PREPARED BY:

RAGHUDATHESH G P
Asst Prof
[redacted]
Cell: +917411459249
Mail: datheshraghubooks@gmail.com

Quotes from Great People:

- The greatest discovery of all time is that a person can change his future by merely changing his attitude.
- It is curious that physical courage is so common in the world and moral courage so rare.
- Education is when you read the fine print. Experience is what you get if you don't.
- Education is the Best Friend. An Educated Person is Respected Everywhere. Education beats the Beauty and the Youth.
- Time perfects men as well as destroys them.

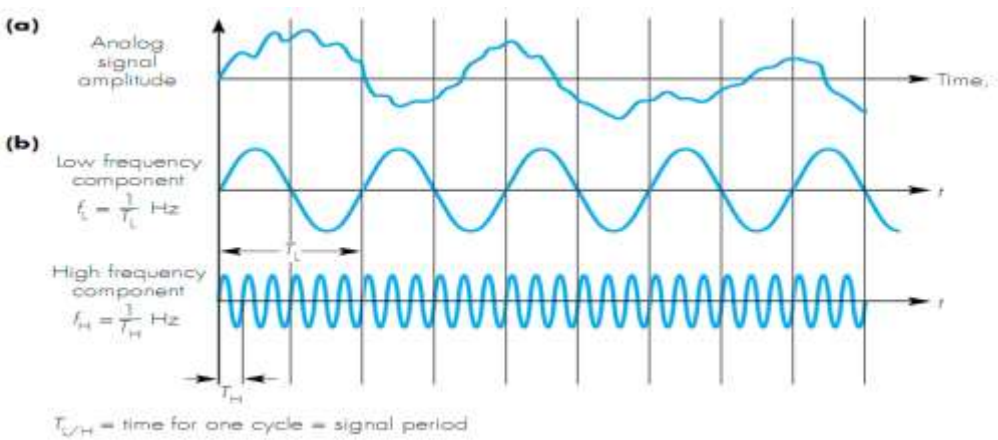
MULTIMEDIA INFORMATION REPRESENTATION

INTRODUCTION:

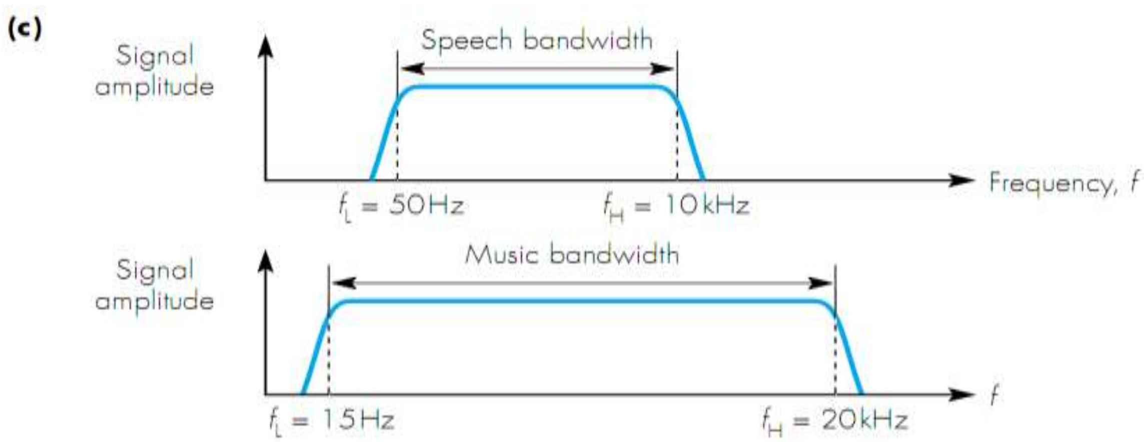
- All types of multimedia information are process and store within the computer in a digital form.
- **Textual information**: contains strings of characters entered through keyboard. Codeword: each character represented by a unique combination of fixed number of bits. Complete text hence, can be represented by strings of codewords.
- **Image**: computer-generated graphical images made up of a mix of lines, circles, squares, and so on each represented in a digital form. Ex.: line - represented by start and end co-ordinates of the line, each coordinate being defined in the form of a pair of digital values relative to the complete image.
- **Audio and video**: microphone and video cameras produce electrical signals, whose amplitude varies continuously with time amplitude indicating the magnitude of the sound wave/image-intensity at that instant.
- **Analog Signal**: signal whose amplitude varies continuously with time. In order to store and process analog signal type of media in a computer we should convert any time-varying analog signals into a digital form is necessary.
- for speech and audio - in like, loud speakers, and for display of digitized images in like, computer monitors - digital values of media types must be converted back again into a corresponding time-varying analog form on output – from the computer.
- for a particular media type:
- Conversion of analog signal into digital signal is carried out using an electrical circuit known as **Signal Encoder**, it includes following steps:
    1. **Sampler:** It samples the amplitude of analog signals at repetitive time intervals.
    2. **Quantization:** converting amplitude of each sample into a corresponding digital value.
- Conversion of stored digital sample relating to a particular media type into their corresponding time-varying analog form is performed by a electrical circuit is known as a **signal decoder**.
- All media types associated with the various multimedia applications stored and processed within a computer in an all-digital form so, different media types can be readily integrated together resulting integrated bitstream can be transmitted over a single all-digital communication network.

Digitization principles:

Analog signals:

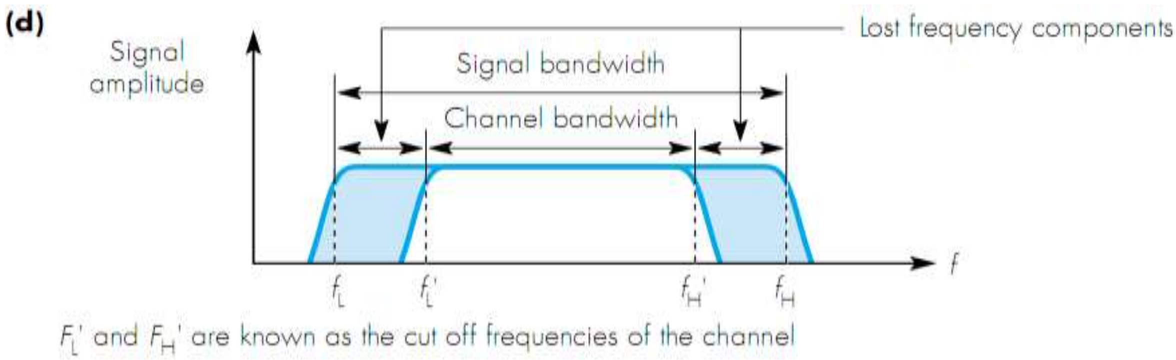


- Figure above shows general properties relating to any time-varying analog signal. In figure amplitude of signals varies continuously with time.
- **Fourier analysis**: A mathematical technique used to show that any analog signal is made up of a possibly infinite number of single-frequency sinusoidal signals, whose amplitude and phase vary continuously with time relative to each other.
- Ex.: highest and lowest frequency components of the signal shown in Figure a. maybe those shown in Figure b.
- **Signal bandwidth:** range of frequencies of the sinusoidal components that make up a signal.
- Figure below (c). Shows two examples relate to an audio signal first - relate to a speech signal second - relate to a music signal produced by say, an orchestra.



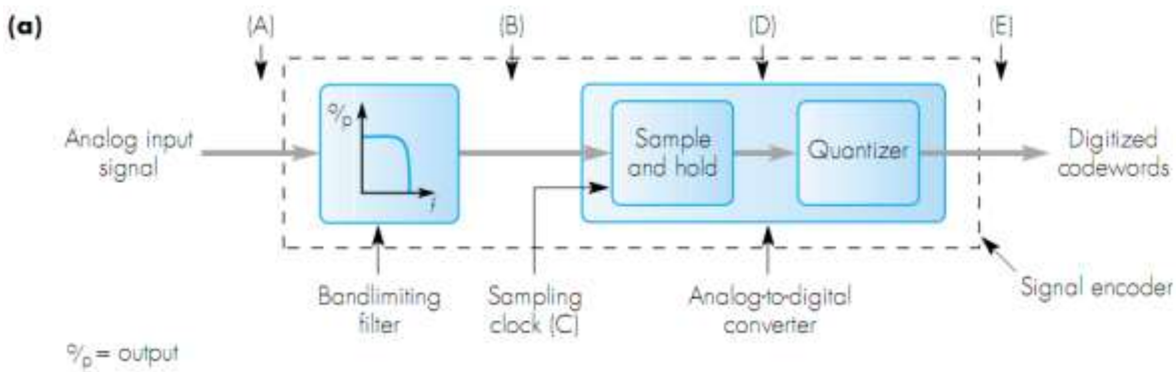
- Speech is a humans produce sounds, which are converted into electrical signals by a **microphone** are made up of a range of sinusoidal signals varying in frequency between **50Hz and 10kHz** and for music range of signals is wider varies between **15kHz to 20kHz** being comparable with the limits of the sensitivity of the ear.

- Analog signal when, being transmitted through a network BW of transmission channel (range of frequencies, channel will pass) ≥ BW of the signal if BW of channel < BW of signal some low and/or high frequency components will be lost, thereby degrading the quality of the received signal. Such, a channel is called the **bandlimiting channel** as in Figure below.

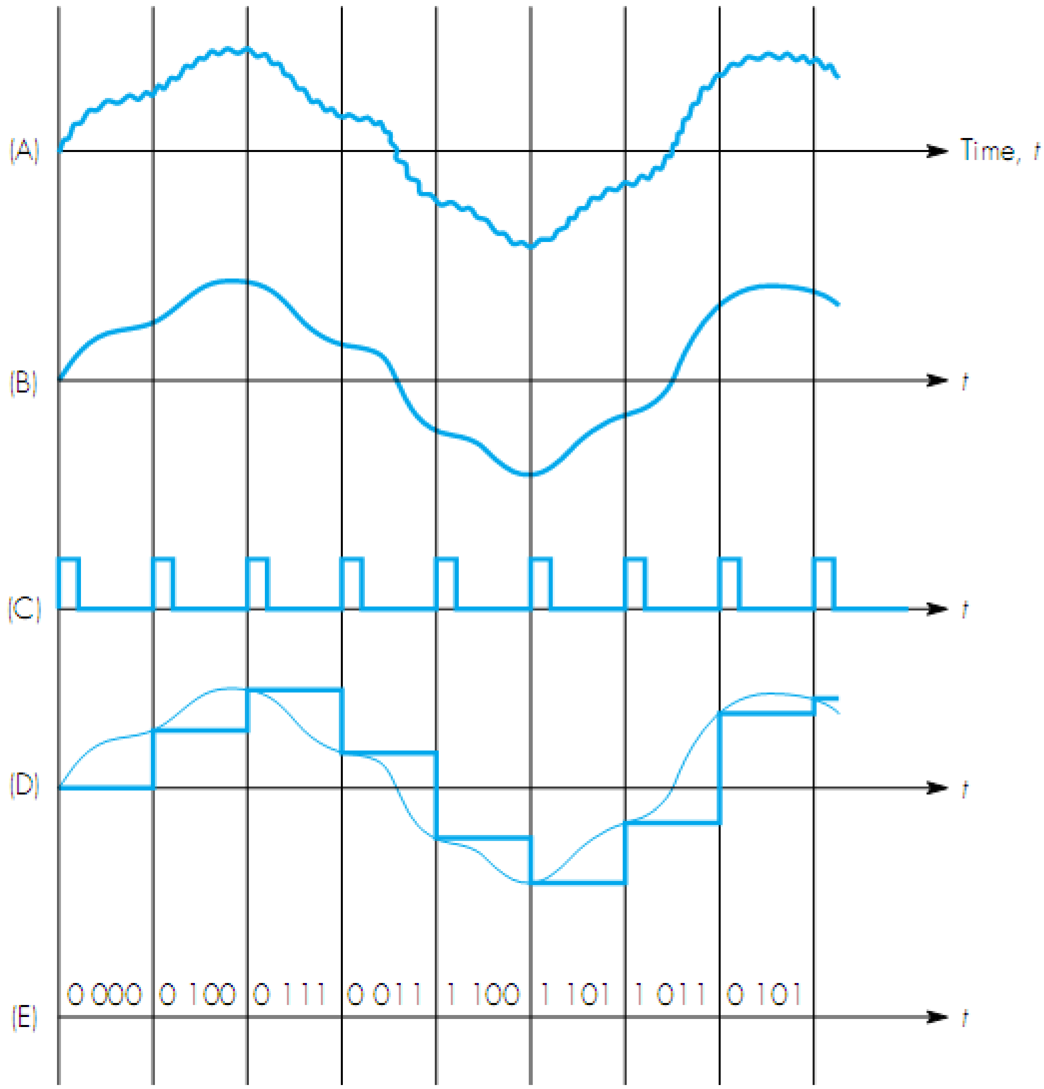


Encoder design:

- Signal encoder is a electronic circuit converts, time-varying analog signals to digital form.



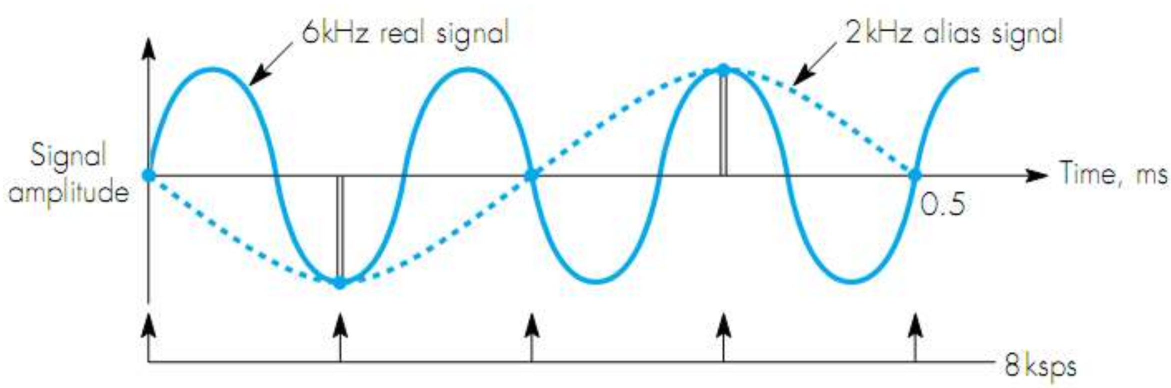
- Figure above. shows - principles of an encoder consists of 2 main circuits:
  1. Bandlimiting filter.
  2. ADC (Analog to Digital Converter) which has 2 components:
     I. Sample-and-hold circuit.
     II. Quantizer
- Figure below. Shows the typical **waveform set** for the signal encoder.



- **Bandlimitng filter**: remove selected higher-frequency components from the source signal **(A).**
- **Sample-And-Hold**: got output of bandlimiting filter, (B) signal used to sample amplitude of the filtered signal at regular time intervals (C) and to hold the sample amplitude constant between samples (D) signal Quantizer circuit got signal (D) which converts each sample amplitude into a binary value known as a **codeword** like (E) signal.
- Polarity (sign) of sample: positive or negative relative to the zero level indicated by most significant bit of each codeword. A binary 0 indicates a positive value and a binary 1 indicates a negative value.
- To represent the amplitude of a time-varying analog signal precisely require 2 things:
    1. Signal should be sampled at a rate > maximum rate of change of signal amplitude.
    2. Number of quantization levels used to be as large as possible.

Sampling Rate:

- **Nyquist Sampling Theorem:** states that for an accurate representation of a time-varying analog signal it's amplitude must be sampled at a minimum rate that is equal to or greater than twice the highest sinusoidal frequency component that is present in the signal known as Nyquist rate, normally represented as either Hz or, or correctly, samples per second (sps).
- sampling signal at a rate < Nyquist rate results in additional frequency components being generated that are not present in the original signal which, in turn cause original signal to become distorted.
- **Figure below,** shows effect of under sampling single-frequency sinusoidal signal caused by sampling a signal at a rate lower than the Nyquist rate.



