Chapter-4

Multimedia System (Pokhara University)

Chapter 4 - Video and Animation

Video Signal Representation

In conventional TV sets, the video signal is displayed using a cathode ray tube. An electron beam carries corresponding pattern information such as intensity in a viewed scene.
Video signal representation includes three aspects:

- ✓ the visual representation
- ✓ transmission
- ✓ digitization

Visual Representation

The objective is to offer the viewer a sense of presence in the scene and of participation in the events portrayed. To meet the objective, the televised images should convey spatial and temporal content of the scene.
Important measures taken for the purpose include:

Vertical Detail and the Viewing Distance

Geometry of the field occupied by the television image is based on the ratio of picture width and height, called as aspect ratio (W/L=4/3=1.33)
Viewing distance D determines the angle h subtended by the picture height. The angle is usually measured by the ration of the viewing distance to the picture height (D/H). The smallest detail that can be reproduced in the image is pixel
Both the number of pixels per scanned line and the number of lines per frame vary; the actual numbers used being determined by the *aspect ratio.* This is the ratio of the screen width to the screen height. The aspect ratio of current television tubes is 4/3 with older tubes – on which PC monitors are based – and 16/9 with widescreen television tubes.

Horizontal Detail and Picture Width

Picture width chosen for the conventional television services is 4/3 * picture height. Using the aspect ratio, we can determine the horizontal field of view from the horizontal angle.
The horizontal detail or the picture width is again dependent upon the aspect ratio and is given by aspect ratio*height of picture.

Total Detail Content of the Image

The vertical resolution is equal to the number of picture elements separately presented in the picture height, while the number of elements in the picture width is equal to the horizontal resolution times the aspect ratio.

The product of horizontal and vertical elements gives the total picture elements in the image. The total detail content of the image is given by the number of pixels used to represent the image and is given by the resolution. Thus total detail content of the image is the product of Vertical Detail * Horizontal Detail.

Perception of Depth

In natural vision, perception of the third spatial dimension, depth, depends primarily on the angular separation of the images received by the two eyes if the viewer.

In the flat image television, a considerable degree of depth perception is inferred from the perspective appearance of the subject matter.

As the screen is two dimensional a special measure is taken to give the sense of depth. This involves the perspective representation of the image content. This is again governed by the focal length of the lenses and changes in depth of focus in camera.

Luminance and Chrominance

Color vision is achieved through three signals, proportional to the relative intensities of Red, green and blue lights (RGB) in each portion of the scene. During the transmission of the signals from the camera to the television (display), a different division of signals in comparison to the RGB division is often used. This color encoding during transmission uses luminance and chrominance.

The term luminance is used to refer to the brightness of a source, and the color information is referred by the chrominance. Human eye however is more sensitive to luminance than to chrominance. The color encoding during transmission uses luminance and two chrominance signals (for hue and saturation).

Thus luminance (Y_s) and two chrominance blue chrominance (C_b), and the red chrominance (C_r) are then used to represent the video content where

$$Y_s = 0.299\, R_s + 0.587\, G_s + 0.144\, B_s$$

$$C_b = B_s - Y_s \text{ and } Cr = R_s - Y_s$$

Temporal Aspects of Illumination

In contrast to continuous pressure waves of an acoustic signal, a discrete sequence of individual pictures can be perceived as a continuous sequence. This property is used in television and motion pictures. The rate of repetition of the images must be high enough to guarantee smooth motion from frame to frame and the persistence of vision extends over the interval between flashes.

One of the major characteristics of human eye is that if the still images are shown in rapid succession, human eye perceive the motion i.e. it does not notice the brief cut off of the light. Thus if the rate of succession is sufficiently high still images can be used to represent motion as the persistence of vision extends over the interval between flashes.

Continuity of Motion

NTSC specified the frame rate for maintaining the visual-aural carrier separation at 4.5 MHz was 30 frames/sec but has been changed to 29.97 HZ. PAL adopted the frame repetition rate of 25 Hz, and the frame rate is 25frames/sec. Digital motion picture uses frame rate of 50 frames/sec. As stated above, motion can be represented by showing the still images in rapid succession. This can be done by showing the still images 15 frames per second. Video motion seems smooth and is achieved at only 30 frames per second.

