What is computer Graphics?

Computer graphics is an art of drawing pictures, lines, charts, etc. using computers with the help of programming. Computer graphics image is made up of number of pixels. **Pixel** is the smallest addressable graphical unit represented on the computer screen.

Introduction

- Computer is information processing machine. User needs to communicate with computer and the computer graphics is one of the most effective and commonly used ways of communication with the user.
- It displays the information in the form of graphical objects such as pictures, charts, diagram and graphs.
- Graphical objects convey more information in less time and easily understandable formats for example statically graph shown in stock exchange.
- In computer graphics picture or graphics objects are presented as a collection of discrete pixels.
- We can control intensity and color of pixel which decide how picture look like.
- The special procedure determines which pixel will provide the best approximation to the desired picture or graphics object this process is known as **Rasterization.**
- The process of representing continuous picture or graphics object as a collection of discrete pixels is called **Scan Conversion.**

Advantages of computer graphics

- Computer graphics is one of the most effective and commonly used ways of communication with computer.
- It provides tools for producing picture of "real-world" as well as synthetic objects such as mathematical surfaces in 4D and of data that have no inherent geometry such as survey result.
- It has ability to show moving pictures thus possible to produce animations with computer graphics.
- With the use of computer graphics we can control the animation by adjusting the speed, portion of picture in view the amount of detail shown and so on.
- It provides tools called motion dynamics. In which user can move objects as well as observes as per requirement for example walk throw made by builder to show flat interior and surrounding.
- It provides facility called update dynamics. With this we can change the shape color and other properties of object.
- Now in recent development of digital signal processing and audio synthesis chip the interactive graphics can now provide audio feedback along with the graphical feed backs.

Application of computer graphics

- User interface: - Visual object which we observe on screen which communicates with user is one of the most useful applications of the computer graphics.
- Plotting of graphics and chart in industry, business, government and educational organizations drawing like bars, pie-charts, histogram's are very useful for quick and good decision making.
- Office automation and desktop publishing: - It is used for creation and dissemination of information. It is used in in-house creation and printing of documents which contains text, tables, graphs and other forms of drawn or scanned images or picture.
- Computer aided drafting and design: - It uses graphics to design components and system such as automobile bodies structures of building etc.

- Simulation and animation: - Use of graphics in simulation makes mathematic models and mechanical systems more realistic and easy to study.
- Art and commerce: - There are many tools provided by graphics which allows used to make their picture animated and attracted which are used in advertising.
- Process control: - Now a day's automation is used which is graphically displayed on the screen.
- Cartography: - Computer graphics is also used to represent geographic maps, weather maps, oceanographic charts etc.
- Education and training: - Computer graphics can be used to generate models of physical, financial and economic systems. These models can be used as educational aids.
- Image processing: - It is used to process image by changing property of the image.

Display devices

- Display devices are also known as output devices.
- Most commonly used output device in a graphics system is a video monitor.

Cathode-ray-tubes

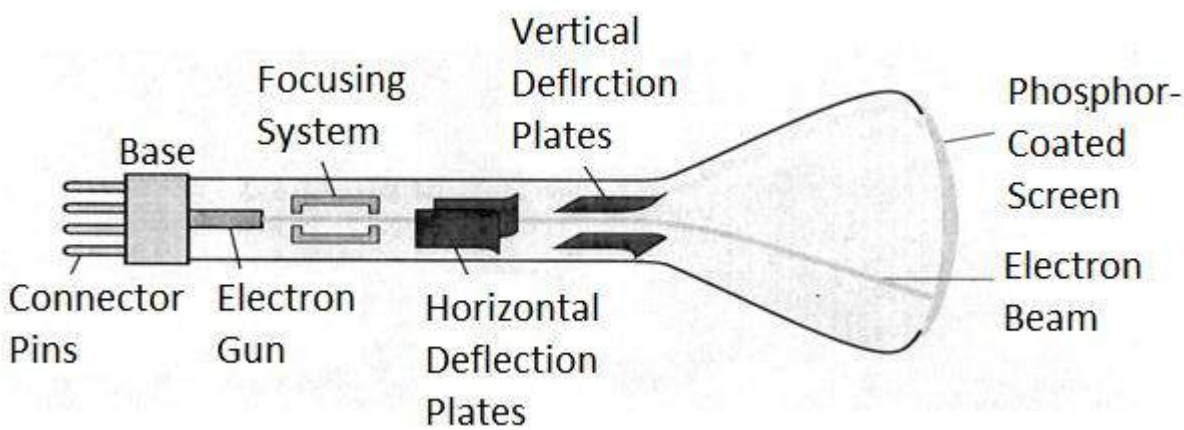


Fig. 1.1: - Cathode ray tube.

- It is an evacuated glass tube.
- An electron gun at the rear of the tube produce a beam of electrons which is directed towards the screen of the tube by a high voltage typically 15000 to 20000 volts
- Inner side screen is coated with phosphor substance which gives light when it is stroked bye electrons.
- Control grid controls velocity of electrons before they hit the phosphor.
- The control grid voltage determines how many electrons are actually in the electron beam. The negative the control voltage is the fewer the electrons that pass through the grid.
- Thus control grid controls Intensity of the spot where beam strikes the screen.
- The focusing system concentrates the electron beam so it converges to small point when hits the phosphor coating.
- Deflection system directs beam which decides the point where beam strikes the screen.
- Deflection system of the CRT consists of two pairs of parallel plates which are vertical and horizontal deflection plates.
- Voltage applied to vertical and horizontal deflection plates is control vertical and horizontal deflection respectively.
- There are two techniques used for producing images on the CRT screen:

1. Vector scan/Random scan display.
2. Raster scan display.

Vector scan/Random scan display

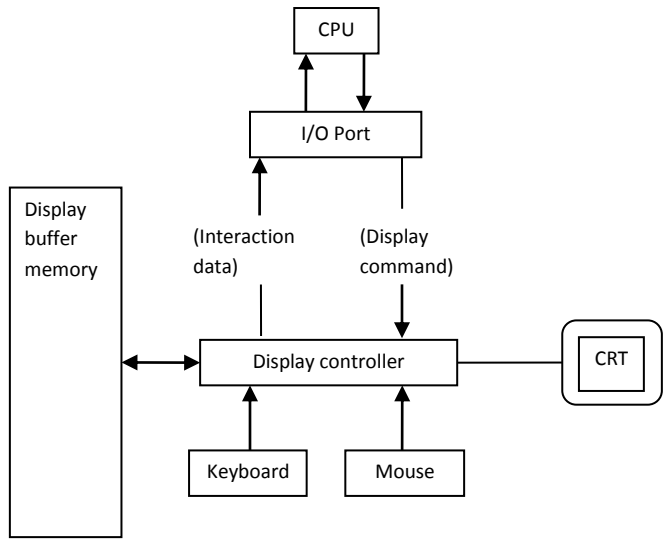


Fig. 1.2: - Architecture of a vector display.

- Vector scan display directly traces out only the desired lines on CRT.
- If we want line between point p1 & p2 then we directly drive the beam deflection circuitry which focus beam directly from point p1 to p2.
- If we do not want to display line from p1 to p2 and just move then we can blank the beam as we move it.
- To move the beam across the CRT, the information about both magnitude and direction is required. This information is generated with the help of vector graphics generator.
- Fig. 1.2 shows architecture of vector display. It consists of display controller, CPU, display buffer memory and CRT.
- Display controller is connected as an I/O peripheral to the CPU.
- Display buffer stores computer produced display list or display program.
- The Program contains point & line plotting commands with end point co-ordinates as well as character plotting commands.
- Display controller interprets command and sends digital and point co-ordinates to a vector generator.
- Vector generator then converts the digital co-ordinate value to analog voltages for beam deflection circuits that displace an electron beam which points on the CRT's screen.
- In this technique beam is deflected from end point to end point hence this techniques is also called random scan.
- We know as beam strikes phosphors coated screen it emits light but that light decays after few milliseconds and therefore it is necessary to repeat through the display list to refresh the screen at least 30 times per second to avoid flicker.
- As display buffer is used to store display list and used to refreshing, it is also called refresh buffer.

Raster scan display

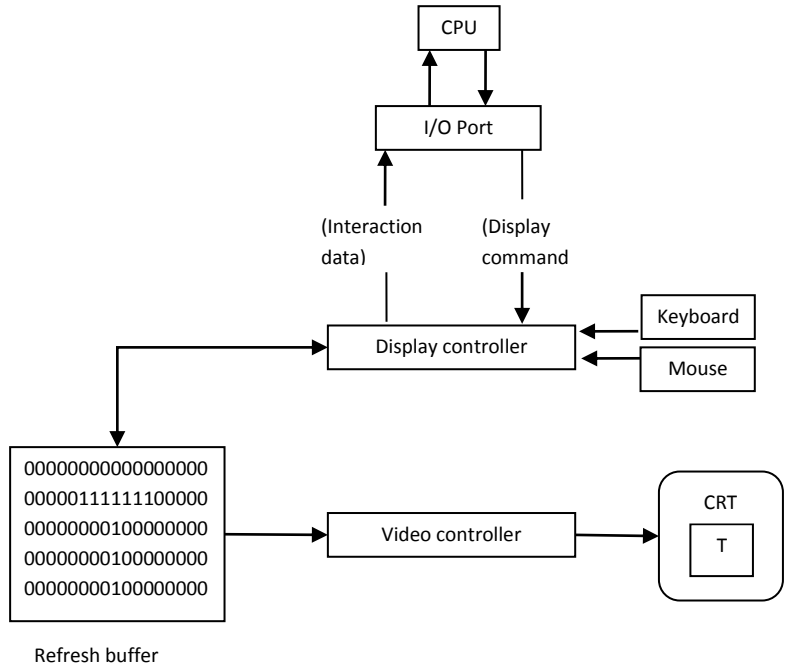


Fig. 1.3: - Architecture of a raster display.

- Fig. 1.3 shows the architecture of Raster display. It consists of display controller, CPU, video controller, refresh buffer, keyboard, mouse and CRT.
- The display image is stored in the form of 1's and 0's in the refresh buffer.
- The video controller reads this refresh buffer and produces the actual image on screen.
- It will scan one line at a time from top to bottom & then back to the top.

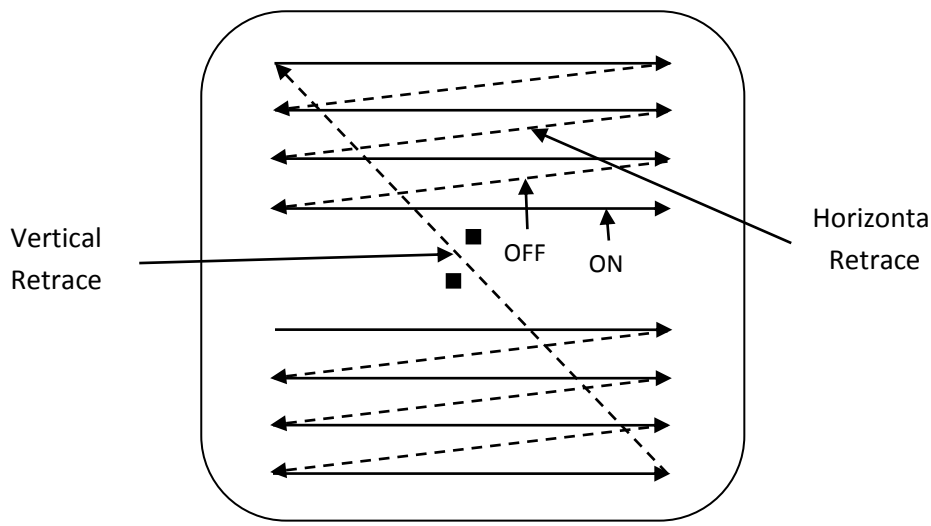


Fig. 1.4: - Raster scan CRT.

- In this method the horizontal and vertical deflection signals are generated to move the beam all over the screen in a pattern shown in fig. 1.4.
- Here beam is swept back & forth from left to the right.
- When beam is moved from left to right it is ON.

- When beam is moved from right to left it is OFF and process of moving beam from right to left after completion of row is known as **Horizontal Retrace.**
- When beam is reach at the bottom of the screen. It is made OFF and rapidly retraced back to the top left to start again and process of moving back to top is known as **Vertical Retrace**.
- The screen image is maintained by repeatedly scanning the same image. This process is known as **Refreshing of Screen.**
- In raster scan displays a special area of memory is dedicated to graphics only. This memory is called **Frame Buffer.**
- Frame buffer holds set of intensity values for all the screen points.
- That intensity is retrieved from frame buffer and display on screen one row at a time.
- Each screen point referred as pixel or **Pel (Picture Element).**
- Each pixel can be specified by its row and column numbers.
- It can be simply black and white system or color system.
- In simple black and white system each pixel is either ON or OFF, so only one bit per pixel is needed.
- Additional bits are required when color and intensity variations can be displayed up to 24-bits per pixel are included in high quality display systems.
- On a black and white system with one bit per pixel the frame buffer is commonly called a **Bitmap.** And for systems with multiple bits per pixel, the frame buffer is often referred as a **Pixmap.**

Difference between random scan and raster scan

Base of Difference	Raster Scan System	Random Scan System
Electron Beam	The electron beam is swept across the screen, one row at a time, from top to bottom.	The electron beam is directed only to the parts of screen where a picture is to be drawn.
Resolution	Its resolution is poor because raster system in contrast produces zigzag lines that are plotted as discrete point sets.	Its resolution is good because this system produces smooth lines drawings because CRT beam directly follows the line path.
Picture Definition	Picture definition is stored as a set of intensity values for all screen points, called pixels in a refresh buffer area.	Picture definition is stored as a set of line drawing instructions in a display file.
Realistic Display	The capability of this system to store intensity values for pixel makes it well suited for the realistic display of scenes contain shadow and color pattern.	These systems are designed for line-drawing and can't display realistic shaded scenes.
Draw an Image	Screen points/pixels are used to draw an image.	Mathematical functions are used to draw an image.

Color CRT monitors

- A CRT monitors displays color pictures by using a combination of phosphors that emit different colored light.
- It produces range of colors by combining the light emitted by different phosphors.
- There are two basic techniques for color display:
    1. Beam-penetration technique
    2. Shadow-mask technique

Beam-penetration technique

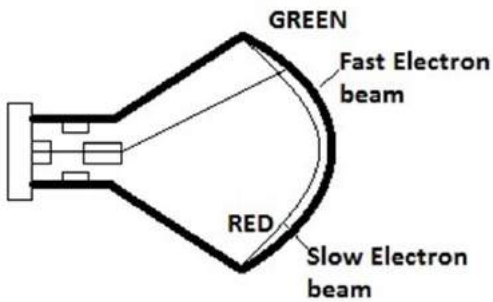


Fig. 1.5: - Beam-penetration CRT

- This technique is used with random scan monitors.
- In this technique inside of CRT coated with two phosphor layers usually red and green. The outer layer of red and inner layer of green phosphor.
- The color depends on how far the electron beam penetrates into the phosphor layer.
- A beam of fast electron penetrates more and excites inner green layer while slow electron excites outer red layer.
- At intermediate beam speed we can produce combination of red and green lights which emit additional two colors orange and yellow.
- The beam acceleration voltage controls the speed of the electrons and hence color of pixel.
- It is a low cost technique to produce color in random scan monitors.
- It can display only four colors.
- Quality of picture is not good compared to other techniques.

Shadow-mask technique

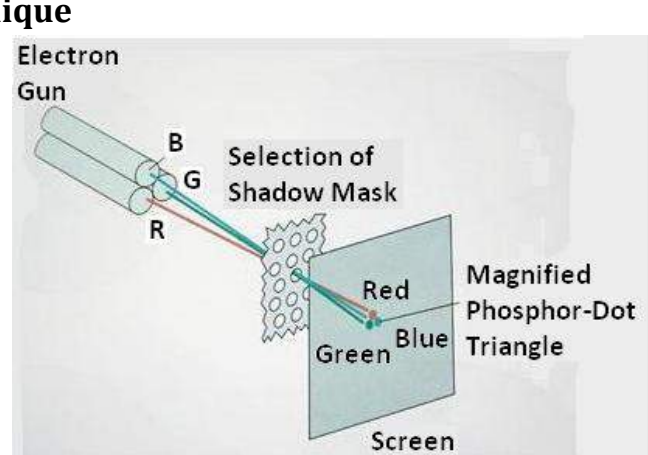


Fig. 1.6: - Shadow-mask CRT.

- It produces wide range of colors as compared to beam-penetration technique.
- This technique is generally used in raster scan displays. Including color TV.
- In this technique CRT has three phosphor color dots at each pixel position. One dot for red, one for green and one for blue light. This is commonly known as **Dot Triangle.**
- Here in CRT there are three electron guns present, one for each color dot. And a shadow mask grid just behind the phosphor coated screen.
- The shadow mask grid consists of series of holes aligned with the phosphor dot pattern.
- Three electron beams are deflected and focused as a group onto the shadow mask and when they pass through a hole they excite a dot triangle.
- In dot triangle three phosphor dots are arranged so that each electron beam can activate only its corresponding color dot when it passes through the shadow mask.
- A dot triangle when activated appears as a small dot on the screen which has color of combination of three small dots in the dot triangle.
- By changing the intensity of the three electron beams we can obtain different colors in the shadow mask CRT.

Direct-view storage tubes (DVST)

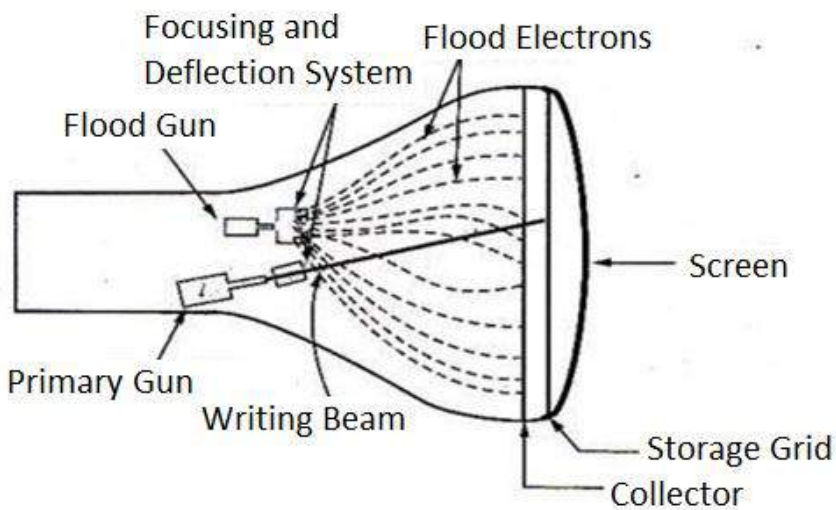


Fig. 1.7: - Direct-view storage tube.

- In raster scan display we do refreshing of the screen to maintain a screen image.
- DVST gives alternative method for maintaining the screen image.
- DVST uses the storage grid which stores the picture information as a charge distribution just behind the phosphor coated screen.
- DVST consists two electron guns a primary gun and a flood gun.
- A primary gun stores the picture pattern and the flood gun maintains the picture display.
- A primary gun emits high speed electrons which strike on the storage grid to draw the picture pattern.
- As electron beam strikes on the storage grid with high speed, it knocks out electrons from the storage grid keeping the net positive charge.
- The knocked out electrons are attracted towards the collector.
- The net positive charge on the storage grid is nothing but the picture pattern.
- The continuous low speed electrons from flood gun pass through the control grid and are attracted to the positive charged area of the storage grid.

- The low speed electrons then penetrate the storage grid and strike the phosphor coating without affecting the positive charge pattern on the storage grid.
- During this process the collector just behind the storage grid smooth out the flow of flood electrons.

Advantage of DVST

- Refreshing of CRT is not required.
- Very complex pictures can be displayed at very high resolution without flicker.
- Flat screen.