- **Ex.:** original signal is assumed to 6kHz sine wave sampling rate(8ksps)< Nyquist rate (12ksps, 2*6ksps) results in a lower frequency 2kHz signal being created in place of the original 6kHz signal such, signals called **alias signals** (since, they replace the corresponding original signals).
- In general, all frequency components present in the original signal higher in frequency than half the sampling frequency being used (in Hz) generate related lower-frequency alias signals which will simply add to those making up the original source signal thereby causing it to become distorted.
- **Bandlimiting filter/Antialiasing filter**: source signal is passed into the bandlimiting filter to pass only those frequency components up to that determined by Nyquist rate any higher-frequency components in the signal which are higher than this are removed before the signal is sampled.
- In practice transmission channel, used/available has a lower bandwidth than that of source signal to avoid distortion bandwidth and hence, frequency range of the transmission channel that determines the sampling rate used rather than the BW of the source signal in such cases, source signal may have higher frequency component, than those dictated by the Nyquist rate of the transmission channel so, it is necessary to

pass the source signal through a bandlimiting filter (designed to pass only those sinusoidal frequency components which are within the BW of the transmission channel) so, generation of any alias signals caused by under sampling source signal is avoided.

Quantization intervals:

- To represent in the digital form the amplitudes of the set of analog samples would require an infinite number of binary digits. When finite numbers of digits are used, each sample can be represented by a corresponding number of discrete levels.
- **Figure below**. Shows effect of using a finite number of bits
- Ex.: here, 3 bits to represent each sample including a sign bit results in 4 positive and 4 negative quantization intervals, the two magnitude bits - being determined by the particular quantization interval the analog input signal is in at the time of each sample.
- Fig a shows: if V_{max}, is the maximum positive and negative signal amplitude.
- n: number of binary bits used
- q: magnitude of each quantization interval is given by

$$q = \frac{2V_{max}}{2^n}$$

- Signal - anywhere within a Quantization intervals will be represented by the same binary codeword. Thus each codeword: corresponds to a nominal amplitude level - which, is at the center of the corresponding quantization interval. Thus, Actual signal level - may different from this by: up to + or -q/2.
- Quantization Error: It is the difference between the actual signal amplitude and the corresponding nominal amplitude.
- Figure below. shows quantization error values shown expanded.
- Usually the error values vary randomly from sample to sample thus quantization error is also known as Quantization noise.
- Noise: This term used in electrical circuits to refer to a signal whose, amplitude varies randomly with time.
- Smallest amplitude relative to its peak amplitude of the signal is the influencing factor for the choice of the number of quantization intervals for a particular signal.
- With high-fidelity music: It is important to be able to hear very quite passages without any distortion created by quantization noise.
- Dynamic range, D (of the signal) is the ratio of the peak amplitude of a signal to its minimum amplitudes.
- decibels (dB): D is normally quantified using logarithmic scale.

$$D = 20log_{10}\left({}^{V_{max}}/_{V_{min}}\right) dB$$

- Determining the quantization intervals, and number of bits to be used it is necessary to ensure level of quantization noise relative to the smallest signal amplitude is acceptable.

Problem: An analog signal has a dynamic range of 40dB. Determine the magnitude of the quantization noise relative to the minimum signal amplitude if the quantizer uses:
i. 6 bits **ii. 10 bits:**
Solution:
Dynamic range of the signal is given by:

$$D = 20log_{10}\left({}^{V_{max}}/_{V_{min}}\right) dB$$

Quantization noise is:

$$q = \frac{2V_{max}}{2^n}$$

Thus,

$$40 = 20log_{10}\left({}^{V_{max}}/_{V_{min}}\right) dB$$

$$V_{min} = \left({}^{V_{max}}/_{100}\right)$$

i. For n = 6 bits

$$\frac{q}{2} = \pm\frac{V_{max}}{2^6} = \pm\frac{V_{max}}{64}$$

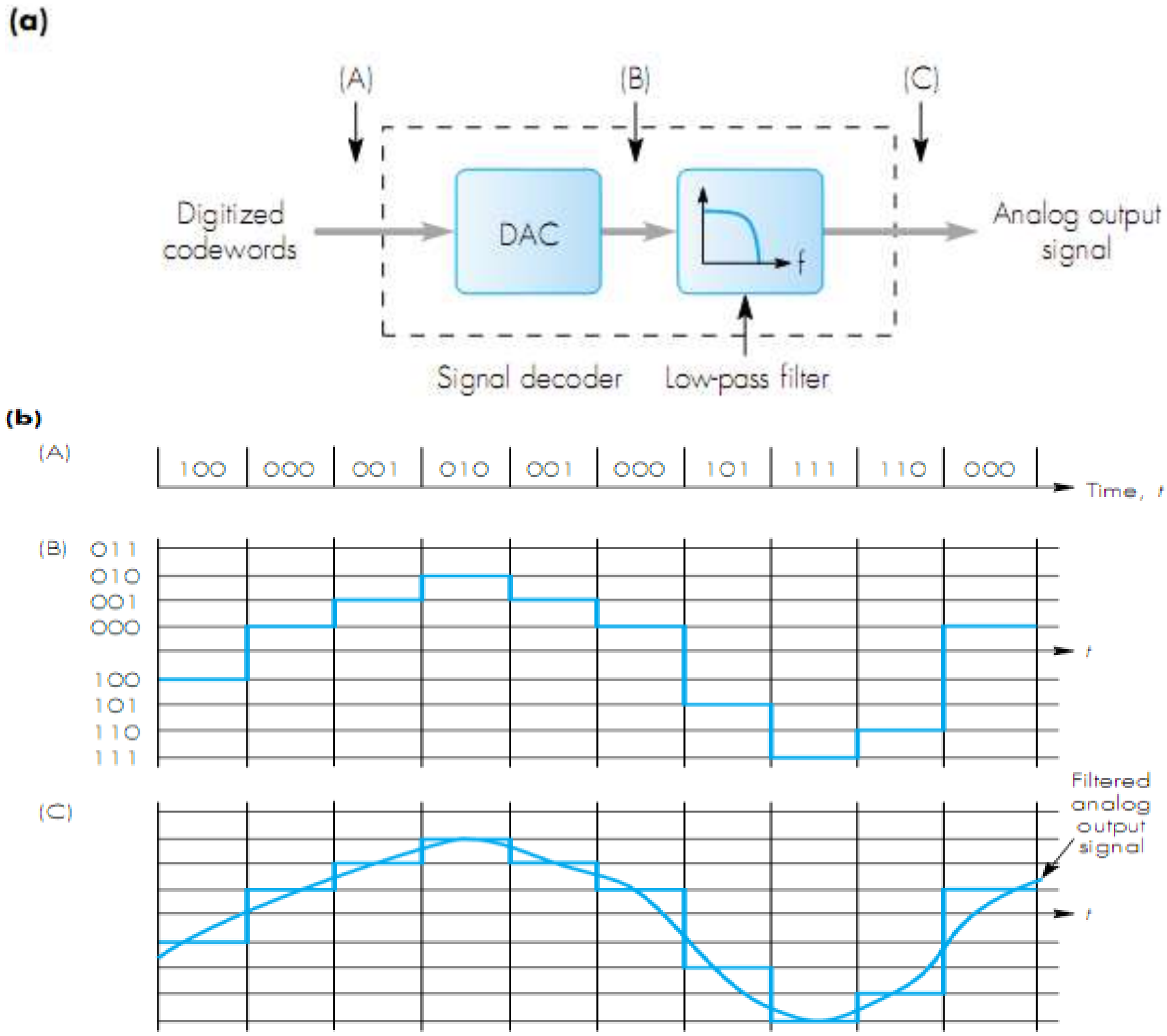
ii. For n = 6 bits

$$\frac{q}{2} = \pm\frac{V_{max}}{2^{10}} = \pm\frac{V_{max}}{1024}$$

Conclusion:
- With 6 bits the quantization noise is greater than V_{min} hence is unacceptable.
- With 10 bits the quantization noise is lesser than V_{min} hence is unacceptable.

Decoder design:

- Analog signals are store, process and transmitted in the digital form, prior to their output, normally analog signals must be converted back again into their analog form.
- Ex.: loudspeakers - are driven by an analog current signal.
- Signal decoder is electronic circuit which performs the conversion of digital to analog form.
- Figure below. shows principles of signal decoder



- Digital To Analog Converter(DAC) is a circuit which converts each digital codeword (A) into an equivalent analog sample (B), amplitude of each level being determined by, corresponding codeword.
- Fourier analysis: used to show that output of DAC comprises sinusoidal frequency components make up the original (filtered) analog signal + an infinite number of additional higher-frequency components.
- For original signal to reproduce DAC output is passed through a LPF, which only passes those frequency components that made up the original filtered signal (C).
- Normally, high-frequency cut-off of the LPF is made same as that used in band limiting filter of the encoder so, LPF is known as - recovery (reconstruction filter).
- Most multimedia application involving audio and video communications channels is 2-way simultaneous. TEs hence, support both input and output simultaneously. So, audio/video signal encoders and decoders in each TE, are often combined into a single unit audio/video encode-decoder or audio/video codec.

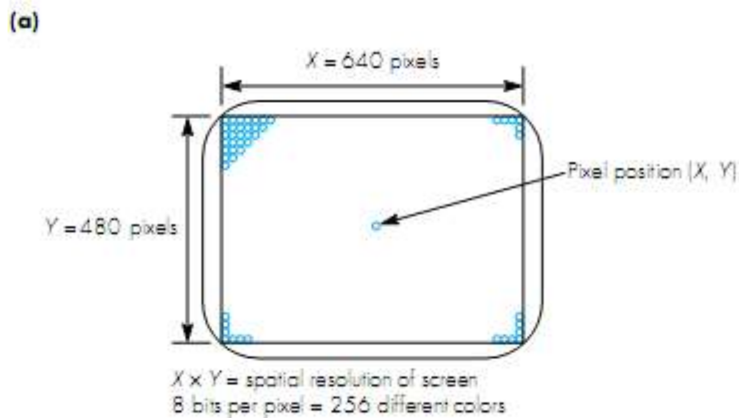
Images:

- **Definition**: “An image is an artifact that depicts or records visual perception”.
- Ex: 2-D picture.
- Different **mode of Image generation**:
  1. Computer-generated images generally referred to as **computer graphics** or simply, graphics.
  2. **Digitized images** of both documents and pictures.
- There are 3 **types of images**:
  1. Graphics
  2. Digitized documents
  3. Digitized pictures
- These images are displayed and printed in 2-D matrix form of individual picture elements.
- **Pixels (Pels):** are individual picture elements.
- Each of the 3 types of images is represented differently within the computer memory, or more generally, in a computer file each type of image is created differently.

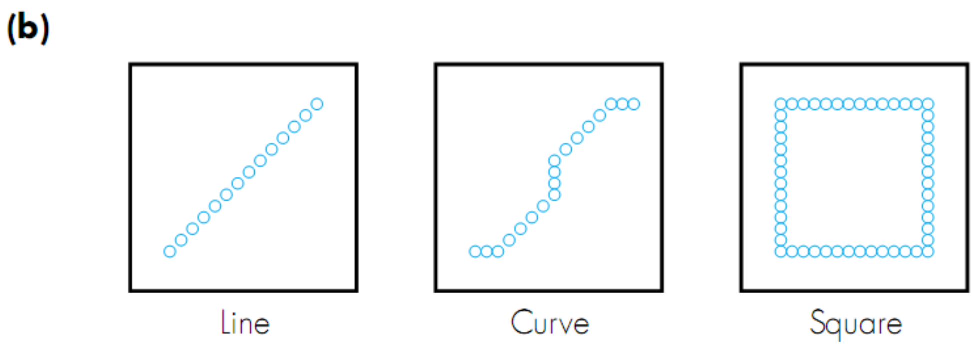
Graphics:

- A range of software packages and programs are available for creation of computer graphics.
- They provide easy-to-use tools to create graphics which, are composed of all kinds of visual objects including lines, arcs, squares, rectangles, circles, ovals, diamonds, stars, and so on, as well as any form of hand-drawn (normally referred to as freeform) objects produced by drawing desired shape on the screen by means of a combination of a cursor symbol on the screen.
- The mouse facilities are also provided to edit these objects. Ex.: to change their shape, size, or color and, to introduce complete predrawn images, either previously created by the author of the graphic or **clip-art** (selected from a gallery of images that come with the package).
- Better packages provide many hundreds of such images.
- Textual information can be included into a graphic together with precreated tables, graphs, digitized pictures and photographs which have been previously obtained. Objects can overlap each other with selected object nearer to the front than another with add fill and add shadows to objects to give the complete 3-D effect.
- **Computer's Display Screen**: It can be considered as made up of 2-D matrix of individual pixels each of which can have a range of colors associated with it.
- Ex.: **VGA (Video Graphics Array)** common type of display

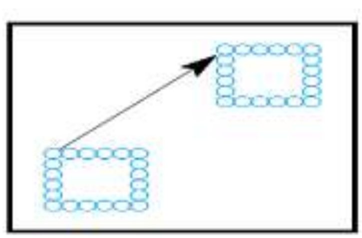
- **Figure below.** Shows a matrix of 640 horizontal pixels by 480 vertical pixels. 8 bits/pixel which allows each pixel to have one of 256 different colors.



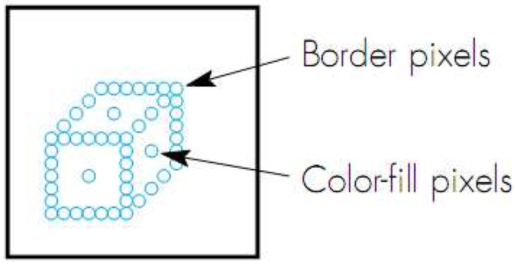
- All objects including the free-form objects made up of a series of lines connected to each other. curved line as what may appear in practice, is a series of very short lines each made up of a string of pixels which, in the limit, have the resolution of a pair of adjacent pixels on the screen.
- **Figure below**. shows some examples



- **Attributes:** each object has a number of attributes associated with it they include:
    1. Its shape - a line, a circle, a square, and so on.
    2. Its size - in terms of pixel positions of its border coordinates.
    3. Color of border.
    4. its shadow, and so on
- Editing of an object involves simply, changing selected attributes associated with the object.
- Ex.: As in **Figure below**, square can be moved to different location on the screen by simply, changing its border coordinates and leaving the remaining attributes unchanged.



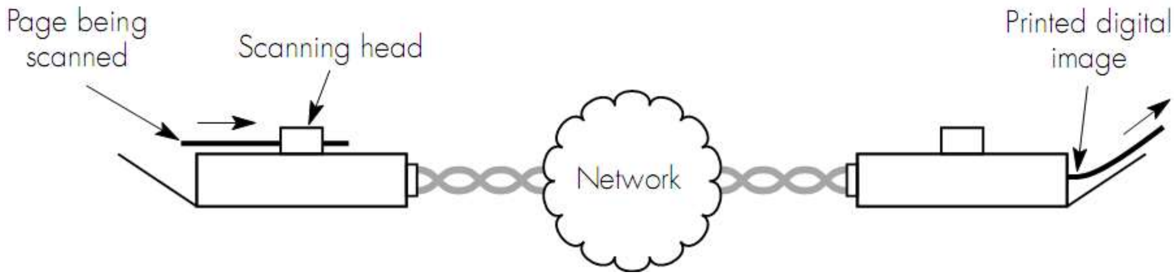
- **Object shape**: It can be either open or closed.
- **Open object**: start of the first line and end of the last line that make up the object's border are not connected i.e., they do not start and end on the same pixel.
- **Closed object**: start of the first line and end of the last line that make up the object's border are connected i.e., they start and end on the same pixel.
- For closed objects **color-fill** can be done i.e., pixels enclosed by its border can all be assigned the same color to **rendering** (creation of the solid objects, by color-fill) as in **Figure below**.



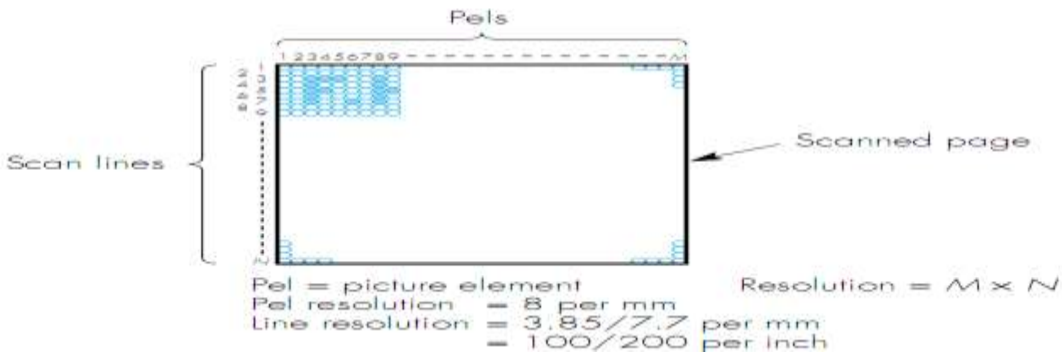
- All objects are drawn on the screen by the user simply specifying name of the objects and its attributes including its color-fill and shadow effect if required set of more basic lower-level commands are then used to determine both the pixel locations that are involved and the color that should be assigned to each pixel.
- Representation of a complete graphic is analogous to the structure of a program written in a high-level programming language.
- 2 forms of representation of computer graphic
    1. **High-level version** (similar to the source code of a high-level program).
    2. **Actual pixel-image of the graphic** (similar to the byte-string, generally, as bit-map format).
- Graphic can be transferred over a network in either form
- High-level program form much more compact requires less memory to store the image requires less BW for its transmission destination must be able to interpret various high-level commands.
- Bit-map form used to avoid above requirements there are a number of standardized forms of representation such as:
    1. **GIF** (Graphical Interchange Format)
    2. **TIFF** (Tagged Image File Format)
- **SRGP (Simple Raster Graphics Package)** convert the high-level language into a pixel-image form.

