Unit 4

Video and animation

Discussion Topics

- Introduction to video and animation
- Digital video Representation
- video Format
- Computer-based video animation
- Animation language
- Methods of controlling animation
- Display of animation
- Transmission of animation

Prepared by: Shiv Raj Pant

What is video ?

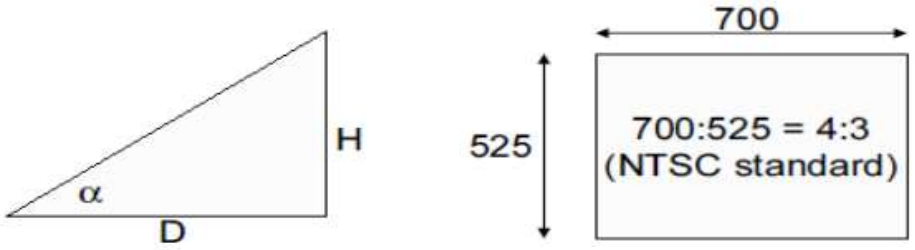
- Video is simply a moving picture.
- A sequence of still pictures (called frames) change in a certain speed so that our eyes get the illusion of a continuous motion.
- The minimum frame rate to achieve a comfortable illusion of a moving image is about sixteen frames per second.
- Motion video and computer-based animation have become basic media for multimedia systems.

Some terms

Aspect ratio

It is the geometry of field occupied by the image. Aspect ratio is the ratio of the picture width to height.

There are different aspect ratios such as 4:3, 16:9 etc used in televisions and monitors.



Vertical Viewing distance (vertical field of view)

The viewing distance (D) determines the angle subtended by the picture height(h).

The angle is usually measured as the ratio D/H.

Pixel

The smallest detail of a scene that can be reproduced in the image is pixel.

Video signal representation

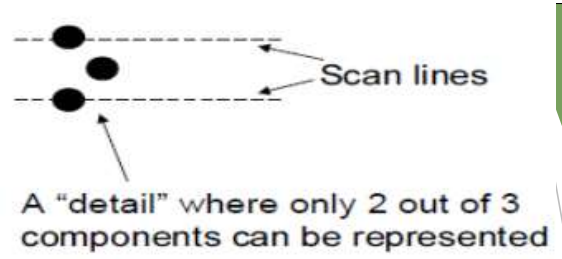
- Video signal representation includes three aspects:
  1. ***Visual representation***
  2. ***Transmission***
  3. ***Digitization***

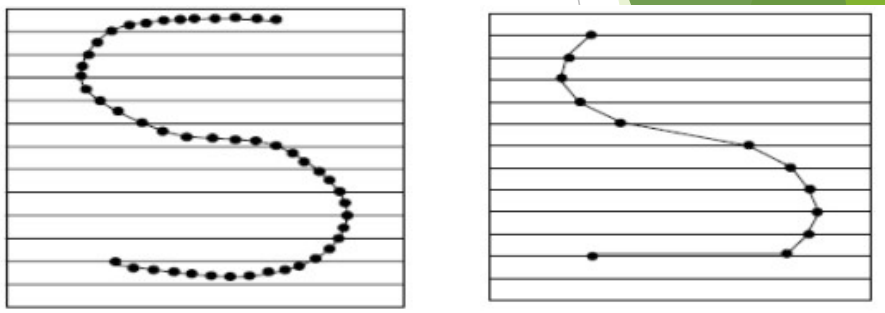
Visual representation

- Visual representation deals with providing the viewer a good feeling of view.
- For this, the visual representation should convey spatial and temporal content of the scene.
- Following are some factors to be considered for visual representation :-

Vertical detail

- Ideally, each detail of the scene should be reproduced by one pixel.
- But practically, some of the details in the scene fall between scanning lines of display device.
- Thus some vertical resolution is lost.





- It has been shown that only 70% of the vertical detail is presented by the scanning lines. It is called Kell factor.

Horizontal detail

The picture width (W) is the aspect ratio*height.

The horizontal angle subtended by the picture width is the ratio D/W

Total detail and total content of image

- Vertical resolution = number of elements separately presented in the picture height.
- Horizontal resolution = number of elements in the picture width = resolution * aspect ratio
- Total number of picture elements = number of elements vertically * number of elements horizontally

Perception of Depth

- Depth is the third spatial dimention.
- In natural vision, the perception of depth depends on the angular separation of images received by two eyes of the viewer.
- In television image (flat image), the depth perception depends on perspective appearance of objects and the choice of focal lenth and focus depth of camera.

