document or a database record. Let us discuss it in detail.

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 takeup storage space within the hypermedia document. If the creator, or authorised 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 object :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, bitmaps. voice and 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.
4. Sound dips.
5. 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.

AMSCE - 1101

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 hyperrnedia message. The final step in this process is mailing the hypermedia message.

5.5 Integrated Multimedia Message Standards

Let us review some of the Integrated Multimedia MessageStandards in detail.

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 caIled 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-defmed 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 send to the sender of the message if the sender requested the delivery 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.

The VIM interface defines two calls for message sending. The SMISendDocuments() call may be used by applications such as spread sheets and word processing to send spreadsheets or word processing documents to one or more users.

AMSCE - 1101

The VIM interface also provides more extensive support for programmers to develop mail-aware and messaging applications that use application specific messages containing text graphic, images, audio , video and file attachenments.

MAPI Support (Multimedia Application Programmable Interface)

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 applic·ations. Support message-reliant workgroup applications.

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.

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.

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 as I, an 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 combine 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.

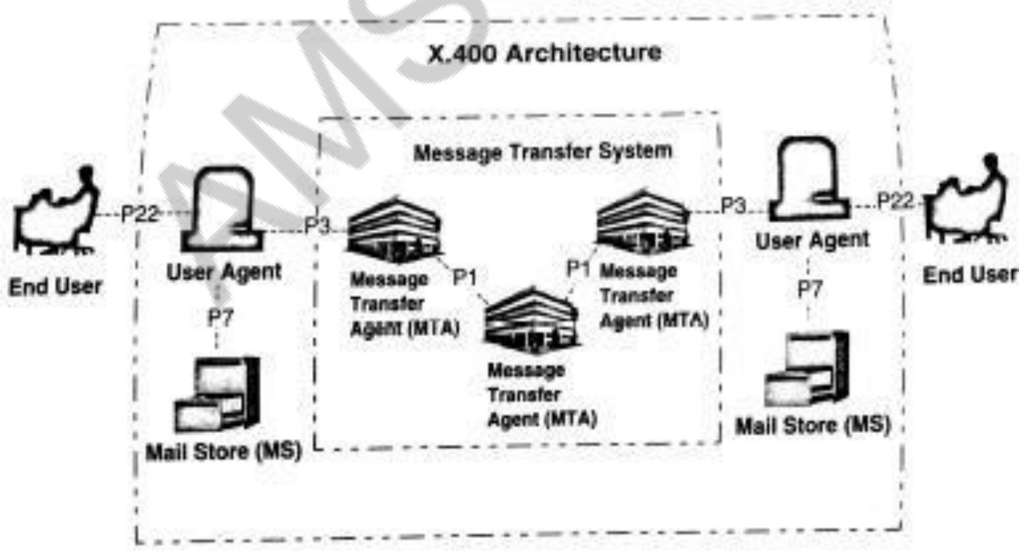


Fig. 9–2 X.400 Architecture

X·500 Directory 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.

5 X·500 Directory System Architecture

Directory System Agents carryout updates and management operations. X ·500 defines a structured information model, an objectoriented 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 unauthorised 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. **Diretory 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.

5.6 Integrated Document Management

It is for managing integrated documents.

Integrated document Management for Messaging Specialized messaging system such as Lotus Notes provide Integrated document management for messaging. The user can attach embed or link a variety of multimedia objects.

When document is forwarded to other users, all associated multimedia objects are also forwarded and avaiiable to the new receivers of the forward message.

Integrated Document management for Messaging:
Specialized messaging systems such as Lotus Notes provide integrated document management for messaging. This means the user can attach , embed, or link a variety of multimedia objects such as graphics, images, audio and video.This also implies that when the document is forwarded to other users all associated multimedia objects are also forwarded and available to the new recipients.

Multimedia Object Server and Mail Server Interactions:

The mail server is used to store i'lll e-mail messages. It consists of a file server with mail files for each user recipient. This file server act as a mail box.

All received mail is dropped in the user's mail file. The user can review or delete these mails. When mail messages include references to multimedia objects, mail file contains only link information.

AMSCE - 1101

DISTRIBUTED MULTIMEDIA SYSTEMS

5.7.1.Components of a Distributed multimedia systems:

If the multimedia systems are supported by multiuser system, then we call those multimedia systems as distributed multimedia systems.

A multi user system designed to support multimedia applications for a large number of users consists of a number of system components. A typical multimedia application environment consists of the following components:

1. Application software.
2. Container object store.
3. Image and still video store.
4. Audio and video component store.
5. Object directory service agent.
6. component service agent.
7. User interface and service agent.
8. Networks (LAN and WAN).

Application Software

The application software perfom1s a number of tasks related to a specific business process. A business process consists ofa series of actions that may be performed by one or more users.

The basic tasks combined to form an application include the following:

(1) Object Selection - The user selects a database record or a hypermedia document from a file system, database management system, or document server.

(2) Object Retrieval- The application ret:ieves the base object.

(3) Object Component Display - Some document components are displayed automatically when the user moves the pointer to the field or button associated with the multimedia object.

(4) User Initiated Display - Some document components require user action before playback/display.

(5) Object Display Management and Editing: Component selection may invoke a component control subapplication which allows a user to control playback or edit the component object.

Document store

A document store is necessary for application that requires storage oflarge volume of documents. The following describes some characteristics of document stores.

1. Primary Document Storage: A file systems or database that contains primary document objects (container objects). Other attached or embedded documents and multimedia objects may be stored in the document server along with the container object.

2. Linked Object Storage: Embedded components, such as text and formatting information, and linked information, and linked components, such as pointers to image, audio, and video. Components contained in a document, may be stored on separate servers.

3. Linked Object Management: Link information contains the name of the component, service class or type, general attributes such as size, duration of play for isochronous objects and hardware, and software requirements for rendering.

Image and still video store

An image and still video is a database system optimized for storage of images. Most systems employ optical disk libraries. Optical disk libraries consist of multiple optical disk platters that are played back by automatically loading the appropriate platter in the drive under device driver control.

The characteristics of image and still video stores are as follows:

(i) Compressed information (ii) Multi-image documents

(iii) Related annotations (iv) Large volumes

(v) Migration between high-volume such as an optical disk library and high-speed media such as magnetic cache storages(vi) Shared access: The server software managing the server has to be able to manage the different requirements.