Digitized documents:

- It is produced by the scanner associated with a facsimile (fax) machine.
- **Figure below**. shows principles of facsimile (fax)



- Scanner associated with the fax machine operated by scanning each complete page from left to right to produce a sequence of scan lines that start at the top of the page and end at the bottom vertical resolution of scanning procedure is either 3.85 or 7.7 lines/mm which is equivalent to approximately 100 or 200 lines/inch.
- As each line is scanned output of the scanner is digitized to a resolution of approximately 8pels with fax machines/mm.
- Fax machines use just a single binary digit to represent each pel:
    1. 0 for white pel
    2. 1 for black pel
- Figure below. Shows digital representation of the scanned page.



- For a typical page which produces a stream of about two million bits.
- Printer of fax then, reproduces original image by printing out the received stream bits to a similar resolutions.
- Use of a single binary digit per pel means, fax machines are best suited to scanning **bitonal (black-and-white) images** such as printed documents comprising mainly textual information.

Digitized pictures:

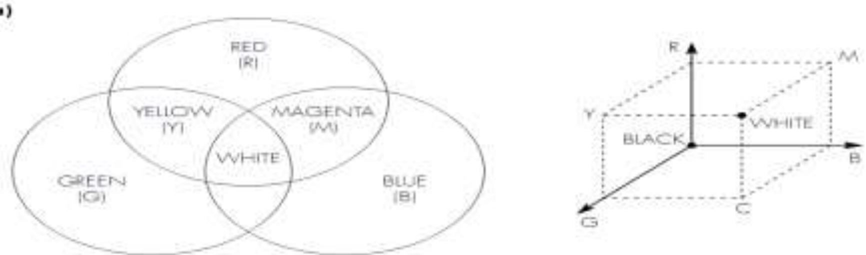
- Scanners used for digitizing continuous tone monochromatic images (such as, printed picture, scene) normally, more than a single bit is used to digitize each pel.
- Ex.: good quality black and white pictures can be obtained by using 8bits/pel yields 256 different levels of gray per element varying between white and black which gives substantially increased picture quality over a facsimile image when reproduced.

- For color images to understand digitization format used, it is necessary to understand the principles of how color is produced and how the picture tubes used in computer monitors (on which the images are eventually displayed) operate.

Color Principles:

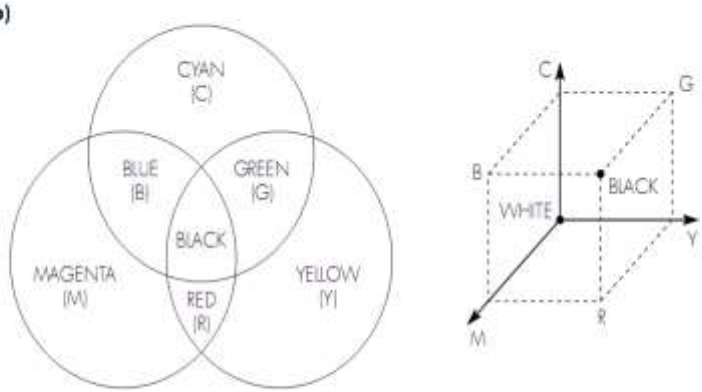
- Studies have shown that human eye sees just a single color when a particular set of 3 primary colors are mixed and displayed simultaneously.
- **Color gamut** is a whole spectrum of colors which is produced by mixing different proportions of 3 primary colors red (R), green (G), and blue (B).
- Figure. A below shows mixing technique used is called **additive color mixing**.

(a)



- **Additive Color Mixing:** Black is produced when all three primary colors are zero particularly useful for producing a color image on a black surface, as is the case in display application.
- **Subtractive Color Mixing:** complementary to additive color mixing produces - similar to additive color mixing range of colors.
- Fig. b shows - as - in subtractive color mixing white is produced, when the 3 chosen primary colors cyan (C), Magenta (M), and Yellow (Y) are all zero these colors are particularly useful for producing a color image on a white surface as in, printing applications.

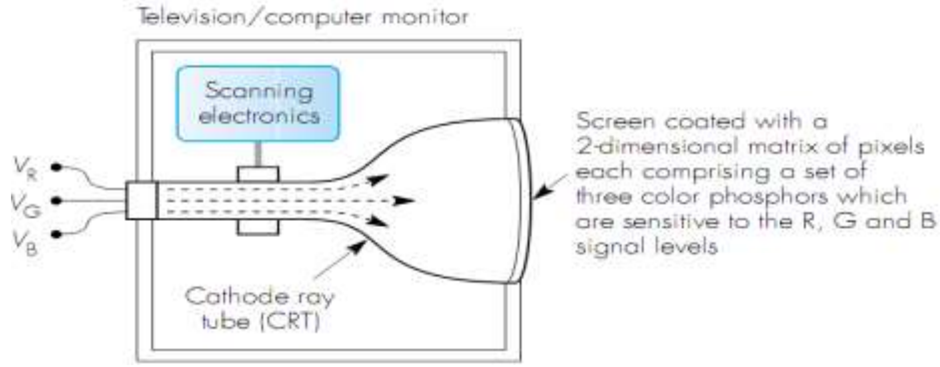
(b)



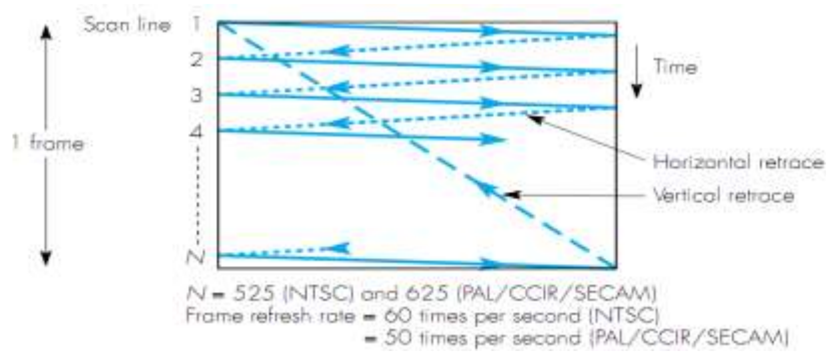
Raster-scan principles:

- Picture tube used in most television sets operates using raster-scan.

- It involves raster a finely focused electron beam i.e., raster scan over the complete screen as shown in the figure below.

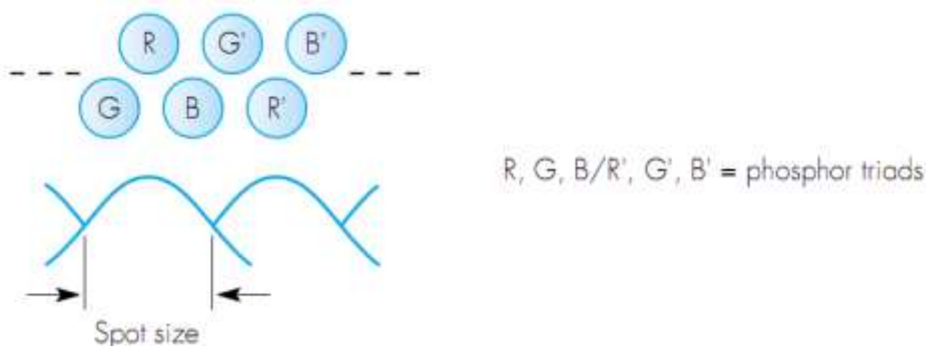


- **Progressive scan:** Here each complete scan comprises a number of discrete horizontal lines first of which starts at the top left corner of the screen and, the last of which ends at the bottom right corner at this point, the beam is deflected back again to the top left corner. Scanning operation repeats in the same way this type of the scanning is called as progressive scanning.
- Figure below, shows progressive scanning.



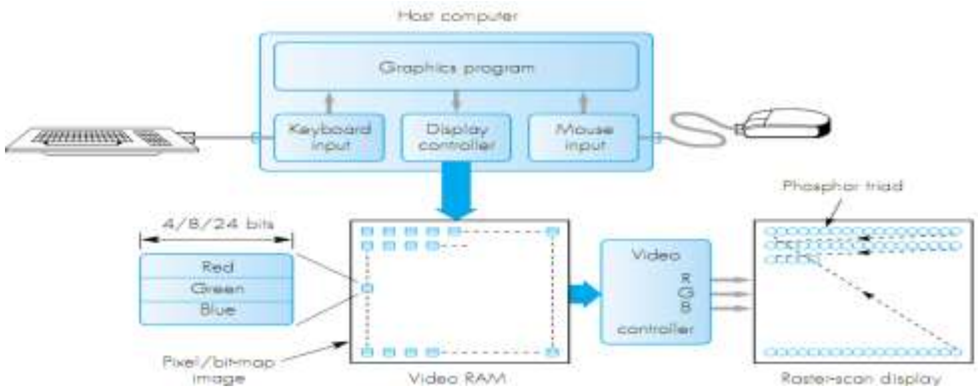
- Frame is a compete set of horizontal scan lines and is made up of N individual scan lines.
- **N** is either **525** (North and South America and most of Asia) or **625** (Europe and number of other countries).
- Inside the display screen of the picture tube is coated with a light sensitive phosphor which emits light when energized by the electron beam.
- **Brightness:** It is the amount of light emitted which is determined by the power in the electron beam at that instant.
- During each horizontal (line) and vertical (frame) retrace period electron beam is turned off to create an image on the screen level of power in the beam is changed as each line is scanned.

- In black-and-white picture tubes a single electron beam is used with a white-sensitive phosphor but in color tubes it use three separate closely located beams, and a 2-D matrix of pixels.
- Each pixel comprises set of 3 related color-sensitive phosphors one each for R,G, and B signals.
- **Phosphor triad** is a set of 3 phosphors associated with each pixel.
- Figure below. Shows a typical arrangement of the triads on each scan line in theory each pixel represents an idealized rectangular area which is independent of its neighboring pixels.



- **Spot** is a practical shape of each pixel which merges with its neighbors when viewed from a sufficient distance a continuous color image is seen.
- Television picture tubes are designed to display moving images persistence of light/color produced by the phosphor is designed to decay very quickly so, continuous refresh of thee screen is needed.
- For moving image light signals associated with each frame change to reflect the motion that has taken place during the time required to scan the preceding frame. For a static/still image same set of light signals are used for each frame.
- **Frame refresh rate:** must be high enough to ensure that the eye is not aware the display is continuously being refreshed.
- **Flicker** is caused by a low refresh rate caused by the previous image fading from the eye retina before the following image is displayed.
- To avoid Flicker a refresh rate of at least 50 times/s is required.
- Frame refresh rate: determined by frequency of the mains electricity supply which is either 60Hz in North and South America and most of Asia and 50 Hz in Europe and a number of other countries.
- Current picture tubes operate in analog mode i.e., amplitude of each of 3 color signals is continuously varying as each line is scanned.
- In case of Digital television digitized pictures are stored within the computer memory color signals are in the digital form comprise a string of pixels with a fixed number of pixels per scan line.

- To display the stored image pixels that make up each line are read from memory in time-synchronism with the scanning process and, converted into a continuously varying analog form by means of DAC.
- **Video RAM:** IS a separate block of memory used to store the pixel image. Area of computer memory that holds the sting of pixels that make up the image the pixel image must be accessed continuously as each line is scanned.
- **Graphics program:** Are needs to write the pixel images into video RAM whenever, either selected pixels or the total image changes.
- Figure below. Shows the architecture of various steps involved used to create a high-level version of the image interactively using, either the keyboard or a mouse.



- **Display controller (frame/display/refresh buffer):** A part of the program interprets sequences of display commands converts them into displayed objects by writing appropriate pixel values into the video RAM.
- **Video controller:** A hardware subsystem that reads the pixel values stored in the video RAM in time-synchronism with the scanning process converts for each set of pixel values into equivalent set of R, G, and B analog signals for output to the display.

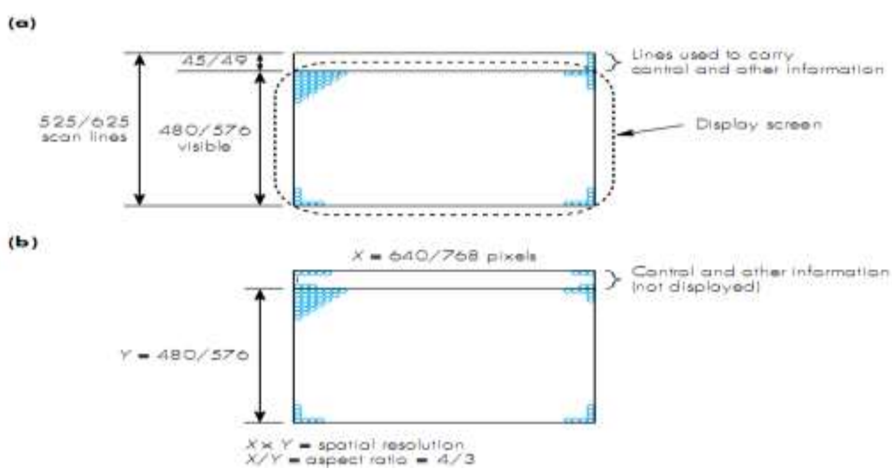
Pixel depth:

- It is defined as **number of bits/pixel**.
- Determines the range of different colors that can be produced by a pixel.
- Ex.: 12 bits - 4 bits per primary color - yielding 4096 different colors and 24 bits - 8 bits per primary color - yielding in excess of 16 million colors.
- Eye cannot distinguish such a range of colors so, in some instances a selected subset of this range of colors been used.
- For the above following steps are followed:
- Selected colors in the subset are then stored in the table.
- CLUT (Color Look-Up Table) is a table where each pixel value is used as an address to a location within the table (color look-up table, CLUT) which, contain the corresponding 3 color values.

- Ex.: if each pixel is 8 bits and the CLUT contains 24 bit entries, then, CLUT had 24 bit entries will provide a subset of 256 (28) different colors selected from the palette of 16 million (224) colors.
- Advantage: amount of memory required to store an image can be reduced significantly.

Aspect ratio:

- It is the ratio of screen width to screen height.
- It is used to determine number of pixels/scanned line and number of lines/frame of a display screen.
- In current television tubes aspect ratio is 4/3 of older tubes (on which the PC monitors are based) and is 16/9 for the wide-screen television tubes.
- Standards for color television
    1. **US: NTSC** (National Television Standards Committee)
        - NTSC: uses 525 scan lines/frame some lines carry information and some lines carry control all lines - are not displayed on the screen.
- Europe: 3 color standards exists:
    1. **PAL: of UK**
    2. **CCIR: of Germany**
    3. **SECAM: of France**
        - PAL, CCIR, SECAM uses 625 scan lines some lines carry information and some lines carry control all lines are not displayed on the screen.
- Number of visible lines/frame = vertical resolution in terms of pixels i.e., 480 for NTSC monitor and 576 with the other 3 standards.
- **Figure below.** Shows diagrammatic form of square lattice structure.



- To produces a square picture avoiding distortion on the screen with 4/3 aspect ratio it is necessary for displaying a square of (N X N) pixels to have :

1. 640 pixels (480 * 4/3) per line, with an NTSC monitor
2. 768 pixels (576 * 4/3) per line, with a European monitor

- **Memory requirements** to store a single digital image can be high, vary between 307.2 Kbytes for an image displayed on a VGA screen with 8bits/pixel through to approximately 2.36Mbytes for a SVGA (Super VGA) screen with 24 bits/pixel as shown in the **table below**.