Luminance and chrominance

- Color perception can be achieved by mixing the intensities of Red, Green and Blue light.
- However, we can use another method called YUV system where the color encoding uses luminance and chrominance signals.

Temporal aspects of illumination

- Motion is the presentation of a rapid succession of slightly different still pictures(frames)
- Between frames, the light is cutoff briefly.
- To represent visual reality, two conditions must be met:

  Rate of frame change must be high enough to guarantee smooth motion

  The rate of frame change must be hight enough so that the vision persists over the interval between flashes.

Contunity of motion

- We perceive a continuous motion to happen at any frame rate faster than 15 frames per second
- Smooth video motion is achieved at 30 frames/second
- There are different standards used in movies, televisions and monitors about how many frames/sec to use.

Flicker

- Flicker is periodic fluctuation of brightness in display systems.
- The marginal value to avoid flicker is 50 refresh cycles per second.
- To achieve continuous fliker-free motion, we need a relatively high refresh frequency.
- Movies and televisions apply some technical measure to work with lower motion frequencies.

How to avoid flicker?

Suppose we have 16 pictures per second for a video. This gives annoying flicker.

To avoid the flicker, same picture can be displayed several times.

In traditional movies system, each picture is displayed 3 times so that 16*3 = 48 frame/sec is achieved

In monitors, refresh buffer is used to store pictures to display repeatedly to achieve 70Hz

Temporal aspect of video bandwidth

- **See next slide.**

Temporal aspect of video bandwidth

- To transmit a video, we must determine the video bandwidth.
- Video bandwidth is determined by temporal specification of video (pixel scan rate and resolution).

> Temporal specification also depends on human eye's scanning characteristic. Human eye requires that every video frame be scanned every 1/25 second to avoid flicker effect.

*Horizontal_scan_frequency = vertical lines * frame_rate*

*Video bandwidth = ½ * horizontal lines * horizontal_scan_frequency*

> The division by 2 is because 2 horizontally adjacent pixels can be scanned during one cycle of video frequency.

Example: consider a computer monitor with resolution 1312*800 and frame_rate = 100 Hz

Then, it requires video bandwidth = ½*1312*800*100 = 52.48 MHz

- Sometimes the scanning rate is given in pixels per second. This is equivalent to the resolution * frame_rate. (e.g. in above example, the product 1312*800*100 is the pixel per second)

Transmission

- Video signals are transmitted through single channel.
- For color encoding, several approaches can be used:

RGB signal

You know this already!

YUV signal

It has been observed that human perception is more sensitive to brightness than color.

So we can separate luminance information(Y) from chrominance information (U and V)

YUV color model is based on above idea.

The component division for YUV signal is:

$$Y = 0.30\,R + 0.59\,G + 0.11\,B$$

$$U = (B-Y)*0.493$$

$$Q = (R-Y)*0.877$$

Y,U and V are not abbreviations.
Y = luminance
U=Blue projection
V = Red projection

since, Luminance is more important than chrominance, the luminance values are codded using higher bandwidth.

The CD-I and DVI used YUV model.

YIQ signal

YIQ is the basis for NTSC format

The component division for YIQ signal is:

$$Y = 0.30\,R + 0.59\,G + 0.11\,B$$

$$U = 0.60R - 0.28\,G - 0.32B$$

$$Q = 0.21R - 0.52G + 0.31B$$

Composite signal

This method composes all information into one signal

i.e. RGV, YUV, YIQ are combined into one signal

The basic video bandwidth required to transmit luminance and chrominance signals is 4.2MHz for NTSC standard.

In HDTV, it is at least twice.

Digitization

- Before a video can be processed or transmitted over a computer network, it needs to be converted from analog to digital representation
- Since, the video is just a sequence of pictures, the digitization is similar to digitizing individual images.
- Thus sampling and quantization of pixel color levels needs to be performed to represent each image in the video as 2D array of pixels.
- However, since video has a temporal aspect, the images must be digitized in strict time to get a sequence of digital images per second that approximates analog video.