Disadvantage of DVST

- They do not display color and are available with single level of line intensity.
- For erasing it is necessary to removal of charge on the storage grid so erasing and redrawing process take several second.
- Erasing selective part of the screen cannot be possible.
- Cannot used for dynamic graphics application as on erasing it produce unpleasant flash over entire screen.
- It has poor contrast as a result of the comparatively low accelerating potential applied to the flood electrons.
- The performance of DVST is somewhat inferior to the refresh CRT.

Flat Panel Display

- The term flat panel display refers to a class of video device that have reduced volume, weight & power requirement compared to a CRT.
- As flat panel display is thinner than CRTs, we can hang them on walls or wear on our wrists.
- Since we can even write on some flat panel displays they will soon be available as pocket notepads.
- We can separate flat panel display in two categories:
    1. **Emissive displays**: - the emissive display or **emitters** are devices that convert electrical energy into light. For Ex. Plasma panel, thin film electroluminescent displays and light emitting diodes.
    2. **Non emissive displays**: - non emissive display or **non emitters** use optical effects to convert sunlight or light from some other source into graphics patterns. For Ex. LCD (Liquid Crystal Display).

Plasma Panels displays

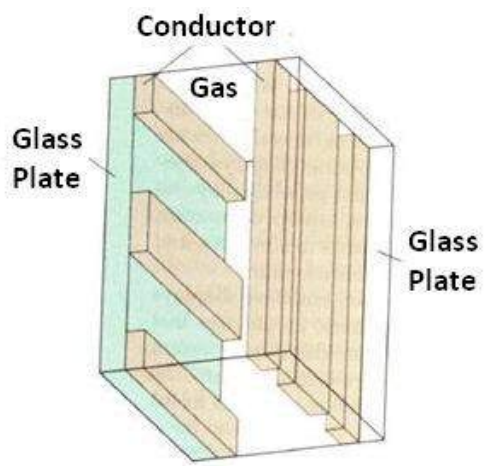


Fig. 1.8: - Basic design of a plasma-panel display device.

- This is also called gas discharge displays.
- It is constructed by filling the region between two glass plates with a mixture of gases that usually includes neon.
- A series of vertical conducting ribbons is placed on one glass panel and a set of horizontal ribbon is built into the other glass panel.
- Firing voltage is applied to a pair of horizontal and vertical conductors cause the gas at the intersection of the two conductors to break down into glowing plasma of electrons and ions.
- Picture definition is stored in a refresh buffer and the firing voltages are applied to refresh the pixel positions, 60 times per second.
- Alternating current methods are used to provide faster application of firing voltages and thus brighter displays.
- Separation between pixels is provided by the electric field of conductor.
- One disadvantage of plasma panels is they were strictly monochromatic device that means shows only one color other than black like black and white.

Thin Film Electroluminescent Displays.

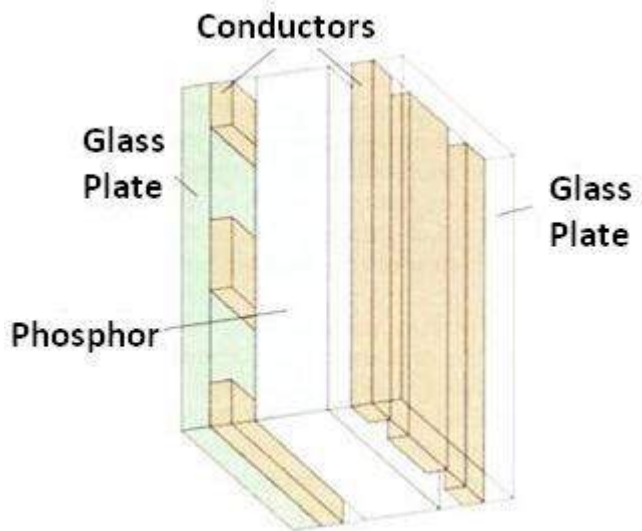


Fig. 1.9: - Basic design of a thin-film electro luminescent display device.

- It is similar to plasma panel display but region between the glass plates is filled with phosphors such as zinksulphide doped with magnesium instead of gas.
- When sufficient voltage is applied the phosphors becomes a conductor in area of intersection of the two electrodes.
- Electrical energy is then absorbed by the manganese atoms which then release the energy as a spot of light similar to the glowing plasma effect in plasma panel.
- It requires more power than plasma panel.
- In this good color and gray scale difficult to achieve.

Light Emitting Diode (LED)

- In this display a matrix of multi-color light emitting diode is arranged to form the pixel position in the display. And the picture definition is stored in refresh buffer.
- Similar to scan line refreshing of CRT information is read from the refresh buffer and converted to voltage levels that are applied to the diodes to produce the light pattern on the display.

Liquid Crystal Display (LCD)

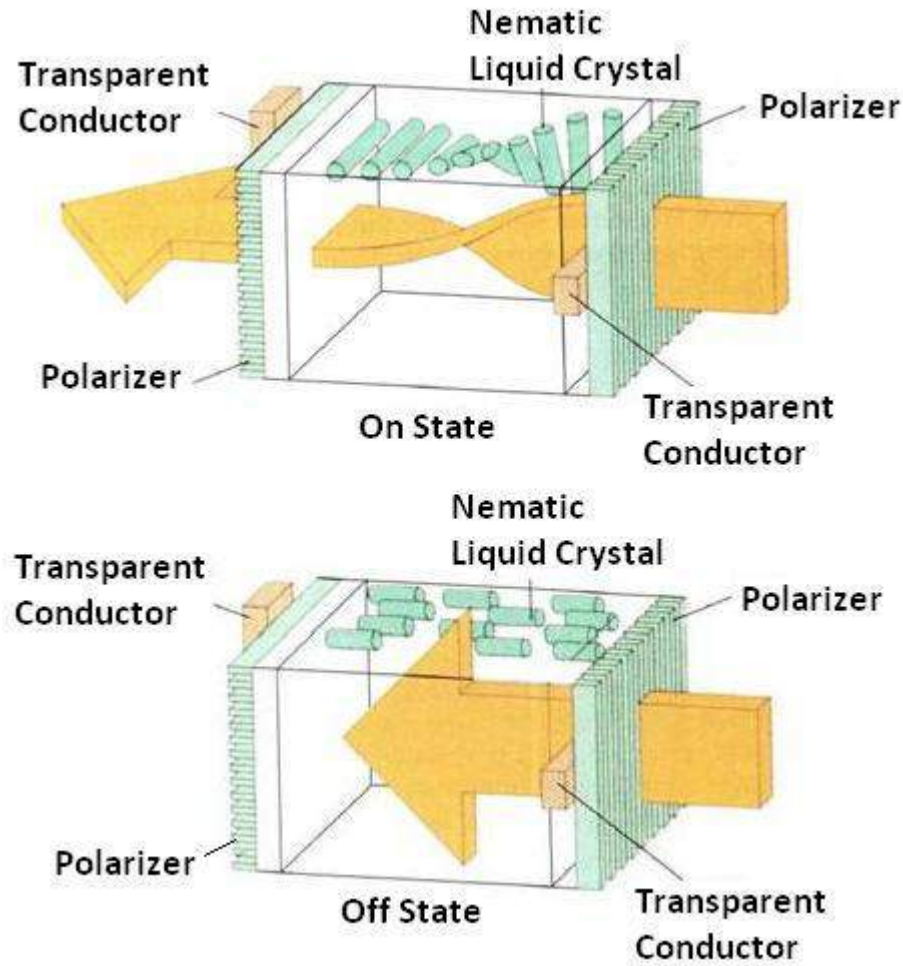


Fig. 1.10: - Light twisting shutter effect used in design of most LCD.

- It is generally used in small system such as calculator and portable laptop.
- This non emissive device produce picture by passing polarized light from the surrounding or from an internal light source through liquid crystal material that can be aligned to either block or transmit the light.
- The liquid crystal refreshes to fact that these compounds have crystalline arrangement of molecules then also flows like liquid.
- It consists of two glass plates each with light polarizer at right angles to each other sandwich the liquid crystal material between the plates.
- Rows of horizontal transparent conductors are built into one glass plate, and column of vertical conductors are put into the other plates.
- The intersection of two conductors defines a pixel position.
- In the ON state polarized light passing through material is twisted so that it will pass through the opposite polarizer.
- In the OFF state it will reflect back towards source.
- We applied a voltage to the two intersecting conductor to align the molecules so that the light is not twisted.
- This type of flat panel device is referred to as a passive matrix LCD.
- In active matrix LCD transistors are used at each (x, y) grid point.

- Transistor cause crystal to change their state quickly and also to control degree to which the state has been changed.
- Transistor can also serve as a memory for the state until it is changed.
- So transistor make cell ON for all time giving brighter display then it would be if it had to be refresh periodically

Advantages of LCD display

- Low cost.
- Low weight.
- Small size
- Low power consumption.

Three dimensional viewing devices

- The graphics monitor which are display three dimensional scenes are devised using a technique that reflects a CRT image from a vibrating flexible mirror.

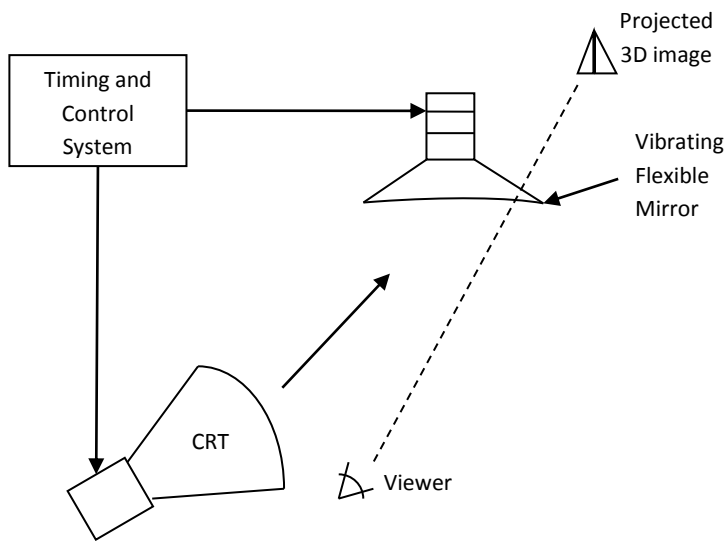


Fig. 1.11: - 3D display system uses a vibrating mirror.

- Vibrating mirror changes its focal length due to vibration which is synchronized with the display of an object on CRT.
- The each point on the object is reflected from the mirror into spatial position corresponding to distance of that point from a viewing position.
- Very good example of this system is GENISCO SPACE GRAPH system, which use vibrating mirror to project 3D objects into a 25 cm by 25 cm by 25 cm volume. This system is also capable to show 2D cross section at different depth.

Application of 3D viewing devices

- In medical to analyze data from ultra-sonography.
- In geological to analyze topological and seismic data.
- In designing like solid objects viewing and 3D viewing of objects.

Stereoscopic and virtual-reality systems

Stereoscopic system

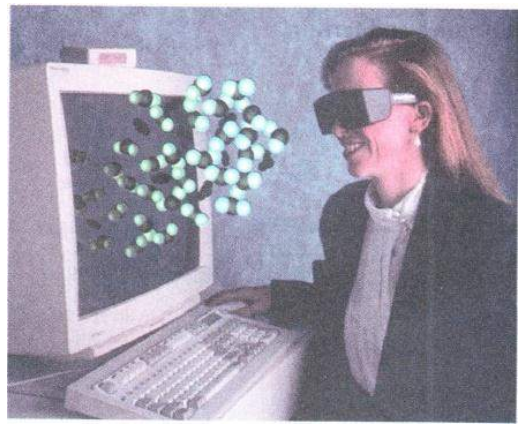


Fig. 1.12: - stereoscopic views.

- Stereoscopic views does not produce three dimensional images, but it produce 3D effects by presenting different view to each eye of an observer so that it appears to have depth.
- To obtain this we first need to obtain two views of object generated from viewing direction corresponding to each eye.
- We can construct the two views as computer generated scenes with different viewing positions or we can use stereo camera pair to photograph some object or scene.
- When we see simultaneously both the view as left view with left eye and right view with right eye then two views is merge and produce image which appears to have depth.
- One way to produce stereoscopic effect is to display each of the two views with raster system on alternate refresh cycles.
- The screen is viewed through glasses with each lance design such a way that it act as a rapidly alternating shutter that is synchronized to block out one of the views.

Virtual-reality

Fig. 1.13: - virtual reality.

- Virtual reality is the system which produce images in such a way that we feel that our surrounding is what we are set in display devices but in actually it does not.
- In virtual reality user can step into a scene and interact with the environment.

- A head set containing an optical system to generate the stereoscopic views is commonly used in conjunction with interactive input devices to locate and manipulate objects in the scene.
- Sensor in the head set keeps track of the viewer's position so that the front and back of objects can be seen as the viewer "walks through" and interacts with the display.
- Virtual reality can also be produce with stereoscopic glass and video monitor instead of head set. This provides low cost virtual reality system.
- Sensor on display screen track head position and accordingly adjust image depth.

Raster graphics systems

Simple raster graphics system

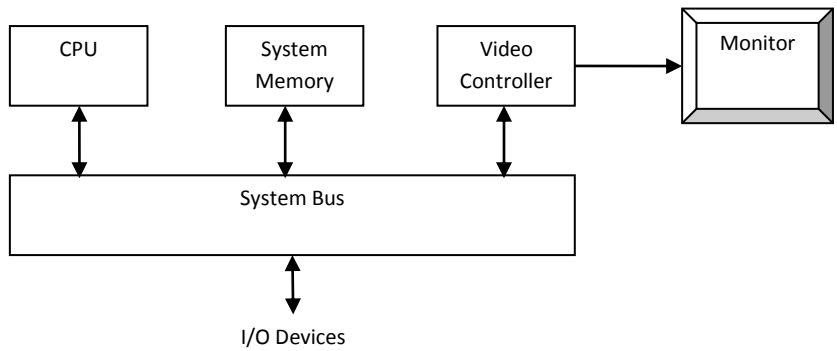


Fig. 1.14: - Architecture of a simple raster graphics system.

- Raster graphics systems having additional processing unit like video controller or display controller.
- Here frame buffer can be anywhere in the system memory and video controller access this for refresh the screen.
- In addition to video controller more processors are used as co-processors to accelerate the system in sophisticated raster system.

Raster graphics system with a fixed portion of the system memory reserved for the frame buffer

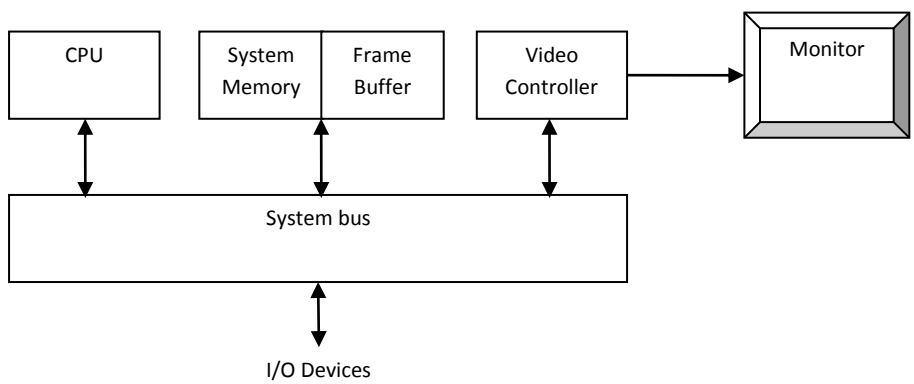


Fig. 1.15: - Architecture of a raster graphics system with a fixed portion of the system memory reserved for the frame buffer.

- A fixed area of the system memory is reserved for the frame buffer and the video controller can directly access that frame buffer memory.
- Frame buffer location and the screen position are referred in Cartesian coordinates.
- For many graphics monitors the coordinate origin is defined at the lower left screen corner.
- Screen surface is then represented as the first quadrant of the two dimensional systems with positive X-value increases as left to right and positive Y-value increases bottom to top.

Basic refresh operation of video controller

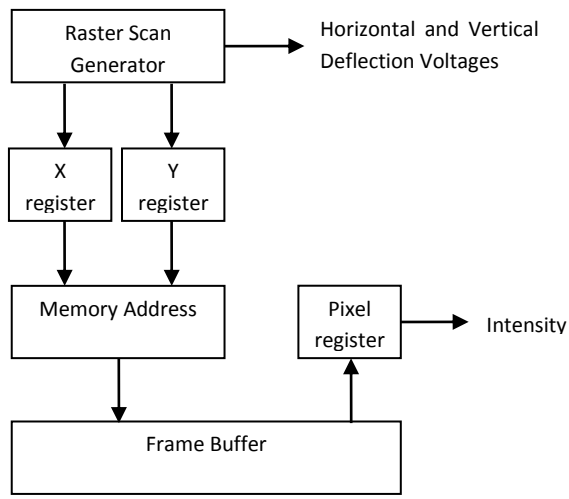


Fig. 1.16: - Basic video controller refresh operation.

- Two registers are used to store the coordinates of the screen pixels which are X and Y
- Initially the X is set to 0 and Y is set to Ymax.
- The value stored in frame buffer for this pixel is retrieved and used to set the intensity of the CRT beam.
- After this X register is incremented by one.
- This procedure is repeated till X becomes equals to Xmax.
- Then X is set to 0 and Y is decremented by one pixel and repeat above procedure.
- This whole procedure is repeated till Y is become equals to 0 and complete the one refresh cycle. Then controller reset the register as top –left corner i.e. X=0 and Y=Ymax and refresh process start for next refresh cycle.
- Since screen must be refreshed at the rate of 60 frames per second the simple procedure illustrated in figure cannot be accommodated by typical RAM chips.
- To speed up pixel processing video controller retrieves multiple values at a time using more numbers of registers and simultaneously refresh block of pixel.
- Such a way it can speed up and accommodate refresh rate more than 60 frames per second.

Raster-graphics system with a display processor

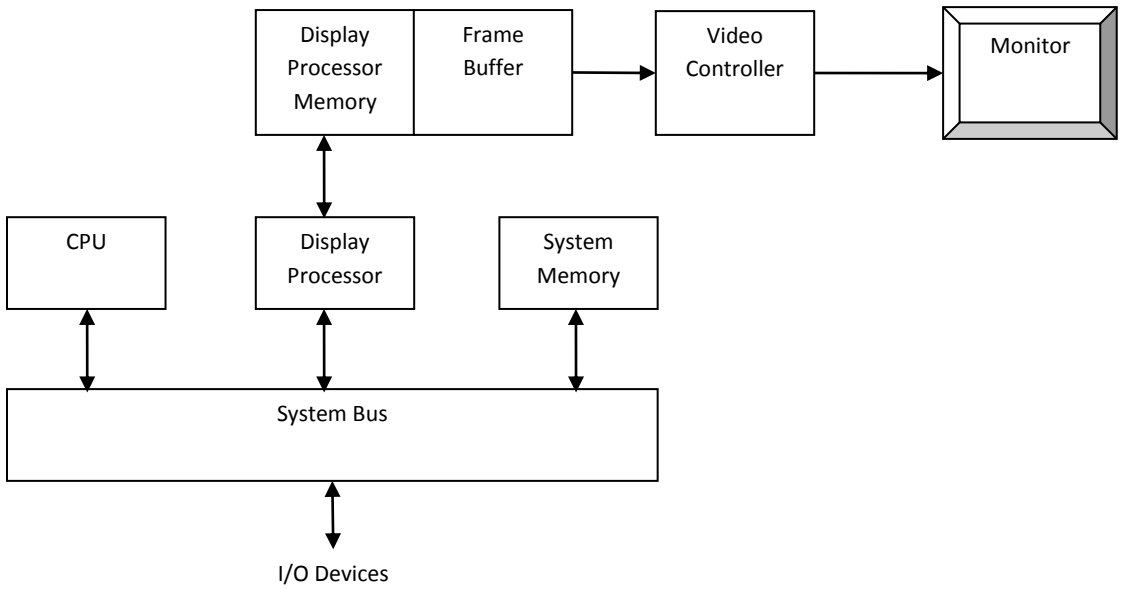


Fig. 1.17: - Architecture of a raster-graphics system with a display processor.