AMSCE - 1701

Audio and video Full motion video store

Audio and Video objects are isochronous. The following lists some characteristIcs of audio and full-motion video object stores:

(i) Large-capacity file system: A compressed video object can be as large as six to ten M bytes for one minute of video playback.

(ii) Temporary or permanent Storage: Video objects may be stored temporarily on client workstations, servers Providing disk caches, and multiple audio or video object servers.

(iii) Migration to high volume/lower-cost media: migration and management of online storage are much of greater importance and more complex than of images.

(iv) Playback isochronocity: Playing back a video object requires consistent speed without breaks. Multiple shared access objects being played back in a stream mode must be accessible by other users.

Object Directory Service Agent

The directory service agent is a distributed service that providea directory of all multimedia objects on the server tracked by that element of the directoryy service agent.

The following describes various services provided by a directory service Agent.

(1) Directory Service: It lists all multimedia objects by class and server location.

(2) Object Assignment: The directory service agent assigns unique identification to each multimedia object.

(3) Object Status Management: The directory service must track the current usage status of each object.

(5) Directory Service Domains: The directory service should be modular to allow setting up Directory Service Server Elements: Each multimedia object server must have directory service element that reside on either server or some other resources.

(6) Network Access: The directory service agent must be accessible from any workstation on the network.

- **Component Service Agent**
- **A service is provided to the multimedia used workstation by each multimedia component.**
- **This service consists of retrieving objects, managing playback of objects, storing objects, and so on.**
- **The characteristics of services provided by each multimedia component are**

1. **object creating service: It obtains a identification for creating a new object from the directory service agents and provides user interfaceservice agent access for storing the new object.**
2. **playback service : It provides services like play, seek,search ,copy, delete and so on.**
3. **component object service agent :This is the code that provides these services for specific object type such as vide component.**
4. **service agents on servers : multiple component agents may co resident on a server if the server stores multiple component object.**

   and
5. **multifaceted services - (multifaceted services component objects may exist in several forms, such as compressed Or uncompressed).**

AMSCE - 1101

- **User Interface Service Agent**
- It resides on each user workstation.
- It provides direct services to the application software for the management of the multimedia object display windows, creation and storage of multimedia objects, and scaling and frame shedding for rendering of multimedia objects.
- The services provided by user interface service agents are
- **windows management**: creates a new window for multimedia object when invoked and registers it. handles messaes for that window.
- **object creation and capture:** requests component service agent to set up a new object and captures and stores new object.
- **object display and playback**: sets up object for decompression, scales and adjusts frame speed for display or playback of object.

Distributed client server operation

The agents so far we have discussed combine to form a distributed client-server system for multimedia applications. Multimedia applications require functionality beyond the traditional client server architecture.

Most client-server systems were designed to connect a client across a network to a server that provided database functions. In this case, the client-server link was firmly established over the network. There was only one copy of the object on the specified server. With the development of distributed work group computing, the picture has changed for the clients and servers. Actually in this case, there is a provision of custom views in large databases. The advantage of several custom views is the decoupling between the physical data and user.

The physical organization of the data can be changed without affecting the conceptual schema by changing the distributed data dictionary and the distributed data repository.

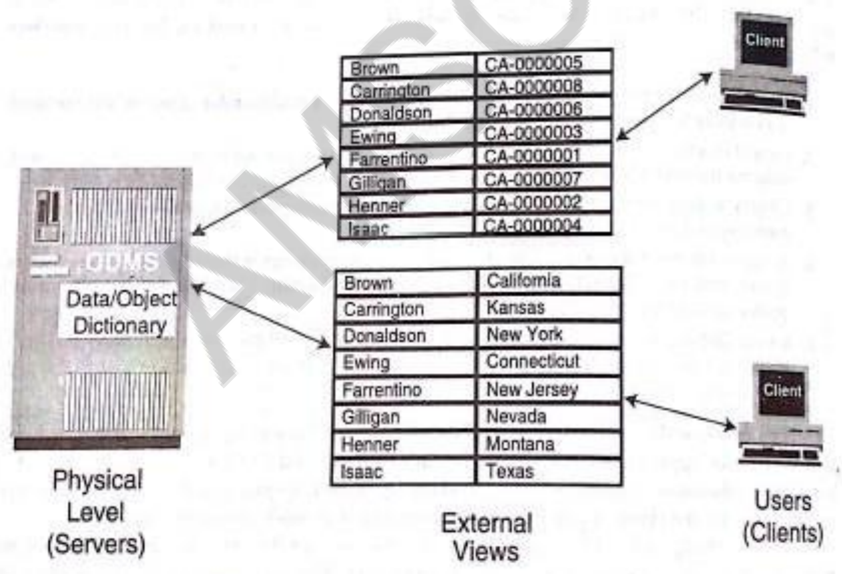


Fig. 10–1 Client-Server Custom Views in Large Databases

Clients in Distributed Work Group Computing

Clients in distributed workgroup computing are the end users with workstations running multimedia applications. The client systems interact with the data servers in any of the following *ways*

1. Request specific textual data.

2. Request specific multimedia objects embedded or linked in retrieved container objects.
3. Require activation of a rendering server application to display/ playback multimedia objects.
4. Create and store multimedia-objects on servers.

Request directory information. on locations of objects on servers

Servers in Distributed Workgroup Computing

Servers are storing data objects. Includes the following functions,

- storage for a variety of object classes,
- transfer objects on demand on clients.
- hierarchical storage for moving unused objects to optical disk lirbaries or optical tape libraries.
- system Administration functions for backing up stored data.
- direct high-speed LAN and WAN server-to-server Tansport for copying multimedia objects.

Database Operation:

Most database systems are used to perform a basic set of operations these includes the following,

Search- find the object in record in response to a query.

Browse- not only retrieve attribute information about object but also render frames of the object contents.

Retrieve- Retrieval functions are different for images , audio, video from symbolic text only database because all three multimedia objects require the retrieved data to be processed by specialized decompression engines before being rendered.

Create and Store- concerned with finding the tables in which the data has to be stored and updating the distributed storage indexing information.

Update- A modification to the existing or previously stored data.