Standard	Resolution	Number of colors	Memory/frame (bytes)
VGA	640 x 480 x 8	256	307.2kB
XGA	640 x 480 x 8	64k	614.4kB
	1024 x 768 x 8	256	786.432kB
SVGA	800 x 600 x 16	64k	960kB
	1024 x 768 x 8	256	786.432kB
	1024 x 768 x 24	16M	2359.296kB

- Computer monitors of expensive computes since are not based on television picture tubes 4/3 aspect ratio for need not to constrain.
- **Ex**.: 1280 X 1024 X 24 may have the refresh rate as high as 75 frames/sec for sharp image.

Problem: Derive the time to transmit the following digitized images at both 64kbps and 1.5Mbps.
A 640 x 480 x 8 VGA – compatible images,
A 1024 x 768 x 24 SVGA-compatible images

Solution:

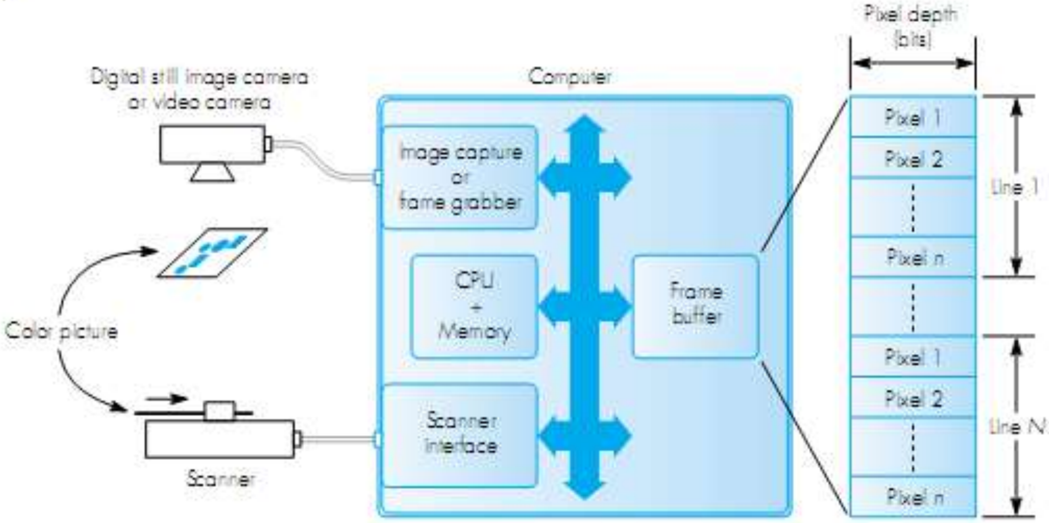
The size of each image in bits is:
VGA = 640 x 480 x 8 = 2.457600 Mbits
SVGA = 1024 x 768 x 24 = 18.874368 Mbits

Time to transmit each image is:
At 64kbps: VAG = $2.4576 \times 10^6 / 64 \times 10^3$ = 38.4 sec
SVGA = $18.874368 \times 10^6 / 64 \times 10^3$ = 294.912 sec

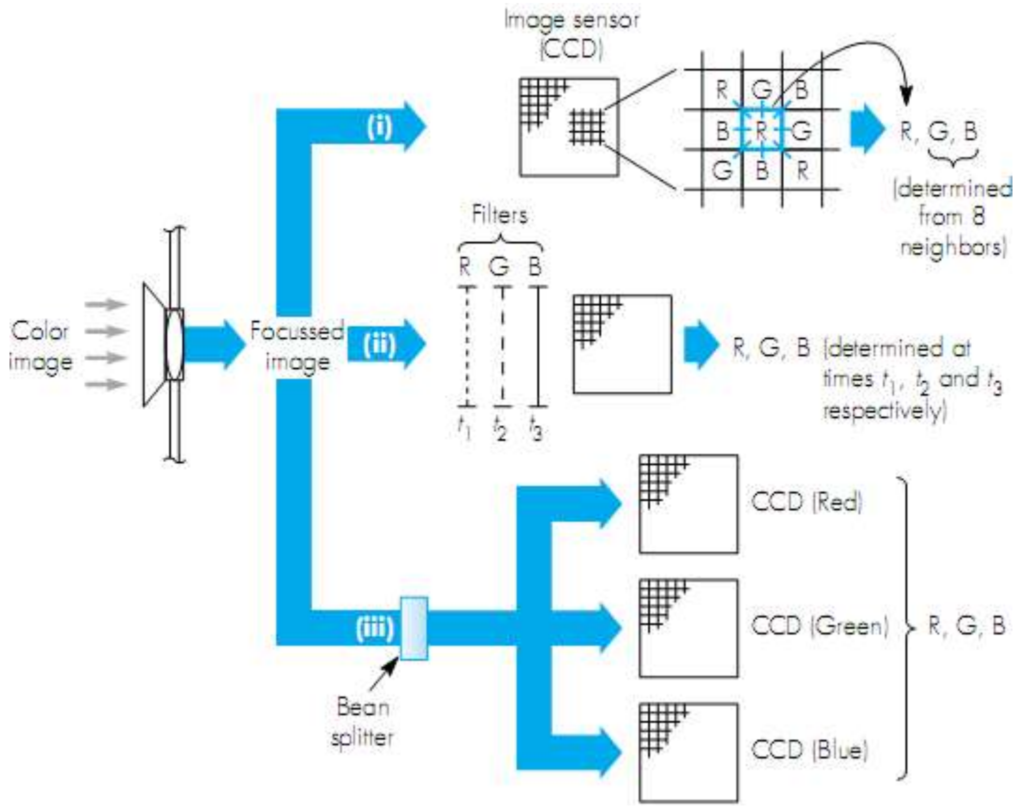
At 1.5Mbps: VAG = $2.4576 \times 10^6 / 1.5 \times 10^6$ = 1.6384 sec
SVGA= $18.874368 \times 10^6 / 1.5 \times 10^6$ = 12.5829 sec

Digital cameras and scanners:

(a)



- **Figure above**. Shows atypical arrangement used to capture and store a digital image produced by a scanner or a digital camera (a still-image camera or a video camera).
- It is assumed captured image is transferred to the computer directly as it is produced or as with digital cameras set of digitized images can be stored within the camera itself then, downloaded into the computer at a later time.
- Image capture within the camera/scanner using solid-state device known as **Image sensor.** It is a 2-D grid of light-sensitive cells known as **photosites**.
- Each photosites stores the level of intensity of the light that falls on it when the camera shutter is activated.
- **CCD (Charge-Coupled Device):** a widely used image sensor comprises of an array of photosites on its surface operates by converting the level of light intensity that falls on each photosite into an equivalent electrical charge.
- The level of charge (i.e., light intensity) are stored at each photosite position is read out then, it is converted into a digital value using an ADC.
- With Scanners CCD like technique is used except the image sensor, comprises a single line of photosites exposed in time-sequence with the scanning operation and each row of stored charges are read out and digitized before the next scan occurs.
- For color images color associated with each photosite and hence the pixel position is obtained in 3 methods as shown in the figure below:



- **Method 1**: Surface of **each photosite coated with either R, G, or B filter** so, that its charge is determined only by the level of R, G, and B lights that falls on it. Coatings are arranged on the 3 X 3 grid structure as in Figure above then, color associated with each photosite/pixel determined by the output of photosite R, G, and B together with each of its 8 immediate neighbors. The levels of two other colors in each pixel are estimated by **interpolation procedure involving all 9 values**.
  - **Application**: most consumer-friendly cameras.

- **Method 2**: It involves **use of 3 separate exposures of a single image sensor** 1st through a Red filter, 2nd through a Green filter, and 3rd through a Blue filter. Color associated with each pixel position is then, determined by the charge obtained with each of 3 filters - R, G, and B.
  - **Application**: used primarily with high-resolution still-image cameras in locations such as photographic studios where, cameras can be attached to a tripod but cannot be used with video cameras since, 3 separate exposures are required for each image.

- **Method 3**: **uses 3 separate image sensors** one with all photosites coated with a red filter, 2nd with a green filter, 3rd with a blue filter. **A single exposure is**

used with the incoming light split into 3 beams each of which exposes a separate image sensor.

- **Application**: in professional-quality-high-resolution still and moving image camera as they are more costly owing to use of the 3 separate sensors and associated signal processing circuits.

- Each image/frame once, captured and stored on the image sensor(s), then charge stored at each photosite location is read and digitized.
- Using CCD set of charges on the matrix of photosites are read a single row at a time.
- Each of the **photosites are read on a row by row basis**.
- Once in **readout register** the charge on each photosite position is shifted out, amplified and digitized using ADC.
- Low-resolution image (640 X 480 pixels) and pixel depth of 24 bits - 8 bits each for R, G, and B amount of memory required to store each image is 921600 bytes.
- If output of this is directed to computer bit-map can be loaded straight into the frame buffer ready to be displayed.
- If required to store within the camera multiple images of this size need to be stored prior to them being output to a computer.
- set of images are stored in an IC memory either on:
  1. removal card
  2. fixed within camera
- Once within the computer, software can be used to insert digital image(s) into a document, sent it by e-mail and so on or photo-editing software can be used to manipulate a stored image. Ex.: to change its size or colors.
- Numbers of file formats are used to store sets of images. One of the most popular is a version of the TIFF called **TIFF for electronic photography (TIFF/EP).**

Audio:

- **Definition1:** sound within the range of human hearing.
- **Definition2:** Audio is sound within the **acoustic range** available to humans.
- An audio frequency is an electrical alternating current within the **20 to 20,000 hertz range** that can be used to produce acoustic sound.
- Two types of audio signals:
  1. **Speech signal**: used in a variety of interpersonal applications like telephony, video telephony.
  2. **Music-quality audio** used in applications such as CD-on-demand and broadcast television.
- Audio can be produced:

1. **Naturally** by means of microphone generates a time-varying analog signal audio signal encoder take analog signals, convert them into digital form to store signals in the memory of the computer and to transmit them over a digital network, at the destination audio signal encoder convert them back into analog form as audio output devices are analogous in nature.
2. Synthesizer **Create audio in the digital form** can be readily stored in the computer memory output.

- BW of a typical **speech signal** is from **50Hz to 10kHz** thus sampling rate of 20ksps (2 X 10 kHz) is used for speech.
- BW of a typical **music signal** is from **15Hz to 20kHz** thus sampling rate of 40ksps (2 X 20 kHz) is used for speech.
- Number of bits/sample must be chosen such that, quantization noise generated by the sampling process is at an acceptable level relative to the minimum signal level.
- Tests have shown that linear (equal) **quantization intervals** for **audio** require use of minimum of **12 bits/sample** and for **music 16 bits/sample**.
- Applications involving music stereoscopic (stereo) sound is utilized (thus 2 such signals must be digitized) this result in bit rate doubles that of a monaural (mono) signal.
- In practice both these sampling rate used and number of bits/sample less than these values.
- **For speech:** BW of the network used in many interpersonal applications is often less than BW of source signal so, lower sampling rate with fewer bits/sample is needed.
- **For music**: sampling rate is lowered to reduce the amount of memory required to store the particular passage of music.

Problem: Assuming the B/W of a speech is from 50 Hz through to 10kHz and that of a music signal is from 15Hz through to 20kHz, derive the bit rate that is generated by the digitization procedure in each case assuming that Nyquist sampling rate is used with 12 bits/sample for the speech signal and 16 bits/sample for the music signal. Derive the memory required to store a 10 minute passage of stereophonic music.

Solution:

i. Bit rates: $Nyquist\ sampling\ rate = 2f_{max}$

Speech: $Nyquist\ rate = 2\ \times 10kHz = 20kHz\ or\ 20ksps$

$$with\ 12\ bit\ per\ sample, bit\ rate\ genetrated = 20k\ \times 12 = 240kbps$$

Music: $Nyquist\ rate = 2\ \times 20kHz = 40kHz\ or\ 40ksps$

$$with\ 16\ bit\ per\ sample, bit\ rate\ genetrated = 40k\ \times 16 = 640kbps\ (mono)$$

$$2\ \times 640k = 1280kbps\ (stereo)$$

ii. Memory required:

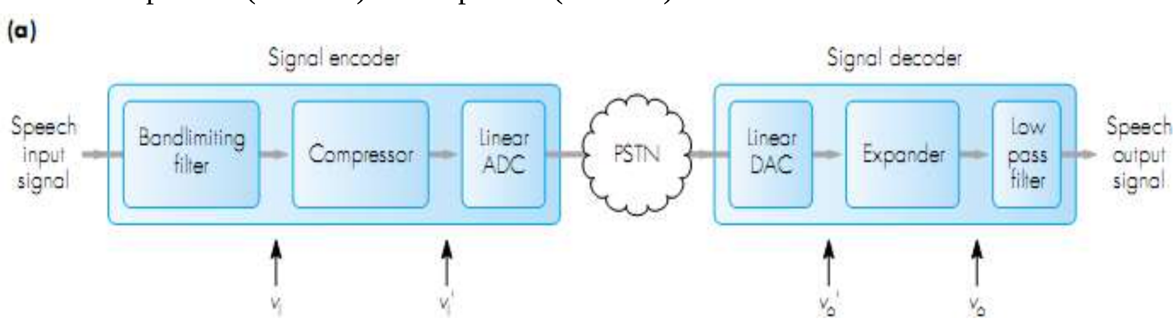
$$\text{Memory required} = {\text{bit rate (bps)} \times \text{time (s)}}\Big/{8\ \text{bytes}}$$

Thus for a bit rate of 1280kbps and for a passage of 10 minutes,

$$Memory\ required = \frac{1280 \times 10^3 \times 600}{8} = 96Mbytes$$

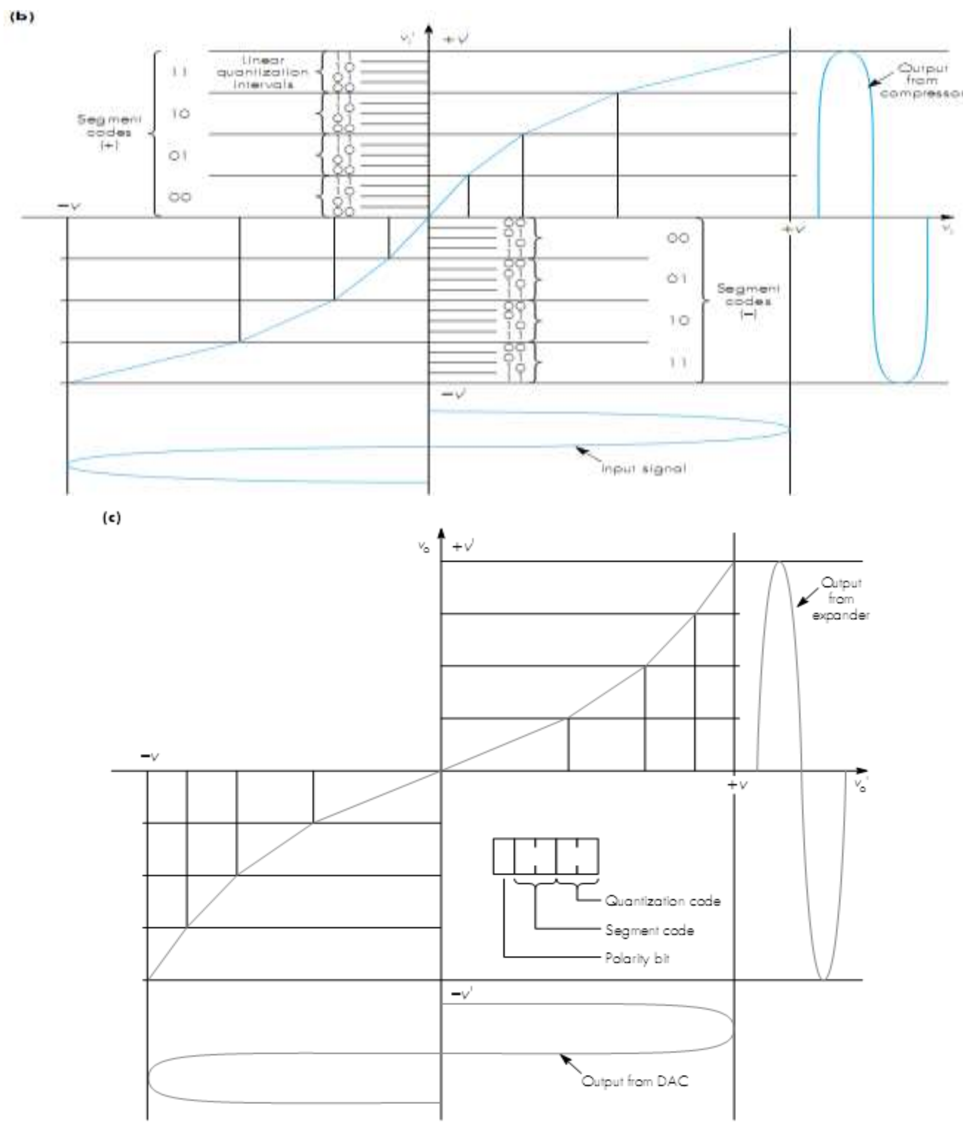
PCM speech:

- Interpersonal applications involving speech uses PSTN for communication purpose.
- Initially PSTN operated with analog signal throughput, the source speech signal being transmitted and switched (routed) unchanged in its original analog form progressively older analog transmission circuits were replaced by digital circuits.
- The above was carried out over a number of years and because of the need to interwork between earlier analog and newer digital equipments during the transition period design of digital equipment was based on operating parameters of the earlier analog network.
- BW of the speech circuit in this network was limited to 200Hz through to 3.4 kHz. Nyquist rate is 6.8 kHz due to poor quality of the band limiting filters used meant that a sampling rate of 8kHz was required to avoid aliasing.
- To minimize the resulting bit rate:
    1. 7 bits/sample were selected for use in North America and Japan with bit rate of 56 kbps.
    2. 8 bits/sample in Europe (both including a sign bit) with bit rate of 64kbps.
- **PCM (Pulse Code Modulation)** is a digitization procedure and the international standard relating to this is ITU-T Recommendation G.711.
- **Figure below**. Shows circuits that make up a PCM encoder and decoder consisting of compressor (encoder) and expander (decoder) circuits.