- One way to designing raster system is having separate display coprocessor.
- Purpose of display processor is to free CPU from graphics work.
- Display processors have their own separate memory for fast operation.
- Main work of display processor is digitalizing a picture definition given into a set of pixel intensity values for store in frame buffer.
- This digitalization process is scan conversion.
- Display processor also performs many other functions such as generating various line styles (dashed, dotted, or solid). Display color areas and performing some transformation for manipulating object.
- It also interfaces with interactive input devices such as mouse.
- For reduce memory requirements in raster scan system methods have been devised for organizing the frame buffer as a line list and encoding the intensity information.
- One way to do this is to store each scan line as a set of integer pair one number indicate number of adjacent pixels on the scan line that are having same intensity and second stores intensity value this technique is called run-length encoding.
- A similar approach is when pixel. Intensity is changes linearly, encoded the raster as a set of rectangular areas (cell encoding).
- Disadvantages of encoding is when run length is small it requires more memory then original frame buffer.
- It also difficult for display controller to process the raster when many sort runs are involved.

Random- scan system

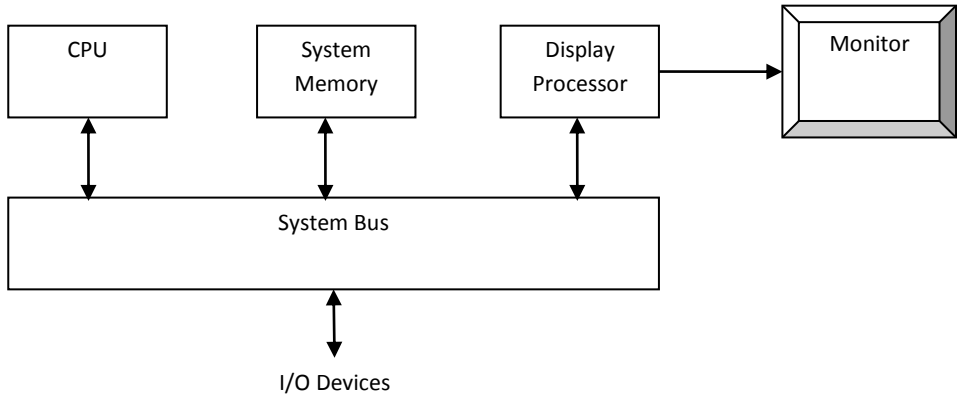


Fig. 1.18: - Architecture of a simple random-scan system.

- An application program is input & stored in the system memory along with a graphics package.
- Graphics commands in the application program are translated by the graphics package into a display file stored in the system memory.
- This display file is used by display processor to refresh the screen.
- Display process goes through each command in display file. Once during every refresh cycle.
- Sometimes the display processor in random scan system is also known as display processing unit or a graphics controller.
- In this system graphics platform are drawn on random scan system by directing the electron beam along the component times of the picture.
- Lines are defined by coordinate end points.
- This input coordinate values are converts to X and Y deflection voltages.
- A scene is then drawn one line at a time.

Graphics input devices

Keyboards

- Keyboards are used as entering text strings. It is efficient devices for inputting such a non-graphics data as picture label.
- Cursor control key's & function keys are common features on general purpose keyboards.
- Many other application of key board which we are using daily used of computer graphics are commanding & controlling through keyboard etc.

Mouse

- Mouse is small size hand-held box used to position screen cursor.
- Wheel or roller or optical sensor is directing pointer on the according to movement of mouse.
- Three buttons are placed on the top of the mouse for signaling the execution of some operation.
- Now a day's more advance mouse is available which are very useful in graphics application for example Z mouse.

Trackball and Spaceball

- Trackball is ball that can be rotated with the finger or palm of the hand to produce cursor movement.

- Potentiometer attached to the ball, measure the amount and direction of rotation.
- They are often mounted on keyboard or Z mouse.
- Space ball provide six-degree of freedom i.e. three dimensional.
- In space ball strain gauges measure the amount of pressure applied to the space ball to provide input for spatial positioning and orientation as the ball is pushed or pulled in various directions.
- Space balls are used in 3D positioning and selection operations in virtual reality system, modeling, animation, CAD and other application.

Joysticks

- A joy stick consists of small vertical lever mounted on a base that is used to steer the screen cursor around.
- Most joy sticks selects screen positioning according to actual movement of stick (lever).
- Some joy sticks are works on pressure applied on sticks.
- Sometimes joy stick mounted on keyboard or sometimes used alone.
- Movement of the stick defines the movement of the cursor.
- In pressure sensitive stick pressure applied on stick decides movement of the cursor. This pressure is measured using strain gauge.
- This pressure sensitive joy sticks also called as isometric joy sticks and they are non movable sticks.

Data glove

- Data glove is used to grasp virtual objects.
- The glow is constructed with series of sensors that detect hand and figure motions.
- Electromagnetic coupling is used between transmitter and receiver antennas which used to provide position and orientation of the hand.
- Transmitter & receiver Antenna can be structured as a set of three mutually perpendicular coils forming 3D Cartesian coordinates system.
- Input from the glove can be used to position or manipulate object in a virtual scene.

Digitizer

- Digitizer is common device for drawing painting or interactively selecting coordinates position on an object.
- One type of digitizers is graphics tablet which input two dimensional coordinates by activating hand cursor or stylus at selected position on a flat surface.
- Stylus is flat pencil shaped device that is pointed at the position on the tablet.

Image Scanner

- Image Scanner scan drawing, graph, color, & black and white photos or text and can stored for computer processing by passing an optical scanning mechanism over the information to be stored.
- Once we have internal representation of a picture we can apply transformation.
- We can also apply various image processing methods to modify the picture.
- For scanned text we can apply modification operation.

Touch Panels

- As name suggest Touch Panels allow displaying objects or screen-position to be selected with the touch or finger.
- A typical application is selecting processing option shown in graphical icons.

- Some system such as a plasma panel are designed with touch screen
- Other system can be adapted for touch input by fitting transparent touch sensing mechanism over a screen.
- Touch input can be recorded with following methods.
    1. Optical methods
    2. Electrical methods
    3. Acoustical methods

Optical method

- Optical touch panel employ a line of infrared LEDs along one vertical and one horizontal edge.
- The opposite edges of the edges containing LEDs are contain light detectors.
- When we touch at a particular position the line of light path breaks and according to that breaking line coordinate values are measured.
- In case two line cuts it will take average of both pixel positions.
- LEDs operate at infrared frequency so it cannot be visible to user.

Electrical method

- An electrical touch panel is constructed with two transparent plates separated by small distance.
- One is coated with conducting material and other is coated with resistive material.
- When outer plate is touch it will come into contact with internal plate.
- When both plates touch it creates voltage drop across the resistive plate that is converted into coordinate values of the selected position.

Acoustical method

- In acoustical touch panel high frequency sound waves are generated in horizontal and vertical direction across a glass plates.
- When we touch the screen the waves from that line are reflected from finger.
- These reflected waves reach again at transmitter position and time difference between sending and receiving is measure and converted into coordinate values.

Light pens

- Light pens are pencil-shaped device used to select positions by detecting light coming from points on the CRT screen.
- Activated light pens pointed at a spot on the screen as the electron beam lights up that spot and generate electronic pulse that causes the coordinate position of the electron beam to be recorded.

Voice systems

- It is used to accept voice command in some graphics workstations.
- It is used to initiate graphics operations.
- It will match input against predefined directory of words and phrases.
- Dictionary is setup for a particular operator by recording his voice.
- Each word is speak several times and then analyze the word and establishes a frequency pattern for that word along with corresponding function need to be performed.
- When operator speaks command it will match with predefine dictionary and perform desired action.

Graphics software and standard

- There are mainly two types of graphics software:
  1. General programming package
  2. Special-purpose application package

General programming package

- A general programming package provides an extensive set of graphics function that can be used in high level programming language such as C or FORTRAN.
- It includes basic drawing element shape like line, curves, polygon, color of element transformation etc.
- Example: - GL (Graphics Library).

Special-purpose application package

- Special-purpose application package are customize for particular application which implement required facility and provides interface so that user need not to vory about how it will work (programming). User can simply use it by interfacing with application.
- Example: - CAD, medical and business systems.

Coordinate representations

- Except few all other general packages are designed to be used with Cartesian coordinate specifications.
- If coordinate values for a picture are specified is some other reference frame they must be converted to Cartesian coordinate before giving input to graphics package.
- Special-purpose package may allow use of other coordinates which suits application.
- In general several different Cartesian reference frames are used to construct and display scene.
- We can construct shape of object with separate coordinate system called modeling coordinates or sometimes local coordinates or master coordinates.
- Once individual object shapes have been specified we can place the objects into appropriate positions called world coordinates.
- Finally the World-coordinates description of the scene is transferred to one or more output device reference frame for display. These display coordinates system are referred to as "**Device Coordinates"** or "**Screen Coordinates".**
- Generally a graphic system first converts the world-coordinates position to normalized device coordinates. In the range from 0 to 1 before final conversion to specific device coordinates.
- An initial modeling coordinates position (Xmc,Ymc) in this illustration is transferred to a device coordinates position(Xdc,Ydc) with the sequence (Xmc,Ymc)→ (Xwc,Ywc)→ (Xnc,Ync)→ (Xdc,Ydc).

Graphic Function

- A general purpose graphics package provides user with Varity of function for creating and manipulating pictures.
- The basic building blocks for pictures are referred to as output primitives. They includes character, string, and geometry entities such as point, straight lines, curved lines, filled areas and shapes defined with arrays of color points.
- Input functions are used for control & process the various input device such as mouse, tablet, etc.
- Control operations are used to controlling and housekeeping tasks such as clearing display screen etc.
- All such inbuilt function which we can use for our purpose are known as graphics function

Software Standard

- Primary goal of standardize graphics software is portability so that it can be used in any hardware systems & avoid rewriting of software program for different system
- Some of these standards are discuss below

Graphical Kernel System (GKS)

- This system was adopted as a first graphics software standard by the international standard organization (ISO) and various national standard organizations including ANSI.
- GKS was originally designed as the two dimensional graphics package and then later extension was developed for three dimensions.

PHIGS (Programmer's Hierarchical Interactive Graphic Standard)

- PHIGS is extension of GKS. Increased capability for object modeling, color specifications, surface rendering, and picture manipulation are provided in PHIGS.
- Extension of PHIGS called **"PHIGS+"** was developed to provide three dimensional surface shading capabilities not available in PHIGS.

Points and Lines

- Point plotting is done by converting a single coordinate position furnished by an application program into appropriate operations for the output device in use.
- Line drawing is done by calculating intermediate positions along the line path between two specified endpoint positions.
- The output device is then directed to fill in those positions between the end points with some color.
- For some device such as a pen plotter or random scan display, a straight line can be drawn smoothly from one end point to other.
- Digital devices display a straight line segment by plotting discrete points between the two endpoints.
- Discrete coordinate positions along the line path are calculated from the equation of the line.
- For a raster video display, the line intensity is loaded in frame buffer at the corresponding pixel positions.
- Reading from the frame buffer, the video controller then plots the screen pixels.
- Screen locations are referenced with integer values, so plotted positions may only approximate actual line positions between two specified endpoints.
- For example line position of $(12.36, 23.87)$ would be converted to pixel position $(12, 24)$.
- This rounding of coordinate values to integers causes lines to be displayed with a stair step appearance ("the jaggies"), as represented in fig 2.1.

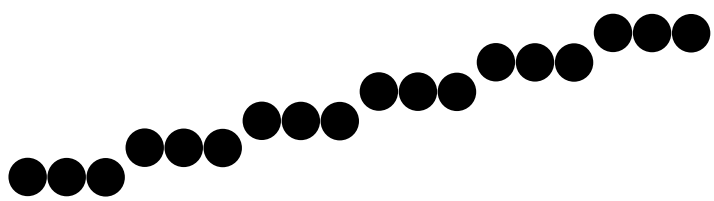


Fig. 2.1: - Stair step effect produced when line is generated as a series of pixel positions.

- The stair step shape is noticeable in low resolution system, and we can improve their appearance somewhat by displaying them on high resolution system.
- More effective techniques for smoothing raster lines are based on adjusting pixel intensities along the line paths.
- For raster graphics device-level algorithms discuss here, object positions are specified directly in integer device coordinates.
- Pixel position will referenced according to scan-line number and column number which is illustrated by following figure.

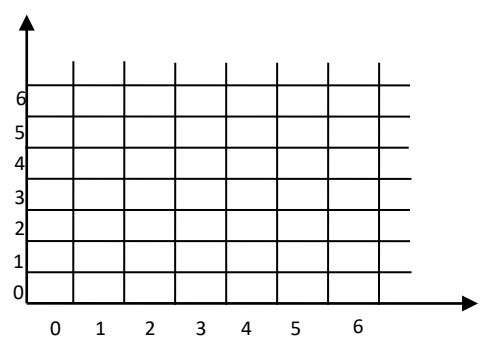


Fig. 2.2: - Pixel positions referenced by scan-line number and column number.

- To load the specified color into the frame buffer at a particular position, we will assume we have available low-level procedure of the form $setpixel(x, y)$.

- Similarly for retrieve the current frame buffer intensity we assume to have procedure $getpixel(x, y)$.

Line Drawing Algorithms

- The Cartesian slop-intercept equation for a straight line is "$y = mx + b$" with 'm' representing slop and 'b' as the intercept.
- The two endpoints of the line are given which are say $(x1, y1)$ and $(x2, y2)$.

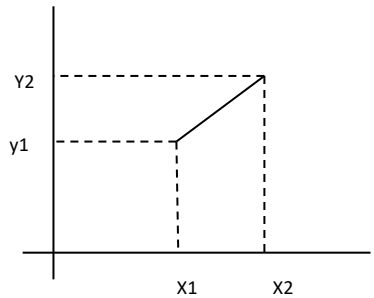


Fig. 2.3: - Line path between endpoint positions.

- We can determine values for the slope m by equation:
  $m = (y2 - y1)/(x2 - x1)$
- We can determine values for the intercept b by equation:
  $b = y1 - m * x1$
- For the given interval Δx along a line, we can compute the corresponding y interval Δy as:
  $\Delta y = m * \Delta x$
- Similarly for Δx:
  $\Delta x = \Delta y/m$
- For line with slop $|m| < 1$, Δx can be set proportional to small horizontal deflection voltage and the corresponding vertical deflection voltage is then set proportional to Δy which is calculated from above equation.
- For line with slop $|m| > 1$, Δy can be set proportional to small vertical deflection voltage and the corresponding horizontal deflection voltage is then set proportional to Δx which is calculated from above equation.
- For line with slop $m = 1$, $\Delta x = \Delta y$ and the horizontal and vertical deflection voltages are equal.

DDA Algorithm

- Digital differential analyzer (DDA) is scan conversion line drawing algorithm based on calculating either Δy or Δx using above equation.
- We sample the line at unit intervals in one coordinate and find corresponding integer values nearest the line path for the other coordinate.
- Consider first a line with positive slope and slope is less than or equal to 1:
  We sample at unit x interval ($\Delta x = 1$) and calculate each successive y value as follow:
  $y = m * x + b$
  $y_k = m * (x + 1) + b$
  In general $y_k = m * (x + k) + b$, &
  $y_{k+1} = m * (x + k + 1) + b$
  Now write this equation in form:
  $y_{k+1} - y_k = (m * (x + k + 1) + b) - (m * (x + k) + b)$
  $y_{k+1} = y_k + m$
  So that it is computed fast in computer as addition is fast compare to multiplication.

- In above equation k takes integer values starting from 1 and increase by 1 until the final endpoint is reached.
- As m can be any real number between 0 and 1, the calculated y values must be rounded to the nearest integer.
- Consider a case for a line with a positive slope greater than 1:
  We change the role of x and y that is sample at unit y intervals ($\Delta y = 1$) and calculate each succeeding x value as:
  $x = (y - b)/m$
  $x_1 = ((y + 1) - b)/m$
  In general $x_k = ((y + k) - b)/m$, &
  $x_{k+1} = ((y + k + 1) - b)/m$
  Now write this equation in form:
  $x_{k+1} - x_k = (((y + k + 1) - b)/m) - (((y + k) - b)/m)$
  $x_{k+1} = x_k + 1/m$
- Above both equations are based on the assumption that lines are to be processed from left endpoint to the right endpoint.
- If we processed line from right endpoint to left endpoint than:
  If $\Delta x = -1$ equation become:
  $y_{k+1} = y_k - m$
  If $\Delta y = -1$ equation become:
  $x_{k+1} = x_k - 1/m$
- Above calculated equations also used to calculate pixel position along a line with negative slope.
- **Procedure for DDA line algorithm.**

```
Void lineDDA (int xa, int ya, int xb, int yb)
{
    int dx = xb – xa, dy = yb – ya, steps, k;
    float xincrement, yincrement, x = xa, y = ya;
    if (abs(dx)>abs(dy))
    {
            Steps = abs (dx);
    }
    else
    {
            Steps = abs (dy);
    }
    xincrement = dx/(float) steps;
    yincrement = dy/(float) steps;

    setpixel (ROUND (x), ROUND (y));
    for(k=0;k<steps;k++)
    {
            x += xincrement;
            y += yincrement;
            setpixel (ROUND (x), ROUND (y));
    }
}
```

Advantages of DDA algorithm

- It is faster algorithm.
- It is simple algorithm.

Disadvantage of DDA algorithm

- Floating point arithmetic is time consuming.
- Poor end point accuracy.

Bresenham's Line Algorithm

- An accurate and efficient raster line-generating algorithm, developed by Bresenham which scan converts line using only incremental integer calculations that can be modified to display circles and other curves.
- Figure shows section of display screen where straight line segments are to be drawn.

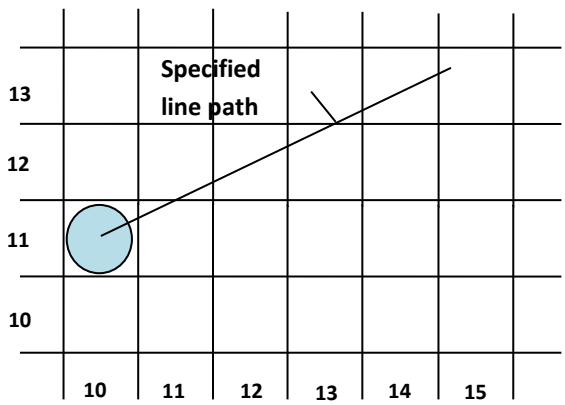


Fig. 2.4: - Section of a display screen where a straight line segment is to be plotted, starting from the pixel at column 10 on scan line 11.

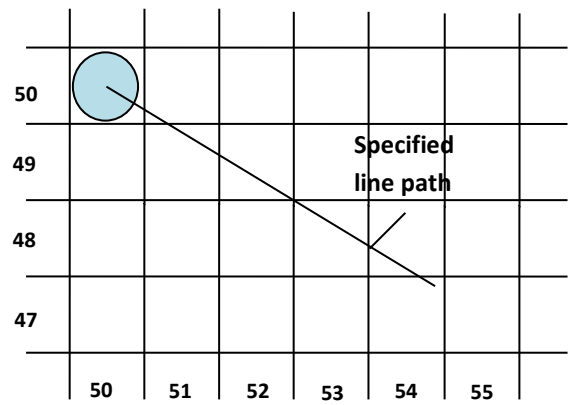


Fig. 2.5: - Section of a display screen where a negative slope line segment is to be plotted, starting from the pixel at column 50 on scan line 50.