Middleware in Distributed Workgroup Computing

The middleware is like interface between back-end database and font-end clients.The role of middleware is to link back end database to front end clients in a highly flexible and loosely connected network model. Middleware provides the glue for dynamically redirecting client requests to appropriate servers that are on-line.

Middleware performs a number of functions in this environment:

1. Provide the user with a local index, an object directory, for objects with which a client is concerned
2. Provide automatic object directory services for locating available copies of objects
3. Provide protocol and data format conversions between the client requests and the stored formats in the server
4. Provide unique identification throughout the enterprise-wide network for every object through time

Figure 10-2 shows the organization of middleware in a distributed client-server operation. Note that the database architecture changes significantly when middleware is introduced in the system. The middleware is capable of accessing multiple databases and combining information for presentation to the user. For example, middleware can perform some or all combinations of the following functions:

- Access a document database to locate a pointer to the required multimedia object
- Locate an object using a distributed object directory database
- Access an object database to retrieve an object
- Retrieve object preprocessing information from an object description database
- Combine all of this information and preprocess the object before passing it on to a client

We call these actions of middleware *content-based processing*. The range and nature of such content-based processing can be changed without affecting either the servers or the clients. Content-based processing allows the middleware to address temporal characteristics of certain multimedia objects such as audio and video. Content-based processing also allows a variety of editing and updating functions on stored multimedia objects.

For now, we just introduce the topic of middleware. Throughout this chapter we will discuss various aspects of middleware.

AMSCE - 1101

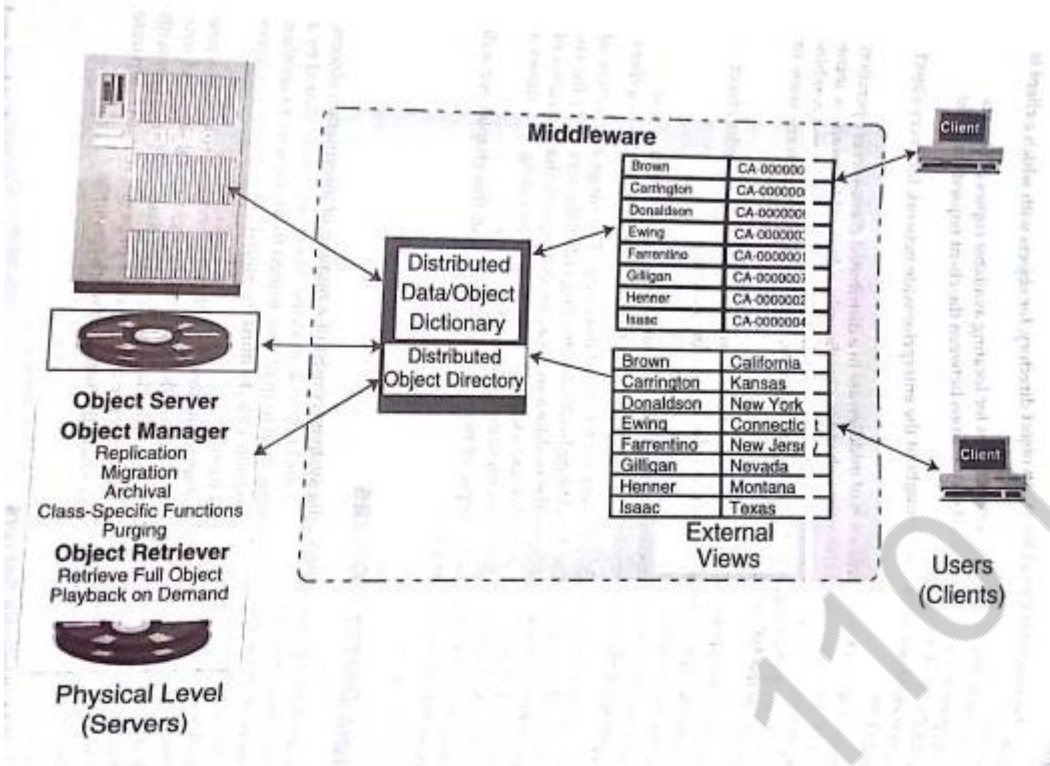


Fig. 10–2 Role of Middleware in Distributed Multimedia Operation

Multimedia Object Servers:

The multimedia system consists of a number of information objects like text, binary files,image,voice and full motion video. Many objects a re shared by a number of users. The information object must be stored on network resourses accessible to all users who need to access them. The resources where information object are stored so that they remain sharable across the network are called servers.

Types of Multimedia Servers

Each object type of multimedia systems would have its own dedicated server optimized for the type of data maintained in the object. A network would consist of some combination of the following types of servers.

(1) Data-processing servers RDBMSs and ODBMSs.
(2) Document database servers.
(3) Document imaging and still-video servers.
(4) Audio and voice mail servers.
(5) Full motion video server.

Data base processing servers are traditional database servers that contain alphanumeric data. In a relational database, data fields are stored in columns in a table. In an object-oriented database these fields become attributes ofthe object. The database serves the purpose of organizing the data and providing rapid indexed access to it. The DBMS can interpret the contents of any column or attribute for performing a search.

Mass Storage for Multimedia Servers

RAID(Redundant Arrays of Inexpensive Disks)

In terms of redundancy,RAID provides a more cost effective solution than disk mirroring.

RAID is a means of increasing disk redundancy, RAID systems use multiple and potentially slower disks to achieve the same task as a single expensive large capacity and high transfer rate disk.

In RAID high transfer rates are achieved by performing operations in parallel on multiple disks. There are different levels of RAID available, namely disk striping(level 0), disk mirroring(level 1, Bit interleaving of date(level 2), Byte interleaving (level 3), sector interleaving(level 4), and block interleaving(level 5)RAID technology is faster than rewritable optical disk and high data volumes can be achieved with RAID. RAID technology provides high performance for disk reads for almost all types of applications.

Write Once Read Many Optical Drives: (WORM)

WORM Optical drives provide very high volumes of storage for very low cost. Some important characteristics of WORM optical disks are:

- Optical drives tend to be slower than magnetic drives by a factor of three to four. .
- WORM drives can write once only; typically 5-10% of disk capacity m left free to provide for changes to existing information.
- They are useful for recording informations that would not change very much. They are virtually indestructible in normal office use and have long shelf lives.
- They an be used in optical disk libraries (Juke boxes). A Juke box may provide anywhere from 50-100 disk platters with two or more drives.
- These characteristics make optical disks ideal candidates for on-line document images (which change very little once scanned and do not have an isochronous requirement) and archived data.

