Chapter 4

Video and Animation

4.1 Basic Concept

Video is the technology of electrically capturing, recording, processing, storing, transmitting, and reconstructing is a sequence of still images representing series in motion. Visual representation is shown an idea or image that is presented in a particular way to have its meaning or symbolism.

Digital video has supplanted analog video as the method of choice for making video for multimedia use. While broadcast stations and professional production and postproduction houses remain greatly invested in analog video hardware, digital video gear produces excellent finished products at a fraction of the cost of analog. A digital camcorder directly connected to a computer workstation eliminates the image-degrading analog-todigital conversion step typically performed by expensive video capture cards, and brings the power of nonlinear video editing and production to everyday users.

Video Signal Representation

Video signal representation consists of three aspects:

Visual Representation:

The main objective of visual representation is to offer the viewer a sense of presence in the scene and of participation in the events portrayed.

Transmission:

Video signals are transmitted to the receiver through a single television channel.

Digitization:

Digitization is the process of analog to digital conversion, sampling of gray (color) level, quantization.

Visual Representation

The main objective of visual representation is to offer the viewer a sense of presence in the scene and of participation in the events portrayed. To meet the main objective, the televised image should convey the spatial and temporal content of the scene. Importance measures are:

1. *Vertical detail and viewing distance:*

   The geometry of the field occupied by the television image is based on the ratio of the picture width W to height H. It is called aspect ratio.

   Aspect ratio: ratio of picture width and height (4/3 = 1.33 is the conventional aspect ratio)

   Viewing angle = Viewing distance/Picture height (D/H)

2. *Horizontal detail and picture width:*

   Picture width (Conventional TV service) = 4/3 x picture height

3. *Total detail content of the image:*

   Number of pixels presented separately in the picture height = vertical resolution

   Number of pixels in the picture width = vertical resolution x aspect ratio

   The product of the number of elements vertically and horizontally equals the total number of picture elements in the image.

4. *Perception of depth:*

   In natural vision, this is determined by angular separation of images received by the two eyes of the viewer.

   In the flat image of TV, focal length of lenses and changes in depth of focus in a camera influence depth perception.

5. *Luminance and chrominance:*

   Color-vision - achieved through 3 signals, proportional to the relative intensities of RED, GREEN and BLUE.

   Color encoding during transmission uses one LUMINANCE and two CHROMINANCE signals

6. *Temporal aspect of resolution:*

   Motion resolution is a rapid succession of slightly different frames. For visual reality, repetition rate must be high enough (a) to guarantee smooth motion and (b) persistence of vision extends over interval between flashes (light cutoff b/w frames).

7. *Continuity of motion:*

   Motion continuity is achieved at a minimal 15 frames per second; is good at 30 frames/sec; some technologies allow 60 frames/sec.

   NTSC standard provides 30 frames/sec - 29.97 Hz repetition rate.

   PAL standard provides 25 frames/sec with 25Hz repetition rate.

8. *Flicker effect:*

   Flicker effect is a periodic fluctuation of brightness perception. To avoid this effect, we need 50 refresh cycles/sec. Display devices have a display refresh buffer for this.

9. *Temporal aspect of video bandwidth:*

   Temporal aspect of video bandwidth depends on rate of the visual system to scan pixels and on human eye scanning capabilities.

Transmission

Video signals are transmitted to receivers through a single television channel. To encode color, a video signal is a composite of three signals. For transmission purposes, a video signal consists of one luminance and two chrominance signals.

- Video bandwidth is computed as follows:
  - 700/2 pixels per line X 525 lines per picture X 30 pictures per second.
  - Visible number of lines is 480.
- Intermediate delay between frames is
  - 1000ms/30fps = 33.3ms
- Display time per line is

- 33.3ms/525 lines = 63.4 microseconds

❖ The transmitted signal is a composite signal
  - Consists of 4.2Mhz for the basic signal and 5Mhz for the color, intensity and synchronization information.

Color Encoding:

❖ A camera creates three signals
  - RGB (red, green and blue)

❖ For transmission of the visual signal, we use three signals: 1 luminance (brightness-basic signal) and 2 chrominances (color signals).
  - In NTSC, luminance and chrominance are interleaved
  - Goal at receiver
    - separate luminance from chrominance components
    - avoid interference between them prior to recovery of primary color signals for display.

RGB signal

- for separate signal coding
- consists of 3 separate signals for red, green and blue colors. Other colors are coded as a combination of primary color. (R+G+B = 1) --> neutral white color.