- Uses linear quantization intervals where quantization intervals are equal irrespective of the magnitude of the input signal same level of quantization noise is produced.
- Effect of above is that noise level is the same for both low amplitude (quite) signals and high amplitude (loud) signals.

- Ear more sensitive to noise on quite signals than it is on loud signals. To reduce the effect of quantization noise with 8 bits/sample, in a PCM system **quantization intervals are made non-linear (unequal)** with narrower intervals used for smaller amplitude signals than, for larger signals these can be achieved by compressor and at the destination, the reverse operation by the expander circuits overall operation is known as **companding.**
- Figures below. Shows input/output relationship of both circuit's characteristic curve for compression - compression characteristics and characteristic curve for expansion - expansion characteristics.



Note that in the G.711 standard a 3-bit segment code and 4-bit quantization code are used.

- Prior to input signal being sampled and converted into a digital form by ADC, it is passed through compression circuit, which effectively compresses the amplitude of the input signal.
- Level of compression and hence, quantization intervals increases as the amplitude of the input signal increases resulting compressed signal is passed through ADC which performs linear quantization of the compressed signal.
- At the destination each received codeword is first fed into a linear DAC Analog output from the DAC is then, passed to the expander circuit which perform the reverse operation of the compressor circuit.
- Modern systems perform compression and expansion operations digitally same principles are applied.
- 2 different compression-expansion characteristics are in use:
    1. **μ-law**: used in North America and Japan.
    2. **A-law**: used in Europe and some other countries.

CD-quality audio:

- **Discs** used in CD players and CD-ROMs are **digital storage devices** for stereophonic music and more general multimedia information streams.
- **CD-DA (CD-digital audio) standard** associated with these devices.
- Music has an audible BW of from 15Hz through 20 kHz. So, minimum sampling rate is **40ksps**.
- In the standard actual rate used is higher this rate because of following reasons:
    1. Allow for imperfections in the band limiting filter used.
    2. So that resulting bit rate is then, compatible with one of the higher transmission channel bit rates available with public networks.
- Sampling rate used one of is 44.1ksps means, signal is sampled at 23 microsecond intervals.
- High number of samples can be used since; BW of a recording channel on a CD is large.
- Recording of stereophonic music requires 2 separate channels - so, total bit rate required is double that for mono.
- Hence, bit rate/channel = sampling rate x bits/sample

  $= 44.1 \times 10^3 \times 16$

  $= 705.6$ kbps (for mono)

Total bit rate = 2 X 705.6 = 1.411 Mbps (for stereo)

- CD-ROMs also uses same bit rate which, are widely used for distribution of multimedia titles. To reduce the access delay multiples of this rate are used within a computers.

- Not feasible to interactively access a 30s portion of a multimedia title over a 64kbps channel with 1.5Mbps channel the time is high for interactive purposes.

Problem: Assume the CD-DA standard is being used, derive:

I. **The storage capacity of a CD-ROM to store a 60 minute multimedia title.**
II. **the time to transmit a 30 second portion of the title using a transmission channel of bit rate: a. 64kbps b. 1.5Mbps**

Solution:

i. The CD-DA digitization procedure yields a bit rate of 1.411Mbps. Thus, storage capacity for 60 minutes is =

= 1.411 x 60 x 60Mb

= 5079.6Mbits or 634.95 Mbytes

ii. One 30 second portion of the title =1.411 x 30 = 42.33 Mbits

Thus time to transmit this data :

At 64kbps = $(42.33 \times 10^6 \div 64 \times 10^3)$

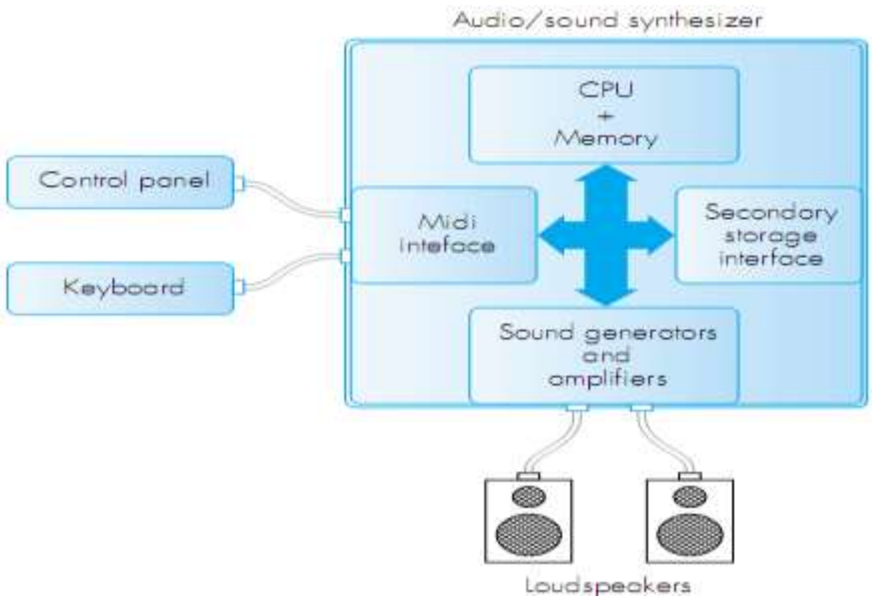
= 661.4 sec

At 1.5kbps = $(42.33 \times 10^6 \div 1.5 \times 10^6)$

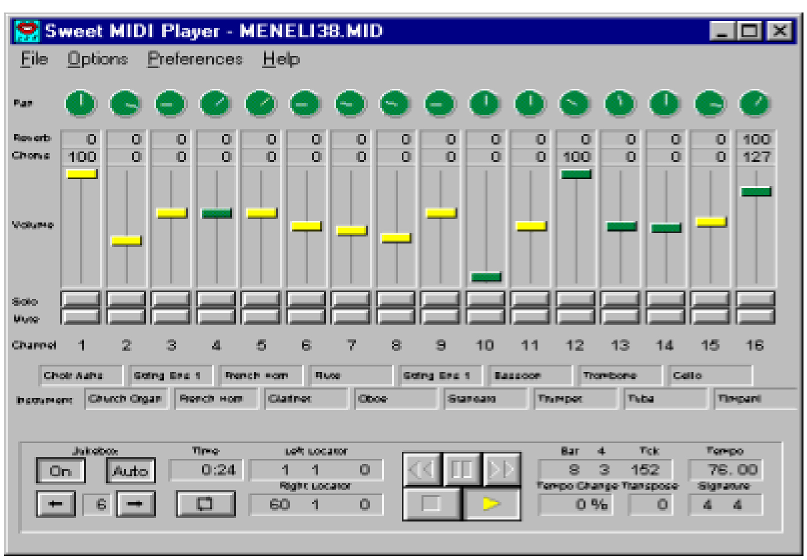
= 28.22 sec

Synthesized audio:

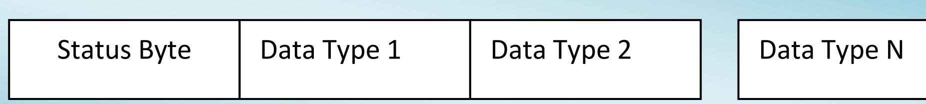
- **Once digitized any form of audio** can be stored within a computer, amount of **memory required** to store a digitized audio waveform can be **very large** even for relatively short passages.
- **Synthesized audio can be defined as** sound generated by electronic signals of different frequencies. Sound can be synthesized by the use of sound synthesizers. The synthesizers use different programmed algorithms to generate sound to different waveform synthesis.
- Synthesized audio is often hence, **used in multimedia applications** since, the amount of memory required to be between two and three orders of magnitude less that required to store the equivalent digitized waveform versions.
- It is **much easier to edit** synthesized audio and mix, several passages together.
- Figure below. Shows components that make up audio synthesizer.



- The 3 main component are:
    1. **Computer**, with various application programs.
    2. **Keyboard**, based on that of a piano.
    3. **Set of sound genera**tors.
- Pressing a key on the keyboard of a synthesizer has similar effect of pressing a key on the keyboard of a computer in asmuch as for each key that is pressed a different codeword known as a message with a synthesizer key board is generated and read by the computer program.
- Messages indicate things such as which key on the keyboard has been pressed and the pressure applied.
- Control panel contains a range of different switches and sliders that collectively allow the user to indicate to the computer program additional information such as the volume of the generated output and selected sound effects to be associated with each key.



- Storage interface allows sequence of messages including those associated with the control panel relating to a particular piece of audio to be saved on secondary storage such as a floppy disk.
- There are programs to allow the use to edit a previously entered passage and if required, to mix the several store passages together. Synthesizer then, ensures that the resulting integrated sequence of messages are synchronized and output to the sound generators to create the merged passage.
- Other possible input instruments other than (piano) keyboard are electric guitar type, all of which generate messages similar to those produced by the keyboard.
- To discriminate between different possible sources a standardized set of messages have been defined for both input and for output to the corresponding set of sound generators defined in the standard known as **MIDI (Music Interface Digital Interface)**.
- **MIDI is a data communications protocol** that describes a means for music systems and related equipment to exchange information and control signals.
- MIDI is not audio signal.
- MIDI define the format of the standardized set of messages used by both synthesizer and the connectors, cables, and by electrical signals used to connect any type of device to the synthesizer.
- Format of message consists of a status byte which defines particular event that has caused the message to be generated, followed by number of data types which collectively define a set of parameters associated with the event.



MIDI Message Format

- Ex.: event is a key being pressed on the board and typical parameters would then be the identity of the key, pressure applied, and so on.
- There can be a variety of instrument types used for input and output. It necessary to identify type of instrument that generated the event so that, when a corresponding message is output to the sound generators appropriate type of sound is produced.
- Different types of devices have a MIDI codes associated with them.
- Ex.: piano has a code of 0 and violin has a code of 40.
- Some codes have been assigned for specific special effects such as, sound from a cannon and applause from an audience.

- Passage of audio produced from the synthesizer consists of a very compact sequence of messages each comprising of bytes which can be played by sequencer program directly heard by the composer or, saved in a file on a floppy disk.
- Typically, in many interactive applications involving multimedia pages comprising text and a passage of music, a synthesizer is first used to create the passage of music which, is then saved in the file. Author of pages links the file contents to the text at the point where music to be played.
- **Sound card needed in the client computer to interpret the sequence of messages** and to generate the appropriate sounds since, the music is in the form of a sequence of MIDI messages.
- **Sound generators use either**:
    1. FM synthesis techniques or
    2. Wavelet synthesis by using samples of sound produced by a real instruments.

MIDI Audio:

- Definition: A protocol that enables computers, synthesizers, keyboards, and other musical devices to communicate with each other.
- MIDI (Musical Instrument Digital Interface) is a communications standard developed in the early 1980s for electronic musical instruments and computers.
- It allows music and sound synthesizers from different manufacturers to communicate with each other by sending messages along cables connected to the devices.
- MIDI provides a protocol for passing detailed descriptions of a musical score, such as the notes, the sequences of notes, and the instrument that will play these notes.
- MIDI data is not digitized sound; it is a shorthand representation of music stored in numeric form.
- Digital audio is a recording, but MIDI is a score.
- The first depends on the capabilities of your sound system, the other on the quality of your musical instruments and the capabilities of your sound system.
- A MIDI file is a list of time-stamped commands that are recordings of musical actions (the pressing down of a piano key or a sustain pedal, or the movement of a control wheel or slider). When sent to a MIDI playback device, this results in sound.
- A concise MIDI message can cause a complex sound or sequence of sounds to play on an instrument or synthesizer; so MIDI files tend to be significantly smaller (per second of sound delivered to the user) than equivalent digitized waveform files.
- The process of creating MIDI music is quite different from digitizing existing recorded audio.
- MIDI music is quite different from digitizing existing recorded audio. If you think of digitized audio as analogous to a bitmapped graphic image (both use sampling of the

original analog medium to create a digital copy), then MIDI is analogous to structured or vector graphics (both involve instructions provided to software to be able to re-create the original on the fly).

- For digitized audio you simply play the audio through a computer or device that can digitally record the sound. To make MIDI scores, however, you will need notation software (Figure A), sequencer software (Figure B), and a sound synthesizer (typically built into the software of multimedia players in most computers and many handheld devices). A MIDI keyboard is also useful for simplifying the creation of musical scores.

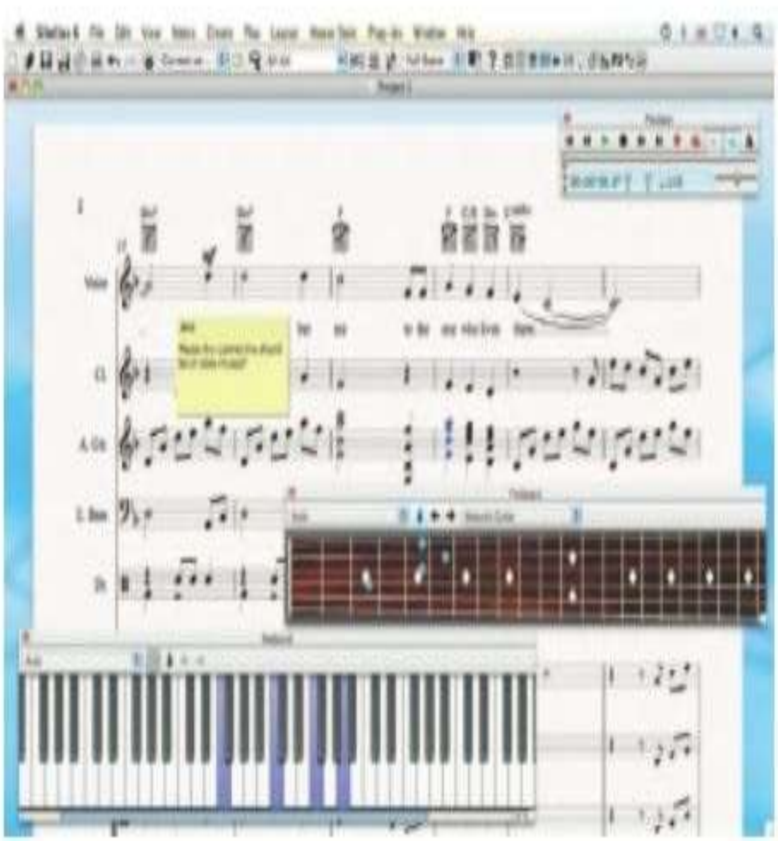


Figure A Notation and composition software such as Sibelius provides a way for composers and musicians to create and arrange scores using MIDI instruments.

Figure B Sequencer software such as ProTools allows you to record, edit, and save music generated from a MIDI keyboard or instruments and blend it with digital audio.