Video Format

- Video format vary according to resolution, quantization and frame rate.
- The formats are implemented on particular input and output devices.
- Different devices may represent videos from 15 frames/second to 30 frames/sec
- However, most display devices refresh rate is 60Hz to avoid flicker. To achieve 60 screens/sec from 30 frames/sec, the same frame are actually displayed repeatedly.
- Some video controller standards are as follows:

Hardware	Resolution	Color	Per-Image size
Color graphics Adaptor (CGA)	320x200	2 bits	320*200*2/8 = 16000 bytes
Enhanced graphics Adaptor(EGA)	640x350	4 bits	640*350*4/8 = 112000 bytes
Video graphics array (VGA)	640x480	8 bits	
Extended graphics array(XGA)	640x480 1024x768	16 bits 8 bits	
Super VGA (SVGA)	1024x768	24 bits	

Video file Formats

Videos can be stored in different file formats:

Some popular formats are:

MP4

MP4 (MPEG-4 Part 14) is the most common type of video file format. MP4 can play on most devices as well. It uses the MPEG-4 encoding algorithm to store video and audio files and text. MP4 provides high quality video while maintaining relatively small file sizes, but it offers lower definition than some others.

FLV

Flash Video is a file format used by Adobe Flash Player. It is one of the most popular and versatile video formats supported by all video platforms and browsers. The FLV format is a good choice for online video streaming. They have a relatively small file size. The only drawback is that it's not compatible with many mobile devices like iPhones.

MKV

Developed in Russia, Matroska Multimedia Container format is free and open source. It supports nearly every video encodings. MKV is suitable to view video on a TV or computer using an open-source media player like VLC.

WMV

WMV (Windows Media Viewer) files offer good video quality and large file size. Microsoft developed WMV for Windows Media Player. YouTube supports WMV, and Apple users can view these videos, but they must download Windows Media Player for Apple. The drawback is that we can't select our own aspect ratio in WMV.

AVI

The AVI file format was introduced in 1992 by Microsoft and is still widely used today. The AVI video format uses less compression than other video formats such as MPEG. This results in very large file sizes.

Television

- Television is the most important application that has driven the development of digital videos.
- Since 1950s, the TV technology has undergone many changes.
- Different standards have been used in television video displays.
- We can distinguish 3 types of televisions systems:
  1. *Conventional systems*
  2. *Enhanced Definition systems*
  3. *High Definition systems*

Conventional systems

- Old Black and white systems and color systems are termed as conventional systems.
- Different countries in the world use different video standards.

- Conventional TV video formats are:

NTSC:

- ✓ Oldest and very popular TV standard developed by National Television Systems Committee USA.
- ✓ Features: resolution 525*330. motion frequency 30Hz, Amplitude modulation for transmission.

SECAM

- ✓ Swquential Couleur Avec Memoire standard was mainly used in france.
- ✓ Employs frequency modulation
- ✓ Motion freq. 25Hz
- ✓ 625*465

PAL

- ✓ Phase Alternating Line was developed in 1963. It was used in some parts of Europe.
- ✓ Employs amplitude modulation
- ✓ It is similar to NTSC but uses different technique for color carrier (chrominance information)

Enhanced Definition systems

- EDTV are conventional systems modified to offer improved resolution.
- Some of ED systems are:

IDTV

- ✓ Improved-Definition TV was an intermediate technology between NTSC and HDTV
- ✓ It uses almost double scanning lines than NTSC and image is more detailed.
- ✓ Vertical resolution is also enhanced with 60Hz refresh frequency

D2-MAC

- ✓ Duobinary multiplexed analogue components was mainly used in Europe as an intermediate technology between conventional systems and HDTV

High Definition systems

- These are the latest TV technology in use today.
- They are fully digital systems which do not use analog system(as it was in old TVs).
- Most countries have already abandoned analog TV signal transmission systems.
- HDTV mostly use separate set-up box to receive digital TV signal.
- Lately, HDTV have become smart TVs with own operating system and internet connection capabilities.
- Characteristics of HDTV are:

  Resolution: very high resolution (more than 1000*600). Increased luminance bandwidth

  Aspect ratio: 16:9

  Viewing distance: HDTV has highly detailed images. But because of the eye's limited ability to distinguish details, HDTV should be viewed closer than conventional TV.