YUV signal

- separate brightness (luminance) component Y and color information (2 chrominance signals U and V)
  - Y = 0.3R + 0.59G + 0.11B
  - U = (B-Y) * 0.493
  - V = (R-Y) * 0.877

Resolution of the luminance component is more important than U, V

Coding ratio of Y, U, V is 4:2:2

YIQ signal

❖ similar to YUV - used by NTSC format
  - Y = 0.3R + 0.59G + 0.11B
  - U = 0.60R - 0.28G + 0.32 B

- V = 0.21R -0.52g + 0.31B

Composite signal

- All information is composed into one signal
- To decode, need modulation methods for eliminating interference b/w luminance and chrominance components.

Digitization

Before a picture or motion video can be processed by a computer or transmitted over a computer network, it need to be converted from analog to digital representation. Digitization is the representation of an object, image, sound, document or a signal (usually analog signal) by a discrete set of its points or samples.

Digitization = Sampling + Quantization

Sampling is the reduction of a continuous signal to a discrete signal.

- Refers to sampling the gray/color level in the picture at MXN array of points.
- Once points are sampled, they are quantized into pixels sampled value is mapped into an integer
  - quantization level is dependent on number of bits used to represent resulting integer, e.g. 8 bits per pixel or 24 bits per pixel.
  - Need to create motion when digitizing video digitize pictures in time obtain sequence of digital images per second to approximate analog motion video.

Computer Video Format

The computer video format depends on the input and output devices for the motion video medium.

Current video digitizers differ in digital image (frame) resolution, quantization and frame rate (frames/s).

The output of the digitalized motion video depends on the display device. The most often used displays are raster displays, which store display primitives in a refresh buffer in terms of their component pixels.

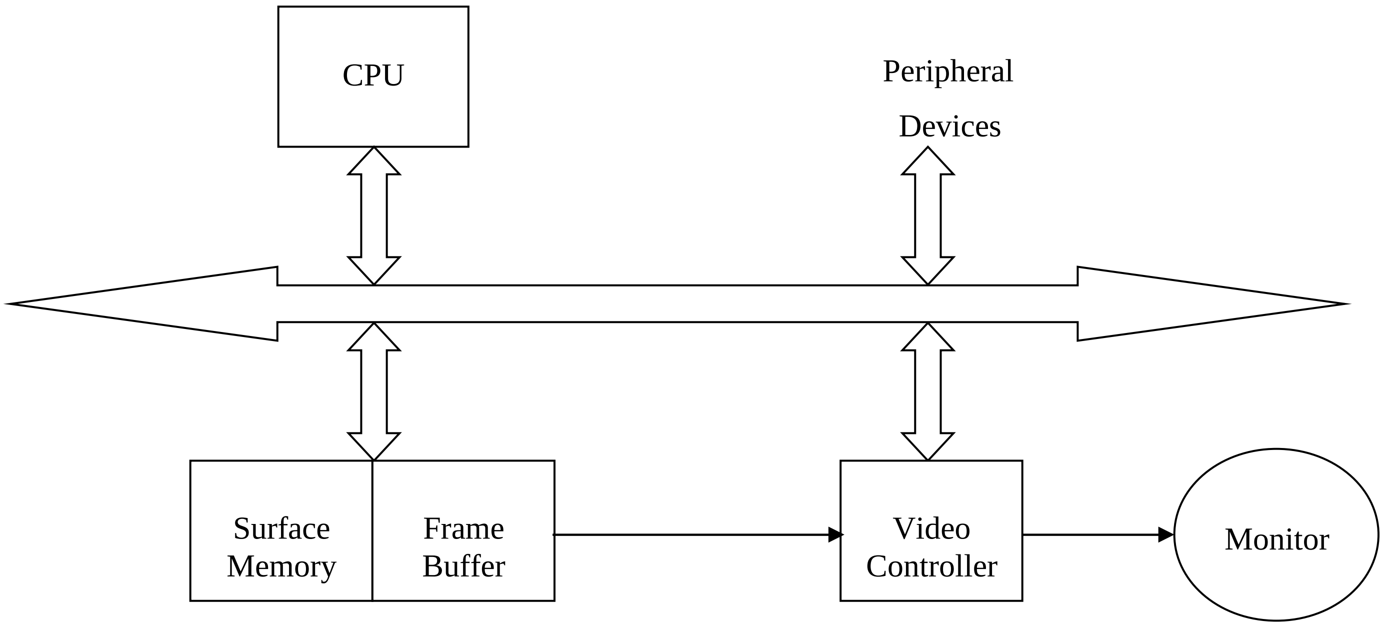

Figure 4.1: A common raster display system architecture

The video controller displays the image stored in the frame buffer, accessing the memory through a separate access port as often as the raster scan rate dictates. The constant refresh of the display is its most important task. Because of the disturbing flicker effect, the video controller cycles through the frame buffer, one scan line at a time, typically 60 times/second.