Flicker

Through a slow motion, a periodic fluctuation of brightness perception, a flicker effect arises. The marginal value to avoid flicker is at least 50 refresh cycles per second. To achieve continuous flicker-free motion, we need relatively high refresh frequency. Movies and television apply some technical measures to work with lower motion frequencies.
When the refresh rate of frames is smaller than the eye will notice the cutoff of light between the frames and there arises the flickering rate. The marginal value to avoid flicker is at least 50 refresh cycles per second.

Temporal Aspect of Video Bandwidth
The most important factor to determine at which bandwidth the video can be transmitted is its temporal specification. Temporal specification depends on the rate of the visual system to scan pixels, as well as on the human eye's scanning capabilities. In a HDTV (High Definition TV) device, a pixel can be scanned in less than a tenth of a millionth of a second. From the human visual perspective, the eye requires that a video frame to be scanned every 1/25 second, this time is equivalent to the time during which human eye does not see the flicker effect.
The choice of bandwidth for transmission of the video signal depends on the rate of the visual system to scan pixels, as well as on the human eye's scanning capabilities. From the human visual perspective, the eye requires that a video frame be scanned every 1/25 second.

Transmission

Standards for transmission of video are: NTSC (National ***Television System Committee)*** *and* PAL *(Phase* ***Alternating Line).*** Approaches of color encoding are: RGB, YUV (luminance Y, two chrominance channels U and V), YIQ, Composite, etc.

Different color encoding techniques can be employed while transmitting the video signal and they are:

RGB signal
In this case separate signal coding is done for the individual R, G and B components of the image and these components are transmitted separately.

YUV signal
This model exploits the fact that the human eye is more sensitive to brightness i.e. luminance than to color information i.e. chrominance. Thus instead of separating the color component the brightness and the coloration of the image is separated. Thus luminance (Y_S) and two

chrominance blue chrominance (C_b), and the red chrominance (C_r) are then used to represent the video content where

$$Y=0.30\,R_s + 0.59\,G_s + 0.11\,B_s$$
$$U=(B_s-Y_s)*0.493$$
$$V=(R_s-Y_s)*0.877$$

As the bandwidth for color broadcasts must be the same as monochrome, in order to fit the Y, C_b, and C_r signals in the same bandwidth, the three signals must be combined together for transmission. The resulting signal is then known as *composite video signal.* However due to the fact that chrominance is less informative the values of two chrominance is scaled down by certain scaling factor (this is done as a measure to save bandwidth).

YIQ signal
This model is very much similar to YUV model and is used in NTSC system. The image information is broken down into three components a luminance component and two chrominance component given by following relation.

$$Y=0.30\,R_s + 0.59\,G_s + 0.11\,B_s$$
$$I=0.60\,R_s - 0.28\,G_s - 0.32\,B_s$$
$$Q=0.21\,R_s - 0.52\,G_s + 0.31\,B_s$$

Composite signal
The alternative to component coding composes all information into one signal; consequently, the individual components (RGB, YUV, and YIQ) must be combined into one signal. The basic information consists of luminance information and chrominance difference signals.

Digitalization
Digitalization is a process of changing the continuously varying signal into discrete components. This basically uses mathematical process like *Fourier analysis* or a series of step consisting of sampling and quantizing. Sampling is the process that actually digitizes the spatial position of the image while quantizing digitizes its color information. Sampling involves dividing the picture at M*N array of points while quantizing involves dividing the signal into a range of gray level values. Finally the digital motion video is created by digitizing the pictures temporally.

Computer Video Format
In order to standardize the process of digitization the international body for television standards, *the International Telecommunications Union – Radio communications Branch (ITU-R)* defined a standard for the digitization of video pictures known as *Recommendation CCIR-601.* In addition, a number of variants of this standard have been defined for use in other application domains such as television broadcasting, video telephony, and videoconferencing. The video digitizers differ in digital image resolution, quantization and frame rate. For e.g. VINO's resolution for NTSC is 640 by 480 with 8 bits/pixel and frame rate of 4 frames/second.