- Rather than recording the sound of a note, MIDI software creates data about each note as it is played on a MIDI keyboard (or another MIDI device)—which note it is, how much pressure was used on the keyboard to play the note, how long it was sustained, and how long it takes for the note to decay or fade away.
- **Ex.:** This information, when played back through a MIDI device, allows the note to be reproduced exactly. Because the quality of the playback depends upon the end user's MIDI device rather than the recording, MIDI is device dependent. The sequencer software quantizes your score to adjust for timing inconsistencies (a great feature for those who can't keep the beat), and it may also print a neatly penned copy of your score to paper.
- An advantage of structured data such as MIDI is the ease with which you can edit the data.
- **Scenario:** Let's say you have a piece of music being played on a honky-tonk piano, but your client decides he wants the sound of a soprano saxophone instead. If you had the music in digitized audio, you would have to re-record and redigitize the music. When it is in MIDI data, however, there is a value that designates the instrument to be used for playing back the music. To change instruments, you just change that value.
- Instruments that you can synthesize are identified by a General MIDI numbering system that ranges from 0 to 127 as shown in the table below

TABLE 3 General MIDI Instrument Sounds Listing

Prog# Instrument name	Prog# Instrument Name	Prog# Instrument Name	Prog# Instrument Name
1. Acoustic Grand Piano	33. Acoustic Bass	65. Soprano Sax	97. FX 1 (rain)
2. Bright Acoustic Piano	34. Elec. Bass (finger)	66. Alto Sax	98. FX 2 (soundtrack)
3. Electric Grand Piano	35. Elec. Bass (pick)	67. Tenor Sax	99. FX 3 (crystal)
4. Honky–tonk Piano	36. Fretless Bass	68. Baritone Sax	100. FX 4 (atmosphere)
5. Electric Piano 1	37. Slap Bass 1	69. Oboe	101. FX 5 (brightness)
6. Electric Piano 2	38. Slap Bass 2	70. English Horn	102. FX 6 (goblins)
7. Harpsichord	39. Synth Bass 1	71. Bassoon	103. FX 7 (echoes)
8. Clavi	40. Synth Bass 2	72. Clarinet	104. FX 8 (Sci–fi)
9. Celesta	41. Violin	73. Piccolo	105. Sitar
10. Glockspspiel	42. Viola	74. Flute	106. Banjo
11. Music Box	43. Cello	75. Recorder	107. Shamisen
12. Vibraphone	44. Contrabass	76. Pan Flute	108. Koto
13. Marimba	45. Tremolo Strings	77. Blown Bottle	109. Kalimba
14. Xylophone	46. Pizzicato Strings	78. Shakuhachi	110. Bag pipe
15. Tubular Bells	47. Orchestral Harp	79. Whistle	111. Fiddle
16. Dulcimer	48. Timpani	80. Ocarina	112. Shanai
17. Drawbar Organ	49. String Ensemble 1	81. Lead 1 (square)	113. Tinkle Bell
18. Percussive Organ	50. String Ensemble 2	82. Lead 2 (sawtooth)	114. Agogo
19. Rock Organ	51. SynthStrings 1	83. Lead 3 (calliope)	115. Steel Drums
20. Church Organ	52. SynthStrings 2	84. Lead 4 (chiff)	116. Wood block
21. Reed Organ	53. Choir Aahs	85. Lead 5 (charang)	117. Taiko Drum
22. Accordion	54. Voice Oohs	86. Lead 6 (voice)	118. Melodic Tom
23. Harmonica	55. Synth Voice	87. Lead 7 (fifths)	119. Synth Drum
24. Tango Accordion	56. Orchestra flit	88. Lead 8 (bass+lead)	120. Reverse Cymbal
25. Acoustic Guitar (nylon)	57. Trumpet	89. Pad 1 (new age)	121. Guitar Fret Noise
26. Acoustic Guitar (steel)	58. Trombone	90. Pad 2 (warm)	122. Breath Noise
27. Electric Guitar (jazz)	59. Tuba	91. Pad 3 (polysynth)	123. Seashore
28. Electric Guitar (clean)	60. Muted Trumpet	92. Pad 4 (choir)	124. Bird T'weet
29. Electric Guitar (muted)	61. French Horn	93. Pad 5 (bowed)	125. Telephone Ring
30. Overdrive Guitar	62. Brass Section	94. Pad 6 (Metallic)	126. Helicopter
31. Distortion Guitar	63. SynthBrass 1	95. Pad 7 (halo)	127. Applause
32. Guitar harmonics	64. SynthBrass 2	96. Pad 8 (sweep)	128. Gunshot

Video:

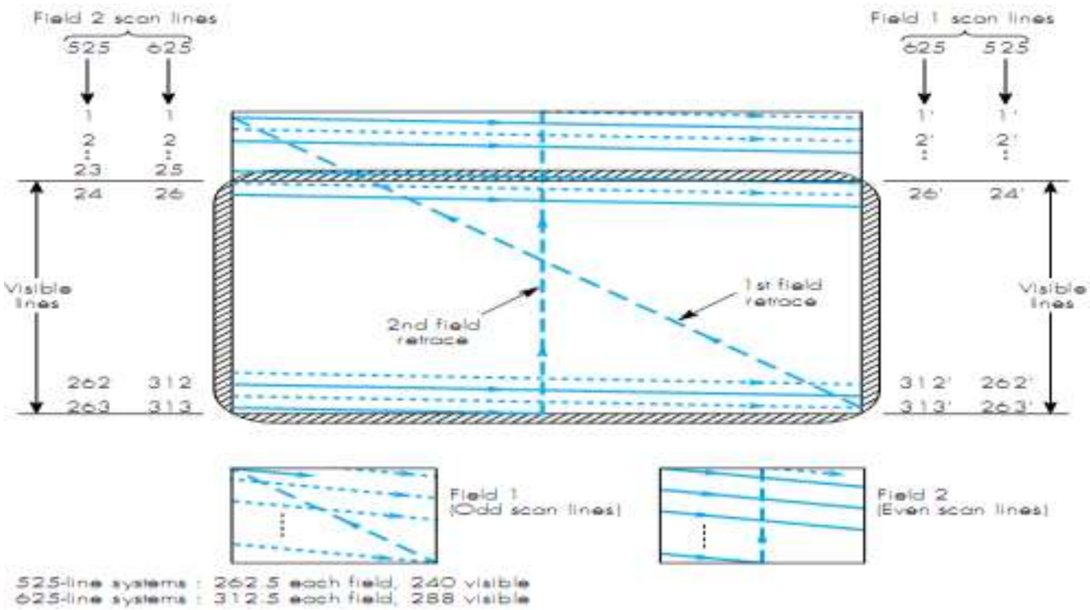
- **Definition**: Correlated sequence of images with respect to time.
- Futures in range of multimedia application.
    1. **Entertainment**: broadcast television and vcr/ dvd recordings;
    2. **Interpersonal**: video telephony and video conferencing;
    3. **Interactive**: window containing short video clips.
- The quality of video required, however varies considerably from one application to another.
- Ex.: in case of video telephony, a small window on the screen of a PC is acceptable while for a movie, a large screen format is preferable.
- In practice there is not just a single standard associated with video but set off standards each targeted at a particular video application.
- All the different standards are based on basic principles that are associated with Broadcast television.

Broadcast television:

- A color picture/image is produced from varying mixtures of the three primary colors red, green and blue. The screen of the picture tube is coated with a set of three different phosphors one for each color, each of which is activated by a separate electron beam.
- The three electron beams are scanned in unison across the screen from left to right with a resolution of either 525 lines (NTSC) or 625 lines (PAL/CCIR/SECAM). The total screen contents are then refreshed at a rate of either 60 or 50 frames per second respectively, the rate being determined by the frequency of the mains electricity supply used in the different countries.
- The computer monitors uses the same picture tubes as those in broadcast television receivers and hence operate in a similar way.
- The three digitized color signals that make up a stored picture/image are read from the computer memory in time-synchronism with the scanning operation of the display tube and, after each complete scan of the display, the procedure repeats so producing a flicker-free color image on the screen.
- **Practically Broadcast television operates slightly different than the above logic** in terms of :
    1. Scanning sequence used
    2. Choice of color signal

Scanning Sequence:

- It is necessary to use a minimum refresh rate of 50 times per second to avoid flicker, but to produce a smooth motion, a refresh rate of 25 times per second is sufficient.
- In order to minimize the amount of transmission bandwidth that is required to broadcast the television signal, the characteristic of the eye is exploited by transmitting the image/picture associated with each frame in two halves.
- Each is known as a **field**, the first comprising only the odd scan lines and the second the even scan lines.
- The two fields are then integrated together in the television receiver using a technique known as **interlaced scanning**, the principles of which are shown in Figure below
    1. In a 525-line system each field comprises 262.5 lines out of which 240 are visible.
    2. In a 625-line system each field comprises 312.5 lines out of which 288 are visible.



- Each field is refreshed alternately at 60/50 fields per second and hence the resulting frame refresh rate is only 30/25 frames per second.
- In this way, a refresh rate equivalent to 60/50 frames per second is achieved but with only half the transmission bandwidth.

Color signals:

- **To support backward compatibility**, the received signals associated with a color television broadcast had to be such that they could be used by an existing (unmodified) monochrome (black-and-white) television set to produce the same picture in high-quality monochrome also a color television had to be able to produce black-and-white pictures from monochrome broadcasts..
- For above reasons a different set of color signals from R, G, and B were selected for color television broadcasts.

Property of Color Source That The Eye Make of Are:

1. **Brightness**: It represents the amount of energy that stimulates the eye and varies on a gray scale from black (lowest) through to white. It is independent of the color of the source.
2. **Hue**: It represents the actual color of the source, each color has a frequency/wavelength and the eye determines the color from that.
3. **Saturation**: this represents the strength or vividness of the color.

- Term **luminance** is used to refer to the **brightness of a source**.
- Term **chrominance** is used to refer to the **hue and saturation**, because they are concerned with its color.

- A range of colors can be produced on a television display screen by varying the magnitude of the three electrical signals that energize the red, green, and blue phosphors.
- If the magnitude of the three signals are in the proportion **0.299R+ 0.587G+ 0.114B** then **white color** is produced on the display screen.
- the luminance of a source is only a function of the amount of white light it contains, for any color source its luminance can be determined by summing together the three primary components that make up the color in this proportion i.e.,

$$Ys = 0.299Rs + 0.587Gs + 0.144Bs$$

- Here
  - Ys = amplitude of the luminance signal.
  - Rs, Gs, and Bs = magnitudes of the three color component signals that make up the source.
- From above analysis:
  - Luminance signal is a measure of the amount of white light it contains; it is the same as the signal used by a monochrome television.
  - Two other signals, the blue chrominance (Cb), and the red chrominance (Cr) are then used to represent the coloration hue and saturation of the source.
- hue and saturation of the source are obtained from the two color difference signals:

$$Cb = Bs - Ys \qquad \text{And} \qquad Cr = Rs - Ys$$

- As Y signal has been subtracted in both cases, above contain no brightness information.
- As Y is a function of all three colors, then G can be readily computed from these two signals.
- In this way, the combination of the three signals Y, Cb, and Cr contains all the information that is needed to describe a color signal while at the same time being compatible with monochrome televisions which use the luminance signal only.

Chrominance Components:

- All color television systems use this same basic principle to represent the coloration of a source; there are some small differences between the two systems in terms of the magnitude used for the two chrominance signals.
- The reason for the above is due to the constraint that the bandwidth of the transmission channel for color broadcasts must be the same as that used for monochrome.

- Thus, in order to fit the Y, Cb and Cr signals in the same bandwidth, the three signals must be combined together for transmission. The resulting signal is then known as the **composite video signal**.
- If the two color difference signals are transmitted at their original magnitudes, the amplitude of the luminance signal can become greater than that of the equivalent monochrome signal. This leads to a degradation in the quality of the monochrome picture and hence is unacceptable.
- Solution:
  1. To overcome this effect, the magnitude of the two color difference sign are both scaled down.
  2. Since they both have different levels of luminance associated with them, the scaling factor used for each signal is different.
- The two color difference signals are referred to by different symbols in each system:
- **PAL system**: C_b, and C_r are referred to as U and V respectively, scaling factors used for the three signals are:

$$Y = 0.299R + 0.587G + 0.114B$$
$$U = 0.493\ (B - Y)$$
$$V = 0.877\ (R - Y)$$

- **NTSC system**: two color difference signals are combined to form two different signals referred to as I and Q, scaling factor used are:

$$Y = 0.299R + 0.587G + 0.114B$$
$$I = 0.74(R - Y) - 0.27(B - Y)$$
$$Q = 0.48(R - Y) + 0.41(B - Y)$$

Problem: Derive the scaling factor used for both the U and V (as used in PAL) and I and Q (as used in NTSC) color difference signals in terms of the three R,G,B color signals.
Solution:

PAL:

$$Y = 0.299R + 0.587G + 0.114B$$
$$U = 0.493\ (B - Y)$$
$$V = 0.877\ (R - Y)$$

Thus, $U = 0.493\,B - 0.493\,(0.299R + 0.587G + 0.114B)$

$= -0.147R - 0.289G + 0.437B$

and $V = 0.877\,R - 0.877(0.299R + 0.587G + 0.114B)$

$= 0.615R - 0.515G - 0.100B$

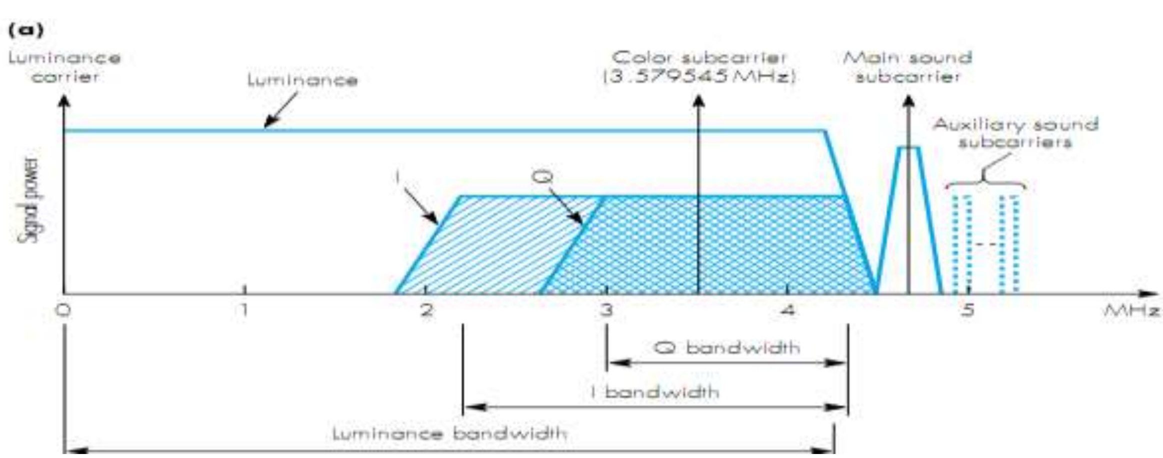
NTSC:

$$Y = 0.299R + 0.587G + 0.114B$$
$$I = 0.74(R - Y) - 0.27(B - Y)$$
$$Q = 0.48(R - Y) + 0.41(B - Y)$$

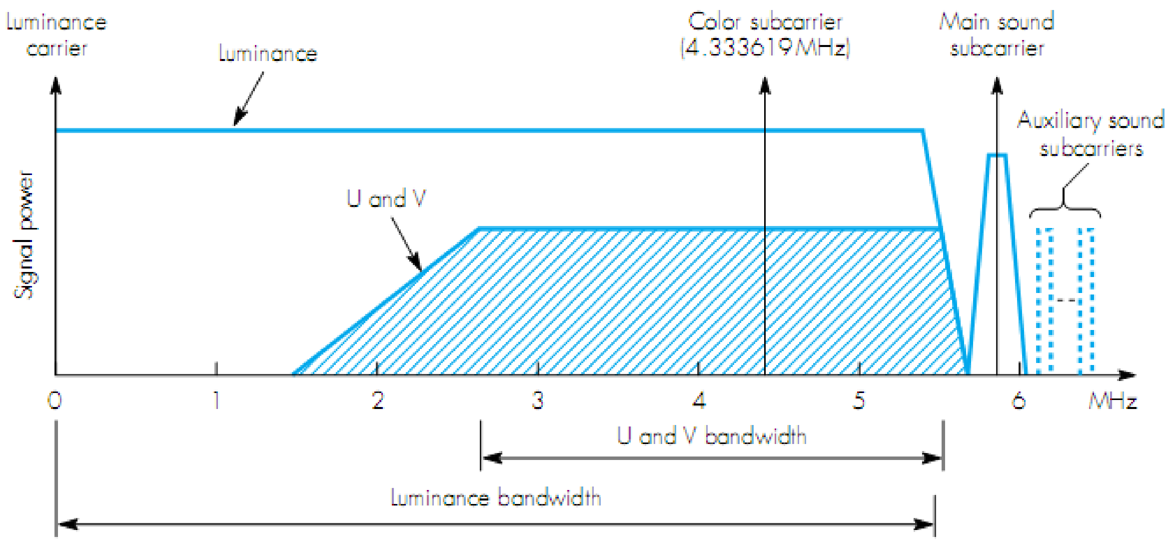
Thus,
$$I = 0.74(R - Y) - 0.27\,(B - Y)$$
$$= 0.74R - 0.27B - 0.47Y$$
$$= 0.599R - 0.276G - 0.324B$$
$$Q = 0.48(R - Y) + 0.41\,(B - Y)$$
$$= 0.48R + 0.41B - 0.89Y$$
$$= 0.212R - 0.528G + 0.311B$$

Signal bandwidth:

- The **bandwidth** of the transmission channel used for **color broadcasts** must be the **same as** that used for **a monochrome broadcast**.
- Thus, for transmission, the two chrominance signals must occupy the same bandwidth as that of the luminance signal.
- Most of the energy associated with the luminance signal is in the lower part of its frequency spectrum.
- Thus, in order to minimize the level of interference between the luminance and two chrominance signals following steps are followed:
    1. The chrominance signals are transmitted in the upper part of the luminance frequency spectrum using two separate subcarriers.
    2. To restrict the bandwidth used to the upper part of spectrum, a smaller bandwidth is used for both chrominance signals.
    3. Both of the two chrominance subcarriers have the same frequency they are 90 degrees out of phase with each other.
    4. Each is modulated independently in both amplitude and phase by the related chrominance signal.
- Using this technique, the two signals can use the same portion of the luminance frequency spectrum.
- The baseband spectrum of a color television signal in NTSC systems is shown in Figure below.