- The vertical axes show scan-line positions and the horizontal axes identify pixel column.
- Sampling at unit x intervals in these examples, we need to decide which of two possible pixel position is closer to the line path at each sample step.
- To illustrate bresenham's approach, we first consider the scan-conversion process for lines with positive slope less than 1.
- Pixel positions along a line path are then determined by sampling at unit x intervals.
- Starting from left endpoint (x_0, y_0) of a given line, we step to each successive column and plot the pixel whose scan-line y values is closest to the line path.
- Assuming we have determined that the pixel at (x_k, y_k) is to be displayed, we next need to decide which pixel to plot in column $x_k + 1$.
- Our choices are the pixels at positions $(x_k + 1, y_k)$ and $(x_k + 1, y_k + 1)$.
- Let's see mathematical calculation used to decide which pixel position is light up.
- We know that equation of line is:
  $y = mx + b$
  Now for position $x_k + 1$.
  $y = m(x_k + 1) + b$
- Now calculate distance bet actual line's y value and lower pixel as d_1 and distance bet actual line's y value and upper pixel as d_2.
  $d_1 = y - y_k$

$d_1 = m(x_k + 1) + b - y_k$.....(1)

$d_2 = (y_k + 1) - y$

$d_2 = (y_k + 1) - m(x_k + 1) - b$.....(2)

- Now calculate $d_1 - d_2$ from equation (1) and (2).

  $d_1 - d_2 = (y - y_k) - ((y_k + 1) - y)$

  $d_1 - d_2 = \{m(x_k + 1) + b - y_k\} - \{(y_k + 1) - m(x_k + 1) - b\}$

  $d_1 - d_2 = \{mx_k + m + b - y_k\} - \{y_k + 1 - mx_k - m - b\}$

  $d_1 - d_2 = 2m(x_k + 1) - 2y_k + 2b - 1$.....(3)

- Now substitute $m = \Delta y/\Delta x$ in equation (3)

  $d_1 - d_2 = 2\left(\frac{\Delta y}{\Delta x}\right)(x_k + 1) - 2y_k + 2b - 1$ .....(4)

- Now we have decision parameter p_k for k^{th} step in the line algorithm is given by:

  $p_k = \Delta x(d_1 - d_2)$

  $p_k = \Delta x(2\Delta y/\Delta x(x_k + 1) - 2y_k + 2b - 1)$

  $p_k = 2\Delta y x_k + 2\Delta y - 2\Delta x y_k + 2\Delta x b - \Delta x$

  $p_k = 2\Delta y x_k - 2\Delta x y_k + 2\Delta y + 2\Delta x b - \Delta x$ .....(5)

  $p_k = 2\Delta y x_k - 2\Delta x y_k + C\ (Where\ Constant\ C = 2\Delta y + 2\Delta x b - \Delta x)$.....(6)

- The sign of p_k is the same as the sign of $d_1 - d_2$, since $\Delta x > 0$ for our example.
- Parameter c is constant which is independent of pixel position and will eliminate in the recursive calculation for p_k.
- Now if p_k is negative then we plot the lower pixel otherwise we plot the upper pixel.
- So successive decision parameters using incremental integer calculation as:

  $p_{k+1} = 2\Delta y x_{k+1} - 2\Delta x y_{k+1} + C$

- Now Subtract pk from p_{k+1}

  $p_{k+1} - p_k = 2\Delta y(x_{k+1} - x_k)$ -$2\Delta x(y_{k+1} - y_k)$

  $p_{k+1} - p_k = 2\Delta y x_{k+1} - 2\Delta x y_{k+1} + C - 2\Delta y x_k + 2\Delta x y_k - C$

  But $x_{k+1} = x_k + 1$, so that $(x_{k+1} - x_k) = 1$

  $p_{k+1} = p_k + 2\Delta y - 2\Delta x(y_{k+1} - y_k)$

- Where the terms $y_{k+1} - y_k$ is either 0 or 1, depends on the sign of parameter p_k.
- This recursive calculation of decision parameters is performed at each integer x position starting at the left coordinate endpoint of the line.
- The first decision parameter p_0 is calculated using equation (5) as first time we need to take constant part into account so:

  $p_k = 2\Delta y x_k - 2\Delta x y_k + 2\Delta y + 2\Delta x b - \Delta x$

  $p_0 = 2\Delta y x_0 - 2\Delta x y_0 + 2\Delta y + 2\Delta x b - \Delta x$

  Now $Substitute\ b = y_0 - mx_0$

  $p_0 = 2\Delta y x_0 - 2\Delta x y_0 + 2\Delta y + 2\Delta x(y_0 - mx_0) - \Delta x$

  $Now\ Substitute\ m = \Delta y/\Delta x$

  $p_0 = 2\Delta y x_0 - 2\Delta x y_0 + 2\Delta y + 2\Delta x(y_0 - (\Delta y/\Delta x)x_0) - \Delta x$

  $p_0 = 2\Delta y x_0 - 2\Delta x y_0 + 2\Delta y + 2\Delta x y_0 - 2\Delta y x_0 - \Delta x$

  $p_0 = 2\Delta y - \Delta x$

- Let's see Bresenham's line drawing algorithm for $|m| < 1$
  1. Input the two line endpoints and store the left endpoint in (x_0, y_0).
  2. Load (x_0, y_0) into the frame buffer; that is, plot the first point.
  3. Calculate constants Δx, Δy, $2\Delta y$, and $2\Delta y - 2\Delta x$, and obtain the starting value for the decision parameter as

$$p_0 = 2\Delta y - \Delta x$$

4. At each x_k along the line, starting at $k = 0$, perform the following test:

   If $p_k < 0$, the next point to plot is $(x_k + 1, y_k)$ and

   $$p_{k+1} = p_k + 2\Delta y$$

   Otherwise, the next point to plot is $(x_k + 1, y_k + 1)$ and

   $$p_{k+1} = p_k + 2\Delta y - 2\Delta x$$

5. Repeat step-4 Δx times.

- Bresenham's algorithm is generalized to lines with arbitrary slope by considering symmetry between the various octants and quadrants of the xy plane.
- For lines with positive slope greater than 1 we interchange the roles of the x and y directions.
- Also we can revise algorithm to draw line from right endpoint to left endpoint, both x and y decrease as we step from right to left.
- When $d_1 - d_2 = 0$ we choose either lower or upper pixel but once we choose lower than for all such case for that line choose lower and if we choose upper the for all such case choose upper.
- For the negative slope the procedure are similar except that now one coordinate decreases as the other increases.
- The special case handle separately. Horizontal line ($\Delta y = 0$), vertical line ($\Delta x = 0$) and diagonal line with $|\Delta x| = |\Delta y|$ each can be loaded directly into the frame buffer without processing them through the line plotting algorithm.

Parallel Execution of Line Algorithms

- The line-generating algorithms we have discussed so far determine pixel positions sequentially.
- With parallel computer we can calculate pixel position along a line path simultaneously by dividing work among the various processors available.
- One way to use multiple processors is partitioning existing sequential algorithm into small parts and compute separately.
- Alternatively we can go for other ways to setup the processing so that pixel positions can be calculated efficiently in parallel.
- Important point to be taking into account while devising parallel algorithm is to balance the load among the available processors.
- Given $\boldsymbol{n_p}$ number of processors we can set up parallel Bresenham line algorithm by subdividing the line path into $\boldsymbol{n_p}$ partitions and simultaneously generating line segment in each of the subintervals.
- For a line with slope $0 < m < 1$ and left endpoint coordinate position (x_0, y_0), we partition the line along the positive x direction.
- The distance between beginning x positions of adjacent partitions can be calculated as:

  $\Delta x_p = (\Delta x + n_p - 1)/n_p$

  Were Δx is the width of the line. And value for partition with Δx_p is computed using integer division.
- Numbering the partitions and the processors, as $0, 1, 2$, up to $\boldsymbol{n_p - 1}$, we calculate the starting x coordinate for the k^{th} partition as:

  $x_k = x_0 + k\Delta x_p$
- To apply Bresenham's algorithm over the partitions, we need the initial value for the y coordinate and the initial value for the decision parameter in each partition.
- The change Δy_p in the y direction over each partition is calculated from the line slope m and partition width Δx_p:
- $\Delta y_p = m\Delta x_p$

- At the k^{th} partition, the starting y coordinate is then
- $y_k = y_0 + round(k\Delta y_p)$
- The initial decision parameter for Bresenham's algorithm at the start of the k^{th} subinterval is obtained from Equation(6):

  $p_k = 2\Delta y x_k - 2\Delta x y_k + 2\Delta y + 2\Delta x b - \Delta x$

  $p_k = 2\Delta y(x_0 + k\Delta x_p) - 2\Delta x(y_0 + round(k\Delta y_p)) + 2\Delta y + 2\Delta x(y_0 - \frac{\Delta y}{\Delta x}x_0) - \Delta x$

  $p_k = 2\Delta y x_0 - 2\Delta y k\Delta x_p - 2\Delta x y_0 - 2\Delta x round(k\Delta y_p) + 2\Delta y + 2\Delta x y_0 - 2\Delta y x_0 - \Delta x$

  $p_k = 2\Delta y k\Delta x_p - 2\Delta x round(k\Delta y_p) + 2\Delta y - \Delta x$
- Each processor then calculates pixel positions over its assigned subinterval.
- The extension of the parallel Bresenham algorithm to a line with slope greater than 1 is achieved by partitioning the line in the y direction and calculating beginning x values for the positions.
- For negative slopes, we increment coordinate values in one direction and decrement in the other.

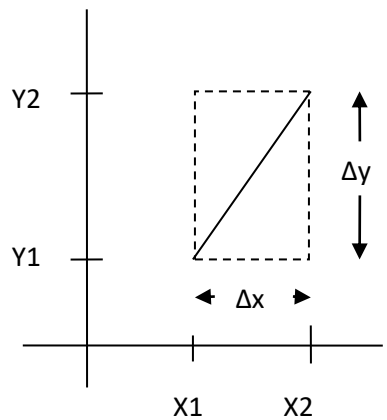


Fig. 2.6: - Bounding box for a line with coordinate extents Δx and Δy.

- Another way to set up parallel algorithms on raster system is to assign each processor to a particular group of screen pixels.
- With sufficient number of processor we can assign each processor to one pixel within some screen region.
- This approach can be adapted to line display by assigning one processor to each of the pixels within the limit of the bounding rectangle and calculating pixel distance from the line path.
- The number of pixels within the bounding rectangle of a line is $\Delta x \times \Delta y$.
- Perpendicular distance d from line to a particular pixel is calculated by:

  $d = Ax + By + C$

  Where

  $A = -\Delta y/linelength$

  $B = -\Delta x/linelength$

  $C = (x_0\Delta y - y_0\Delta x)/linelength$

  With

  $linelength = \sqrt{\Delta x^2 + \Delta y^2}$
- Once the constant A, B, and C have been evaluated for the line each processors need to perform two multiplications and two additions to compute the pixel distance d.
- A pixel is plotted if d is less than a specified line thickness parameter.
- Instead of partitioning the screen into single pixels, we can assign to each processor either a scan line or a column a column of pixels depending on the line slope.

- Each processor calculates line intersection with horizontal row or vertical column of pixels assigned to that processor.
- If vertical column is assign to processor then x is fix and it will calculate y and similarly is horizontal row is assign to processor then y is fix and x will be calculated.
- Such direct methods are slow in sequential machine but we can perform very efficiently using multiple processors.

Circle

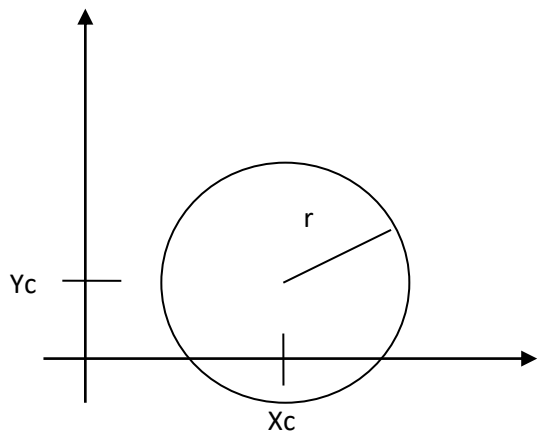


Fig. 2.7: - Circle with center coordinates (x_c, y_c) and radius r.

- A circle is defined as the set of points that are all at a given distance r from a center position say (x_c, y_c).

Properties of Circle

- The distance relationship is expressed by the Pythagorean theorem in Cartesian coordinates as:
  $(x - x_c)^2 + (y - y_c)^2 = r^2$
- We could use this equation to calculate circular boundary points by incrementing 1 in x direction in every steps from $x_c - r$ to $x_c + r$ and calculate corresponding y values at each position as:
  $(x - x_c)^2 + (y - y_c)^2 = r^2$
  $(y - y_c)^2 = r^2 - (x - x_c)^2$
  $(y - y_c) = \pm\sqrt{r^2 - (x_c - x)^2}$
  $y = y_c \pm \sqrt{r^2 - (x_c - x)^2}$
- But this is not best method for generating a circle because it requires more number of calculations which take more time to execute.
- And also spacing between the plotted pixel positions is not uniform as shown in figure below.

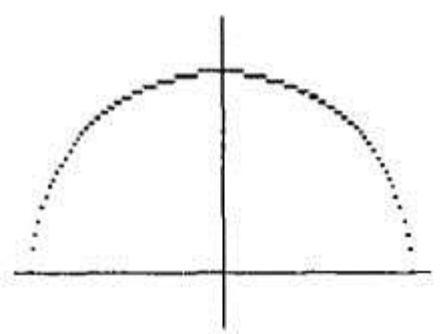


Fig. 2.8: - Positive half of circle showing non uniform spacing bet calculated pixel positions.

- We can adjust spacing by stepping through y values and calculating x values whenever the absolute value of the slop of the circle is greater than 1. But it will increases computation processing requirement.
- Another way to eliminate the non-uniform spacing is to draw circle using polar coordinates 'r' and 'θ'.
- Calculating circle boundary using polar equation is given by pair of equations which is as follows.
  $x = x_c + r\cos\theta$
  $y = y_c + r\sin\theta$
- When display is produce using these equations using fixed angular step size circle is plotted with uniform spacing.
- The step size 'θ' is chosen according to application and display device.
- For a more continuous boundary on a raster display we can set the step size at $1/r$. This plot pixel position that are approximately one unit apart.
- Computation can be reduced by considering symmetry city property of circles. The shape of circle is similar in each quadrant.
- We can obtain pixel position in second quadrant from first quadrant using reflection about y axis and similarly for third and fourth quadrant from second and first respectively using reflection about x axis.
- We can take one step further and note that there is also symmetry between octants. Circle sections in adjacent octant within one quadrant are symmetric with respect to the 45^0 line dividing the two octants.
- This symmetry condition is shown in figure below where point (x, y) on one circle sector is mapped in other seven sector of circle.

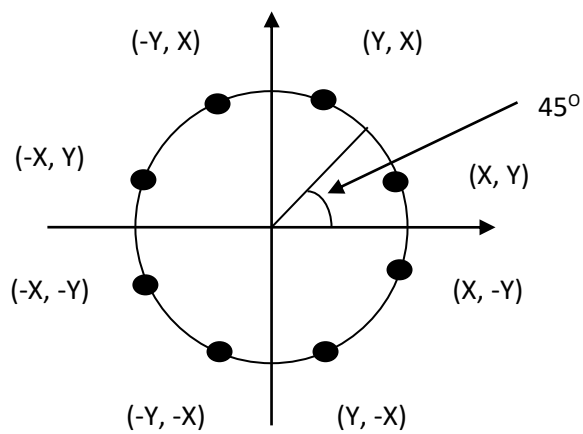


Fig. 2.9: - symmetry of circle.

- Taking advantage of this symmetry property of circle we can generate all pixel position on boundary of circle by calculating only one sector from $x = 0$ to $x = y$.
- Determining pixel position along circumference of circle using any of two equations shown above still required large computation.
- More efficient circle algorithm are based on incremental calculation of decision parameters, as in the Bresenham line algorithm.
- Bresenham's line algorithm can be adapted to circle generation by setting decision parameter for finding closest pixel to the circumference at each sampling step.
- The Cartesian coordinate circle equation is nonlinear so that square root evaluations would be required to compute pixel distance from circular path.
- Bresenham's circle algorithm avoids these square root calculation by comparing the square of the pixel separation distance.

- A method for direct distance comparison to test the midpoint between two pixels to determine if this midpoint is inside or outside the circle boundary.
- This method is easily applied to other conics also.
- Midpoint approach generates same pixel position as generated by bresenham's circle algorithm.
- The error involve in locating pixel positions along any conic section using midpoint test is limited to one-half the pixel separation.

Midpoint Circle Algorithm

- Similar to raster line algorithm we sample at unit interval and determine the closest pixel position to the specified circle path at each step.
- Given radius 'r' and center (x_c, y_c)
- We first setup our algorithm to calculate circular path coordinates for center (0, 0). And then we will transfer calculated pixel position to center (x_c, y_c) by adding x_c to x and y_c to y.
- Along the circle section from $x = 0$ to $x = y$ in the first quadrant, the slope of the curve varies from 0 to -1 so we can step unit step in positive x direction over this octant and use a decision parameter to determine which of the two possible y position is closer to the circular path.
- Position in the other seven octants are then obtain by symmetry.
- For the decision parameter we use the circle function which is:
  $f_{circle}(x, y) = x^2 + y^2 - r^2$
- Any point which is on the boundary is satisfied $f_{circle}(x, y) = 0$ if the point is inside circle function value is negative and if point is outside circle the function value is positive which can be summarize as below.

$$f_{circle}(x,y)\begin{cases} < 0 & if\ (x,y) is\ inside\ the\ circle\ boundary \\ = 0 & if\ (x,y) is\ on\ the\ circle\ boundary \\ > 0 & if\ (x,y) is\ outside\ the\ circle\ boundary \end{cases}$$

- Above equation we calculate for the mid positions between pixels near the circular path at each sampling step and we setup incremental calculation for this function as we did in the line algorithm.
- Below figure shows the midpoint between the two candidate pixels at sampling position $x_k + 1$.

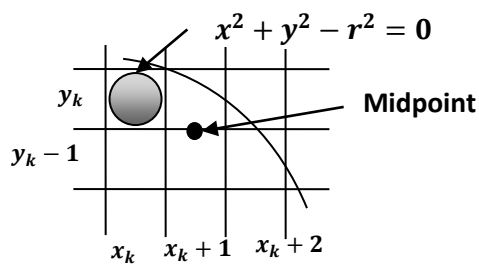


Fig. 2.10: - Midpoint between candidate pixel at sampling position $x_k + 1$ along circle path.

- Assuming we have just plotted the pixel at (x_k, y_k) and next we need to determine whether the pixel at position '$(x_k + 1, y_k)$' or the one at position' $(x_k + 1, y_k - 1)$' is closer to circle boundary.
- So for finding which pixel is more closer using decision parameter evaluated at the midpoint between two candidate pixels as below:

  $p_k = f_{circle}\left(x_k + 1, y_k - \frac{1}{2}\right)$

  $p_k = (x_k + 1)^2 + \left(y_k - \frac{1}{2}\right)^2 - r^2$

- If $p_k < 0$ this midpoint is inside the circle and the pixel on the scan line y_k is closer to circle boundary. Otherwise the midpoint is outside or on the boundary and we select the scan line $y_k - 1$.
- Successive decision parameters are obtain using incremental calculations as follows:

$p_{k+1} = f_{circle}\left(x_{k+1}+1, y_{k+1}-\frac{1}{2}\right)$

$p_{k+1} = [(x_k+1)+1]^2 + \left(y_{k+1}-\frac{1}{2}\right)^2 - r^2$

- Now we can obtain recursive calculation using equation of p_{k+1} and p_k as follow.