Computer video format depends on the input and output devices for the motion video medium
- ✓ Current video digitizers differ in digital image resolution, quantization and frame rate
- ✓ Most often used display is raster display
- ✓ The raster display architecture (as shown below)

Typically, the graphics program is used to create the high-level version of the image interactively and the *Display controller* part of the program interprets sequences of display commands and converts them into displayed objects by writing the appropriate pixel values into the *refresh buffer.*

Normally the *video controller* is a hardware subsystem that reads the pixel values stored in the refresh buffer in time-synchronism with the scanning process and, for each set of pixel values, converts these into the equivalent set of red, green, and blue analog signal for output of the display.

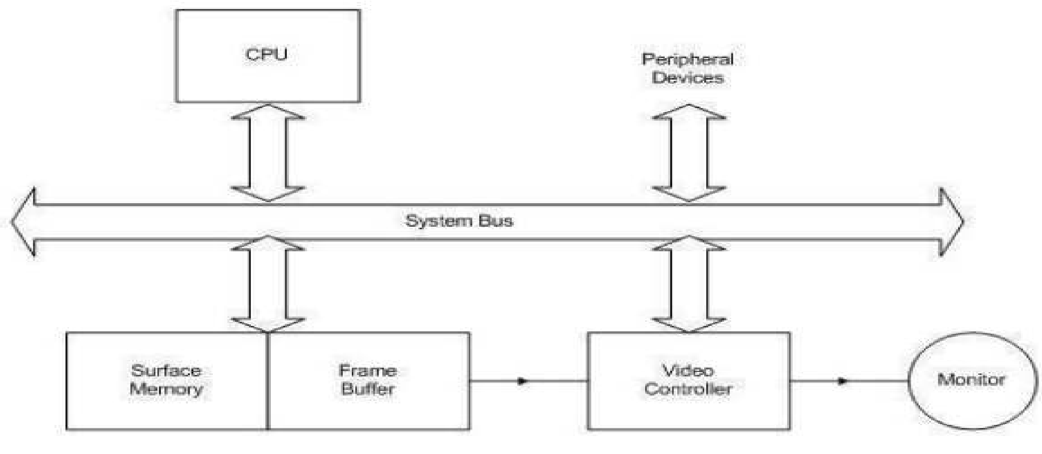

Fig. A common raster display system

A range of color values are stored in a table and each pixel value is used as an address to a location within the table which contains the corresponding three color values. The table is known as *color look up table.*

Some computer video controller standards are given below:

Color Graphics Adapter (CGA)
Resolution: 320*200
Color depth: 2 bits/pixel
Image size: 16 KB
No. of colors: 4
Storage Capacity: 320*200*4/8=16,000 bytes

Enhanced Graphics Adapter (EGA)
Resolution: 640*480
Color depth: 4 bits/pixel
Image size: 112 KB
No. of colors: 16
Storage Capacity: 640*480*4/8=112,000 bytes

Video Graphics Array (VGA)
Resolution: 640*480
Color depth: 8 bits/pixel
Image size: 307.2 KB
No. of colors: 256
Storage Capacity: 640*480*8/8=307,200 bytes

8514/A Display Adapter Mode
Resolution: 1024*768
No. of colors: 256
Storage Capacity: 1024*768*8/8=786,432 byte

Extended Graphics Array (XGA)
Resolution: 640*480 / 1024*768
Color depth: 65,000 colors / 256 colors
No. of colors: 256
Storage Capacity: 1024*768*8/8=786,432 bytes

Super VGA (SVGA)
Resolution: 1024*768
Color depth: 24 bits/pixel
Image size: 2.35 MB
No. of colors: 2^{24}
Storage Capacity: 1024*768*24/8=2,359,296 bytes

Television

Conventional Systems

Conventional television systems employ the following standards:

- ✓ NTSC
- ✓ SECAM
- ✓ PAL

NTSC

- ✓ National Television Systems Committee
- ✓ Developed in the U.S.
- ✓ Oldest and most widely used television standard
- ✓ Color carrier is used with approximately 4.429 MHz or with approximately 3.57 MHz.
- ✓ Uses a quadrature amplitude modulation with a suppressed color carrier
- ✓ Works with a motion frequency of approximately 30 Hz
- ✓ A picture consists of 525 lines
- ✓ 4.2 MHz is used for the luminance and 1.5 MHz is used for each of the two chrominance channels.