Some computer video controller standards are given here an example. Each of these systems supports different resolution and color presentation.

The Color Graphics Adapter (CGA):

The CGA has a resolution of 320x200 pixels with simultaneous presentation of four colors. Therefore, the storage capacity per image is:

$$300 \times 200\ pixels \times \frac{2\,\frac{bits}{pixels}}{8\,\frac{bits}{byte}}$$

$$= 16{,}000\ bytes$$

The Enhanced Graphics Adapter (EGA):

The EGA has a resolution of 640×350 pixels with 16-color presentation. Therefore, the storage capacity per image is:

$$640 \times 350\ pixels \times \frac{4\,\frac{bits}{pixels}}{8\,\frac{bits}{byte}}$$
$$= 112{,}000\ bytes$$

The Video Graphics Adapter (VGA):

The EGA has a resolution of 640×480 pixels with 16-color presentation. In this case, 256 colors can be displayed simultaneously. The monitor is controlled through an RGB output. The storage capacity per image is:

$$640 \times 480\ pixels \times \frac{8\,\frac{bits}{pixels}}{8\,\frac{bits}{byte}}$$
$$= 307{,}200\ bytes$$

The 8514/ A Display Adapter Mode:

A Display Adapter Mode can present 256 colors with a resolution of 1024×768 pixels. The storage capacity per image is:

$$1024 \times 768\ pixels \times \frac{8\,\frac{bits}{pixels}}{8\,\frac{bits}{byte}}$$
$$= 786{,}432\ bytes$$

The Extended Graphics Array (XGA):

The XGA supports a resolution of 640×480 pixels and 65,000 different colors. With the resolution of 1024×768 pixels, 256 colors can be presented. In this case, we have the same storage capacity per image as the 8514/A adapter.

The Super VGA (SVGA):

The SVGA offers resolutions up to 1024×768 pixels and color formats up to 24 bits per pixel. The storage capacity per image is:

$$1024 \times 768\ pixels \times \frac{24\,\frac{bits}{pixels}}{8\,\frac{bits}{byte}}$$

$$= 2{,}359{,}296\ bytes$$

Television

Television is the most important application that has driven the development of motion video. Television is a telecommunication medium for transmitting and receiving moving images that can be monochrome (black and white) or colored, with or without accompanying sound. Television may also refer specifically to a television set, television programming or television transmission.

Conventional Systems

Conventional system used in black and white and color television. Conventional television systems employ the following standards:

NTSC (National Television Systems Committee)

- NTSC developed in U.S., is the oldest and most widely used television standard.
- The color carrier is used with approximately 4.429 MHZ or with approximately 3.57 MHZ.
- NTSC uses a quadrature amplitude modulation with a suppressed color carrier and work with a motion frequency of approximately 30 Hz.
- 4×3 Aspect ratio.
- 525 lines
- 30 frames per second.
- Scanned in fields.

PAL (Phase Alternating Line)

- Invented by W.Bruch (Telefunken) in 1963.
- It is used in parts of Western Europe.
- The basic principle of PAL is a quadrature amplitude modulation similar to NTSC, but the color carrier is not suppressed.

- PAL is an analogue television color encoding system used in broadcast television systems in many countries.
- 4×3 Aspect ratio.
- 625 lines
- 25 frames per second.
- Scanned in fields.
- There are slight variations: PAL-B, PAL-G, PAL-H and PAL-N.
- Used in continental Europe and parts of Africa, Middle East and South America.
- More Lines = Better Resolution
- Fewer Frame/fields = More Flicker

SECAM (Sequential Color and Memory)

- SECAM is a standard used in France and Eastern Europe.
- In contrast to NTSC and PAL, it is based on frequency modulation.
- It uses a motion frequency of 25 Hz and each picture has 625 lines.
- SECAM is an analog color television system first used in france.

