Parag Havaldar and Gérard Medioni

Multimedia Systems

Algorithms, Standards, and Industry Practices

Chapter 3

Media Representation and Formats

Mohd Saufy Rohmad

EE UiTM

Media Representation and Formats

- The digital media forms need to be represented and stored so that they can be viewed, exchanged,edited and transmitted in a standard manner.
- We will deal with representation of digital media and commonly used media format.
- Evolution of analog media format also and how they still influence the current digital format also discuss.

1. Digital Images

- Images may be photographs images,gray scale or color, use with text in documents and fax.
- Image is the basic element of video

Figure 3.1 Upper left is image with single channel (8 bit per pixel). Right image is 3 channel image. (24 bit per pixel). Bottom image is large mosaic created by combining different image.

Digital Representation of Images

- All images are represented digitally as pixels.
- Image is defined by image width,height and pixel depths.
- Image width is the number of pixel that span the image horizontally and image height gives the number of lines in the image.
- Pixel depth is the representation of each pixel.
- The number of pixel is depend on the color space representation (gray or color) and segregated into channel.
- Sometimes a additional fourth channel called alpha channel is used. (add 8 bits for alpha – make total 32 bits)
- For 640 x 480 image with R,G,B color representations, the size of color image = 640 x 480 x 3 x 8 =7.37Mbits (921.6Kbytes)
- If this were a gray image, its size would be 640 x 480 x 8 = 2.45Mbits (307.2Kbytes)

Digital Representation of Images..

- Printing technology prefer to print halftone printing process where number of colors used is minimized to lower printing costs.
- Halftone printing creates ranges of grays or colors using variables sized dots. Measure in dots per inch.
- The dots control how much ink is deposited at a specific location while printing.
- For a process color image,four halftone channels are used: cyan,magenta,yellow and black.

Digital Representation of Images...

Figure 3.2 Halftone printed image with different number of dots

Digital Representation of Images....

- Image with alpha channel is for the transparenc y for that pixel value and is used in image compositing application
- The 8-bit alpha channel value for each pixel associates the degree of importance of each pixel in a compositing operation
- A value 0 indicates that the pixel will not be composite, a high value of 255 indicates that the pixel will be entirely composited.

Alpha Channel

- Image with alpha channel is for the transparency for that pixel value and is used in image compositing application
- The 8-bit alpha channel value for each pixel associates the degree of importance of each pixel in a compositing operation
- A value 0 indicates that the pixel will not be composite, a high value of 255 indicates that the pixel will be entirely composited.

Alpha Channel..

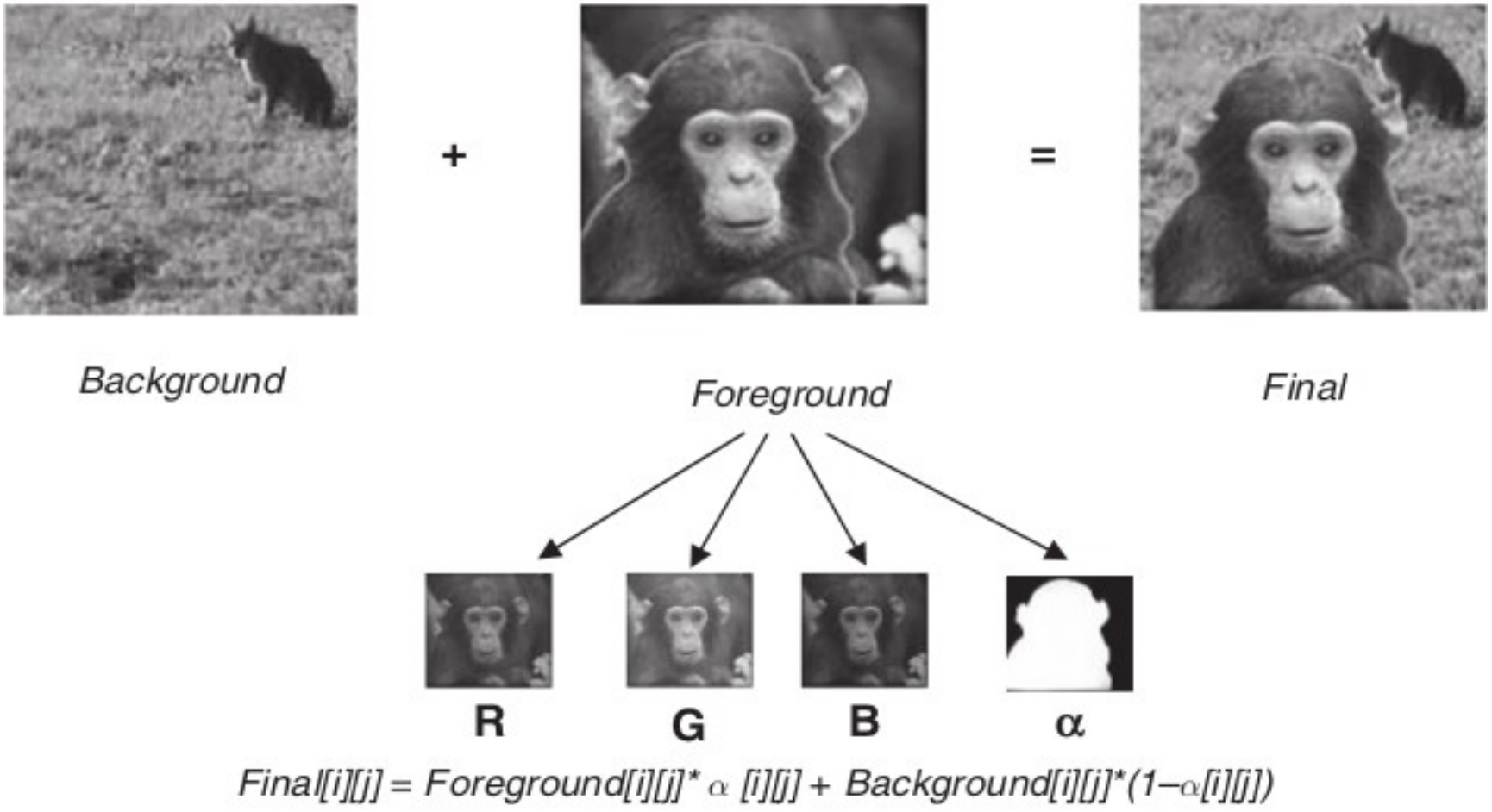


Figure 3.3 The alpha channel is used to composite the pixels on the foreground with the background image, producing final image on the far right.