NTSC **N**ational **T**elevision **S**ystem **C**ommittee	
Lines/Field	525/60
Horizontal Frequency	15.734 kHz
Vertical Frequency	60 Hz
Color Subcarrier Frequency	3.579545 MHz
Video Bandwidth	4.2 MHz
Sound Carrier	4.5 MHz

SECAM

- ✓ Sequential Couleur Avec Memoire
- ✓ Used in France and Eastern Europe
- ✓ Unlike NTSC and PAL, it is based on frequency modulation
- ✓ Uses a motion frequency of 25 Hz
- ✓ Each picture has 625 lines

SECAM **S**equential **C**ouleur **A**vec **M**emoire *or* **S**equential **C**olor with **M**emory		
SYSTEM	**SECAM B,G,H**	**SECAM D,K,K1,L**
Line/Field	625/50	625/50
Horizontal Frequency	15.625 kHz	15.625 kHz
Vertical Frequency	50 Hz	50 Hz
Video Bandwidth	5.0 MHz	6.0 MHz
Sound Carrier	5.5 MHz	6.5 MHz

PAL

- ✓ Phase Alternating Line
- ✓ Invented in 1963 by W. Bruch
- ✓ Used in parts of Western Europe
- ✓ Uses a quadrature amplitude modulation similar to NTSC, but the color carrier is not Suppressed

P A L **P**hase **A**lternating **L**ine			
SYSTEM	**PAL**	**PAL N**	**PAL M**
Line/Field	625/50	625/50	525/60
Horizontal Freq.	15.625 kHz	15.625 kHz	15.750 kHz
Vertical Freq.	50 Hz	50 Hz	60 Hz
Color Sub Carrier	4.433618 MHz	3.582056 MHz	3.575611 MHz
Video Bandwidth	5.0 MHz	4.2 MHz	4.2 MHz
Sound Carrier	5.5 MHz	4.5 MHz	4.5 MHz

Enhanced Definition System

EDTV are conventional systems modified to offer improved vertical and/or horizontal resolution. This stands for Enhanced-Definition TV, and usually it describes a television that can display HDTV signals but doesn't have enough resolution to really do them justice. Most often it applies to plasma TVs and denotes 852x480 pixels.

Used in NTSC broadcasts, they essentially make full use of the 4.2 MHz bandwidth of luminance, producing better than 30% improvement in horizontal resolution. Vertical resolution improvements were achieved by introducing progressive scanning into receivers designed for NTSC, PAL and SECAM services. Some of the improved intermediate TV stations are: IDTV and D2-MAC

IDTV

- ✓ Improved-Definition Television
- ✓ Attempts to improve NTSC image by using digital memory to double the scanning lines from 525 to 1050
- ✓ The pictures are only slightly more detailed than NTSC images
- ✓ Vertical resolution is enhanced
- ✓ Normally, 525-line TV pictures are presented in two interlaced half pictures, each 1/60 second apart

D2-MAC

- ✓ Duobinary Multiplexed Analogue Components
- ✓ Uses a time-multiplexing mechanism for component transmission
- ✓ 625 lines are transmitted, but only 574 lines are visible
- ✓ Width to height ratios (aspect ratios) of 4:3 and 16:9 are supported.
- ✓ The audio and further data are transmitted in a duobinary coding
- ✓ Audio channels can be coded as two high-quality stereo signal channels, or up to eight low audio quality channels

High Definition Television

HDTV is high-resolution digital television (DTV) combined with Dolby Digital surround sound (AC-3). HDTV is the highest DTV resolution in the new set of standards. This combination creates a stunning image with stunning sound. **HDTV** requires new production and transmission equipment at the **HDTV stations,** as well as new equipment for reception by the consumer. The higher resolution picture is the main selling point for HDTV. Imagine 720 or 1080 lines of resolution compared to the 525 lines people are used to in the United States (or the 625 lines in Europe) -- it's a huge difference!