Rewritable Optical Disks:

Rewritable optical drives are produced by using the technologies like magneto-optical. It has the advantage ofrewritability over the WORM where rewritable is not possible. It can be used as primary or secondary media for storage of large objects, which are then archieved. (Placed where documents are preserved) on WORM disks.

If it is used as primary media, it should be accompanied by highspeed magnetic disk cache. This is to achieve acceptable video performance.

Optical Disk Libraries:

Optical disk libraries are nothing but juke boxes. Work disks and rewritables can be used in optical disk libraries to achieve very high volumes of near-lines storage. Optical disk libraries range from desk top juke boxes with one 5' 1/4" drive and I O-slot optical disk stack for upto lOG Bytes of stroage of large libraries using as many as four 12" drives with an 80-s10t optical disk stack for upto terabytes of storage. The disadvantage of optical disk library is the time taken for a platter to be loaded into a drive and span to operating speed.

Network Topologies for Multimedia Object Servers

A number of network topologies are available Network topology is the geometric arrangement of nodes and cable links in a network. We still study three different approaches to setting up multimedia servers.

(i) **Centralized Multimedia Server**: A centralized multimedia object server performs as a central store for multimedia objects. All user requests for multimedia objects are forwarded by the applications to the centralized server and are played back from this server. The centralized server may serve a particular site of the corporation or the entire enterprise. Every multimedia object has a unique identity across the enterprise and can be accessed from any workstation. The multimedia object identifier is referenced in every data that embeds or links to it.

(ii) **Dedicated Multimedia Servers**: This is the approach where a video server is on a separate dedicated segment In this approach, when a workstation dumps a large video, the other servers on the networks are not affected. Provides high performance for all local operations. The isochronocity of the objects is handled quite well in a dedicated mode.

Disadvantage of this approach is that the level of duplication of objects.

(iii) Distributed multimedia servers:

In this approach multimedia object servers are distributed in such a manner that they are placed in starategic locations on different LANs.They are replicated on a programmed basis to provide balanced serviceto all users.

AMSCE - 1101

Multiserver Network Topologies

To distribute the full functionality of multimedia network wide there are vareity of network topologies available. ' The primary topologies are Traditional LANs (Ethernet or Token Ring Extended LANs (Using network switching hubs bridges and routers). ' High speed LANs (ATM and FDDI II). WANs (Including LANs, dial-up links-including ISDN T1 and T_3 lines-etc.). ' I

Traditional LANS (Ethernet or Token Ring) Ethernet:

Ethernet: It is a Local Area Network hardware, communication, and cabling standard originally developed by Xerox corporation that link up to 1024 nodes in a bus network. It is ahigh speed standard using a baseband (single-channel) communication technique. It provides for a raw data transfer rate of 10 Mbps, with actual throughput in the range of 2-3 Mbps. It support a number of sessions in a mix of live video, audio electronic mail and so on.

The following figure shows the LAN topology for small work group.

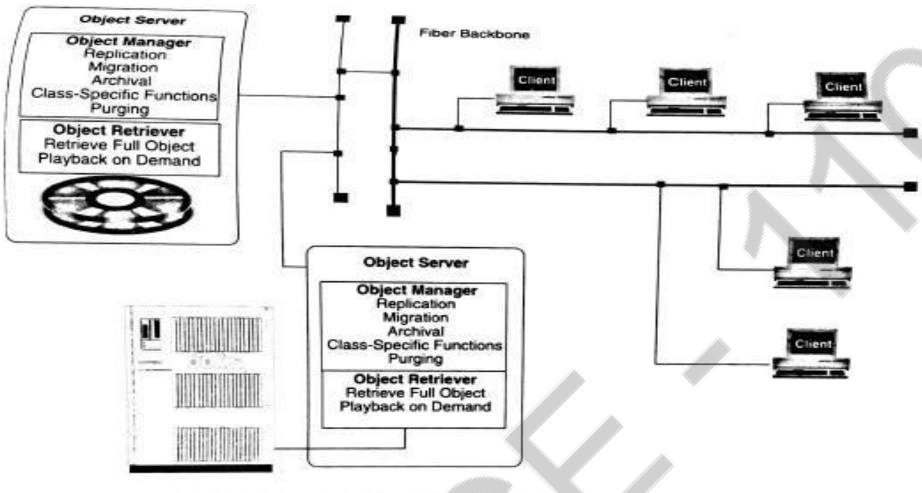


Fig. 10–3 LAN Topology for Small Workgroup

Token Ring: It is a Local Area Network architecture that combines token passing with a hybrid star/ring topology. It was developed by IBM. Token Ring Network uses a multistation Access unit at its hub ..

Extended LANs:

Ethernet and token rings are extended to get high transfer rates. For extension they use hubs , bridges and routers. In extended LANs each segment operates at the normal LAN bandwidth.

Switching Hubs:

- These have fast switching and these network can support the requirements for full motion video standard such as MPEG2.
- The important advantage of this approach is that workstation do not require additional LAN hardwareif they are already connected to a LAN.
- The user workstation continues to operate on low-cost, low speed LAN connection.