Image Aspect Ratios

- Refers to the width/height ratio of the images and play important role in standard.
- Commonly use aspect ratio for images are 3:2 (when developing and printing photographs), 4:3(television images), 16:9(high-definition images) and 47:20(anomorphic formats used in cinemas).
- We can change image aspect ratios with digital image processing software.
- Changing image aspect ratio also change the pixel size, known as pixel aspect ratio (PAR) or sample aspect ratio (SAR).
- Change from 4:3 to 16:9 cause the content to stretch horizontally. It makes the PAR change from 1:1 to 1.333:1
- If anamorphic is remapped and view on standard definition television, the PAR is changed to 1:0.75.

Digital Image Formats

- Uncompress and raw image format is bitmap. (.bmp). Image is represented in bitmaps and stored in a binary file with no compression.
- Raw format is used when we dont want any artifact in our file and when we want to perform specific customized operation on images such as convolution,filtering,dithering and so on.
- Commonly use compress formats are .jpg, .gif and .png.
- Raw and compress format are a result of a raster representation of images.
- Raster images are images stored as row of pixels.

Digital Image Formats..

- Beside raster image, vector image are also used. Vector graphic files store image data in terms of geometric objects.
- The objects are specified by parameters such as line styles, side lengths, radius, color, gradients, and so on, store in binary or text files.
- Some file formats combine vector graphics with bitmap images, refer as metafiles.
- Example of metafiles is computer graphics metafile (CGM).

Vector Image File

```
Object myRectangle
LineType Dotted
LineWidth 4
LineColor  0 0 0
FillColor 255 0 0
Rectangle (100,200) (200,220)
EndObject
Object myCircle
LineColor 0, 0, 0
FillColor 0 0 255
Circle (200, 200), 50
EndObject
```

Figure 3.4 Vector file description showing two objects – a red rectangle and a blue circle

Digital Image File Formats

File suffix	File name	File type	Features
.bmp	Windows bitmap	Uncompressed raster	Represents from 1 to 24 bits per pixel. Normally uncompressed but can use lossless run length encoding (RLE)
.pcx	Windows Paintbrush	Uncompressed/ compressed raster	Used only on Microsoft Windows platforms. Has similar features to .bmp.
.gif	Graphics Interchange Format	Compressed raster	Predominantly used on the Web. Allows 256 indexed colors and simple animations. Alpha channel supported. Uses LZW compression Proprietary to CompuServe

Figure 3.5 Digital Image File formats with their main features

Digital Image File Formats..

.jpg, .jpeg	Joint Photographic Experts Group	Compressed raster	For continuous tone pictures (photographs). Lossy and lossless compression supported. No alpha channel supported. Level of compression can be specified. Commonly used on the Web
.png	Portable Network Graphics	Compressed raster	Allows 1–48 bits of color. Supports alpha channel. Designed to replace proprietary .gif files. File format approved by W3C
.psd	Adobe Photoshop	Uncompressed layered raster	Used for image editing. Supports a variety of color models. Supports varying pixel bit depths Image can be organized into layers. Commonly used processing file format.
.psp	Paint Shop Pro	Uncompressed layered raster	Similar to .psd
.tif, .tiff	Tagged Image File Format	Uncompressed raster, also compressed raster	Used in traditional print graphics. Can be compressed using lossless and lossy methods of compression, including RLE, JPEG, and LZW. TIFF comes in many flavors
.fh	Macromedia Freehand	Compressed vector format	Proprietary to Macromedia, used by Flash Players. Supports animation
.cdr	CorelDRAW	Uncompressed vector format	Proprietary to Corel
.swf	Macromedia Shockwave Flash format	Uncompressed vector format	Proprietary format created by Macromedia (now Adobe). Contains vector representations and animations that can be put on the Web.
.dxf	AutoCAD ASCII Drawing Interchange Format	Uncompressed vector format	ASCII text stores vector data. Used for 2D/3D graphical images.
.ps or .eps	Postscript, or Encapsulated Postscript	Uncompressed metafile	Supports text, fonts, vectors, and images.
.ai	Adobe Illustrator	Metafile format	Proprietary format. Similar to .eps.
.pdf (portable document format)	Adobe PDF document	Compressed metafile	Supports text, fonts, and images. Commonly used document format. Supports hyperlinks. Supports authorized access.
.pict	Macintosh Quickdraw	Compressed metafile	Used predominantly on Macintosh platforms. Can use RLE or JPEG compression. Supports grayscale, RGB, CMYK, or indexed color.

Figure 3.5 Digital Image File formats with their main features

Digital Video

- Digital Video is a vehicle of most of our art, information and entertainment today.
- Video is represented by a sequence of discrete images shown in quick succession.
- Each image in the video called a frame.
- The image attributed remain constant for all the images in the length of the video.
- Two important properties of video are : frame rate and scanning format.

Digital Video..

- Film is displayed at 24 frames per second.
- Television standards use 30 frames per second (NTSC) or 25 frames per second (PAL).
- If frame rate is too slow, the human eye perceives an unevenness of motion called flickering.
- Although digital video can be considered a three dimensional signal – a 2D image changing over time – analog video is converted to a 1D signal of scan lines.

Digital Video...