- **NTSC system**: the eye is more responsive to the I signal than the Q signal. Hence to maximize the use of the available bandwidth while at the same time minimizing the level of interference with the luminance signal, the I signal has a modulated bandwidth of about 2MHz and the Q signal a bandwidth of about I MHz.
- The baseband spectrum of a color television signal in PAL systems is shown in Figure below.



- **PAL system**: the larger luminance bandwidth about 5.5 MHz relative to 4.2 MHz allows both the U and V chrominance signals to have the same modulated bandwidth which is about 3 MHz.
- As show in the figures above, the audio/sound signal is transmitted using one or in separate subcarriers which are all just outside of the luminance signal bandwidth. Typically, the main audio subcarrier is for mono sound and the auxiliary subcarriers are used for stereo sound.
- When above signals are added to the baseband video signal, the composite signal is called the **complex baseband signal**.

Digital video:

- In most multimedia applications the video signals need to be digital form since it then becomes possible to store them in the memory of the computer and to readily edit and integrate them with other media type.
- For transmission reasons the three component signals have to be combined for analog television broadcasts, with digital television it is more usual to digitize the three component signals separately prior to transmission.
- The above is done to enable editing and other operations readily performed.

- Since the three component signals are treated separately in digital transmission, in principle it is possible simply to digitize the three RGB signals make up the picture.
- **Disadvantage of this approach**:
  - Same resolution in terms of sampling rate and bits per sample must be used three signals i.e., RGB.
- Visual perception of the eye is less sensitive for color than it is for luminance this means that the two chrominance signals can tolerate a reduced resolution relative to that used for the luminance signal.
- Thus by using luminance and two color difference signals instead of the RGB signals we can achieve significant saving in terms of:
  1. Resulting bit rate.
  2. Transmission bandwidth.
- Digitization of video signals has been carried out in television studios for many years like to perform conversions from one video format into another.
- To standardize this process: **The International Telecommunications Union Radiocommunications Branch (ITU-R)** formerly known as the **Consultative Committee for International Radiocommunications (CCIR)** defined a standard for the digitization of video pictures known as **Recommendation CCIR-601**.
- A number of variants of this standard have been defined for use in other application domains such as digital television broadcasting, video telephony, and videoconferencing. Collectively these are known as **digitization formats**.
- They all exploit the fact that the two chrominance signals can tolerate a reduced resolution relative to that used for the luminance signal.

4:2:2 Format:

- This is the original digitization format used in **television studios**.
- **Noninterlaced scanning** is used.
- The three component (analog) video signals from a source in the studio can have bandwidths of up to 6 MHz for the luminance signal and less than half this for the two chrominance signals.
- To digitize these signals, it is necessary to use band-limiting filters of 6 MHz for the luminance signal and 3 MHz for the two chrominance signals with a minimum sampling rate of 12MHz (12 Msps) and 6 MHz respectively.
- In the standard, however, a line sampling rate of **13.5 MHz** for luminance and **6.75 MHz** for the two chrominance signals was selected, both of which are independent of the particular scanning standard - NTSC, PAL and so on being used.

- The 13.5MHz rate was chosen since it is the nearest frequency to 12 MHz which results in a **whole number of samples per line** for both 525 and 625 line systems.
- The number of samples per line chosen is 702 and can be derived as follows:

- For 525-line system:

$$\text{the total line sweep time} = 63.56 \text{ microseconds}$$
$$\text{retrace time (electron beam is turned off)} = 11.56 \text{ microseconds}$$
$$\text{Thus active line sweep time} = 52 \text{ microseconds}$$

- For 625-line system:

$$\text{the total line sweep time} = 64 \text{ microseconds}$$
$$\text{retrace time (electron beam is turned off)} = 12 \text{ microseconds}$$
$$\text{Thus active line sweep time} = 52 \text{ microseconds}$$

- In both cases, a sampling rate of 13.5MHz yields:

$$52 \times 10^{-6} \times 13.5 \times 10^{6} = 702 \; samples \; per \; line$$

- In practice, the number of **samples per line is increased to 720(for luminance)** by taking a slightly longer active line time which results in a small number of black samples at the beginning and end of each line for reference purposes. The corresponding number of samples for each of the **two chrominance signals** is set at half this value; that is, **360 samples per active line**.
- This result in 4Y samples for every $2C_b$ and $2C_r$, samples which is the origin of the term 4:2:2, the normally indicating the digitization is based on the R, G, B signals.
- The **number of bits per sample** was chosen to be **8** for all three signals which correspond to 256 quantization intervals.
- The **vertical resolution** for all three signals was also chosen to be the same, the number being determined by the scanning system in use;
    1. With 525-line system (NTSC): 480 lines (number of active/visible lines).
    2. With 625-line system (PAL) : 576 lines (number of active/visible lines).
- Thus total resolution is:
    - ❖ 525 line system:

$$Y = 720 \times 480$$
$$Cb = Cr = 360 \times 480$$

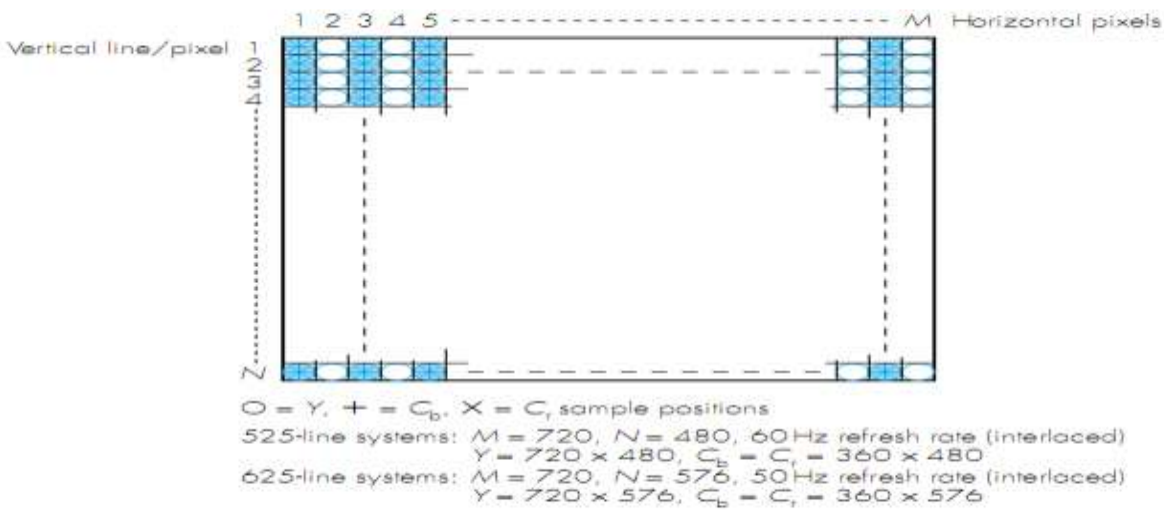
    - ❖ 625 line system:

$$Y = 720 \times 576$$
$$Cb = Cr = 360 \times 576$$

- Bit rate in both the system with this format is:

$$13.5 \times 10^{6} \times 8 + 6.75 \times 10^{6} \times 8 + 6.75 \times 10^{6} \times 8 = 216 \; Mbps$$

- **Frame refresh rate selected**:
    1. With 525-line system: 60Hz
    2. With 625-line system: 50Hz
- Since each line is sampled at a constant rate (13.5 and 6.75 MHZ) with a fixed number of samples per line (720 and 360), the **samples for each line are in a fixed position which repeats from frame to frame**. The samples are then said to be orthogonal and the sample method **orthogonal sampling**.
- the sampling positions for each of the three signals relative to a rectangular grid are as shown in Figure below:



Problem: Derive the bit rate and the memory requirements to store each frame that result from the digitization of both a 525 and 625 -line system assuring a 4:2:2 format. Also find the total memory required to store a 1.5 hour movie/video.

Solution

525-line system: The number of samples per line is 720 and the number of visible lines is 480. Hence the resolution of the luminance (Y) and two chrominance {C_b and C_r) signals are:

Y= 720 x 480

C_b= C_r= 360 x 480

Bit rate: Line sampling rate is fixed at 13.5 MHz for Y and 6.75 MHz for both C_b and C_r all with 8 bits per sample.

Thus,

Bit rate = $13.5 \times 10^6 \times 8 + 2(6.75 \times 10^6 \times 8)$ = 216Mbps.

Memory required: Memory required per frame = 720 x 480 x 8 + 360 x 8 + 360 x 8

= 11520 bits or 1440 bytes

Hence memory per frame, each of 480 lines = 480 x 11520

= 5.5296Mbits or 69l.2kbytes

And memory to store 1.5 hours assuming 60 frames per second:

= 691.2 x 60 x 1.5 x 3600kbytes

= 1.7915 Tbits or 223.9488Gbytes

625-line system: The number of samples per line is 720 and the number of visible lines is 576. Hence the resolution of the luminance (Y) and two chrominance {C_b and C_r) signals are:

$$\mathbf{Y = 720 \times 576}$$

$$\mathbf{C_b = C_r = 360 \times 576}$$

Bit rate: Line sampling rate is fixed at 13.5 MHz for Y and 6.75 MHz for both C_b and C_r all with 8 bits per sample.

Thus,

Bir rate = 13.5×10^6 x8 + 2 (6.75 x 10^6 x 8) = 216Mbps.

Memory required: Memory required per line = 720 x 8 + 2 (360 x 8)

= 11520 bits or 1440 bytes

Hence memory per frame, each of 576 lines = 576 x 11520

= 6.63555Mbps or 829.44kbytes

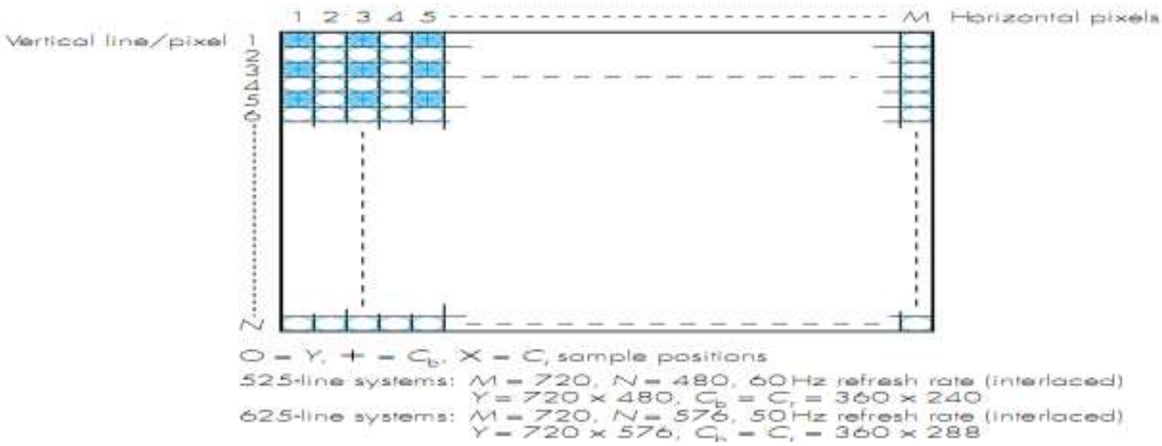
and memory to store 1.5 hours assuming 50 frames per second:

= 829.44 x 50 x 1.5 x 3600kbytes

= 1.7915 Tbits or 223.9488Gbytes

4:2:0 format

- **Derivative of the 4:2:2** format and is used in digital video broadcast applications.
- Gives good picture quality.
- Uses the same set of chrominance samples for two consecutive lines.
- Interlaced scanning is used since it is intended for broadcast applications.
- Due to the absence of chrominance samples in alternative lines is origin of the term 4:2:0.
- The position of the three sample instants per frame are as shown in Figure below.



- Thus it has same luminance resolution as the 4:2:2 format but half the chrominance resolution:
  - 525 line system:

$$Y = 720 \times 480$$
$$Cb = Cr = 360 \times 240$$

  - 625 line system:

$$Y = 720 \times 576$$
$$Cb = Cr = 360 \times 288$$

- Bit rate in both the system with this format is:

$$13.5 \times 10^6 \times 8 + 3.375 \times 10^6 \times 8 + 3.375 \times 10^6 \times 8 = 162\ Mbps$$

- Above value is the worst-case bit rate since it includes samples during the retrace times when the beam is switched off.
- To avoid flicker effects with the chrominance signals, the receiver uses the same chrominance values from the sampled lines for the missing lines.
- With large-screen televisions, flicker effects are often reduced further by the receiver storing the incoming digitized signals of each field in a memory buffer.

Problem: Derive the bit rate and the memory requirements to store each frame that result from the digitization of both a 525 and 625-line system assuring a 4:2:0 format. Also find the total memory required to store a 1.5 hour movie/video. 8 marks

Solution

525-line system: The number of samples per line is 720 and the number of visible lines is 480. Hence the resolution of the luminance (Y) and two chrominance {C_b and C_r) signals are:

Y= 720 x 480

C_b= C_r= 360 x 240

Bit rate: Line sampling rate is fixed at 13.5 MHz for Y and 3.375MHz for both C_b and C_r all with 8 bits per sample.

Thus,

Bit rate = 13.5 x 10^6 x 8 +2 (3.375 x 10^6 x 8) = 162Mbps.

Memory required:

Memory required per frame for Y = 720 x 480 x 8 = 2.7648 Mbits

Memory required per frame for $C_b + C_r =$ 360 x 240 x 8 x 2 = 1.3824 Mbits

Total Memory required per frame = Y + C_b + C_r

= 2.7648M + 1.3824M

= 4.1472 Mbits or 518.4 kbytes

And memory to store 1.5 hours assuming 60 frames per second:

= 518.4 kbytes x 60 x 1.5 x 3600

= 1.3436 Tbits or 167.961 Gbytes

625-line system: The number of samples per line is 720 and the number of visible lines is 572. Hence the resolution of the luminance (Y) and two chrominance {C_b and C_r) signals are:

Y = 720 x 576

C_b= C_r = 360 x 288

Bit rate: Line sampling rate is fixed at 13.5 MHz for Y and 3.375 MHz for both C_b and C_r all with 8 bits per sample.

Thus,

Bit rate = 13.5 x 10^6 x8 + 2 (3.375 x 10^6 x 8) = 162Mbps.