Of the 18 DTV formats, six are HDTV formats, five of which are based on progressive scanning and one on interlaced scanning. Of the remaining formats, eight are SDTV (four wide-screen formats with 16:9 aspect ratios, and four conventional formats with 4:3 aspect ratios), and the remaining four are video graphics array (VGA) formats. Stations are free to choose which formats to broadcast.

The formats used in HDTV are:

- ✓ **720p** - 1280x720 pixels progressive
- ✓ **1080i** - 1920x1080 pixels interlaced
- ✓ **1080p** - 1920x1080 pixels progressive

"Interlaced" or "progressive" refers to the **scanning system.** In an **interlaced** format, the screen shows every odd line at one scan of the screen, and then follows that up with the even lines in a second scan. Since there are 30 frames shown per second, the screen shows one half of the frame every sixtieth of a second. For smaller screens, this is less noticeable. As screens get larger, the problem with interlacing is **flicker. Progressive** scanning shows the whole picture, every line in one showing, every sixtieth of a second. This provides for a much smoother picture, but uses slightly more bandwidth.

High-Definition Television (HDTV) is defined by the image it presents to its viewer, which has the following characteristics:

Resolution
The HDTV image has approximately twice as many horizontal and vertical pixels as conventional systems. Luminance detail is also increased by employing a video bandwidth approximately five times that used in conventional systems.

Aspect Ratio
The aspect ratio of the HDTV images is 16/9 = 1.777.

Viewing Distance
Since the eye's ability to distinguish details is limited, the more detailed HDTV image should be viewed closer than conventional systems.

Digital coding is essential in the design and implementation of HDTV. There are two kinds of possible digital coding: *composite coding* and *component coding.*

Composite coding
The composite analog video signal is sampled. All signal components are converted together into a digital representation. Composite coding of the whole video signal is easier as compared to separately digitizing each signal component. But there are certain problems associated with this approach.

They are:

- ✓ There is often disturbing cross-talk between the luminance and chrominance information.
- ✓ The composite coding of the color television signal depends on the television standard. Hence, the difference among the various standards must be adjusted.
- ✓ Since luminance information is more important than chrominance, this information should be allocated more bandwidth. But this is not possible with composite coding.
- ✓ Using component coding, the sampling frequency is not coupled with the color carrier frequency.

Component coding
The various components are separately digitized. For example, the luminance and the two chrominance signals are separately digitized. The separate can be transmitted together using multiplexing. The luminance signal is sampled with 13.5 MHz as it is more important than the chrominance signal. The chrominance signals are sampled with 6.75 MHz.

Thus, the component bandwidth ratio is 4:2:2. Due to the 4:2:2 component bandwidth ratio, we get 864 sample values per line for the luminance (out of which 720 are visible) and 432 sample values for each chrominance component (out of which 360 are visible). The digitized luminance and chrominance signals are then quantized uniformly with eight bits.

Three HDTV systems have and are being developed worldwide. They are described below.

United States

- ✓ HDTV will be the follower of IDTV.
- ✓ Compatibility with NTSC is a prime concern.
- ✓ Aspect ratio = 16:9.
- ✓ Number of lines = 1025
- ✓ Motion frequency = 59.94
- ✓ Compatibility with NTSC will be achieved by using IDTV.

Europe

- ✓ The transmission method of the European HDTV system is known as HD-MAC (High Definition Multiplexed Analog Components).
- ✓ For compatibility with PAL and D2-MAC, the HD-MAC will use a two-level sampling scheme.
- ✓ Specified in Eureka Project EU 95, which began in 1986

Japan

- ✓ The first HDTV service available to the public was the NHK system, called MUSE.
- ✓ MUSE began operation in late 1989.
- ✓ MUSE is a Direct-Broadcast-from-Satellite (DBS) system.
- ✓ MUSE employs a 1125-line image at its input.
- ✓ The 1125 lines are converted to the narrow channel required by the satellite transponders.
- ✓ The above conversion retains the full detail of the 1125 line image, but only when the image is stationary. When motion occurs, the definition is reduced by approximately 50 %.