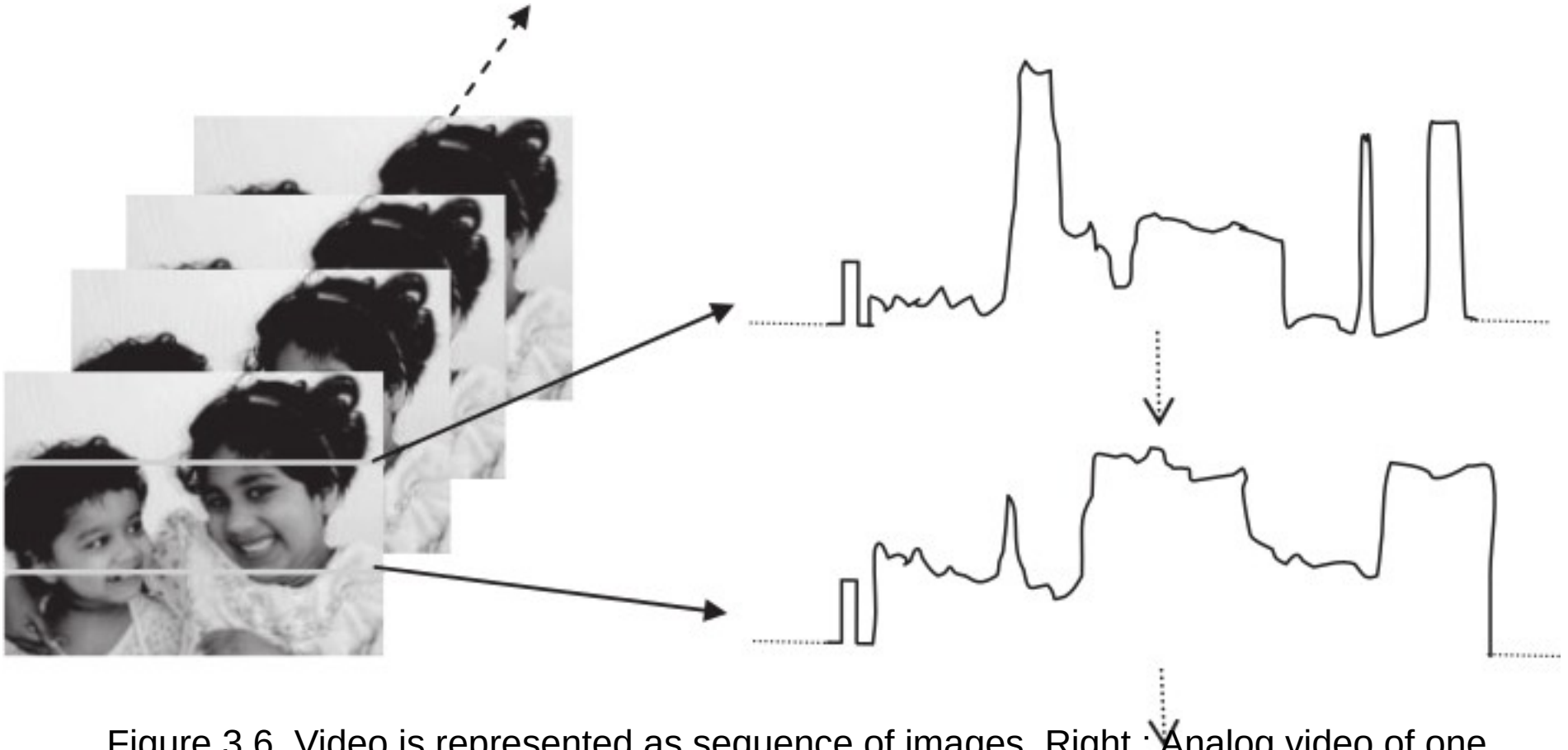


Figure 3.6 Video is represented as sequence of images. Right : Analog video of one frame scanned as a 1D signal. Each signal line is scanned from left to right as an analog signal separated by horizontal sync. Two scan lines are shown, each begins with a horizontal sync and traces through the intensity variation on that scan line.

Digital Video...

- The scan lines conversion introduced to make analog television broadcast technology work and is central to the television display.
- Electron gun in the television projects electrons on the phospher screen.
- The phospher screen glows at each location on a scan line creating a color at all position on the line.
- Scanning have a interlaced or progressive format.
- Digital video display on these devices, such as LCD or plasma, does not require the scanning mechanism.

Analog Television and Television

- Television requires this analog scanned information to be broadcast from a broadcast station to all users.
- There are need to standardized the requirement of analog TV, which is YUV color space and interlaced scanning.
- Analog display will gradually transition into digital display devices.

Analog Television and Television..