Enhanced Systems

Enhanced Definition Television Systems (EDTV) are conventional systems modified to offer improved vertical and/or horizontal resolution. EDTV are an intermediate solution, to digital interactive television system and their coming standards.

HDTV (High-Definition Television)

- The next generation of TV is known as HDTV.
- HDTV is a digital system.
- 16:9 Aspect ratio.
- Permits several levels of picture resolution similar to that of High-Quality Computer Monitors, with 720 or 1080 line (1280×720 pixels or 1920×1080 pixels).
- Range from 24 to 60 frame per second, progressive or interlaced scan.

- Uses MPEG-2 compression to squeeze a 19 Megabit per second data flow so that it can be accommodated by a standard broadcast TV channel of 6 MHz bandwidth.

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 simplest possibility for digitizing video signal is to sample the composite analog video signal. Here, all signal components are converted together into a digital representation. The composite coding of the color television signal depends on the television standard. Using component coding, the sampling frequency is not coupled with the color carrier frequency.

Component Coding:

The principle of component coding consists of separate digitization of various image components or planes; for example, coding of luminance and color difference (chrominance) signals. These digital signals can be transmitted together using multiplexing. The luminance signal is sampled with 13.5 MHz as it is more important than the chrominance signal.

4.3 Basic Concepts of Animation

Animation is the rapid display of a sequence of images of 2-D or 3-D artwork or model positions in order to create an illusion of movement.

It is an optical illusion of motion due to the phenomenon of persistence of vision, and can be created and demonstrated in a number of ways.

The most common method of presenting animation is as a motion picture or video program, although several other forms of presenting animation also exist.

Animation is anything that moves on your screen like a cartoon character. It is the visual art of creating the illusion of motion through the successive display of still images with slightly perceptible changes in positioning of images. Animation is the illusion of movement.

Animating = making something appear to move that doesn’t move itself

Animation = a motion picture made from a series of drawings simulating motion by means of slight progressive changes in the drawings

The result of animation is a series of still images assembled together in time to give the appearance of motion

Animation is the art of movement expressed with images that are not taken directly from reality. In animation, the illusion of movement is achieved by rapidly displaying many still images or frames in sequence.

4.4 Types & Techniques of Animation

When you create an animation, organize its execution into a series of logical steps. First, gather up in your mind all the activities you wish to provide in the animation; if it is complicated, you may wish to create a written script with a list of activities and required objects. Choose the animation tool best suited for the job. Then build and tweak your sequences; experiment with lighting effects. Allow plenty of time for this phase when you are experimenting and testing. Finally, post-process your animation, doing any special rendering and adding sound effects.

Cel Animation

The term cel derives from the clear celluloid sheets that were used for drawing each frame, which have been replaced today by acetate or plastic. Cels of famous animated cartoons have become sought-after, suitable-for-framing collector's items. Cel animation artwork begins with key frames (the first and last frame of an action). For example, when an animated figure of a man walks across the screen, the balances the weight of his entire body on one foot and then the other in a series of falls and recoveries, with the opposite foot and leg catching up to support the body.

The animation techniques made famous by Disney use a series of progressively different on each frame of movie film which plays at 24 frames per second. A minute of animation may thus require as many as 1,440 separate frames. The term cel derives from the clear celluloid sheets that were used for drawing each frame, which is been replaced today by acetate or plastic. Cel animation artwork begins with key frames.

Computer Animation

Computer animation programs typically employ the same logic and procedural concepts as cel animation, using layer, key frame, and tweening techniques, and even borrowing from the vocabulary of classic animators. On the computer, paint is most often filled or drawn with tools using features such as gradients and antialiasing. The word links, in computer animation terminology, usually means special methods for computing RGB pixel values, providing edge detection, and layering so that images can blend or otherwise mix their colors to produce special transparencies, inversions, and effects. Computer Animation is same as that of the logic and procedural concepts as cel animation and use the vocabulary of classic cel animation– terms such as layer, Key frame, and tweening. The primary difference between the animation software program is in how much must be drawn by the animator and how much is automatically generated by the software In 2D animation the animator creates an object and describes a path for the object to follow. The software takes over, actually creating the animation on the fly as the program is being viewed by your user. In 3D animation the animator puts his effort in creating the models of individual and designing the characteristic of their shapes and surfaces. Paint is most often filled or drawn with tools using features such as gradients and anti- aliasing.