Computer-based Animation

- *"To animate"* literally means *"to bring it to life"*.
- Animation makes still pictures alive. Animation covers all changes that have a visual effect.
- Computer-based animation is an animation performed by a computer using graphical tools to provide visual effects.
- Computer-based animation is a fascinating part of multimedia systems.
- The basic idea of creating animation is as follows:
  - Create(compose) individual images (frames) in the motion sequence using graphics (animation) tools
  - While composing frames, we first create key frames and then create in-between frames according to our own choice of frame rate.
  - Different options for color and quantization can be used.

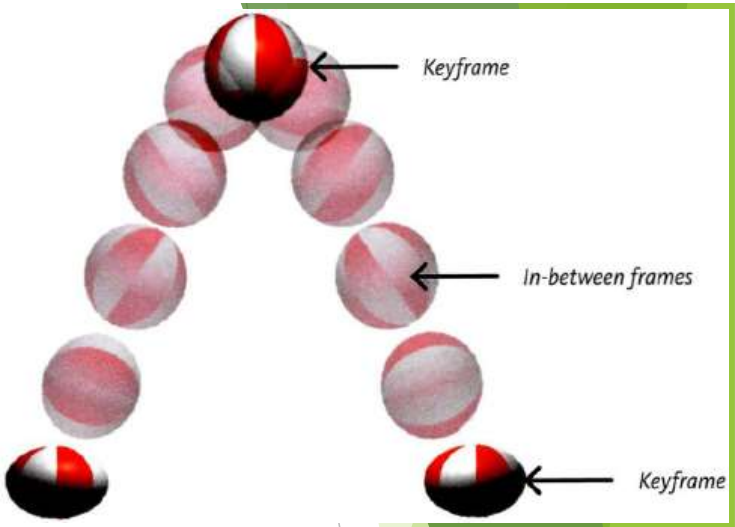
Basic concepts in Animation

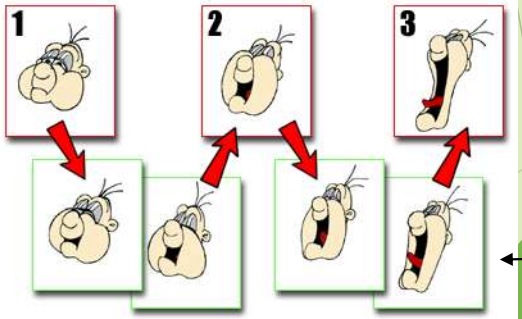
Key frames

- Key frames are main images representing a major change in events.
- The events when entities being animated are at extreme or characteristic positions are candidates for key frames.
- For example: we are creating a scene in which on person is moving his/her hand to pick up a glass. Then key frames will the initial position of person and the position where the hand is touching the glass.

In-between frames

- These are intermediate frames between the key frames.
- Each pair of key frames includes several in-between frames.
- Inbetween frames represent the smooth transition of events from one key frame to adjacent key frame.





Key frames

In-between frames

Stages(processes) in animation

Input process

- First we create and digitize key frames.
- Traditionally, images were first drawn on paper and later shot by video camera.
- But nowadays we can create any image directly in computer using sophisticated graphics tools.

Composition stage

- In this stage, the foreground and background figures are combined to generate individual frames for the final animation.
- Several image composition technique are used for composition.

In-between process

- In order to have a smooth transition from one key frame to another, we need to compose intermediate frames.
- The process of inbetweening can be performed using interpolation technique.
- Several interpolation techniques can be used such as linear interpolation, spline interpolation etc. Each technique has it's own advantage and disadvantage.
- In-betweening also involves interpolating the shapes of an object in the intermediate frames.

Animation Languages

- There are many different languages for animation.
- Many languages are being developed recently and many of the old popular languages have gone out of use.
- We can categorize the animation languages in 3 main categories:
  - **Linear-list notations**: In linear-list notation, each event in the animation is described by a starting and ending frame number and an action that is to take place (event). An example of such language is SCEFO(scene format).

    example: (42,53,B,ROTATE, "PALM", 1,30)

    This notation means "between frames 42 and 53, rotate the object called PALM about axis 1 by 30 degrees, determining the amount of rotation at each frame from table B
  - **General-purpose languages**: Some general purpose languages have embedded animation capability. These languages support vector graphics primitives and their manipulation. Many current high level general purpose language provide graphics libraries to work with graphics and animation.
  - **Graphical languages**: Graphical animation languages describe animation in visual way. i.e. we can create animation by actually working on visual images rather than textual description. MEL is a graphical language used in the software tool Maya.