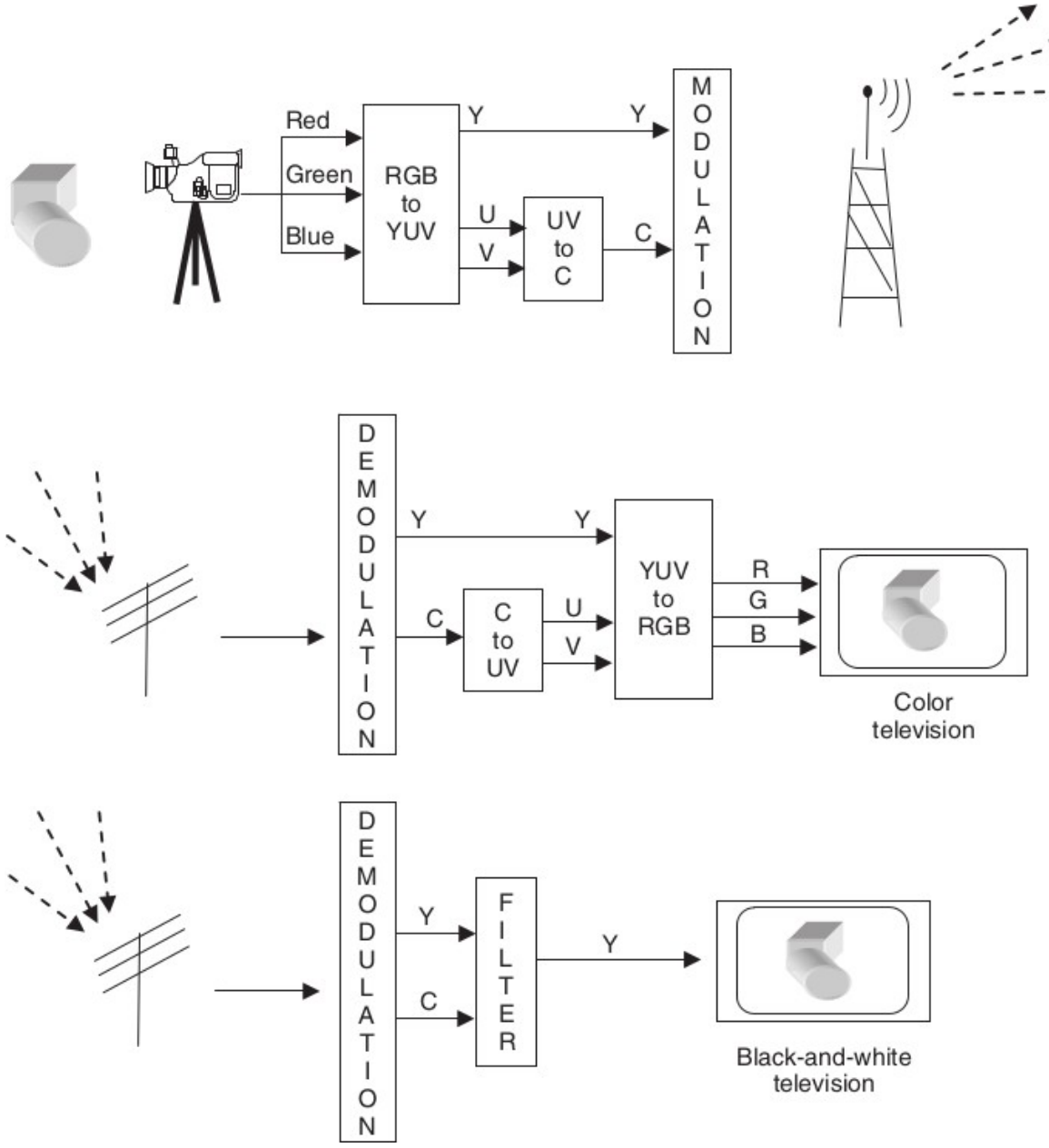


Figure 3.7 Television works by sending scan line information in interlaced YUV format.

Conversion to YUV

- Video frames, like images, are represented using a color format, which is RGB.
- RGB is use by TV to render video signal.
- For transmission,RGB signal is transformed into a YUV signal.
- YUV signal decouple the intensity information (Y or luminance) from the color information (UV or chrominance).
- Human is less sensitive to color distortion than intensity distortion.
- In order words, reducing the color resolution does not affect our perception.

Analog Video Scanning

- Analog line by line raster signal has to be rendered on your TV in a corresponding manner.
- Every 60hz (1/60 second), the electron gun is reset by the vertical sync to draw the beginning of new frame. Each frame is broken into odd and even field.
- Electron gun first draws the odd lines of the screen image, followed by even lines.

Analog Video Scanning..

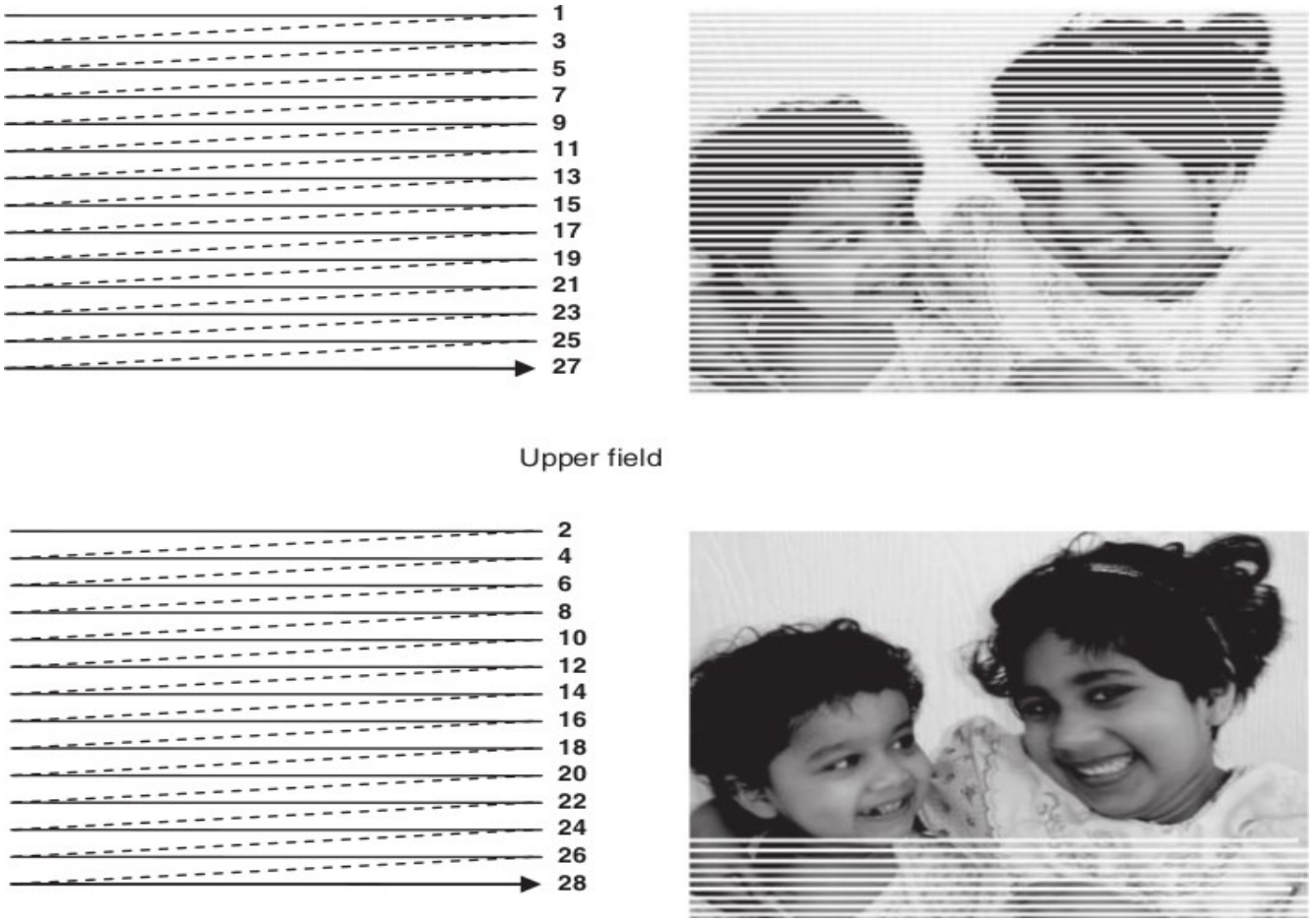


Figure 3.8 Interlaced scanning. Top figure show the odd field of odd numbered lines. The bottom shows a lower even field.