$p_{k+1} - p_k = \left([(x_k+1)+1]^2 + \left(y_{k+1}-\frac{1}{2}\right)^2 - r^2\right) - \left((x_k+1)^2 + \left(y_k-\frac{1}{2}\right)^2 - r^2\right)$

$p_{k+1} - p_k = (x_k+1)^2 + 2(x_k+1) + 1 + {y_{k+1}}^2 - y_{k+1} + \frac{1}{4} - r^2 - (x_k+1)^2 - {y_k}^2 + y_k - \frac{1}{4} + r^2$

$p_{k+1} - p_k = 2(x_k+1) + 1 + {y_{k+1}}^2 - y_{k+1} - {y_k}^2 + y_k$

$p_{k+1} - p_k = 2(x_k+1) + ({y_{k+1}}^2 - {y_k}^2) - (y_{k+1} - y_k) + 1$

$p_{k+1} = p_k + 2(x_k+1) + ({y_{k+1}}^2 - {y_k}^2) - (y_{k+1} - y_k) + 1$

- In above equation y_{k+1} is either y_k or $y_k - 1$ depending on the sign of the p_k.
- Now we can put $2x_{k+1} = 2x_k + 2$ and when we select $y_{k+1} = y_k - 1$ we can obtain $2y_{k+1} = 2y_k - 2$.
- The initial decision parameter is obtained by evaluating the circle function at the start position $(x_0, y_0) = (0, r)$ as follows.

$p_0 = f_{circle}\left(0+1, r-\frac{1}{2}\right)$

$p_0 = 1^2 + \left(r-\frac{1}{2}\right)^2 - r^2$

$p_0 = 1 + r^2 - r + \frac{1}{4} - r^2$

$p_0 = \frac{5}{4} - r$

Algorithm for Midpoint Circle Generation

1. Input radius r and circle center (x_c, y_c), and obtain the first point on the circumference of a circle centered on the origin as
$$(x_0, y_0) = (0, r)$$
2. calculate the initial value of the decision parameter as
$$p_0 = \frac{5}{4} - r$$
3. At each x_k position, starting at $k = 0$, perform the following test:
If $p_k < 0$, the next point along the circle centered on $(0, 0)$ is $(x_k + 1, y_k)$ &
$$p_{k+1} = p_k + 2x_{k+1} + 1$$
Otherwise, the next point along the circle is $(x_k + 1, y_k - 1)$ &
$$p_{k+1} = p_k + 2x_{k+1} + 1 - 2y_{k+1}$$
Where $2x_{k+1} = 2x_k + 2$, & $2y_{k+1} = 2y_k - 2$.
4. Determine symmetry points in the other seven octants.
5. Move each calculated pixel position (x, y) onto the circular path centered on (x_c, y_c) and plot the coordinate values:
$$x = x + x_c, y = y + y_c$$
6. Repeat steps 3 through 5 until $x \geq y$.

Ellipse

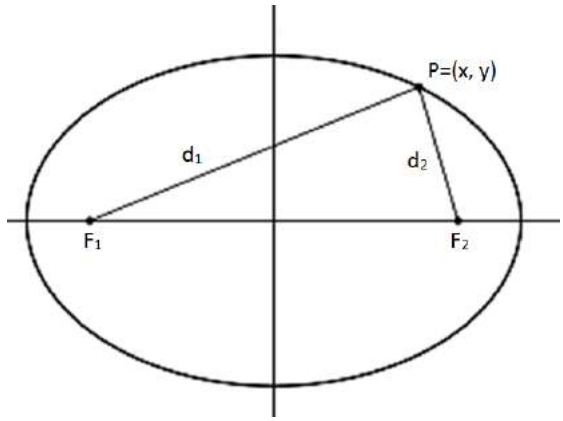


Fig. 2.11: - Ellipse generated about foci f_1 and f_2.

- AN ellipse is defined as the set of points such that the sum of the distances from two fixed positions (**foci**) is same for all points.

Properties of Ellipse

- If we labeled distance from two foci to any point on ellipse boundary as d_1 and d_2 then the general equation of an ellipse can be written as follow.
  $d_1 + d_2 = Constant$
- Expressing distance in terms of focal coordinates $f_1 = (x_1, y_1)$ and $f_2 = (x_2, y_2)$ we have
  $\sqrt{(x - x_1)^2 + (y - y_1)^2} + \sqrt{(x - x_2)^2 + (y - y_2)^2} = Constant$
- An interactive method for specifying an ellipse in an arbitrary orientation is to input two foci and a point on the ellipse boundary.
- With this three coordinates we can evaluate constant in equation:
- $\sqrt{(x - x_1)^2 + (y - y_1)^2} + \sqrt{(x - x_2)^2 + (y - y_2)^2} = Constant$
- We can also write this equation in the form
  $Ax^2 + By^2 + Cxy + Dx + Ey + F = 0$
- Where the coefficients $A, B, C, D, E,$ and F are evaluated in terms of the focal coordinates and the dimensions of the major and minor axes of the ellipse.
- Major axis of an ellipse is straight line segment passing through both foci and extends up to boundary on both sides.
- The minor axis spans shortest dimension of ellipse, it bisect the major axis at right angle in two equal half.
- Then coefficient in $Ax^2 + By^2 + Cxy + Dx + Ey + F = 0$ can be evaluated and used to generate pixels along the elliptical path.
- Ellipse equation are greatly simplified if we align major and minor axis with coordinate axes i.e. $x - axis$ and $y - axis$.
- We can say ellipse is in standard position if their major and minor axes are parallel to $x - axis$ and $y - axis$ which is shown in below figure.

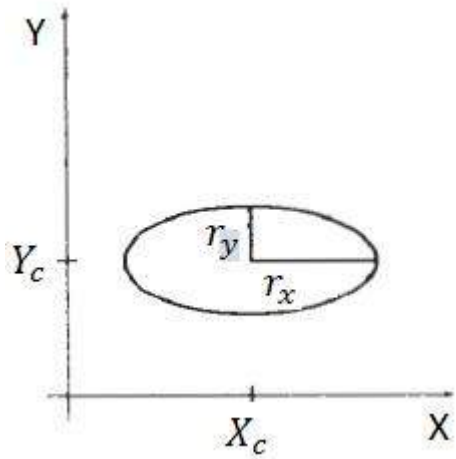


Fig. 2.12: - Ellipse centered at (x_c, y_c) with semi major axis r_x and semi minor axis r_y are parallel to coordinate axis.

- Equation of ellipse shown in figure 2.12 can be written in terms of the ellipse center coordinates and parameters r_x and r_y as.

  $$\left(\frac{x - x_c}{r_x}\right)^2 + \left(\frac{y - y_c}{r_y}\right)^2 = 1$$

- Using the polar coordinates r and θ, we can also describe the ellipse in standard position with the parametric equations:

  $x = x_c + r_x \cos\theta$

  $y = y_c + r_y \sin\theta$

- Symmetry considerations can be used to further reduced computations.
- An ellipse in standard position is symmetric between quadrants but unlike a circle it is not symmetric between octant.
- Thus we must calculate boundary point for one quadrant and then other three quadrants point can be obtained by symmetry as shown in figure below.

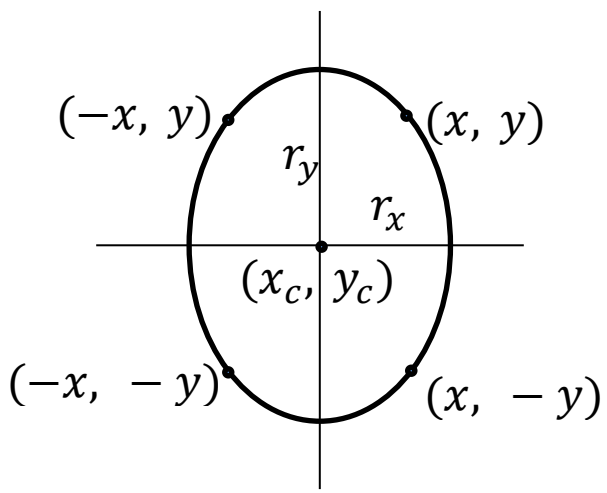


Fig. 2.13: - symmetry of an ellipse.

Midpoint Ellipse Algorithm

- Midpoint ellipse algorithm is a method for drawing ellipses in computer graphics. This method is modified from Bresenham's algorithm.

- The advantage of this modified method is that only addition operations are required in the program loops.
- This leads to simple and fast implementation in all processors.
- Given parameters r_x, r_y and (x_c, y_c) we determine points (x, y) for an ellipse in standard position centered on the origin, and then we shift the points so the ellipse is centered at (x_c, y_c).
- If we want to display the ellipse in nonstandard position then we rotate the ellipse about its center to align with required direction.
- For the present we consider only the standard position.
- In this method we divide first quadrant into two parts according to the slope of an ellipse as shown in figure below.

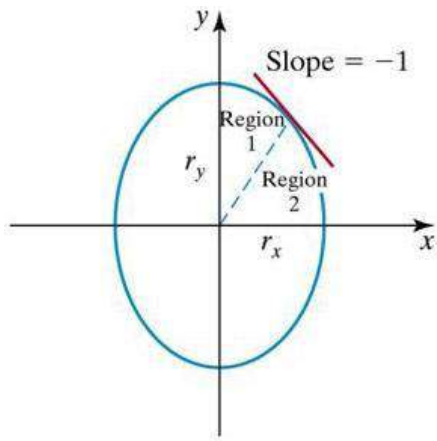


Fig. 2.14: - Ellipse processing regions. Over the region 1 the magnitude of ellipse slope is < 1 and over the region 2 the magnitude of ellipse slope > 1.

- We take unit step in x direction if magnitude of slope is less than 1 in that region otherwise we take unit step in y direction.
- Boundary divides region at $slope = -1$.
- With $r_x < r_y$ we process this quadrant by taking unit steps in x direction in region 1 and unit steps in y direction in region 2.
- Region 1 and 2 can be processed in various ways.
- We can start from $(0, r_y)$ and step clockwise along the elliptical path in the first quadrant shifting from unit step in x to unit step in y when slope becomes less than -1.
- Alternatively, we could start at $(r_x, 0)$ and select points in a counterclockwise order, shifting from unit steps in y to unit steps in x when the slope becomes greater than -1.
- With parallel processors, we could calculate pixel positions in the two regions simultaneously.
- Here we consider sequential implementation of midpoint algorithm. We take the start position at $(0, r_y)$ and steps along the elliptical path in clockwise order through the first quadrant.
- We define ellipse function for center of ellipse at $(0, 0)$ as follows.
  $f_{ellipse}(x, y) = {r_y}^2 x^2 + {r_x}^2 y^2 - {r_y}^2 {r_x}^2$
- Which has the following properties:

$$f_{ellipse}(x, y) \begin{cases} < 0 & if\ (x, y) is\ inside\ the\ ellipse\ boundary \\ = 0 & if\ (x, y) is\ on\ the\ ellipse\ boundary \\ > 0 & if\ (x, y) is\ outside\ the\ ellipse\ boundary \end{cases}$$

- Thus the ellipse function serves as the decision parameter in the midpoint ellipse algorithm.
- At each sampling position we select the next pixel from two candidate pixel.

- Starting at $(0, r_y)$ we take unit step in x direction until we reach the boundary between region 1 and 2 then we switch to unit steps in y direction in remaining portion on ellipse in first quadrant.
- At each step we need to test the value of the slope of the curve for deciding the end point of the region-1.
- The ellipse slope is calculated using following equation.
  $\frac{dy}{dx} = -\frac{2{r_y}^2 x}{2{r_x}^2 y}$
- At boundary between region 1 and 2 $slope = -1$ and equation become.
  $2{r_y}^2 x = 2{r_x}^2 y$
- Therefore we move out of region 1 whenever following equation is false
  $2{r_y}^2 x \leq 2{r_x}^2 y$
- Following figure shows the midpoint between the two candidate pixels at sampling position $x_k + 1$ in the first region.

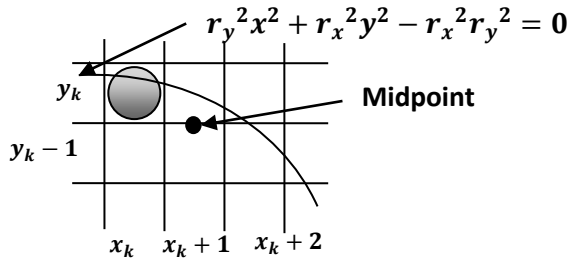


Fig. 2.15: - Midpoint between candidate pixels at sampling position $x_k + 1$ along an elliptical path.

- Assume we are at (x_k, y_k) position and we determine the next position along the ellipse path by evaluating decision parameter at midpoint between two candidate pixels.
  $p1_k = f_{ellipse}\left(x_k + 1, y_k - \frac{1}{2}\right)$
  $p1_k = {r_y}^2(x_k + 1)^2 + {r_x}^2\left(y_k - \frac{1}{2}\right)^2 - {r_x}^2{r_y}^2$
- If $p1_k < 0$, the midpoint is inside the ellipse and the pixel on scan line y_k is closer to ellipse boundary otherwise the midpoint is outside or on the ellipse boundary and we select the pixel $y_k - 1$.
- At the next sampling position decision parameter for region 1 is evaluated as.
  $p1_{k+1} = f_{ellipse}\left(x_{k+1} + 1, y_{k+1} - \frac{1}{2}\right)$
  $p1_{k+1} = {r_y}^2[(x_k + 1) + 1]^2 + {r_x}^2\left(y_{k+1} - \frac{1}{2}\right)^2 - {r_x}^2{r_y}^2$
- Now subtract $p1_k$ from $p1_{k+1}$
  $p1_{k+1} - p1_k = {r_y}^2[(x_k + 1) + 1]^2 + {r_x}^2\left(y_{k+1} - \frac{1}{2}\right)^2 - {r_x}^2{r_y}^2 - {r_y}^2(x_k + 1)^2 - {r_x}^2\left(y_k - \frac{1}{2}\right)^2 + {r_x}^2{r_y}^2$
  $p1_{k+1} - p1_k = {r_y}^2[(x_k + 1) + 1]^2 + {r_x}^2\left(y_{k+1} - \frac{1}{2}\right)^2 - {r_y}^2(x_k + 1)^2 - {r_x}^2\left(y_k - \frac{1}{2}\right)^2$
  $p1_{k+1} - p1_k = {r_y}^2(x_k + 1)^2 + 2{r_y}^2(x_k + 1) + {r_y}^2 + {r_x}^2\left(y_{k+1} - \frac{1}{2}\right)^2 - {r_y}^2(x_k + 1)^2 - {r_x}^2\left(y_k - \frac{1}{2}\right)^2$
  $p1_{k+1} - p1_k = 2{r_y}^2(x_k + 1) + {r_y}^2 + {r_x}^2\left[\left(y_{k+1} - \frac{1}{2}\right)^2 - \left(y_k - \frac{1}{2}\right)^2\right]$
- Now making $p1_{k+1}$ as subject.

$$p1_{k+1} = p1_k + 2r_y{}^2(x_k + 1) + r_y{}^2 + r_x{}^2\left[\left(y_{k+1} - \frac{1}{2}\right)^2 - \left(y_k - \frac{1}{2}\right)^2\right]$$

- Here y_{k+1} is either y_k or $y_k - 1$, depends on the sign of $p1_k$
- Now we calculate the initial decision parameter $p1_0$ by putting $(x_0, y_0) = (0, r_y)$ as follow.

$$p1_0 = f_{ellipse}\left(0 + 1, r_y - \frac{1}{2}\right)$$

$$p1_0 = r_y{}^2(1)^2 + r_x{}^2\left(r_y - \frac{1}{2}\right)^2 - r_x{}^2 r_y{}^2$$

$$p1_0 = r_y{}^2 + r_x{}^2\left(r_y - \frac{1}{2}\right)^2 - r_x{}^2 r_y{}^2$$

$$p1_0 = r_y{}^2 - r_x{}^2 r_y + \frac{1}{4} r_x{}^2$$

- Now we similarly calculate over region 2 by unit stepping in negative y direction and the midpoint is now taken between horizontal pixels at each step as shown in figure below.

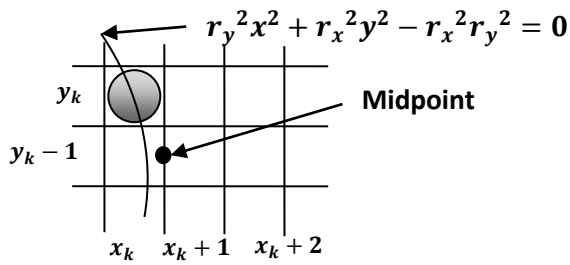


Fig. 2.16: - Midpoint between candidate pixels at sampling position $y_k - 1$ along an elliptical path.

- For this region, the decision parameter is evaluated as follows.

$$p2_k = f_{ellipse}\left(x_k + \frac{1}{2}, y_k - 1\right)$$

$$p2_k = r_y{}^2\left(x_k + \frac{1}{2}\right)^2 + r_x{}^2(y_k - 1)^2 - r_x{}^2 r_y{}^2$$

- If $p2_k > 0$ the midpoint is outside the ellipse boundary, and we select the pixel at x_k.
- If $p2_k \leq 0$ the midpoint is inside or on the ellipse boundary and we select $x_k + 1$.
- At the next sampling position decision parameter for region 2 is evaluated as.

$$p2_{k+1} = f_{ellipse}\left(x_{k+1} + \frac{1}{2}, y_{k+1} - 1\right)$$

$$p2_{k+1} = r_y{}^2\left(x_{k+1} + \frac{1}{2}\right)^2 + r_x{}^2[(y_k - 1) - 1]^2 - r_x{}^2 r_y{}^2$$

- Now subtract $p2_k$ from $p2_{k+1}$

$$p2_{k+1} - p2_k = r_y{}^2\left(x_{k+1} + \frac{1}{2}\right)^2 + r_x{}^2[(y_k - 1) - 1]^2 - r_x{}^2 r_y{}^2 - r_y{}^2\left(x_k + \frac{1}{2}\right)^2 - r_x{}^2(y_k - 1)^2 + r_x{}^2 r_y{}^2$$

$$p2_{k+1} - p2_k = r_y{}^2\left(x_{k+1} + \frac{1}{2}\right)^2 + r_x{}^2(y_k - 1)^2 - 2r_x{}^2(y_k - 1) + r_x{}^2 - r_y{}^2\left(x_k + \frac{1}{2}\right)^2 - r_x{}^2(y_k - 1)^2$$

$$p2_{k+1} - p2_k = r_y{}^2\left(x_{k+1} + \frac{1}{2}\right)^2 - 2r_x{}^2(y_k - 1) + r_x{}^2 - r_y{}^2\left(x_k + \frac{1}{2}\right)^2$$

$$p2_{k+1} - p2_k = -2r_x{}^2(y_k - 1) + r_x{}^2 + r_y{}^2\left[\left(x_{k+1} + \frac{1}{2}\right)^2 - \left(x_k + \frac{1}{2}\right)^2\right]$$

- Now making $p2_{k+1}$ as subject.

$$p2_{k+1} = p2_k - 2r_x{}^2(y_k - 1) + r_x{}^2 + r_y{}^2\left[\left(x_{k+1} + \frac{1}{2}\right)^2 - \left(x_k + \frac{1}{2}\right)^2\right]$$

- Here x_{k+1} is either x_k or $x_k + 1$, depends on the sign of $p2_k$.
- In region 2 initial position is selected which is last position of region one and the initial decision parameter is calculated as follows.

  $$p2_0 = f_{ellipse}\left(x_0 + \frac{1}{2}, y_0 - 1\right)$$

  $$p2_0 = {r_y}^2\left(x_0 + \frac{1}{2}\right)^2 + {r_x}^2(y_0 - 1)^2 - {r_x}^2{r_y}^2$$

- For simplify calculation of $p2_0$ we could select pixel position in counterclockwise order starting at $(r_x, 0)$.
- In above case we take unit step in the positive y direction up to the last point selected in region 1.