Transmission

For multimedia transmission, the data rates created by motion video are important. The data rate for U.S. HDTV is 1.0368×10^9 bits/second in general. The data rate for European HDTV is 1.152×10^9 bits/second in general.

In digital television, different substandards and data reductions can be applied, and hence the data rate can be calculated as follows.

Substandard 1

- ✓ Works with a luminance sampling frequency of 11.25 MHz and with a chrominance sampling frequency of 5.625 MHz
- ✓ Data rate is 190×10^6 bits/second.

Substandard 2

- ✓ Works with a luminance sampling frequency of 10.125 MHz and with a chrominance sampling frequency of 3.375 MHz
- ✓ Data rate is 125×10^6 bits/second.

Substandard 3

- ✓ Works with a luminance sampling frequency of 9.0 MHz and with a chrominance sampling frequency of 2.25 MHz
- ✓ Data rate is 108×10^6 bits/second.

Further reduction of the data rates can be achieved by following three steps:

Removing the sampling gaps; this means that only visible lines are coded. For example, the luminance consists of 648 samples out of which only 540 are visible. The chrominance consists of 216 samples out of which only 180 are visible. Hence, if the picture consists of 575 lines, then the data rate equals 103.5×10^6 bits/second, which is calculated as shown below.

(540 + 180 + 180) sample values per line × 575 visible lies/picture = 517,500 sample values /picture

517,500 sample values/picture × 8 bits/sample value × 25 pictures/second = 103.5×10^6 bits/second.

Reduction of vertical chrominance resolution; Not all but alternating lines of both the chrominance components are digitized. Hence, if the picture consists of 575 lines, then the data rate equals 82.8×10^6 bits/second, which is calculated as shown below.
(540 + 90 + 90) sample values per line × 575 visible lies/picture = 414,000 sample values /picture
414,000 sample values/picture × 8 bits/sample value × 25 pictures/second = 82.8×10^6 bits/second.

Different kinds of source coding can be applied to the components; For example, with an intra-frame working ADPCM, a data rate reduction of 3 bits/sample (instead of 8 bits/sample) for luminance and chrominance can be achieved. Hence, if the picture consists of 575 lines, then the data rate equals 31.050×10^6 bits/second, which is calculated as shown below.
(540 + 90 + 90) sample values per line × 575 visible lies/picture = 414,000 sample values /picture
414,000 sample values/picture × 3 bits/sample value × 25 pictures/second = 31.050×10^6 bits/second.

Computer-Based Animation

Animation is the processes of making something look and feel live. Animation can broadly be categorized into two categories:

Motion Dynamics
Induction of life by changing the position of the object
Ex: animation of a bouncing ball

Update Dynamics
Induction of life by changing the shape, size and/or structure of the object
Ex: animation of flower blooming from bud

Computer based animation used computer based graphical tools to provide visual effects. Advantage of computer based animation includes efficiency, time and simplicity.

In several applications the video can be generated by a computer program rather than a video camera. This type of video content is normally referred to as **computer animation** or sometimes, because of the way it is generated, animated graphics.

A range of special programming languages is available for creating computer animation. To animate something means to bring it life. An animation covers all changes that have a visual effect. Visual effect can be of different nature. They might include time varying positions (motion dynamics), shape, color, transparency, structure and texture of an object (update dynamics), and changes is lighting camera position, orientation and focus. Computer based animation is an animation performed by a computer using graphical tools to provide visual effects. Animation covers all changes that have a visual effect. Visual effect can be of different nature. They might include:

- ✓ Time-varying positions (motion dynamics),
- ✓ Shape, color, transparency, structure and texture of an object (update dynamics), and
- ✓ Changes in lighting, camera position, orientation and focus.