Analog Video Scanning..

- Interlaced scanning technique might be unacceptable and has flickering and artifacts.
- Video is better when it is captured and drawn progressively, which eliminates flickering.

Figure 3.9 Progressive scan resulting in 30 frames per second. All the scan lines are drawn in succession

Types of Video Signals

- For TV broadcasting, color and luminance information are combined into one signal called composite video.
- Other type of better quality video signal are S-Video and component video (used in DVD format).
- Basic idea of differences between type of signal is either YUV signals is send collectively or separately.

Composite Video

- Also called baseband video or RCA video.
- It contains both chrominance (color) and luminance (brightness) information.
- This is older type of video which used NTSC format.
- Interference between the chrominance and luminance is inevitable when the signal is weak.

S-Video

- Also called as super video or Y/C video, is a video signal transmission in which the luminance signal and chrominance signal are transmitted separately for superior picture clarity.
- Here the C (chrominance) data combine U and V signals and at the display time,the signal C can be separated into U and V signal.
- Y and C that separated to reduce problem caused by interference between luminance and chrominance signal.

YUV Subsampling Schemes

- Video signal is process and rendered in RGB but transmit as YUV.
- The color information (UV) will be further subsampled to gain more bandwidth.
- Human still can accept the reduce color quality because human are less sensitive to that.
- In analog, sub sampling is achieved by allocating half as much bandwidth to chrominance as to luminance.
- In digital, subsampling can be done by reducing the number of bits used for color channel.

YUV Subsampling Schemes..

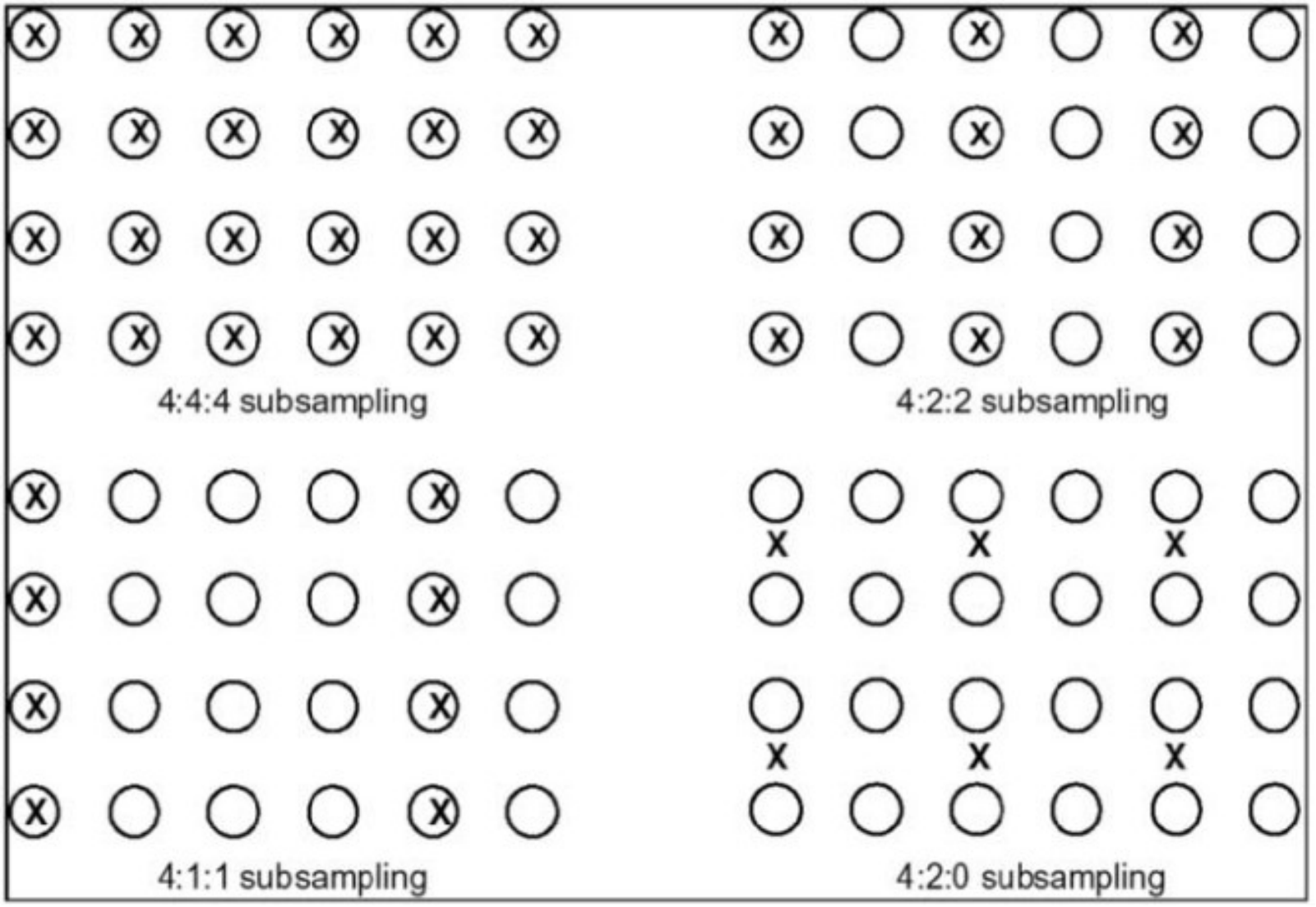


Figure 3.10 Subsampling Ratio between Y:U:V for Analog Video

Analog Video Formats

Property	NTSC	PAL	SECAM
Frame rate	30	25	25
Number of scan lines	525	625	625
Number of active lines	480	576	576
Aspect ratio	4:3	4:3	4:3
Color model	YIQ	YUV	YDbDr
Primary area of usage	North America (USA and Canada), Japan	Asia	France and Russia