Algorithm for Midpoint Ellipse Generation

1. Input r_x, r_y and ellipse center (x_c, y_c), and obtain the first point on an ellipse centered on the origin as
   $$(x_0, y_0) = (0, r_y)$$
2. Calculate the initial value of the decision parameter in region 1 as
   $$p1_0 = {r_y}^2 - {r_x}^2 r_y + \frac{1}{4}{r_x}^2$$
3. At each x_k position in region 1, starting at $k = 0$, perform the following test:

   If $p1_k < 0$, the next point along the ellipse centered on $(0, 0)$ is (x_{k+1}, y_k) and
   $$p1_{k+1} = p1_k + 2{r_y}^2 x_{k+1} + {r_y}^2$$

   Otherwise, the next point along the ellipse is $(x_{k+1}, y_k - 1)$ and
   $$p1_{k+1} = p1_k + 2{r_y}^2 x_{k+1} + {r_y}^2 - 2{r_x}^2 y_{k+1}$$
   With
   $$2{r_y}^2 x_{k+1} = 2{r_y}^2 x_k + 2{r_y}^2$$
   $$2{r_x}^2 y_{k+1} = 2{r_x}^2 y_k - 2{r_x}^2$$
   And continue until $2{r_y}^2 x \leq 2{r_x}^2 y$
4. Calculate the initial value of the decision parameter in region 2 using the last point (x_0, y_0) calculated in region 1 as
   $$p2_0 = {r_y}^2\left(x_0 + \frac{1}{2}\right)^2 + {r_x}^2(y_0 - 1)^2 - {r_x}^2{r_y}^2$$
5. At each y_k position in region-2, starting at $k = 0$, perform the following test:

   If $p2_k > 0$, the next point along the ellipse centered on $(0, 0)$ is $(x_k, y_k - 1)$ and
   $$p2_{k+1} = p2_k - 2{r_x}^2 y_{k+1} + {r_x}^2$$
   Otherwise, the next point along the ellipse is $(x_k + 1, y_k - 1)$ and
   $$p2_{k+1} = p2_k - 2{r_x}^2 y_{k+1} + {r_x}^2 + 2{r_y}^2 x_{k+1}$$
   Using the same incremental calculations for x and y as in region 1.
6. Determine symmetry points in the other three quadrants.
7. Move each calculated pixel position (x, y) onto the elliptical path centered on (x_c, y_c) and plot the coordinate values:
   $$x = x + x_c$$
   $$y = y + y_c$$
   Repeat the steps for region 2 until $y_k \geq 0$.

Filled-Area Primitives

- In practical we often use polygon which are filled with some color or pattern inside it.

- There are two basic approaches to area filling on raster systems.
- One way to fill an area is to determine the overlap intervals for scan line that cross the area.
- Another method is to fill the area is to start from a given interior position and paint out wards from this point until we encounter boundary.

Scan-Line Polygon Fill Algorithm

- Figure below shows the procedure for scan-line filling algorithm.

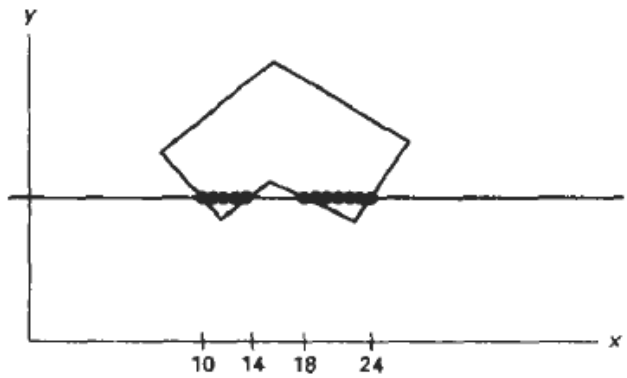


Fig. 2.17: - Interior pixels along a scan line passing through a polygon area.

- For each scan-line crossing a polygon, the algorithm locates the intersection points are of scan line with the polygon edges.
- This intersection points are stored from left to right.
- Frame buffer positions between each pair of intersection point are set to specified fill color.
- Some scan line intersects at vertex position they are required special handling.
- For vertex we must look at the other endpoints of the two line segments of the polygon which meet at this vertex.
- If these points lie on the same (up or down) side of the scan line, then that point is counts as two intersection points.
- If they lie on opposite sides of the scan line, then the point is counted as single intersection.
- This is illustrated in figure below

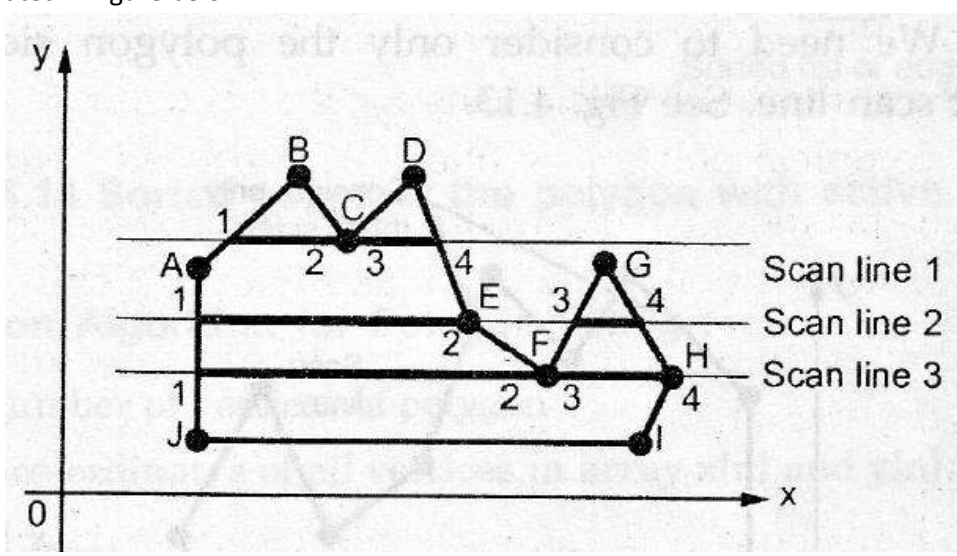


Fig. 2.18: - Intersection points along the scan line that intersect polygon vertices.

- As shown in the Fig. 2.18, each scan line intersects the vertex or vertices of the polygon. For scan line 1, the other end points (B and D) of the two line segments of the polygon lie on the same side of the scan

line, hence there are two intersections resulting two pairs: 1 -2 and 3 - 4. Intersections points 2 and 3 are actually same Points. For scan line 2 the other endpoints (D and F) of the two line segments of the Polygon lie on the opposite sides of the scan line, hence there is a single intersection resulting two pairs: l - 2 and 3 - 4. For scan line 3, two vertices are the intersection points"

- For vertex F the other end points E and G of the two line segments of the polygon lie on the same side of the scan line whereas for vertex H, the other endpoints G and I of the two line segments of the polygon lie on the opposite side of the scan line. Therefore, at vertex F there are two intersections and at vertex H there is only one intersection. This results two pairs: 1 - 2 and 3 - 4 and points 2 and 3 are actually same points.
- Coherence methods often involve incremental calculations applied along a single scan line or between successive scan lines.
- In determining edge intersections, we can set up incremental coordinate calculations along any edge by exploiting the fact that the slope of the edge is constant from one scan line to the next.
- Figure below shows three successive scan-lines crossing the left edge of polygon.

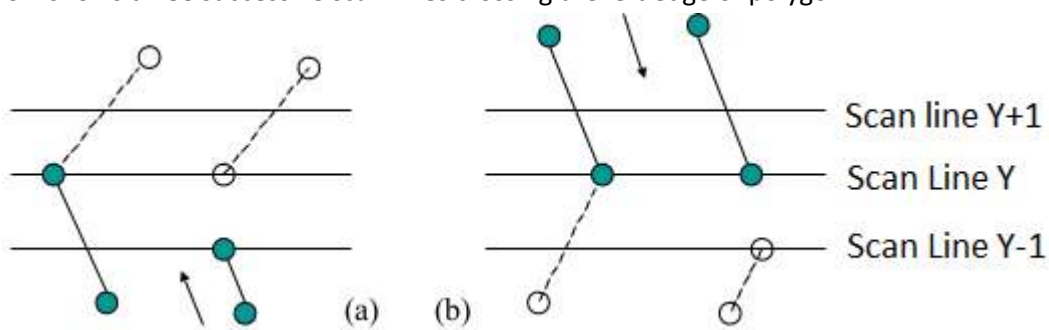


Fig. 2.18: - adjacent scan line intersects with polygon edge.

- For above figure we can write slope equation for polygon boundary as follows.

$$m = \frac{y_{k+1} - y_k}{x_{k+1} - x_k}$$

- Since change in y coordinates between the two scan lines is simply

$$y_{k+1} - y_k = 1$$

- So slope equation can be modified as follows

$$m = \frac{y_{k+1} - y_k}{x_{k+1} - x_k}$$

$$m = \frac{1}{x_{k+1} - x_k}$$

$$x_{k+1} - x_k = \frac{1}{m}$$

$$x_{k+1} = x_k + \frac{1}{m}$$

- Each successive x intercept can thus be calculated by adding the inverse of the slope and rounding to the nearest integer.
- For parallel execution of this algorithm we assign each scan line to separate processor in that case instead of using previous x values for calculation we use initial x values by using equation as.

$$x_k = x_0 + \frac{k}{m}$$

- Now if we put $m = \frac{\Delta y}{\Delta x}$ in incremental calculation equation $x_{k+1} = x_k + \frac{1}{m}$ then we obtain equation as.

$$x_{k+1} = x_k + \frac{\Delta x}{\Delta y}$$

- Using this equation we can perform integer evaluation of x intercept by initializing a counter to 0, then incrementing counter by the value of Δx each time we move up to a new scan line.
- When the counter value becomes equal to or greater than Δy, we increment the current x intersection value by 1 and decrease the counter by the value Δy.
- This procedure is seen by following figure.

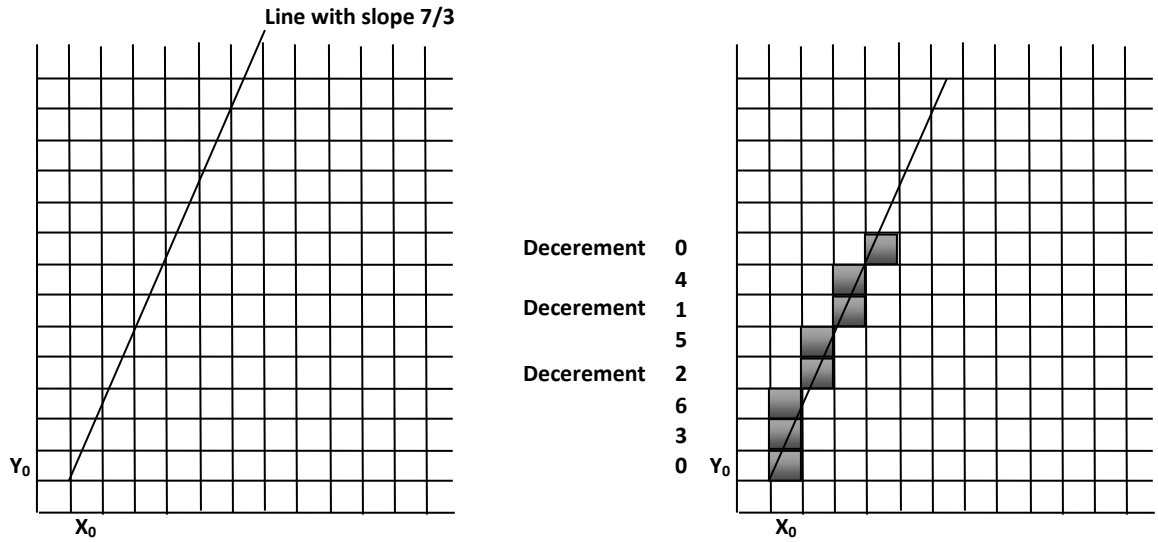


Fig. 2.19: - line with slope 7/3 and its integer calculation using equation $x_{k+1} = x_k + \frac{\Delta x}{\Delta y}$.

- Steps for above procedure

1. Suppose m = 7/3
2. Initially, set counter to 0, and increment to 3 (which is Δx).
3. When move to next scan line, increment counter by adding Δx
4. When counter is equal or greater than 7 (which is Δy), increment the x-intercept (in other words, the x-intercept for this scan line is one more than the previous scan line), and decrement counter by 7(which is Δy).

- To efficiently perform a polygon fill, we can first store the polygon boundary in a sorted edge table that contains all the information necessary to process the scan lines efficiently.
- We use bucket sort to store the edge sorted on the smallest y value of each edge in the correct scan line positions.
- Only the non-horizontal edges are entered into the sorted edge table.
- Figure below shows one example of storing edge table.

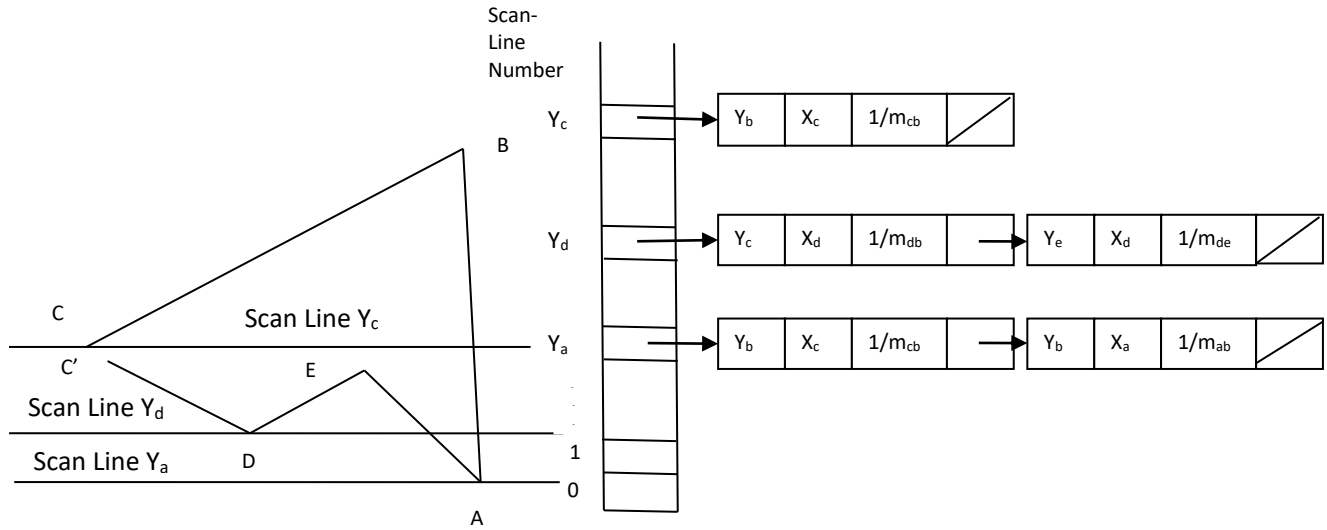


Fig. 2.20: - A polygon and its sorted edge table.

- Each entry in the table for a particular scan line contains the maximum y values for that edges, the x intercept value for edge, and the inverse slope of the edge.
- For each scan line the edges are in sorted order from left to right.
- Than we process the scan line from the bottom to top for whole polygon and produce active edge list for each scan line crossing the polygon boundaries.
- The active edge list for a scan line contains all edges crossed by that line, with iterative coherence calculation used to obtain the edge intersections.

Inside-Outside Tests

- In area filling and other graphics operation often required to find particular point is inside or outside the polygon.
- For finding which region is inside or which region is outside most graphics package use either odd even rule or the nonzero winding number rule.

Odd Even Rule

- It is also called the odd parity rule or even odd rule.
- By conceptually drawing a line from any position p to a distant point outside the coordinate extents of the object and counting the number of edges crossing by this line is odd, than p is an interior point. Otherwise p is exterior point.
- To obtain accurate edge count we must sure that line selected is does not pass from any vertices.
- This is shown in figure 2.21(a).

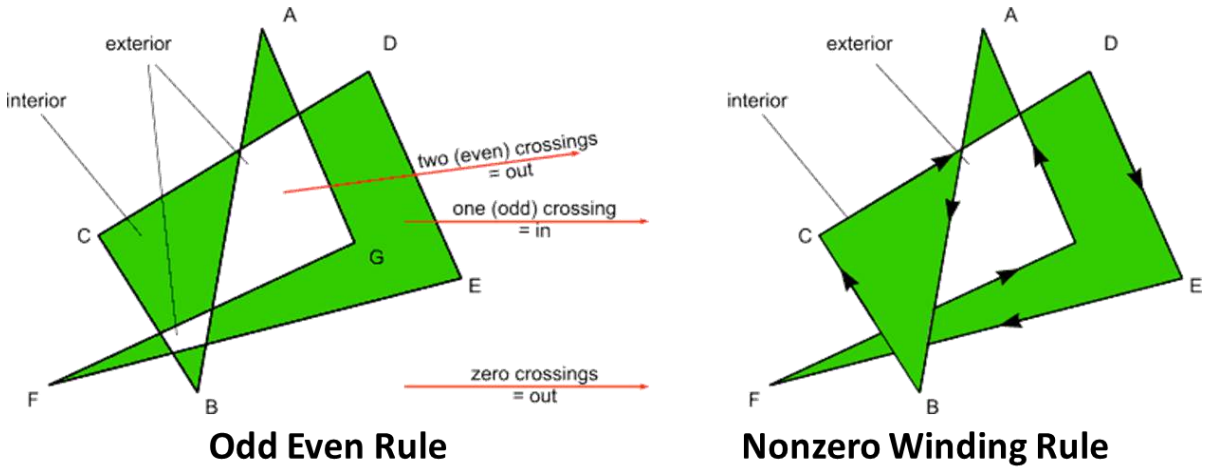


Fig. 2.21: - Identifying interior and exterior region for a self-intersecting polygon.

Nonzero Winding Number Rule

- This method counts the number of times the polygon edges wind around a particular point in the counterclockwise direction. This count is called the winding number, and the interior points of a two-dimensional object are defined to be those that have a nonzero value for the winding number.
- We apply this rule by initializing winding number with 0 and then draw a line for any point p to distant point beyond the coordinate extents of the object.
- The line we choose must not pass through vertices.
- Then we move along that line we find number of intersecting edges and we add 1 to winding number if edge cross our line from right to left and subtract 1 from winding number if edges cross from left to right.
- The final value of winding number is nonzero then the point is interior and if winding number is zero the point is exterior.
- This is shown in figure 2.21(b).
- One way to determine directional edge crossing is to take the vector cross product of a vector U along the line from p to distant point with the edge vector E for each edge that crosses the line.
- If z component of the cross product$U \times E$ for a particular edge is positive that edge is crosses from right to left and we add 1 to winding number otherwise the edge is crossing from left to right and we subtract 1 from winding number.

Comparison between Odd Even Rule and Nonzero Winding Rule

- For standard polygons and simple object both rule gives same result but for more complicated shape both rule gives different result which is illustrated in figure 2.21.

Scan-Line Fill of Curved Boundary Areas

- Scan-line fill of region with curved boundary is more time consuming as intersection calculation now involves nonlinear boundaries.
- For simple curve such as circle or ellipse scan line fill process is straight forward process.
- We calculate the two scan line intersection on opposite side of the curve.
- This is same as generating pixel position along the curve boundary using standard equation of curve.
- Then we fill the color between two boundary intersections.
- Symmetry property is used to reduce the calculation.
- Similar method can be used for fill the curve section.