Kinematics

It is the study of the movement and motion of structures that have joints, such as a walking man. Inverse Kinematics is in high-end 3D programs, it is the process by which you link objects such as hands to arms and define their relationships and limits. Once those relationships are set you can drag these parts around and let the computer calculate the result.

Morphing

Morphing is popular effect in which one image transforms into another. Morphing application and other modeling tools that offer this effect can perform transition not only between still images but often between moving images as well the morphed images were built at a rate of 8 frames per second, with each transition taking a total of 4 seconds.

4.5 Principles of Animation

Animation is possible because of a a biological phenomenon known as *persistence of vision a*nd the psychological phenomenon called *phi.*

An object seen by the human eye remains chemically mapped on the eye's retina for a brief time after viewing. This makes it possible for a series of images that are changed very slightly and very rapidly, one after the other, seem like continuous motion.

1. *Squash and Stretch*: - Defining the rigidity & mass of an object by distorting its shape during an action.
2. *Timing:* - Spacing actions to define the weight & size of objects & the personality of characters.
3. *Anticipation:* - The preparation for an action.
4. *Staging:* - Presenting an idea so that it is unmistakably clear.
5. *Follow Through & Overlapping Action:* - The termination of an action & establishing its relationship to the next action.
6. *Straight Ahead Action & Pose-To-Pose Action:* - The two contrasting approaches to the creation of movement.
7. *Slow In and Out:* - The spacing of in-between frames to achieve subtlety of timing & movements.
8. *Arcs:* - The visual path of action for natural movement.
9. *Exaggeration:* - Accentuating the essence of an idea via the design & the action.
10. *Secondary Action:* - The Action of an object resulting from another action
11. *Appeal:* - Creating a design or an action that the audience enjoys watching.
12. *Solid Drawing:* - Knowing them can dramatically improve one's ability to create good, strong poses and compose them with well-crafted environments.

4.6 Animation Languages:

There are many different languages for describing animation, and new ones are constantly being developed. They fall into three categories:

(i) Linear-list Notations:

In linear-list notations for animation 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, so a statement such as 42, 53, B, ROTATE "PALM", 1, 30 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 tabled B".

(ii) General-purpose Language:

The values of a variables in the language can be used as parameters to the routines, which perform the animation. ASAS is an example of such a language. It is built on top of LISP, and its primitive entities include vectors, colors, polygons, solids, groups, points of view, subworlds and lights. ASAS also includes a wide range of geometric transformations that operate on objects. The ASAS program fragment below describes an animated sequence in which an object called my-cube is spun while the camera pans.

(**grasp** my-cube); The cube becomes the current object
(**cw** 0.05); Spin it clockwise by a small amount
(**grasp** camera); Make the camera the current object
(right panning-speed); Move it to the right

(iii) Graphical Languages:

Graphical animation languages describe animation in a more visual way. These languages are used for expressing, editing and comprehending the simultaneous changes taking place in an animation. The principal notion in such languages is substitution of a visual paradigm for a textual one. Rather than explicitly writing out descriptions of actions, the animator provides a picture of the action. Example of such systems and languages are *GENESYS*, *DIAL* and S-Dynamics System.

4.7 Method of Controlling Animation

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

Full Explicit Control

Explicit control is the simplest type of animation control. The 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.

Procedural Control

Procedural control is based on communication between various objects to determine their properties.

Physical-based systems, the position of one object may influence the motion of another. Action-based systems, the individual actors may pass their position to other actors to affect the other actors' behaviors.

Constraint-based Systems

Support a hierarchy of constraints and to provide motion where constraints are specified by the dynamics of physical bodies and structural characteristics of materials is a subject of active research.

Tracking Live Action

Trajectories of objects in the course of an animation can also be generated by tracking live action. Traditional animation uses rotoscoping. Tracking live-action technique is to attach some sort of indicator to key points on a person's body.

4.8 Transmission of Animation

Transmission over computer networks may be performed using one of two approaches:

The symbolic representation

The symbolic representation of animation objects is transmitted together with the operation commands performed on the object, and at the receiver side the animation is displayed. The transmission time is short because animated object is smaller in byte size than its pixmap representation, but the display time at the receiver takes longer because the scan-converting operation has to be performed at the receiver side.

<u>The pixmap representation</u>

The pixmap representation of the animated objects is transmitted and displayed on the receiver side. The transmission time is longer than the symbolic representation but the display time is shorter because the scan-conversion of the animated objects is avoided at the receiver side.