AMSCE - 1101

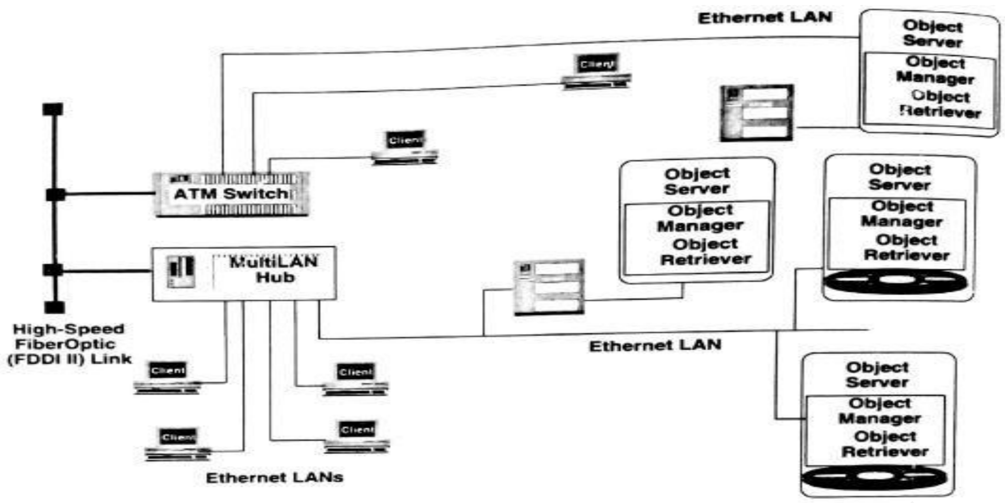


Fig. 10–4 LAN Topology Based on Hubs

Bridges and Routers:

- Bridges and Routers transfer a packet of data from LAN segment to another LAN segment.
- Bridges are devices that interconnects local or remote networks no matter what higher level protocols are involved.
- Bridges form a single logical network, centralized network administration.
- Routewrs are protocol dependent device that connects subnetworks to gether. It is useful in breaking down a very large network into smaller subnetworks.

Switching and Routing Latency:

- Switching latency is defined as the time it takes a switching hub to interconnect one LAN segment to another LAN segment.
- Routing delay is defined as the delay expirenced by a packet of data within the router.

ATM (Asynchronous Transfer Mode)

It is a network architecture that divides messages into fixed size units (called cells) of small size and that establishes a switched connection between the originating and receiving stations.

ATM appears to be a potential technology for multimedia systems for connecting object servers and user workstations. ATM is actually a good candidate for two reasons: as a hub and spoke teclmology, it adapts very well to the wiring closest paradign; and it allows workstations to operate at speeds defined by the workstation. Figure below illustrates LAN topology using an A TM Switching System.

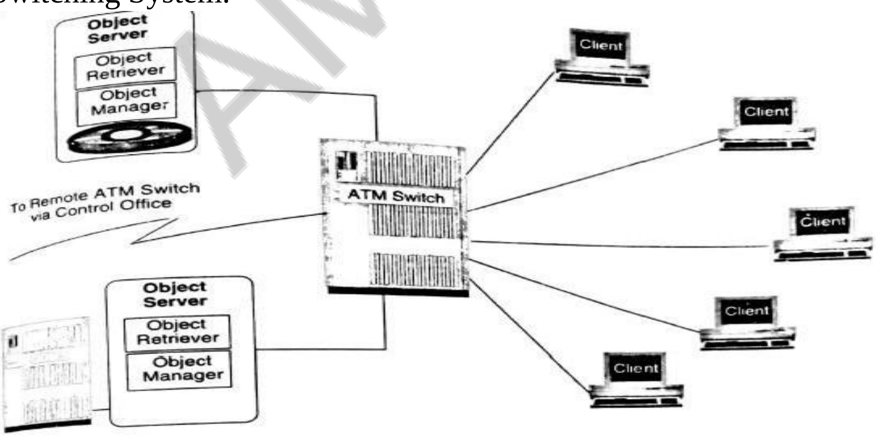


Fig. 10–6 LAN Topology Using ATM Switch

FDDI II (Fiber Distributed Data Interface II)

It is a standard for creating highspeed computer networks that employ fiber-optic cable. FOOl II operates exactly like token ring, . with one difference: FOOl employs two wires through all the hosts in a network.

AMSCE - 1101

FOOl II is a single media LAN and its full bandwidth supports all users.

FOOl II appears to be a very useful high-speed technology for connecting servers on an additional separate network and providing the dedicated high bandwidth necessary for rapid transfer and replication of information objects. Figure 5.13 shows a multiievel network based this approach.

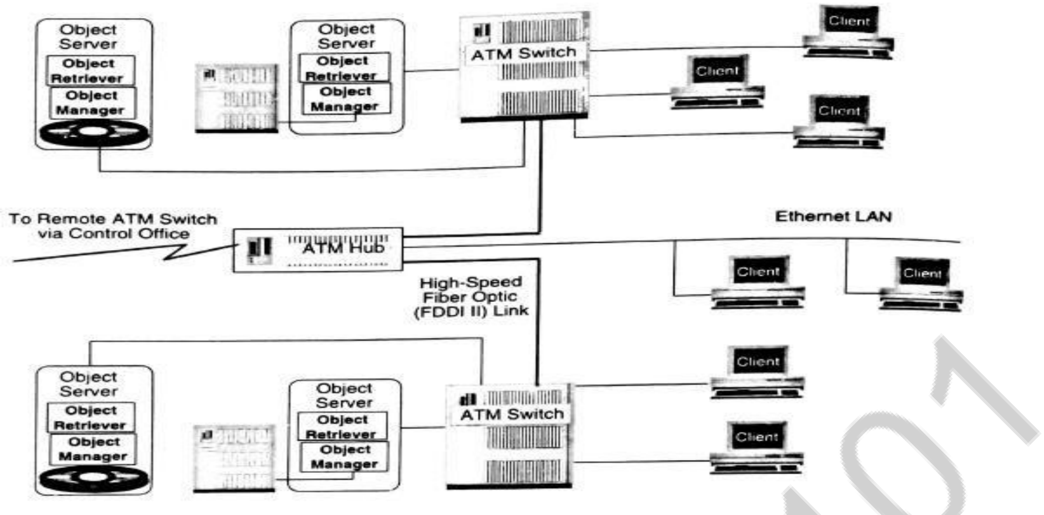


Fig. 10–7 Multilevel LAN Topology

WANS (Wide Area Network)

This includes LANs, dial up ISDN, T1 (1.544 Mbits/sec) and T3 (45.3 Mbits/sec) lines and regular telephone dial-up lines. The two big issues here are:

- :. W ANs may have a mix of networking and communication protocols.
- :. WAN has a variety of speeds at which various parts of it where it communicates.

Protocol Layering: Layering helps to isolate the network from the . application. Layering of protocols started with the release of the ISO model.