Boundary Fill Algorithm/ Edge Fill Algorithm

- In this method, edges of the polygons are drawn. Then starting with some seed, any point inside the polygon we examine the neighbouring pixels to check whether the boundary pixel is reached.
- If boundary pixels are not reached, pixels are highlighted and the process is continued until boundary pixels are reached.
- Boundary defined regions may be either 4-cormected or 8-connected. as shown in the Figure below

(a) Four connected region (b) Eight connected region

Fig. 2.22: - Neighbor pixel connected to one pixel.

- If a region is 4-connected, then every pixel in the region may be reached by a combination of moves in only four directions: left, right, up and down.
- For an 8-connected region every pixel in the region may be reached by a combination of moves in the two horizontal, two vertical, and four diagonal directions.
- In some cases, an 8-connected algorithm is more accurate than the 4-connected algorithm. This is illustrated in Figure below. Here, a 4-connected algorithm produces the partial fill.

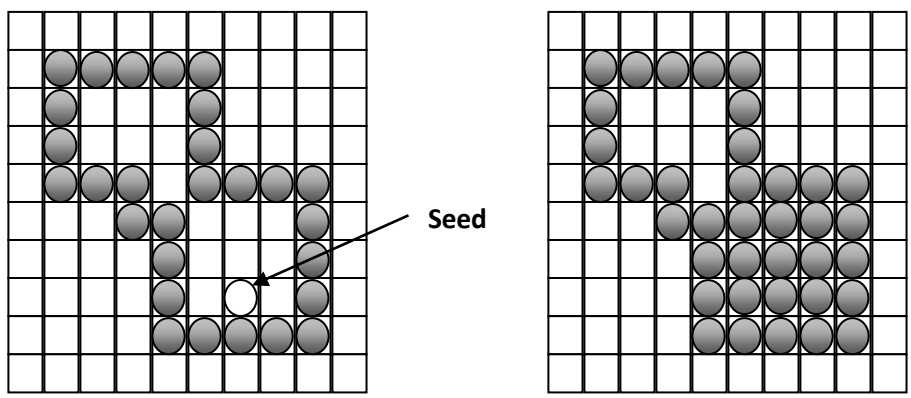


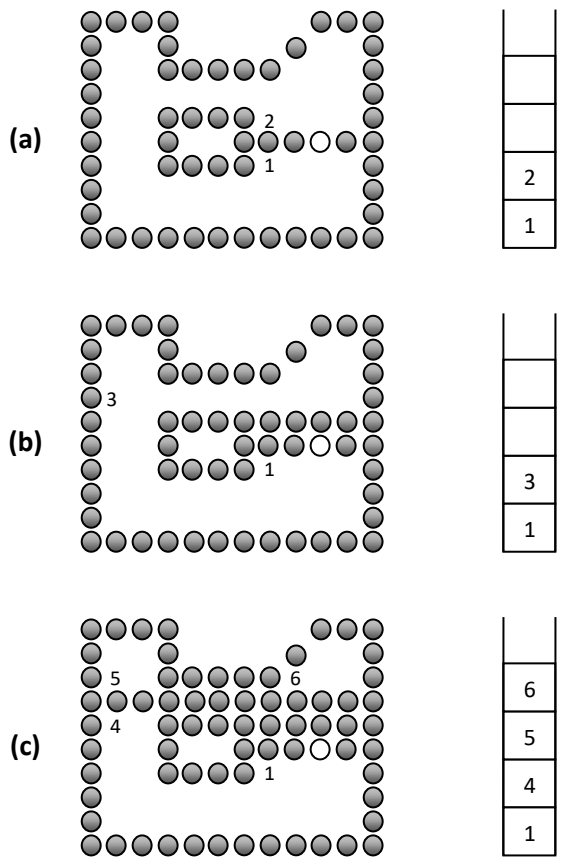
Fig. 2.23: - partial filling resulted due to using 4-connected algorithm.

- The following procedure illustrates the recursive method for filling a 4-connected region with color specified in parameter fill color (f-color) up to a boundary color specified with parameter boundary color (b-color).
- Procedure :

```
boundary-fill4(x, y, f-color, b-color)
{
    if(getpixel (x,y) ! = b-color&&gepixel (x, y) ! = f-color)
    {
            putpixel (x, y, f-color)
            boundary-fill4(x + 1, y, f-color, b-color);
```

```
            boundary-fill4(x, y + 1, f-color, b-color);
            boundary-fill4(x - 1, y, f-color, b-color);
            boundary-fill4(x, y - l, f-color, b-color);
        }
    }
```

- Note: 'getpixel' function gives the color of .specified pixel and 'putpixel' function draws the pixel with specified color.
- Same procedure can be modified according to 8 connected region algorithm by including four additional statements to test diagonal positions, such as $(x + 1, y + 1)$.
- This procedure requires considerable stacking of neighbouring points more, efficient methods are generally employed.
- This method fill horizontal pixel spans across scan lines, instead of proceeding to 4 connected or 8 connected neighbouring points.
- Then we need only stack a beginning position for each horizontal pixel span, instead of stacking all unprocessed neighbouring positions around the current position.
- Starting from the initial interior point with this method, we first fill in the contiguous span of pixels on this starting scan line.
- Then we locate and stack starting positions for spans on the adjacent scan lines, where spans are defined as the contiguous horizontal string of positions bounded by pixels displayed in the area border color.
- At each subsequent step, we unstack the next start position and repeat the process.
- An example of how pixel spans could be filled using this approach is illustrated for the 4-connected fill region in Figure below.



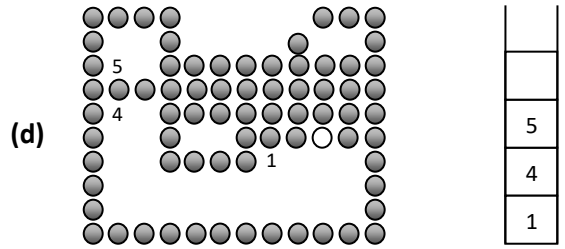


Fig. 2.24: - Boundary fill across pixel spans for a 4-connected area.

Flood-Fill Algorithm

- Sometimes it is required to fill in an area that is not defined within a single color boundary.
- In such cases we can fill areas by replacing a specified interior color instead of searching for a boundary color.
- This approach is called a flood-fill algorithm. Like boundary fill algorithm, here we start with some seed and examine the neighbouring pixels.
- However, here pixels are checked for a specified interior color instead of boundary color and they are replaced by new color.
- Using either a 4-connected or 8-connected approach, we can step through pixel positions until all interior point have been filled.
- The following procedure illustrates the recursive method for filling 4-connected region using flood-fill algorithm.
- Procedure :

```
flood-fill4(x, y, new-color, old-color)
{
    if(getpixel (x,y) = = old-color)
    {
            putpixel (x, y, new-color)
            flood-fill4 (x + 1, y, new-color, old -color);
            flood-fill4 (x, y + 1, new -color, old -color);
            flood-fill4 (x - 1, y, new -color, old -color);
            flood-fill4 (x, y - l, new -color, old-color);
    }
}
```

- Note: 'getpixel' function gives the color of .specified pixel and 'putpixel' function draws the pixel with specified color.

Character Generation

- We can display letters and numbers in variety of size and style.
- The overall design style for the set of character is called typeface.
- Today large numbers of typefaces are available for computer application for example Helvetica, New York platino etc.
- Originally, the term font referred to a set of cast metal character forms in a particular size and format, such as 10-point Courier Italic or 12- point Palatino Bold. Now, the terms font and typeface are often used interchangeably, since printing is no longer done with cast metal forms.
- Two different representations are used for storing computer fonts.

Bitmap Font/ Bitmapped Font

- A simple method for representing the character shapes in a particular typeface is to use rectangular grid patterns.
- Figure below shows pattern for particular letter.

1	1	1	1	1	1	0
0	1	1	0	0	1	1
0	1	1	0	0	1	1
0	1	1	1	1	1	0
0	1	1	0	0	1	1
0	1	1	0	0	1	1
1	1	1	1	1	1	0

Fig. 2.25: - Grid pattern for letter B.

- When the pattern in figure 2.25 is copied to the area of frame buffer, the 1 bits designate which pixel positions are to be displayed on the monitor.
- Bitmap fonts are the simplest to define and display as character grid only need to be mapped to a frame-buffer position.
- Bitmap fonts require more space because each variation (size and format) must be stored in a font cache.
- It is possible to generate different size and other variation from one set but this usually does not produce good result.

Outline Font

- In this method character is generated using curve section and straight line as combine assembly.
- Figure below shows how it is generated.

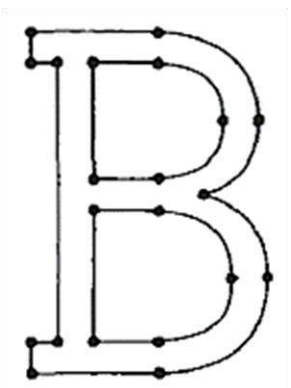


Fig. 2.26: - outline for letter B.

- To display the character shown in figure 2.26 we need to fill interior region of the character.
- This method requires less storage since each variation does not required a distinct font cache.
- We can produce boldface, italic, or different sizes by manipulating the curve definitions for the character outlines.
- But this will take more time to process the outline fonts, because they must be scan converted into the frame buffer.

Stroke Method

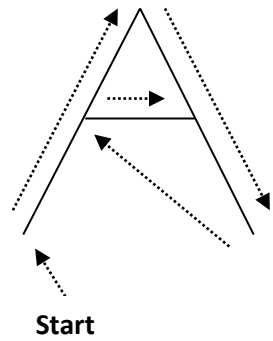


Fig. 2.27: - Stroke Method for Letter A.

- It uses small line segments to generate a character.
- The small series of line segments are drawn like a stroke of a pen to form a character as shown in figure.
- We can generate our own stroke method by calling line drawing algorithm.
- Here it is necessary to decide which line segments are needs for each character and then draw that line to display character.
- It support scaling by changing length of line segment.

Starbust Method

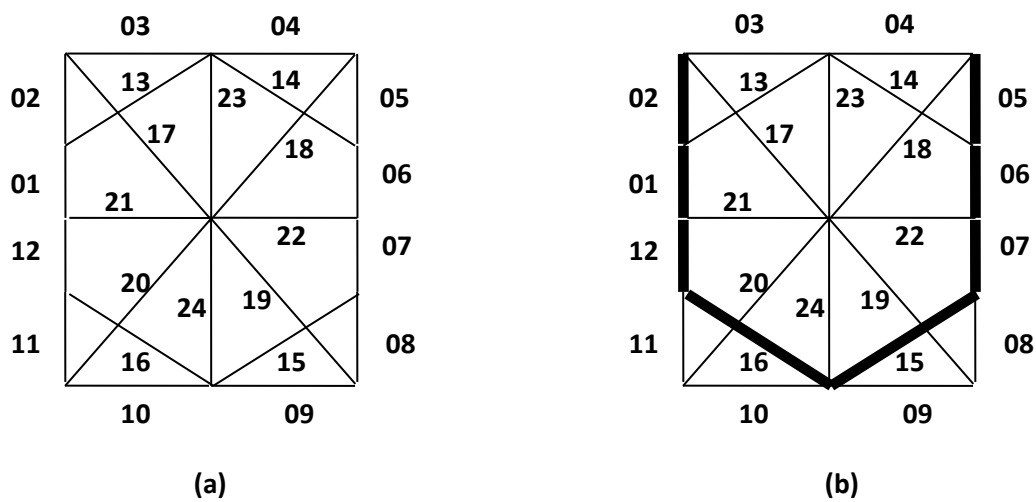


Fig. 2.28: - (a) Starbust Method. (b) Letter V using starbust method

- In this method a fix pattern of line segments are used to generate characters.
- As shown in figure 2.28 there are 24 line segments.
- We highlight those lines which are necessary to draw a particular character.
- Pattern for particular character is stored in the form of 24 bit code. In which each bit represents corresponding line having that number.
- That code contains 0 or 1 based on line segment need to highlight. We put bit value 1 for highlighted line and 0 for other line.
- Code for letter V is
  1 1 0 0 1 1 1 0 0 0 0 1 0 0 1 1 0 0 0 0 0 0 0 0
- This technique is not used now a days because:
  1. It requires more memory to store 24 bit code for single character.

2. It requires conversion from code to character.
3. It doesn't provide curve shapes.

Line Attributes

- Basic attributes of a straight line segment are its type, its dimension, and its color. In some graphics packages, lines can also be displayed using selected pen or brush option.

Line Type

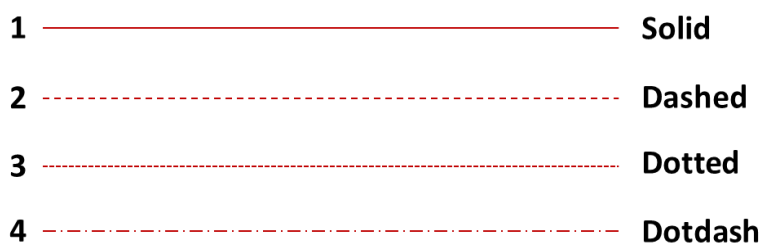


Fig. 2.29: - Line type

- Possible selection for the line-type attribute includes solid lines, dashed lines, and dotted lines etc.
- We modify a line –drawing algorithm to generate such lines by setting the length and spacing of displayed solid sections along the line path.
- A dashed line could be displayed by generating an inter dash spacing that is equal to the length of the solid sections. Both the length of the dashes and the inter dash spacing are often specified as user options.
- To set line type attributes in a PHIGS application program, a user invokes the function:
  setLinetype(lt)
- Where parameter lt is assigned a positive integer value of 1, 2, 3, 4… etc. to generate lines that are, respectively solid, dashed, dotted, or dash-dotted etc.
- Other values for the line-type parameter lt could be used to display variations in the dot-dash patterns. Once the line-type parameter has been set in a PHIGS application program, all subsequent line-drawing commands produce lines with this Line type.
- Raster graphics generates these types by plotting some pixel and some pixel is off along the line path. We can generate different patterns by specifying 1 for on pixel and 0 for off pixel then we can generate 1010101 patter as a dotted line.
- It is used in many application for example comparing data in graphical form.

Line Width

- Implementation of line-width options depends on the capabilities of the output device.
- A heavy line on a video monitor could be displayed as adjacent parallel lines, while a pen plotter might require pen changes.
- To set line width attributes in a PHIGS application program, a user invokes the function:
  setLinewidthScalFactor (lw)
- Line-width parameter lw is assigned a positive number to indicate the relative width of the line to be displayed.
- Values greater than 1 produce lines thicker than the standard line width and values less than the 1 produce line thinner than the standard line width.

- In raster graphics we generate thick line by plotting above and below pixel of line path when slope $|m| < 1$ and by plotting left and right pixel of line path when slope $|m| > 1$ which is illustrated in below figure.

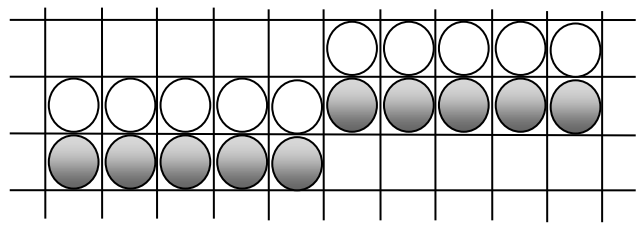


Fig. 2.30: - Double-wide raster line with slope $|m| < 1$ generated with vertical pixel spans.

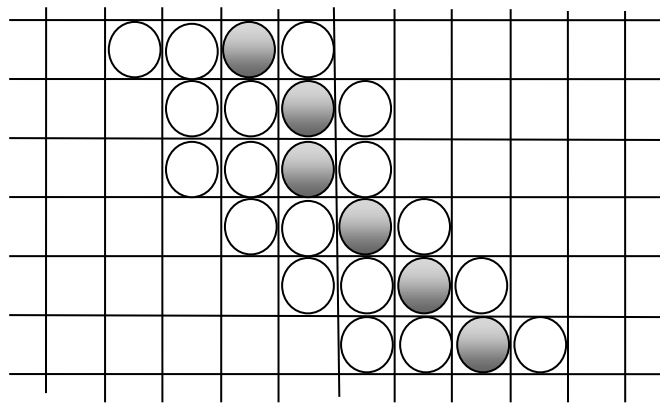


Fig. 2.31: - Raster line with slope $|m| > 1$ and line-width parameter $lw = 4$ plotted with horizontal pixel spans.

- As we change width of the line we can also change line end which are shown below which illustrate all three types of ends.

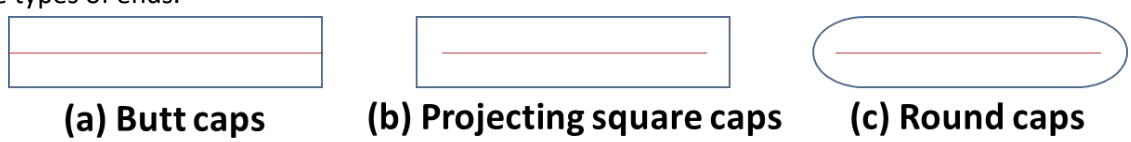


Fig. 2.32: - Thick lines drew with (a) butt caps, (b) projecting square caps, and (c) round caps.

- Similarly we generates join of two lines of modified width are shown in figure below which illustrate all three type of joins.

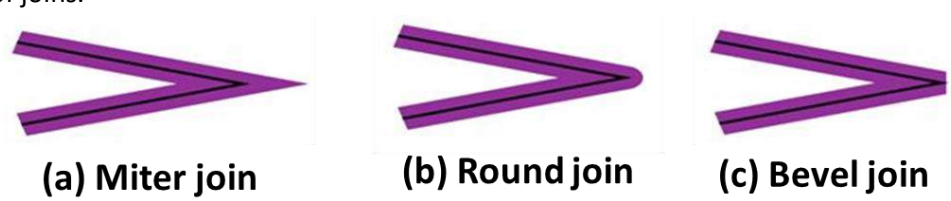


Fig. 2.33: - Thick lines segments connected with (a) miter join, (b) round join, and (c) bevel join.

Line Color

- The name itself suggests that it is defining color of line displayed on the screen.
- By default system produce line with current color but we can change this color by following function in PHIGS package as follows:
  setPolylineColorIndex (lc)
- In this lc is constant specifying particular color to be set.

Pen and Brush Options

- In some graphics packages line is displayed with pen and brush selections.
- Options in this category include shape, size, and pattern.
- Some possible pen or brush are shown in below figure.

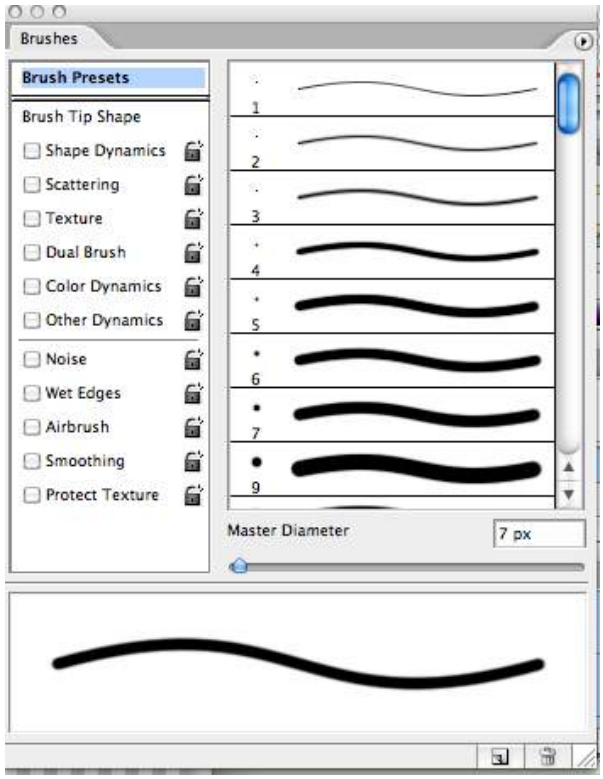


Fig. 2.34: - Pen and brush shapes for line display.