Figure 3.11 Table Illustrating Analog Video Formats

Digital Video Formats

Format name	Lines per frame	Pixels per line	Frames per second	Support for interlaced format	Subsampling scheme	Image aspect ratio
CIF	288	352		N	4:2:0	4:3
QCIF	144	176		N	4:2:0	4:3
SQCIF	96	128		N	4:2:0	4:3
4CIF	576	704		N	4:2:0	4:3
SIF-525	240	352	30	N	4:2:0	4:3
SIF-625	288	352	25	N	4:2:0	4:3
CCIR 601 NTSC (DV, DVB, DTV)	480	720	29.97	Y	4:2:2	4:3
CCIR 601 PAL/SECAM	576	720	25	Y	4:2:0	4:3
EDTV (576p)	480/576	720	29.97	N	4:2:0	4:3/16:9
HDTV (720p)	720	1280	59.94	N	4:2:0	16:9
HDTV (1080i)	1080	1920	29.97	Y	4:2:0	16:9
HDTV (1080p)	1080	1920	29.97	N	4:2:0	16:9
Digital cinema (2K)	1080	2048	24	N	4:4:4	47:20
Digital cinema (4K)	2160	4096	24	N	4:4:4	47:20

Figure 3.12 Table Illustrating Digital Video Formats

Digital Audio

- Physics describe sound as a form of energy that is transmitted through a medium as pressure waves making rarefaction and compression.
- The pressure changes can be captured by an electromechanical device such as microphone and converted into an analog signal.
- Monophonic sounds having only one channel.
- Stereophonic sounds aim to add dimensional to the sound experience by using more than one channel.

Digital Audio..

- Binaural and stereophonic recording
- Binaural is reserved for the use of two channels.
- Stereophonic refers to the use of more than two channels.
- In binaural recording, a pair of microphones records two different sound or channels,which are then played back on two different speakers.
- This creates the perception of spacial sound.
- Multichannel stereo is where three channels created a surround sound experience. (in 1940 by Walt Disney)
- Surround sound use five and more channels since then.
- Another advancement is the creation of spatial audio, which attempt to create surround sound like effect using a few channels.

Digital Representation of Audio

- Analog audio signal represented as waveforms.
- The amplitude of the wave give the strength of the sinosoid at that time.
- Most of the audio that interest us, such as voice and music are composed of multiple frequencies, where the amplitude is the joint combination of the amplitudes of all individual frequencies.
- The process of converting digital sound to analog call as Pulse Code Modulation (PCM).
- Other characteristics to describe audio signal besides sampling and quantization is number of channels, which may be one (mono), two (stereo) or multiple (surround sound).
- Spatial sound also being research and studied nowadays.

Surround Sound

- Surround sound aims to create a three-dimensional sound experience.
- It making feel as the listener are in the middle of the action or concert.
- It plays a large part in producing what movie makers call *suspended disbelief.*
- True surround sound formats rely on multiple dedicated speakers that literally and physically surround the audience.
- Commonly accepted version is 5.1 surround sound, which consist of six speakers.

Surround Sound..

- For 5.1 surround sound, one center speaker carries most of the dialog and part of the soundtrack.
- Left and right sound speakers carry most of the sound trank.
- Pair of the surround sound speakers are placed to the side and slightly above the audience to provide surround sound and ambient effect.
- Subwoofer can be used to reproduce the low and very low frequency effects (LFE).
- 5.1 surround sound concept has also been extended to 7.1 to include two more spatial speakers.

Surround Sound...

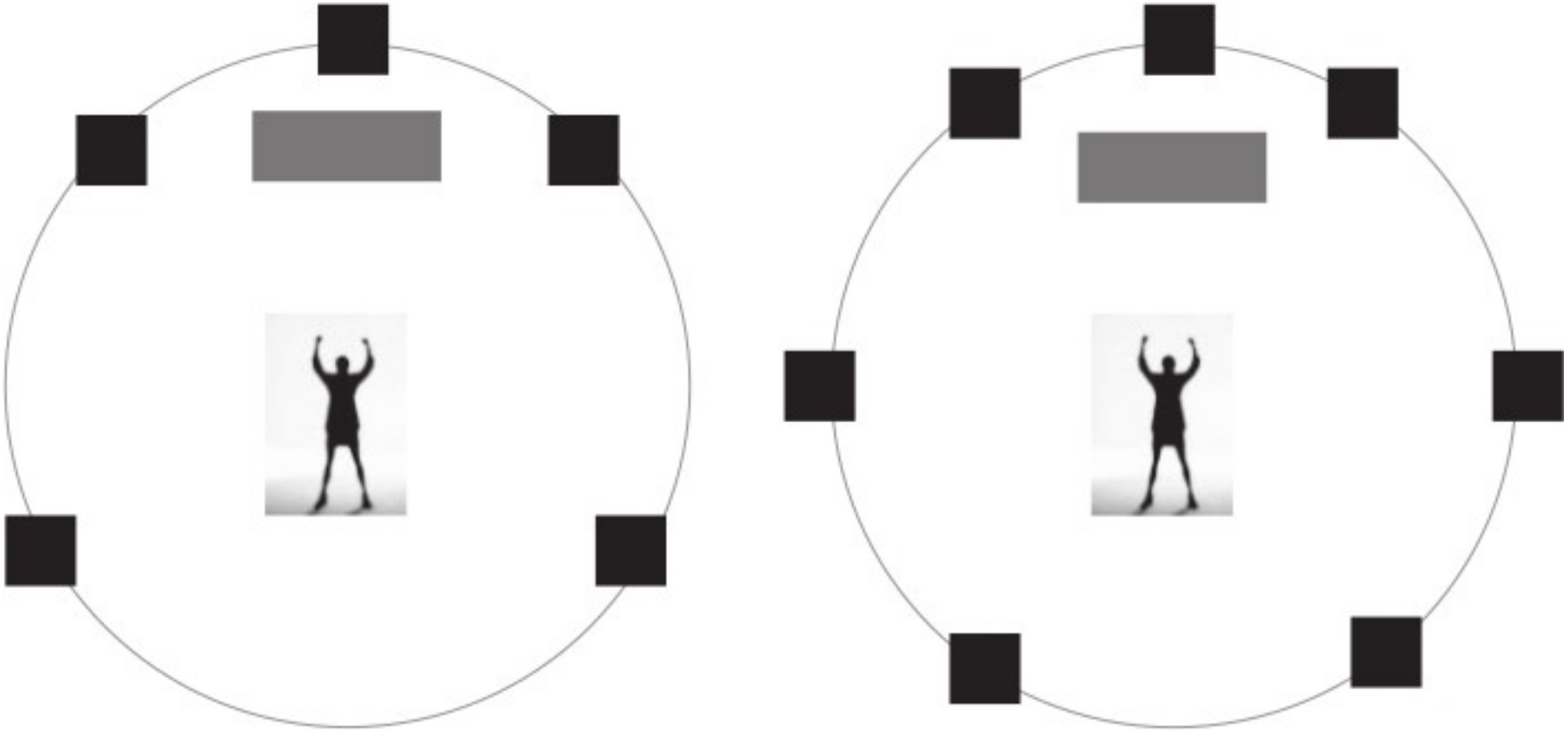


Figure 3.13 5.1 surround sound (left) and 7.1 surround sound system (right). Black square is surround sound speaker and the light gray is the low frequency subwoofer.