A computer-based animation is an animation performed by a computer using graphical tools to provide visual effects.

Basic Concepts:

Input process:
Before computer can be used, drawing must be digitized because key frames, in which the entities being animated are at extreme or characteristic positions, must be drawn. This can be done through optical scanning, tracing the drawings with a data tablet or producing the original drawings with a drawing program. Drawings may need to be post-processed (e.g. filtered) to clean up any glitches arising from the input process

Composition stage:
In this stage, foreground and background figures are combined to generate individual frames for the final animation. This is done by image-composition techniques.

Inbetween process:
Animation of movement from one position (starting frame) to another (ending frame) needs a composition of frames with intermediate positions (intermediate frames) in between the key frames. This is called inbetween process.

This process is performed in computer-based animation through interpolation. The easiest interpolation is linear interpolation (sometimes called lerping-Linear intERPolation). Splines are an alternative to lerping for removing the drawbacks of lerping, but they, too, do not entirely solve the inbetweening problem. Inbetweening also involves interpolating the shapes of objects in intermediate frames.

Changing colors:
For changing colors, computer based animators use CLUT (lut) in a frame buffer and the process of double buffering. The lut animation is generated by manipulating the lut. The simplest method is to cycle the colors in the lut, thus changing the colors of various pieces of the image. Using lut animation is faster than sending an entire new pixmap to the frame buffer for each frame.

Animation Languages:
Animation languages fall into three categories:

Linear-list Notations:
Each event in the animation is described by a starting and ending frame number and an action that is to take place (event). The actions typically take parameters, such as degree of rotation, axis of rotation, etc.

E.g. 42, 53, B, ROTATE "PALM", 1, 30 means "between 42 and 53 rotate object called PALM about axis 1 by 30 degrees, determining the amount of rotation at each frame from table B.
E.g.: Scefo (SCEne FOrmat)

General-purpose languages:
Another way to describe animation is to embed an animation capability within a general purpose programming language. The values of variables in the language can be used as parameters to the routines, which perform the animation.

E.g.: ASAS.

Values in the language can be used as parameters to the routine, which perform the animation

Ex:
(grasp my-cube); /* cube becomes the current object
(cw 0.05); /* spin clockwise by a small amount
(grasp camera); /* make the camera the current object
(right 5); /*move right by 5 units

Graphical languages:
Graphical animation languages describe animation in a more visual way than textual animation languages. These languages are used for expressing, editing and comprehending the simultaneous changes taking place in an animation. Rather than writing out description of actions, the animator provides a picture of the action.

E.g.: GENESYSTM, S-Dynamics System.

Graphical languages provide user the ability to visualize the action clearly. Graphical language describes animation in a more visual way. It can be used for expressing, editing and comprehending simultaneous changes.

Methods of Controlling Animation:
Controlling animation is independent of the language used for describing it. Animation control mechanisms can employ different techniques.

Full Explicit Control:
It is a simplest type of animation control. Animator provides a description of everything that occurs in the animation, either by specifying simple changes, such as scaling, translation, and rotation, or by providing key frame information and interpolation methods to use between key frames.

This interpolation may be given explicitly or by direct manipulation with a mouse, joystick, data glove or other input device (in an interactive system). E.g. BBOP system

Procedural Control:
It is based on communication between various objects to determine their properties. It is a significant part of several other control mechanisms. In physically-based systems, the position of one object may influence the motion of another (eg: balls cannot pass through walls). In actor-based systems, the individual actors may pass their positions to other actors to affect the other actors' behaviors.

Constraint-based Systems:
Many objects move in a manner determined by other objects with which they are in contact, and this compound motion may not be linear at all. Such motion can be modeled by constraints. Specifying an animated sequence using constraints is often much easier to do than using explicit control. E.g. Sutherland’s Sketchpad

Tracking Live Action:
Trajectories of objects in the course of an animation can be generated by tracking live action. Traditional animation uses **rotoscoping.** A film is made in which people/animals act out the parts of the characters in the animation. Then animators draw over the film, enhancing the background and replacing the human actors with their animated equivalents.