Methods of controlling Animation

1. **Full explicit control**

   - The animator provides a description of everything that occurs in animation. E.g. the change in position, shape, color of an object in each frame etc..

2. **Procedural control**

   - Based on communication between various objects to determine their properties.

     Physically-based system: the position of one object may influence the motion of another object (e.g. a ball can not pass through wall).

     Actor-based system: individual actors may pass their positions to other actors to affect the other actors' behaviors.

3. **constraint-based system**

   - Changes in the animated object's shape or position is governed by a set of constraints.
   - Sometimes the movements or other changes are based on physical laws.
   - An animation constraint helps automate the animation process. We can use constraints to control an object's position, rotation, or scale through a binding relationship with another object.
   - A constraint requires an animated object and at least one target object. The target imposes specific animation limits on the constrained object.
   - For example, to quickly animate an airplane flying along a predefined path, we can use a path constraint to restrict the airplane's motion.

4. Tracking Live Action

- Track the movements of actual objects.

Two methods:

Traditional method (rootscooping):

- First, a video is made where people and animals act out as the characters of animation.
- The animators edit the video enhancing the background and changing the real actors with animation objects.

Another method:

- Attach a indicator devices ko key points on a person's or animal's body
- Track the positions of the indicators, we can get corresponding points in animated model.
- E.g. data glove works with this technique.

5. Kinematics and dynamics:

- Kinematics refers to the position and velocity of points.
- Example of kinematic description: "The cube is at origin at time t = 0. Then, it moves with a constant acceleration in the direction (1,1,6)"
- Dynamics takes into account the physical laws that govern kinematics (e.g. newton's laws of motion or Euler-Lagrange equation for fluids etc).
- Example: "At time t , the cube of mass 100 gm is at position (0,100,0). Then the force of gravity acts on it".

  The result of the dynamic simulation of such a model is that the cube falls.

Display of animation:

- To display animation with raster system, animated objects (which may consists of graphical primitives such as lines, polygons, etc..) must be scan-converted into the pixmap in the frame buffer.
- Since animation display has temporal aspect, there is a strict timing requirement for the scan-conversion process as follows:
  - The scan-conversion must be done at least 10 (15-20 is better) times per second to give reasonably smooth effect.
  - It means a new image must be ready in no more than 100 ms, which includes erasing old image, conversion of new image, and display of new image.
  - e.g., if scan-converting of an object takes 75 ms, then we have only 25 ms to erase old image and display new image on screen. This 25 ms time is not enough. It gives *distraction effect*.

Double buffering

- Double buffering is a technique to avoid distracting effect (flicker) in animation display during clearance of old frame.
- Main idea: is to create a virtual screen out of the user's view.
  - divide the frame buffer into two images, one half of the buffer is used as virtual screen to perform the clearing and painting of the animation frame

    (Data in one buffer are being processed while the next set of data is read into the other one)

Advantage:

With double buffering, even if the scan-conversion is slow, the transition from one image to next image appears instantaneous.

This avoids flicker between image transition.

Disadvantage: More memory and CPU consumption.

Transmission of animation

- Since animated objects can be represented symbolically using graphics primitives or pixmap, we can transmit animation using two approaches:

1. Transmission with pixmap representation

- Transfer the animation in pixmap(raster form with pixel representation) form.

  Bandwidth required = (no of pixels*bits per pixel)* frames per second

- Requires high bandwidth
- But display time is shorter because no scan-conversion is needed at the receiver side.

2. Transmission with Symbolic representation

- In this method, the symbolic representation (e.g. circle) of the animated objects (ball) is transmitted together with control commands (e.g. roll the ball).
- In the receiver side, the symbolic representation is converted into pixmap for display
- In this approach, the transmission time is short because it requires less bandwidth.
- But it takes longer display time because of the scan-conversion process on the receiver side.

  Bandwidth required = *The size of symbolic representation structure and operation commands* Number of animated objects and commands per second*

End of unit 4

Thank you!