Protocol Layering:

Layerin helps to isolate the network from the application.Internet Packets can also be layered. Consider the example where the source and destination are connected via ATM using a SONET protocol at the media level.In the following figure each successive layertacks on its own header.Each layer strips off its header before presenting the data to the next layer.

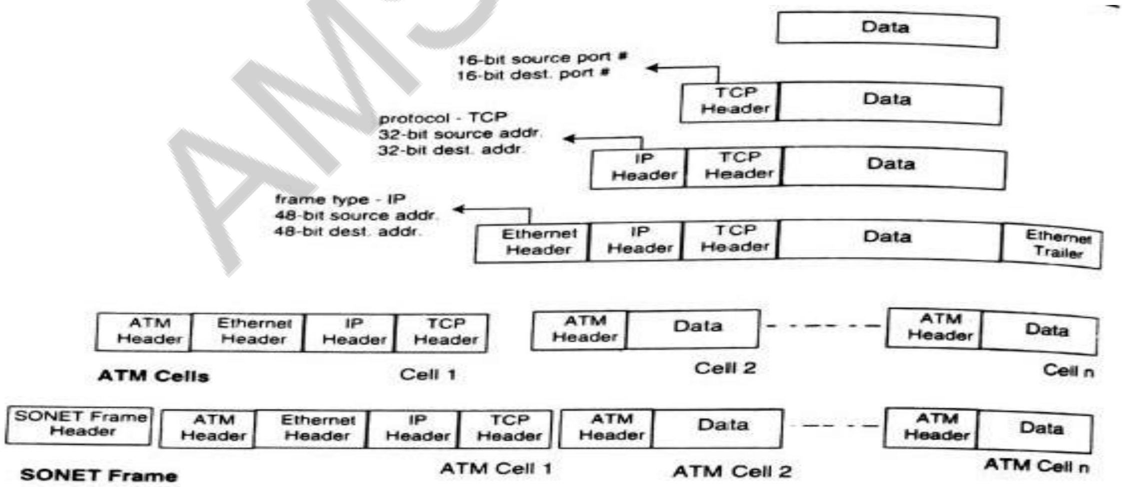


Fig. 10–8 Packet Construction Example

AMSCE - 1101

5.7.5Distributed Multimedia Databases:

A multimedia database consists of a member of different types of multimedia objects. These may include relational database records, object-oriented databases with objects for alphanumeric attributes, and s:orage servers for multimedia objects such &s images, still video, audio, and full-motion video. Itis feasible to include an image or a video object as a binary large object (BLOB) in a relational database. It is also feasible to include such an object as an attribute in an object.

Database Organization for Multimedia Applications

Optical disk storage technology has reduced the cost of multimedia document storage by a significant factor. Distributed architectures have opened the way for a variety of applications distributed around a network accessing the safe database in an independent manner. The following discussion addresses some key issues of the data organization for multimedia systems.

Data Independence: Flexible access to a variety of distributed databases for one or more applications requires that the data be independent from the application so that future applications can access the data without constraints related to a previous application. Important features of data independent design are:

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

Common Distributed Database Architecture: Employment of Common Distributed database architectured is presented by the insulation of data from an application and distributed application access.

Key features of this architecture are:

1.The ability for multiple independent data structures to co-exist in the system (multiple server classes).

2.Uniform distributed access by clients.

3.Single point for recovery of each database erver.

4.Convenient data re-organization to suit requirements

5.Tunability and creation of object classes.

6.Expandibility.

Multiple Data Servers: A database server is a dedicated resource on a network accessible to a number of apphcations, When a large number of users need to access the same resources, problem arises

This problem is solved by setting up multiple data servers that have copies of the same resources,

Transaction management for Multimedia Systems

It is defined as the sequence of events that starts when a user makes a request to create, render, edit, or print a hypermedia document. The transaction is complete when the user releases the hypermedia document and stores back any edited versions or discards the copy in memory or local storage.

Use of object classes provides an excellent way for managing and tracking hypermedia documents, Given that all components of a hypermedia document can be referenced within an object as attributes, we can find a solution for the three-dimensional transaction management problem also in the concept of objects.

Andleigh and Gretzinger expand on the basic concepts developed for the object request broker (ORB) by the Object Management Group (OMG) and combine it with their transaction management approach.

AMSCE - 1101

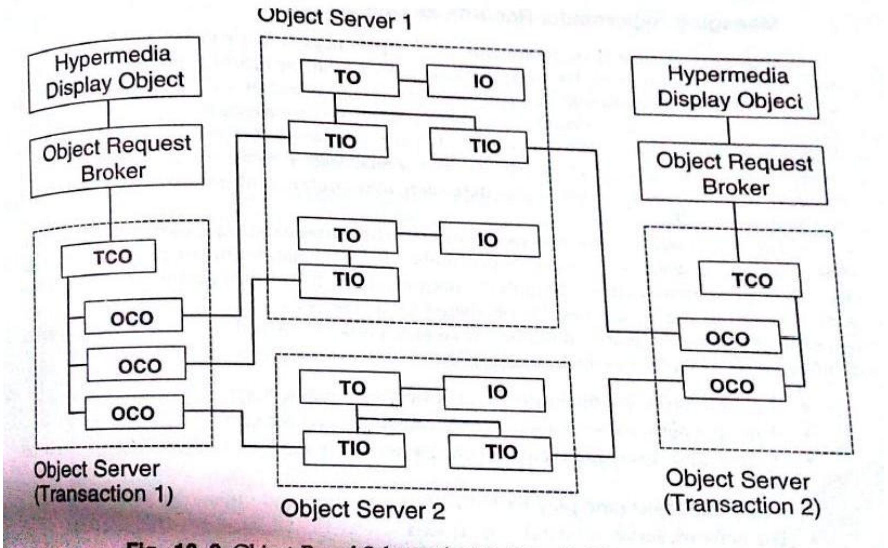


Fig. 10–9 Object-Based Scheme for Multimedia Transactions

Managing Hypermedia Records as Objects

Hypermedia records or documents are complex objects that contain multimedia information objects within them, A hypermedia document can be stored in a document data base, as a BLOB in a relational database, or in an object-oriented data base. A Hyper media document may contain multimedia objects embedded in it as special fields. An number of may be included in an embedded reference, they are,

- The object type for multimedia object (image, text, voice or video)
- A unique network-wide Object ID.
- A file name used for creation and by which it may be known.
- Size of the object.
- The application that created the object.
- Time and date of creation.
- The application or player that required to display or play it back.
- Related objects that must be at the same time for playback.
- Indexing information for indexed objects.