Another live-action technique is to attach some sort of indicator to key points on a person's body. By tracking the positions of the indicators, one can get location for corresponding key points in an animated model.

Eg: data glove, which measures the position and orientation of the wearer's hand as well as the flexion and hyperextension of each finger point.

Kinematics and Dynamics:
Kinematics refers to the position and velocity of points. E.g. a kinematic description of a scene might say, “The cube is at origin at time t=0. It moves with a constant acceleration in the direction (1, 1, 5) thereafter”.

Dynamics takes into account physical laws that govern kinematics (Eg: Newton's laws of motion for large bodies etc.). A particle moves with an acceleration proportional to the forces acting on it, and the proportionality constant is the mass of the particle. Thus, a dynamic description of a scene might be, “At time t=0 seconds, the cube is at position (0 meters, 100 meters, 0 meters). The cube has a mass of 100 grams. The force of gravity acts on the cube.

Having the knowledge of kinematics of kinematics and dynamics of the object at a point, we can determine the next point in motion.

Display of Animation
Animated objects (which may consist of graphical primitives such as lines, polygons, and so on) must be scan-converted into their pixmap in frame buffer. To show a rotating object, we can scan-convert into the pixmap successive views from slightly different locations, one after another. This scan-conversion must be done at least 10 (preferably 15 to 20) times per sec to give a reasonably smooth effect. Hence, a new image must be created in no more than 100 milliseconds.

To remove distracting effect (due to slow scan-conversion), double buffering is used. To display animations with raster systems, animated objects must be scan-converted into their pixmap in the frame buffer. To show a rotating object, we can scan-convert into the pixmap successive views from slightly different locations, one after another. This scan conversion must be done at least 10 times per second to give reasonably smooth effect. Hence, a new image must be created in no more than 100 ms. Let us assume that the two halves of the pixmap are image0 and image1.

Load look-up table to display values as background color
Scan-convert object into image0
Load look-up table to display only image0

Repeat

Scan-convert object into image1
Load look-up table to display only image1
Rotate object data structure description
Scan-convert object into image0
Load look-up table to display only image0
Rotate object data structure description

Until (frame no. is 15)

Transmission of Animation:

One of the following two approaches may be used for the transmission of animation over computer networks:

Symbolic Representation

The symbolic representation (e.g. circle) of animation objects (e.g. ball) is transmitted together with the operation commands (e.g. roll the ball) performed on the object, and at the receiver side the animation is displayed. The transmission time is short because the symbolic representation of an animated object is smaller in byte size than its pixmap representation. However, the display time at the receiver takes longer because the scan-converting operation has to be performed at the receiver side. The transmission rate (bits/sec or bytes/sec) of animated objects depends on:

- ✓ The size of the symbolic representation structure, where the animated object is encoded,
- ✓ The size of the structure, where the operation command is encoded, and
- ✓ The number of animated objects and operation commands sent per second.

Pixmap Representation

The pixmap representation of the animated objects is transmitted and displayed on the receiver side. The transmission time is longer in comparison to the previous because of the size of the pixmap representation. However, the display time is shorter because the scan-conversion of the animated objects is avoided at the receiver side. It is performed at the sender side where animation objects and operation commands are generated. The transmission rate of the animation is equal to the size of the pixmap representation of an animated object (graphical image) multiplied by the number of graphical images per second.

References:

- ✓ Multimedia: Computing, Communications and Applications", Ralf Steinmetz and Klara Nahrstedt, Pearson Education Asia
- ✓ "Multimedia Communications, Applications, Networks, protocols ad Standards", Fred Halsall, Pearson Education Asia
- ✓ "Multimedia Systems", John F. Koegel Buford, Pearson Education Asia

Assignments:

(1) What do you mean by computer based animation? List the different types of animation languages.
(2) List three distinct models of color used in Multimedia. Explain why there are a number of different color models exploited in multimedia data formats.
(3) Explain Tele-services and the implementation of Conversation services in Multimedia communication.
(4) Explain the methods that are used to control animation.
(5) Discuss the YUV model for video transmission.
(7) Describe the television standards.