Spatial Audio

- Spatial audio attempt to create directional effects using fewer channels.
- This can be classified as virtual surround sound processes.
- The spatial audio process widen the stereo effect by modifying phase information among channels.
- It make use of head related transfer function (HRTF) and *reverberation* to simulate the virtual localization of the sound source as coming from left, right, front, back or behind the listener.
- With HRTF, our brain perceives a sound as coming from a particular direction and elevation in large part due to the prefiltering effects of these external structures.
- Reverberation is the persistence of sound in a particular space after the original sound is removed.
- Reverberation plays a role in our perception of where the sound source is and can be simulated to create a sense of a virtual source.

Audio Formats

File suffix or logo	Filename	File type	Features
.wav	WAV	Uncompressed PCM coded	Default standard for audio on PCs. WAV files are coded in PCM format.
.au	G.711 μ-law, or ITU μ-law	Uncompressed audio	Universal support for telephone. Packs each 16-bit sample into 8 bits, by using logarithmic table to encode with a 13-bit dynamic range. Encoding and decoding is very fast.
GSM 06.10	Global System for Mobile Communication	Lossy Compressed mobile audio	International standard for cellular telephone technology. Uses linear predictive coding to substantially compress the data. Compression/decompression is slow. Freely available and, thus, widely used
.mp3	MPEG1 Layer3	Compressed audio file format	Uses psychoacoustics for compression Very good bandwidth savings and, hence, used for streaming and Internet downloads.

Figure 3.14 Audio Format

Audio Formats..

.ra	Real Audio	Compressed format	Proprietary to Real Audio. Capable of streaming and downloading. Comparable quality to mp3 at high data rates but not so at low data rates
AAC	Advanced Audio Codec MPEG4	Compressed format	Superior quality to .mp3.
.mid	MIDI—Musical Instrument Digital Interface	Descriptive format	MIDI is a language of communication among musical instruments. Description achieved by frequencies, decays, transients, and event lists. Sound has to be synthesized by the instrument.
DOLBY DIGITAL	Dolby Digital (formerly called	Compressed 5.1 surround sound	De facto standard of home entertainment (Dolby AC-3) Distributed with DVD, HDTV systems. Provides five discrete channels—center, left, right, surround left, and surround right—plus an additional six for LFE.
DIGITAL dts SURROUND	DTS Surround Sound	Compressed 5.1 surround sound	Alternate to Dolby Digital. Distributed with DVDs, but not HDTV. Has higher data rate compared with Dolby Digital.
LUCASFILM THX ULTRA	THX Surround Sound	Compressed 5.1 surround sound	Designed for movie theaters (THX Ultra) as well home theaters (THX Select). Has become the select brand for surround sound today.
LUCASFILM THX ULTRA	THX Surround Sound Extended	Compressed 6.1 or 7.1 surround sound	Jointly developed by Lucasfilm, THX and Dolby Laboratories. Also known as Dolby Digital ES. Has a surround back channel, placed behind audience achieving 360^0 of sound.

Figure 3.15 Audio Format

Graphics

- Computer Graphics has evolved into an advanced field having applications in high end markets such as visual effects movies,interactive and multiplayer games,scientific visualization and computer aided design.
- Today graphic related applications tend to be three dimensional,interactive and available on variety of platform from computers, game console and smartphones.
- Graphics can be represented as vectors and rasters.
- Advantage of vector representation is that they provide infinite resolution.

Graphics..