Memory required:

Memory required per frame for Y = 720 x 576 x 8 = 3.317 Mbits

Memory required per frame for $C_b + C_r =$ 360 x 288 x 8 x 2 = 1.65Mbytes

Total Memory required per frame = Y + C_b + C_r

= 3.317M + 1.65M

= 4.967 Mbits or 620.875 kbytes

And memory to store 1.5 hours assuming 50 frames per second

= 620.875 k x 50 x 1.5 x 3600

= 1.3436 Tbits or 167.961 Gbytes

HDTV formats:

- There is a number of alternative digitization formats associated with high-definition television (HDTV).
- The resolution of those which relate to the older 4/3 aspect ratio tubes can be up to 1440 x 1152 pixels and the resolution of those which relate to the newer 16/9 wide-screen tubes can be up to 1920 x 1152 pixels.
- In both cases, the number of visible lines per frame is 1080.
- Both use either the 4:2:2 digitization format for studio applications or the 4:2:0 format for broadcast applications.
- The corresponding frame refresh rate is either 50/60Hz with the 4:2:2 format or 25/30 Hz with the 4:2:0 formats.
- The resulting worst case bit rates are four times the values derived in the previous two sections and proportionally higher for the wide-screen format.

SIF format:

- Has been found to give a **picture quality** comparable with that obtained with **video cassette recorders** (VCRs).
- It uses **half the spatial resolution** in both horizontal and vertical directions as that used in the 4:2:0 format a technique known as **subsampling**.
- Uses **half the refresh rate** as that used in the 4:2:0 format known as **temporal resolution**.
- Thus **frame refresh rate** is:
    1. For 525-line system: 30Hz
    2. For 625-line system: 25Hz
- Thus **total resolution** is:
    - ❖ 525 line system:

$$Y = 360 \times 240$$
$$Cb = Cr = 180 \times 120$$

    - ❖ 625 line system:

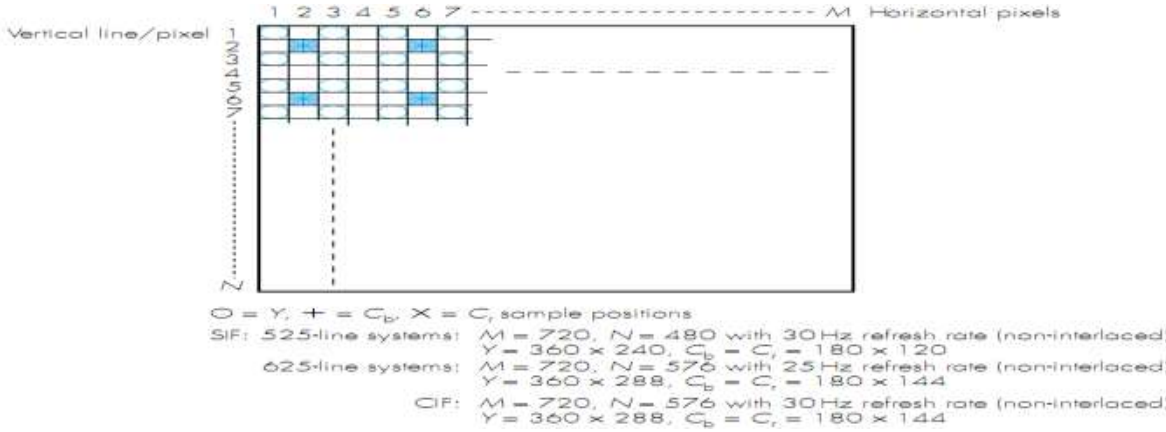
$$Y = 360 \times 288$$
$$Cb = Cr = 180 \times 144$$

- The **worst-case bit rate** in both systems with this format is:

$$6.75 \times 10^6 \times 8 + 1.685 \times 10^6 \times 8 + 1.685 \times 10^6 \times 8 = 81\, Mbps$$

- At the receiver, the missing samples are estimated by interpolating between each pair of values that are sent.

- This digitization format is known as **4:1:1**.
- Intended for storage applications, **progressive (non-interlaced) scanning** is used.
- The positions of the three sampling instants per frame are as shown in Figure below



Problem: Derive the bit rate and the memory requirements to store each frame that result from the digitization of both a 525-line system assuring a 4:1:1 format. Also find the total memory required to store a 1.5 hour movie/video.

Solution

525-line system: The number of samples per line is 360 and the number of visible lines is 240. Hence the resolution of the luminance (Y) and two chrominance {C_b and C_r) signals are:

Y= 360 x 240

C_b= C_r= 180 x 120

Bit rate: Line sampling rate is fixed at 6.75 MHz for Y and 1.6785MHz for both C_b and C_r all with 8 bits per sample.

Thus,

Bit rate = 6.75 x 10^6 x 8 + 2 (1.6785 x 10^6 x 8) = 81Mbps.

Memory required:

Memory required per frame for Y = 360 x 240 x 8 = 691.2 kbits

Memory required per frame for $C_b + C_r$ = 180 x 120 x 8 x 2 =345.6 kbits

Total Memory required per frame = Y + C_b + C_r

= 691.2 k + 345.6 k

= 1.0368 Mbits or 129.6 kbytes

And memory to store 1.5 hours assuming 30 frames per second:

= 1.0368 x 10^6 x 30 x 1.5 x 3600

= 167.96 Gbits or 20.995 Gbytes

625-line system: The number of samples per line is 720 and the number of visible lines is 572. Hence the resolution of the luminance (Y) and two chrominance {C_b and C_r) signals are:

Y = 360 x 288

C_b= C_r = 180 x 144

Bit rate: Line sampling rate is fixed at 6.75 MHz for Y and 1.6785MHz for both C_b and C_r all with 8 bits per sample.

Thus,

Bit rate = 6.75 x 10^6 x 8 + 2 (1.6785 x 10^6 x 8) = 81Mbps.

Memory required:

Memory required per frame for Y = 360 x 288 x 8 = 829.44 kbits

Memory required per frame for $C_b + C_r$ = 180 x 144 x 8 x 2 = 414.72 kbits

Total Memory required per frame = Y + C_b + C_r

= 829.44 k + 414.72 k

= 1.244 Mbits or 155.52 kbytes

And memory to store 1.5 hours assuming 25 frames per second

= 121.77 kb x 50 x 1.5 x 3600

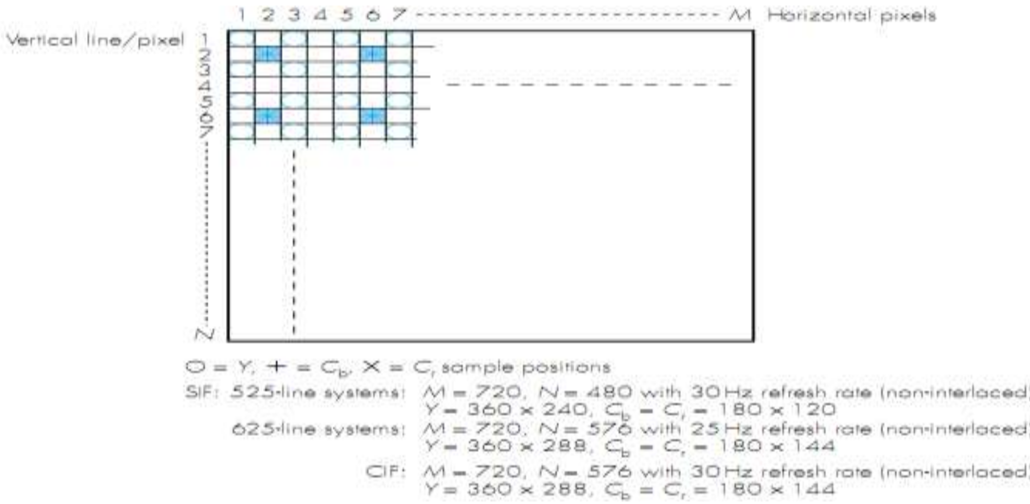
=167.96 Gbits or 20.995 Gbytes

CIF:

- Common intermediate format (CIF) has been defined for videoconferencing applications.
- It is derived from the SIF.
- Uses a combination of the spatial resolution used for the SIF in the 625-line system the temporal resolution used in the 525-line system.
- Thus spatial resolution is:

$$Y = 360 \times 288$$
$$Cb = Cr = 180 \times 144$$

- Uses a temporal resolution of 30 Hz.
- Progressive scanning is employed.
- The position of the sampling instants per frame is the same as for SIF is shown below and hence the digitization format is 4:1:1.



- The worst-case bit rate is the same as that of SIF and is given by:

$$6.75 \times 10^6 \times 8 + 1.685 \times 10^6 \times 8 + 1.685 \times 10^6 \times 8 = 81\ Mbps$$

- We can deduce from this, to convert to the CIF, a 525-line needs a line-rate conversion and a 625-line system a frame-rate conversion.
- a number of higher-resolution derivative the CIF have also been defined since there are a number of different types of videoconferencing applications including those that involve a linked set of desktop PCs and those that involve a linked set of videoconferencing studios.
- Most desktop applications use switched circuits, a typical bit rate used is a single 64 kbps ISDN channel. For linking videoconferencing studios, however, dedicated circuits are normally used that comprise multiple 64 kbps channels.
- As the bit rate of these circuits is much higher typically four or sixteen 64 kbps channels then a higher-resolution version of the basic CIF can be used to improve the quality of the video.
- Examples:

4CIF:

$$Y = 720 \times 576$$
$$Cb = Cr = 360 \times 288$$

16CIF:

$$Y = 1440 \times 1152$$
$$Cb = Cr = 720 \times 576$$

QCIF:

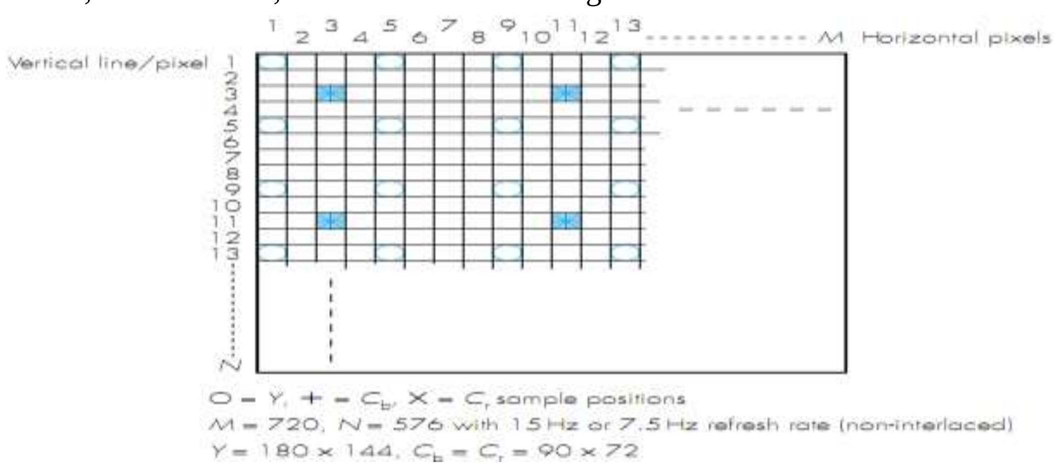
- The **quarter CIF** (QCIF) format has been defined for use in **video telephony** applications.
- It is derived from the CIF.
- Uses **half the spatial resolution of CIF** in both horizontal and vertical directions.
- The temporal resolution is **divided by either 2 or 4 fold of CIF**.
- Spatial resolution is:

$$Y = 180 \times 144$$
$$Cb = Cr = 90 \times 72$$

- Temporal resolution is either **15** or 7.5 Hz.
- The worst-case bit rate with this format is:

$$3.375 \times 10^6 \times 8 + 0.843 \times 10^6 \times 8 + 0.843 \times 10^6 \times 8 = 40.5\ Mbps$$

- The positions of the three sampling instants per frame are as shown in Figure below and, as we can see, it has the same 4:1:1 digitization format as CIF.



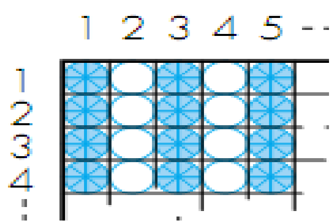
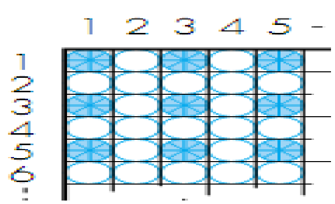
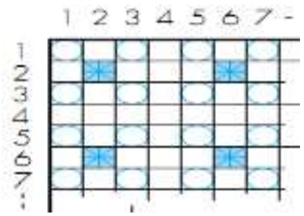
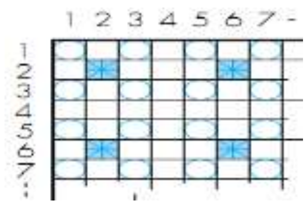
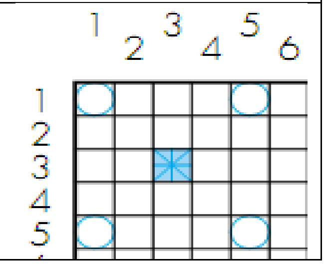
- A QCIF is intended for use with video telephony application involves a single switched 64 kbps channels.
- In addition, there are lower-resolution versions of the QCIF which are intended for use in applications that use lower bit rate channels such as that provided by a modem and the PSTN.
- These lower-resolution versions are known as **sub-QCIF or S-QCIF** and is given by:

$$Y = 128 \times 96$$
$$Cb = Cr = 64 \times 48$$

- Here the sampling matrix appears sparse, in practice, only a small screen (or a small window of a larger screen) is normally used for video telephony and hence the total set of samples may occupy all the pixel positions on the screen or window.

PC video:

- Number of multimedia applications that involve live video, use a window on the screen of a PC monitor for display purposes.
- Ex.: desk video telephony, videoconferencing, and also video-in-a-window.
- for multimedia applications that involve mixing live video with other information on a PC screen, the line sampling rate is normally modified in order to obtain the required horizontal resolution like 640 (480 x pixels per line with a 525-line PC monitor and 768 (576 x 4/3) pixels per with a 625-line PC monitor.
- To achieve the necessary resolution with a 525-line monitor, the line sampling rate is reduced from 13.5 MHz to 12.2727MHz while for a 625-line monitor, the line sampling rate must be increased from 13.5 MHz to 14.75 MHz
- In the case of desktop video telephony and videoconferencing, the video signals from the camera are sampled at this rate prior to transmission and hence can be displayed directly on a PC screen. In the case of a digital television broadcast a conversion is necessary before the video is displayed.
- It is to be remembered that all PC monitors use progressive scanning rather than interlaced scanning.

Formats		**4:2:2**	**4:2:0**	**SIF - 4:1:1**	**CIF - 4:1:1**	**Q CIF**
Spatial Resolution	525 line system (NTSC)	$Y = 720 \times 480$ $Cb = Cr = 360 \times 480$	$Y = 720 \times 480$ $Cb = Cr = 360 \times 240$	$Y = 360 \times 240$ $Cb = Cr = 180 \times 120$	---	---
	625 line system (PAL)	$Y = 720 \times 576$ $Cb = Cr = 360 \times 576$	$Y = 720 \times 576$ $Cb = Cr = 360 \times 288$	$Y = 360 \times 288$ $Cb = Cr = 180 \times 144$	$Y = 360 \times 288$ $Cb = Cr = 180 \times 144$	$Y = 180 \times 144$ $Cb = Cr = 90 \times 72$
Sampling Rate (samples per second)		for Y $\rightarrow 13.5 \times 10^6$ for C_r & C_b $\rightarrow 6.75 \times 10^6$	for Y $\rightarrow 13.5 \times 10^6$ for C_r & C_b $\rightarrow 3.375 \times 10^6$	for Y $\rightarrow 6.75 \times 10^6$ for C_r & C_b $\rightarrow 1.6875 \times 10^6$	for Y $\rightarrow 6.75 \times 10^6$ C_r & C_b $\rightarrow 1.6875 \times 10^6$	for Y $\rightarrow 3.375 \times 10^6$ Cr & Cb $\rightarrow 0.843 \times 10^6$
Temporal Resolution (Frame Refresh Rate)	525 line system (NTSC)	60 Frames Per Second	60	30	30	15 or 7.5
	625 line system (PAL)	50 Frames Per Second	50	25	---	
Image						

Formats	**S - QCIF**	**4 CIF**	**16 CIF**
Spatial Resolution	$Y = 128 \times 96$ $Cb = Cr = 64 \times 48$	$Y = 720 \times 576$ $Cb = Cr = 360 \times 288$	$Y = 1440 \times 1152$ $Cb = Cr = 720 \times 576$