Object linking and embedding: OLE provides an object oriented framework for compound documents. When a user double cliks or click on an icon for an embedded object, the application that created the object starts, and allows the user to view andor the object .

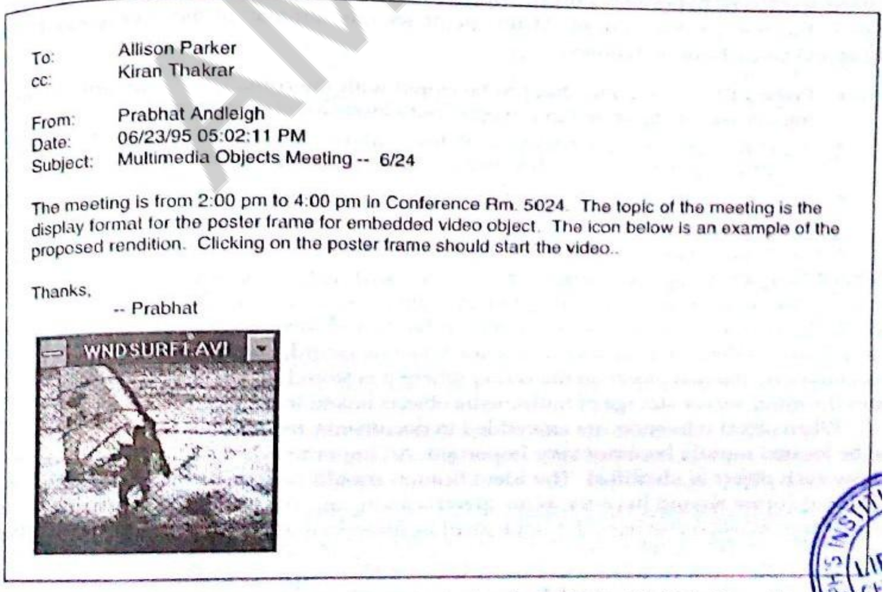


Fig. 10–10 Example of Memo with a Linked/Embedded Multimedia

AMSCE - 101

The difference between embedding and linking is given below,

Embedding	Linking
Embedded causes the object to be stored with the container document	Linking allows it to be stored in specialized object server.
An embedded object is always available with the container	A linked object depends on resolving the link to a copy on an accessible server.
Editing an embedded copy affects only the embedded copy	Editing an Linked copy affects all the container documents that reference it

The following gfigure shows multiserver storage of multimedia objects linked in a compound document**.**

Managing Distributed Objects:

We see the nature of communication between servers and the managing of distributed objects.

Interserver communications: Object replication , object distribution, object recompilation and object management and network resources are some of the design requirements that play a role in defining interserver The following lists the types of communications that one server may make to another server:

1. Obtain a token from an object name server for creating a new multimedia object; the object is not accessible by others users until complete and released.
2. Search the object class directory for the current locations of that object and the least expensive route for accessing it.
3. Perform a shared read lock on the object to ensure that it is n archived or purged while it is being retrieved.
4. Replicate a copy of the object; update the object name server directory.
5. Copy an object for non-persistent use.
6. Test and set an exclusive lock on an object for editing purposes'
7. create new versions.
8. Pause the retrieval of an object to support a user action or to pace the retrieval to the speed supported by the network.
9. A Sound server architecture is necessary for providing these services in a fully distributed environment.

Object Server Architecture

Figure describes an object server architecture that can support multimedia applications for a large number of users.

AMSCE - 1101

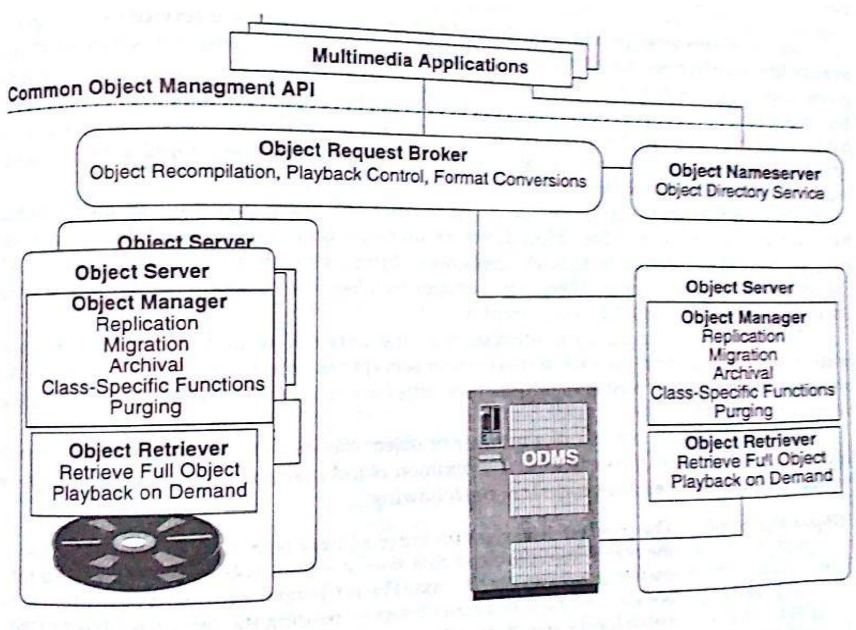


Fig. 10–12 Object Server Architecture

The architecture describes the logical distribution of functions. The following lists the key elements of this architecture:

- **Multimedia Applications**
- **Common Object Management API.**
- **Object Request Broker.**
- **Object Name Server**
- **Object Directory Manager**
- **object Server**
- **Object Manager.**
- **Network Manager**
- **Object Data Store.**

1. Any **multimedia application** designed to operate on the common object management API can function in this architecture irrespective of whether the application is electronic mail, hypermedia document management, a medical application or any other application.
2. The **common object management** API is a programming interface definition that provides a library of functions the applications can call.
3. The **common Broker Architecture API** provides a uniform interface to all applications and a standardized method for managing all information objects in a corporate network.
4. A **common Object Request Broker Architecture (CORBA**) h been defined by a Object Management Group. An object request broker performs the following functions:
   - **(i) Object recompilation.**
   - **(ii) Playback control.**
   - **(iii) Format conversions.**
5. The object **name server** provides an object directory service. That is a service to look up the availability of a information object and the list of object servers on which it is located.
6. The **object directory manager** may exist in a distributed form within an object server. The object directory manager updates the object directory when changes take place.
7. The **object server** is a logical subsystem in the network responsIble for storing and retrieving objects on demand.
8. The **object manager** consists of a number of object classes that performs a number of specialized services. They are: (i)Object retrieval. (ii) Replication(Ui) Migration. (iv) Transaction and Lock Management. (v) User Preferen'ce. (vi) Versioning. (vii) System Administration. (ix) Archival. (x) Purging. (xi) Class-Specific functions.