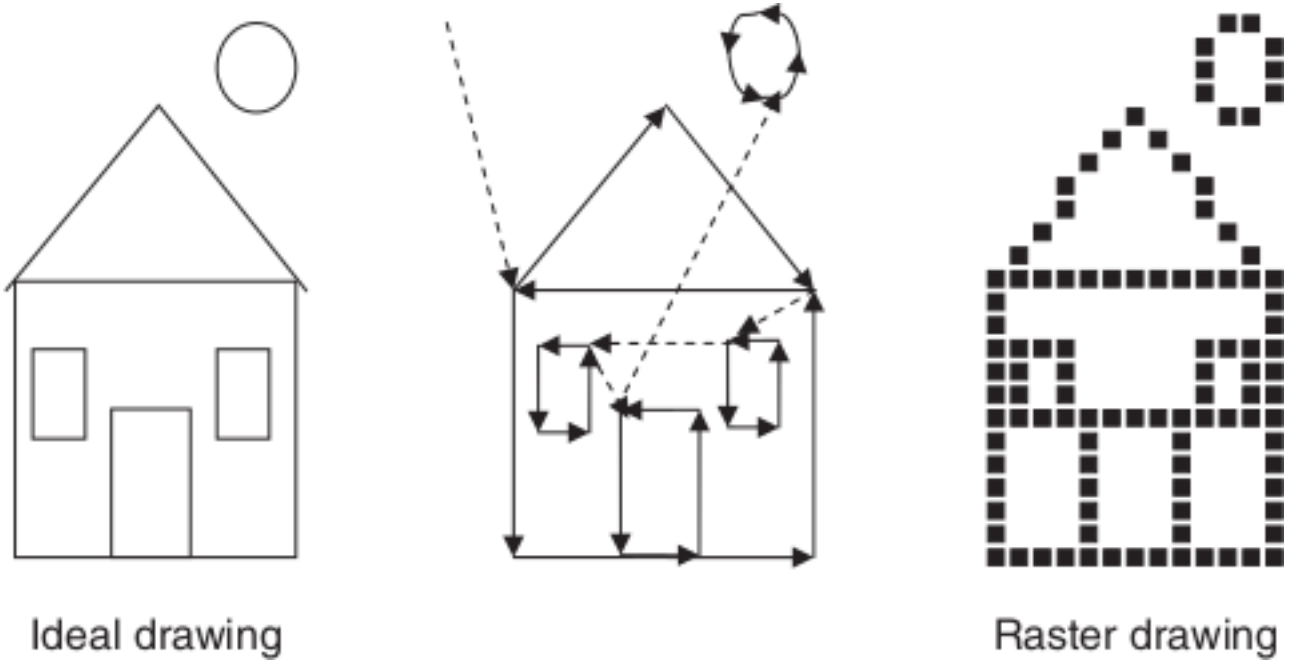


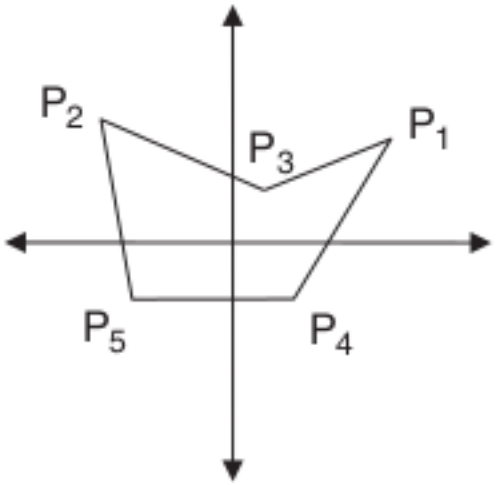
Figure 3.16 Most left shows drawing that needs to be represented graphically. Middle image shows the corresponding representation using vectors. This is shown as a sequence of points joined by lines. The right image shows a raster representation Of the drawing.

2D Vector Graphic Representation

- In vector graphics, objects are represented as smooth or discrete primitives.
- Smooth 2D primitives include such as circles, ellipses, and other control point–defined curves such as splines.
- Discrete primitives, such as triangles and polygons have now become a choice for representing vector graphics because all hardware graphics cards have been designed to render such primitives.
- In 2D, points are represented by their x and y coordinates.

2D Vector Graphic Representation..

$$P_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \quad P_2 = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \quad P_3 = \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} \quad P_4 = \begin{bmatrix} x_4 \\ y_4 \end{bmatrix} \quad P_5 = \begin{bmatrix} x_5 \\ y_5 \end{bmatrix}$$



$$Polygon\ O = \begin{bmatrix} x_1 & x_4 & x_5 & x_2 & x_3 \\ y_1 & y_4 & y_5 & y_2 & y_3 \end{bmatrix}$$

Figure 3.17 An example of a polygon and its representation. Five 2D points are connected to form a polygon O.

Animation Using 2D Graphics

- Animations are created by showing different images over time.
- These different images are raster representation of a vector objects that at each image frame has undergone change.
- The change can be translation,rotation, scaled up or down, or a variety of above operation.

Animation Using 2D Graphics..

$$\begin{bmatrix} x_i + t_x \\ y_i + t_y \end{bmatrix} = \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}$$

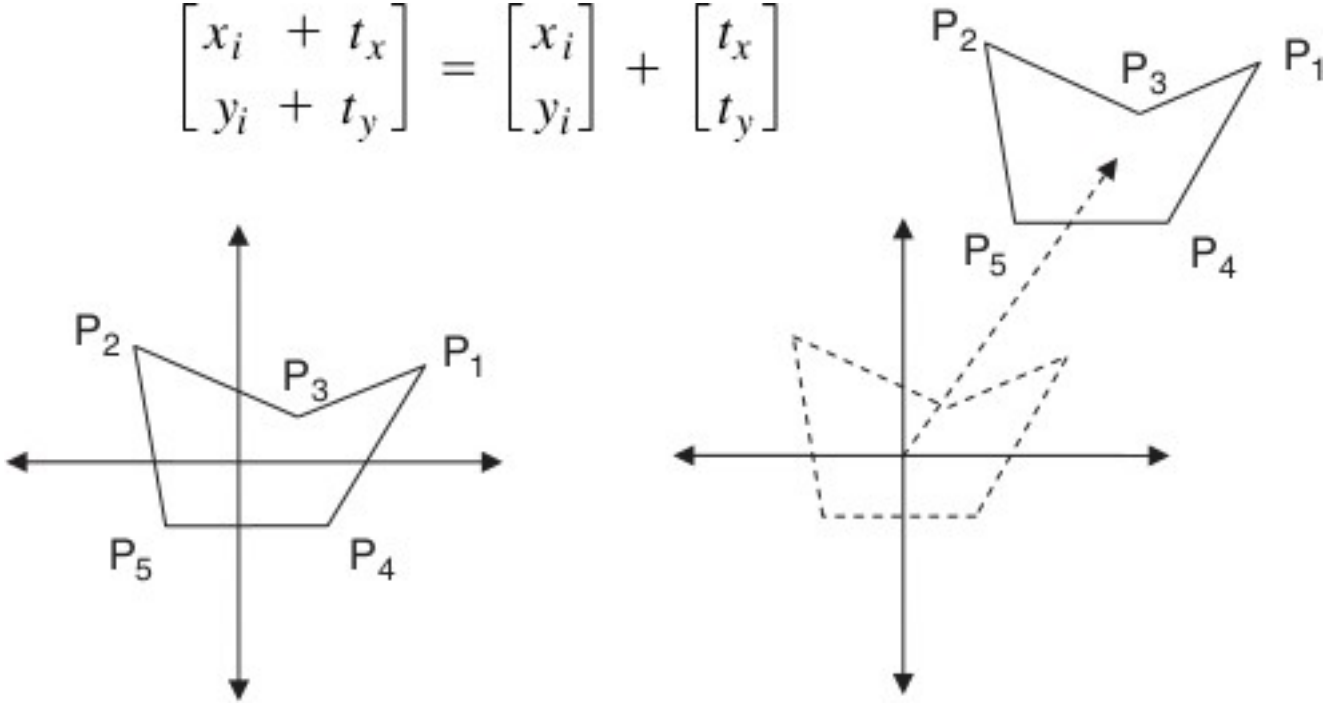


Figure 3.18 Graphic Translation. The polygon shown on the left is translated by a vector illustrated on the right

Conclusion

- This chapter present various multimedia formats.
- Technical fundamentals inside each multimedia element also presented.