AMSCE - 1101

9. A **network** manager is essential for all interactions that require communicating via the network or transferring an object over the network.
10. An **object data store** is a term used for the actual storage media for specific classes of objects. The data storage may be magnetic disk or optical disk or even array of disk or jke boxes.

Identification method: Objects can be distinguished from one another in many potential ways. Identification of objects in a persistent state is different from non-persistent objects. At the highest level, persistent objects are distinguished by the class of objects. Andleigh and Gretzinger defined a rule for unique object identification as follows:

RULE: An object must have an identifier that is unique in a time dimension as well as with location such that it cannot be modified by any programmed action. An alternative approach is to divide the network into domains and have a name server in each domain be responsible for assigning new object IDs for all objects created in that domain. An object identification algorithm can be made unique by combining several of the following components.

- Network domain name.
- Address and server ID of the name server node.
- A time stamp of creating time.-- An object class identifier

Object Directory services

A multimedia object directory manager is the name server for all multimedia objects in a LAN. It has an entry for every multimedia object on all servers on the LAN, or in a domain if a LAN or WAN is subdivided into domains. The object directory manager manages changes to the object directory resulting from object manager actions.

Multimedia Object Retrieval

The multimedia object manager performs the functions of managing all requests from the multimedia applications for retrieving existing multimedia objects *or'* storing new or edited multimedia objects created by the user. In systems actively designed using an object request broker, this request is channeled through the object request broker. Data structure maintained by the multimedia object manager:

Database Replication Techniques In the simplest form of data management, the databases are set up as duplicates of the databases. Database duplication ensures that the multiple copies are identical.

There is an approach to allow each copy of the database to be modified as needed and to synchronize them by comparing them and copying the changes to all other database copies on a very frequent basis, this process is called replication.

Types of Database Replication: There are eight types of modes available. They are:

i)Round Robin replication- Each database copy replicates with the rest in a round robin manner.
ii).Manual replication- Requires the database administrator to type in a replication command at the server and define the servers with which it should replicate.
iii) Scheduled replication- Each database copy is set on its own schedule for replication with each other database copy.
(iv) Immediate replication- It is used where the multimedia object is expected to be required almost immediately.
V)Replication-on-demand- This method replicates the object only when they are required by a user logged into a specific database.
Vi) Predictive replication- It is a complex approach. The replication algorithm develops a prediction criteriafor replicating selected objects.
Vii) Replication references- It is a different approach. A reference to a replicated copy of the object is replicated.
Viii)No replication- This method is selected when the user needs to view the multimedia object just once and knows definitely that there is no need for replication.

AMSCE - 1101

The figure below shows the architecture for a distributed system providing replication service.

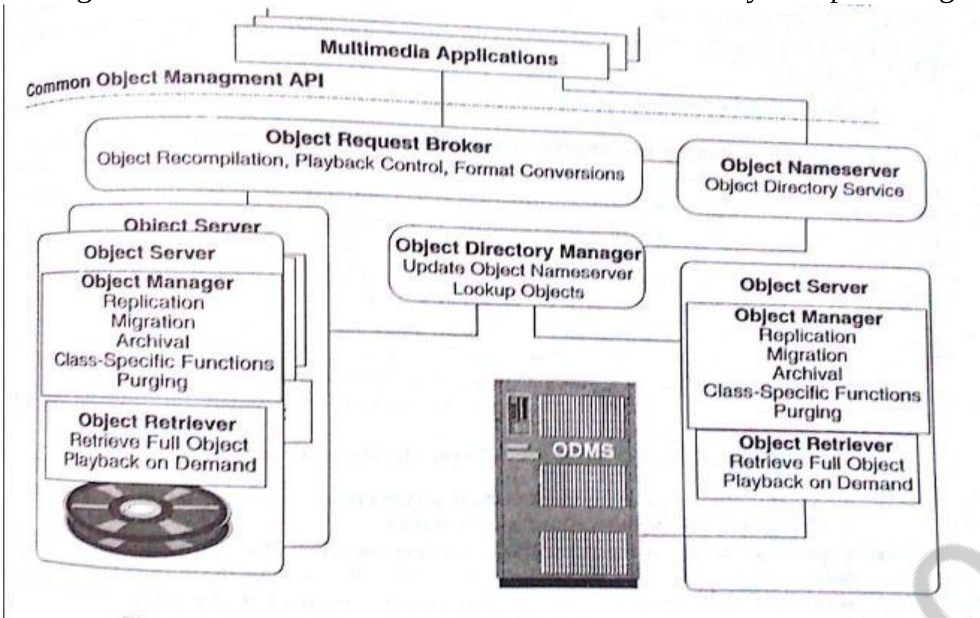


Fig. 10–14 Object Replication Architecture

Object Migration Schemes

Optimizing Object Storage A number of techniques are available for optimizing data storage for multimedia objects. Let us consider the three design approaches

1. **Optimizing Servers by Object Type:**

   The mechanism for optimizing storage is to dedicate a server to a particular type of object. The object server may ne designed to provide specialized services for specific object classes related to rendering

2. **Automatic Load Balancing:** It can be achieved by programming the replication algorithm to monitor use counts for each copy of a replicated object.

3. **Versioned Object Storage:**

   The storage problem will be more complex if multiple vcrsions need to be stored. Hence, we should follow the technique which is based on saving changes rather than storing whole new objects. New versions of the object can be complex objects,.

AMSCE - 1101

----- END-----