- These shapes can be stored in a pixel mask that identifies the array of pixel positions that are to be set along the line path.
- Lines generated with pen (or brush) shapes can be displayed in various widths by changing the size of the mask.
- Also, lines can be displayed with selected patterns by superimposing the pattern values onto the pen or brush mask.

Color and Greyscale Levels

- Various color and intensity-level options can be made available to a user, depending on the capabilities and design objectives of a particular system.
- General purpose raster-scan systems, for example, usually provide a wide range of colors, while random-scan monitors typically offer only a few color choices, if any.
- In a color raster system, the number of color choices available depends on the amount of storage provided per pixel in the frame buffer
- Also, color-information can be stored in the frame buffer in two ways: We can store color codes directly in the frame buffer, or we can put the color codes in a separate table and use pixel values as an index into this table
- With direct storage scheme we required large memory for frame buffer when we display more color.
- While in case of table it is reduced and we call it color table or color lookup table.

Color Lookup Table

- Color values of 24 bit is stored in lookup table and in frame buffer we store only 8 bit index which gives index of required color stored into lookup table. So that size of frame buffer is reduced and we can display more color.
- When we display picture using this technique on output screen we look into frame buffer where index of the color is stored and take 24 bit color value from look up table corresponding to frame buffer index value and display that color on particular pixel.

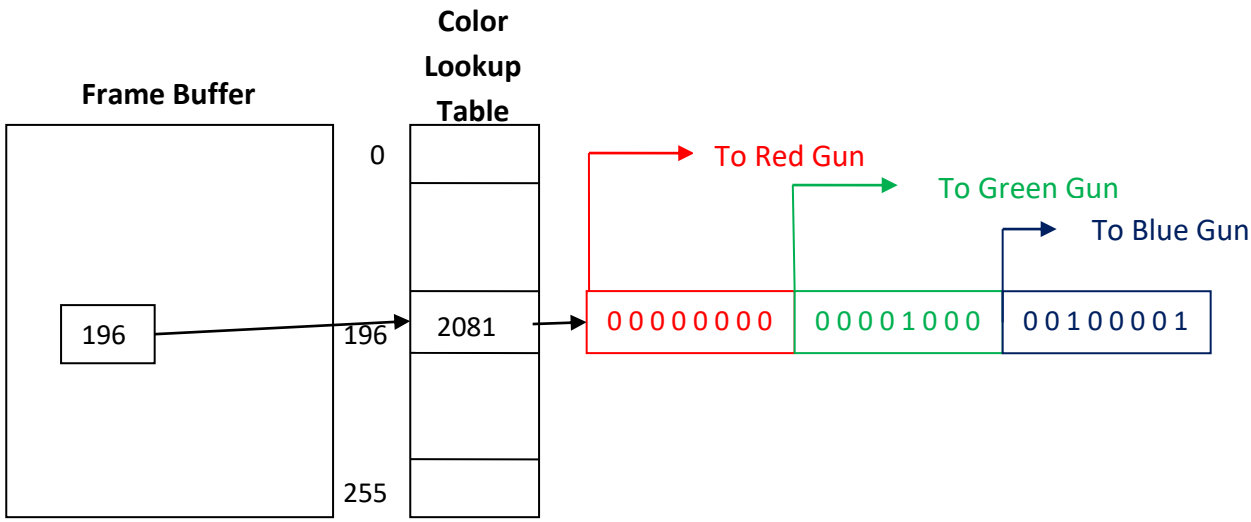


Fig. 2.35: - color lookup table.

Greyscale

- With monitors that have no color capability, color function can be used in an application program to set the shades of grey, or greyscale for display primitives.
- Numeric values between 0-to-1 can be used to specify greyscale levels.
- This numeric values is converted in binary code for store in raster system.
- Table below shows specification for intensity codes for a four level greyscale system.

Intensity Code	Stored Intensity Values In The		Displayed Greyscale
	Frame Buffer	Binary Code	
0.0	0	00	Black
0.33	1	01	Dark grey
0.67	2	10	Light grey
1.0	3	11	White

Table 2.1: - Intensity codes for a four level greyscale system.

- In this example, any intensity input value near 0.33 would be stored as the binary value 01 in the frame buffer, and pixels with this value would be displayed as dark grey.
- If more bits are available per pixel we can obtain more levels of grey scale for example with 3 bit per pixel we can achieve 8 levels of greyscale.

Area-Fill Attributes

- For filling any area we have choice between solid colors or pattern to fill all these are include in area fill attributes.
- Area can be painted by various burses and style.

Fill Styles

- Area are generally displayed with three basic style hollow with color border, filled with solid color, or filled with some design.
- In PHIGS package fill style is selected by following function.
  setInteriorStyle (fs)
- Value of fs include hollow, solid, pattern etc.
- Another values for fill style is hatch, which is patterns of line like parallel line or crossed line.
- Figure bellow shows different style of filling area.

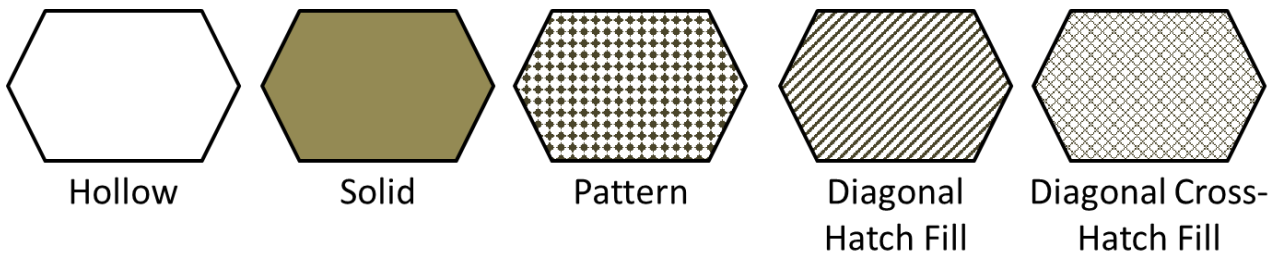


Fig. 2.36: - Different style of area filling.

- For setting interior color in PHIGS package we use:
  setInteriorColorIndex (fc)
- Where fc specify the fill color.

Pattern Fill

- We select the pattern with
  setInteriorStyleIndex (pi)
- Where pattern index parameter pi specifies position in pattern table entry.
- Figure below shows pattern table.

Index(pi)	Pattern(cp)
1	$\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$
2	$\begin{bmatrix} 2 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix}$

Table 2.2: - Pattern table.

- For example, the following set of statements would fill the area defined in the fillArea command with the second pattern type stored in the pattern table:
  SetInteriorStyle(pattern) ;
  setInteriorStyleIndex (2) ;
  fillArea (n, points);
- Separate table can be maintain for hatch pattern and we can generate our own table with required pattern.
- Other function used for setting other style as follows
  setpatternsize (dx, dy)
  setPaternReferencePoint (positicn)
- We can create our own pattern by setting and resetting group of pixel and then map it into the color matrix.

Soft Fill

- Soft fill is modified boundary fill and flood fill algorithm in which we are fill layer of color on back ground color so that we can obtain the combination of both color.
- It is used to recolor or repaint so that we can obtain layer of multiple color and get new color combination.
- One use of this algorithm is soften the fill at boundary so that blurred effect will reduce the aliasing effect.
- For example if we fill t amount of foreground color then pixel color is obtain as:
  $p = tF + (1 - t)B$
- Where F is foreground color and B is background color
- If we see this color in RGB component then:
  $p = (p_r, p_g, p_b)$ $f = (f_r, f_g, f_b)$ $b = (b_r, b_g, b_b)$
- Then we can calculate t as follows:
  $t = \frac{P_k - B_k}{F_k - B_k}$
- If we use more than two color say three at that time equation becomes as follow:
  $p = t_0 F + t_1 B_1 + (1 - t_0 - t_1) B_2$
- Where the sum of coefficient t_0, t_1, and $(1 - t_0 - t_1)$ is 1.

Character Attributes

- The appearance of displayed characters is controlled by attributes such as font, size, color, and orientation.
- Attributes can be set for entire string or may be individually.

Text Attributes

- In text we are having so many style and design like italic fonts, bold fonts etc.
- For setting the font style in PHIGS package we have one function which is:
  setTextFont (tf)
- Where tf is used to specify text font
- It will set specified font as a current character.
- For setting color of character in PHIGS we have function:
  setTextColorIndex (tc)
- Where text color parameter tc specifies an allowable color code.
- For setting the size of the text we use function.
  setCharacterheight (ch)
- For scaling the character we use function.
  setCharacterExpansionFacter (cw)
- Where character width parameter cw is set to a positive real number that scale the character body width.
- Spacing between character is controlled by function
  setCharacterSpacing (cs)
- Where character spacing parameter cs can be assigned any real value.
- The orientation for a displayed character string is set according to the direction of the character up vector:
  setCharacterUpVector (upvect)
- Parameter upvect in this function is assigned two values that specify the x and y vector components.

- Text is then displayed so that the orientation of characters from baseline to cap line is in the direction of the up vector.
- For setting the path of the character we use function:
  setTextPath (tp)
- Where the text path parameter tp can be assigned the value: right, left, up, or down.
- It will set the direction in which we are writing.
- For setting the alignment of the text we use function.
  setTextAlignment (h, v)
- Where parameter h and v control horizontal and vertical alignment respectively.
- For specifying precision for text display is given with function.
  setTextPrecision (tpr)
- Where text precision parameter tpr is assigned one of the values: string, char, or stroke.
- The highest-quality text is produced when the parameter is set to the value stroke.

Marker Attributes

- A marker symbol display single character in different color and in different sizes.
- For marker attributes implementation by procedure that load the chosen character into the raster at defined position with the specified color and size.
- We select marker type using function.
  setMarkerType (mt)
- Where marker type parameter mt is set to an integer code.
- Typical codes for marker type are the integers 1 through 5, specifying, respectively, a dot (.), a vertical cross (+), an asterisk (*), a circle (o), and a diagonal cross (x). Displayed marker types are centred on the marker coordinates.
- We set the marker size with function.
  SetMarkerSizeScaleFactor (ms)
- Where parameter marker size ms assigned a positive number according to need for scaling.
- For setting marker color we use function.
  setPolymarkerColorIndex (mc)
- Where parameter mc specify the color of the marker symbol.

Antialiasing

- Primitives generated in raster graphics by various algorithms discussed in unit 2 have stair step shape or jagged appearance because in sampling process we calculate only integer calculation by rounding actual values.
- This distortion of actual information due to low frequency sampling is called **aliasing**.
- We can compensate this aliasing effect by applying antialiasing methods.

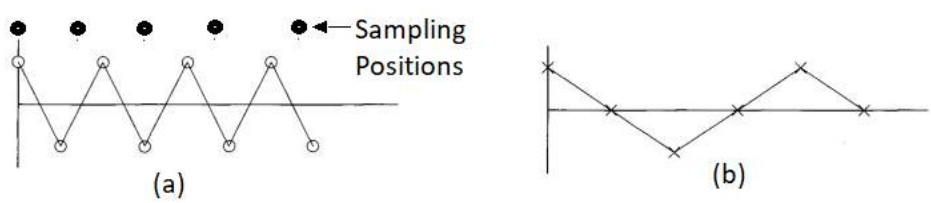


Fig. 2.34: - Sampling the periodic shape in (a) at the marked positions produces the aliased lower-frequency representation in (b).

- An example of the effects of undersampling is shown in Figure above To avoid losing information from such periodic objects, we need to set the sampling frequency to at least twice that of the highest

frequency occurring in the object, referred to as the Nyquist sampling frequency (or Nyquist sampling rate) f_s:

$$f_s = 2f_{max}$$

- In other way we can say this is by sampling interval should be no larger than one-half the cycle. Which is called nyquist sampling interval.

$$\Delta x_s = \frac{\Delta x_{cycle}}{2}$$

- Where $\Delta x_{cycle} = 1/f_{max}$.
- One way to solve this problem is to display image on higher resolution screen but it has also limit that how much large frame buffer you can maintain along with maintaining refresh rate 30 to 60 frame per second.
- And also on higher resolution jaggies will remains up to some extents.
- With raster systems that are capable of displaying more than two intensity levels (color or gray sc,lle), we can apply antialiasing methods to modify pixel intensities.
- By appropriately varying the intensities of pixels along the boundaries of primitives, we can smooth the edges to lessen the jagged appearance.

Antialiasing Methods

Supersampling Straight Line Segments

- Supersampling straight line can be performed in several ways.
- For the greyscale display of a straight-line segment, we can divide each pixel into a number of sub pixels and determine the number of sub pixel along the line path.
- Then we set intensity level of each pixel proportional to number of sub pixel along the line path.
- For example in figure 2.35 (A) area of each pixel is divided into nine sub pixel and then we determine how many number of sub pixel are along the line (it can be 3 or 2 or 1 as we divide into 9 sub pixel).
- Based on number 3 or 2 or 1 we assign intensity value to that pixel.
- We can achieve four intensity levels by dividing pixel into 16 sub pixels and five intensity levels by dividing into 25 sub pixels etc.
- Lower intensity gives blurred effect and hence performs antialiasing.

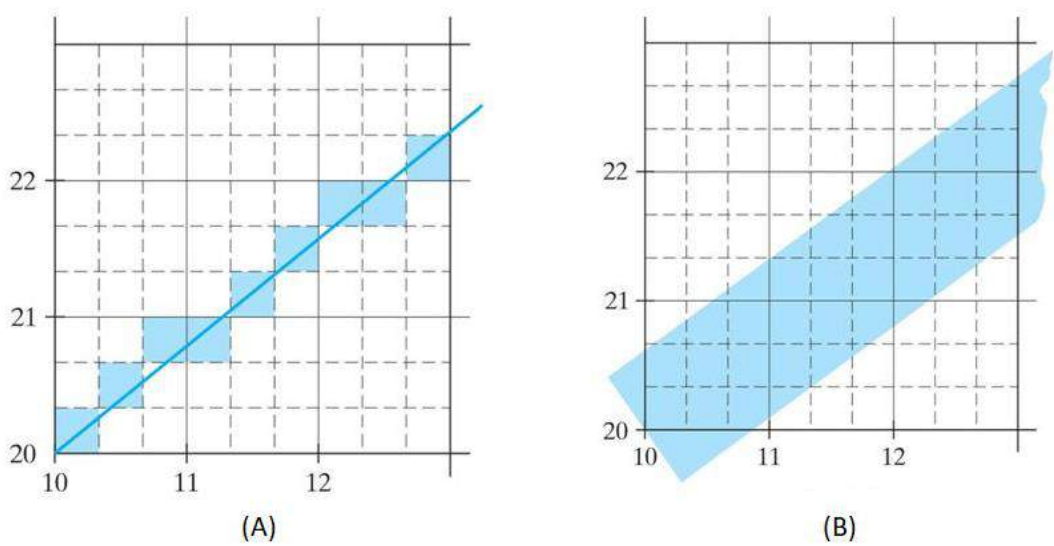


Fig. 2.35: - (A) Super sampling sub pixel positions along a straight line segment. (B) Super sampling sub pixel positions in relation to the interior of a line of finite width.

- In the supersampling example of Fig. 2.35 (A), we considered pixel areas of finite size, but we treated the line as a mathematical entity with zero width. Actually, displayed lines have a width approximately equal to that of a pixel.
- If we take the finite width of the line into account (figure 2.35 (B)), we can perform super sampling by setting each pixel intensity proportional to the number of sub pixels inside the polygon representing the line area.
- A sub pixel can be considered to be inside the line if its lower left corner is inside the polygon boundaries.
- Advantage of this is that it having number of intensity equals to number of sub pixel.
- Another advantage of this is that it will distribute total intensity over more pixels. For example in figure 2.35 pixel below and left to (10, 20) is also assigned some intensity levels so that aliasing will reduce.
- If we have color display then we modified to take background color into account and we develop levels of color by mixing background color and line color.

Pixel-Weighting Masks

1	2	1
2	4	2
1	2	1

Fig. 2.36: - Relative weights for a grid of 3 by 3 sub pixels.

- Supersampling method are often implemented by giving more weight to sub pixel near the center of pixel area, since we expect these sub pixel to be more important in determining the overall intensity of a pixel.
- For the 3 by 3 pixel subdivisions we have considered so far, a weighting scheme as in Fig. 2-36 could be used.
- The center sub pixel here is weighted four times that of the corner sub pixels and twice that of the remaining sub pixels.
- By averaging the weight of sub pixel which are along the line and assign intensity proportional to average weight will antialiasing stair step effect.

Area Sampling Straight Line Segments

- In this scheme we treat line as finite width rectangle, and the section of the line area between two adjacent vertical or two adjacent horizontal screen grid lines is then a trapezoids.
- Overlap areas for pixels are calculated by determining how much of the trapezoid overlaps each pixel in that vertical column (or horizontal row).
- In Fig. 2.35 (B), the pixel with screen grid coordinates (10, 20) is about 90 percent covered by the line area, so its intensity would be set to 90 percent of the maximum intensity. Similarly, the pixel at (10 21) would be set to an intensity of about 15-percent of maximum.
- With color displays, the areas of pixel overlap with different color regions is calculated and the final pixel color is taken as the average color of the various overlap areas.

Filtering Techniques

- A more accurate method for antialiasing lines is to use filtering techniques.
- Common example of filter is rectangular, conical and Gaussian filters.

- Methods for applying the filter function are similar to applying a weighting mask, but now we integrate over the pixel surface to obtain the weighted average intensity.
- For reduce computation we often use table look up.

Pixel Phasing

- On raster system if we pass the electron beam from the closer sub pixel so that overall pixel is shifted by factor ¼, ½, or ¾ to pixel diameter.
- Where beam strike at that part of pixel get more intensity then other parts of the pixel and gives antialiasing effect.
- Some systems also allow the size of individual pixels to be adjusted as an additional means for distributing intensities.

Compensating For Line Intensity Differences

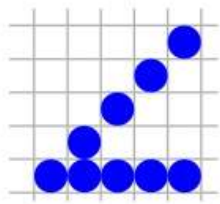


Fig. 2.38: - Unequal-length lines displayed with the same number of pixels in each line.

- Antialiasing a line to soften the stairstep effect also compensates for another raster effect, illustrated in above figure.
- As both lines are display with same number of pixel and then also length of diagonal line is greater than horizontal line by factor$\sqrt{2}$. So that diagonal line is display with less intensity then horizontal line.
- For compensating this we display diagonal line with high intensity and horizontal line with low intensity so that this effect is minimise.
- In general we set intensity according to slope of the line.

Antialiasing Area Boundaries

- Methods we discuss for antialiasing line can be applied for area boundary so that boundary areas can remove their jagged appearance.
- If system capabilities permit the repositioning of pixels, area boundaries can be smoothed by adjusting boundary pixel positions so that they are along the line defining an area boundary.
- Other method is to adjust each pixel intensity at boundary position according to percent of area inside the boundary.

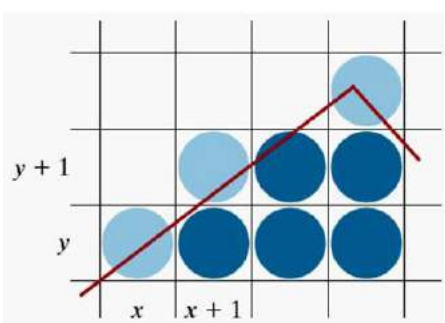


Fig. 2.39: - Adjusting pixel intensities along an area boundary.

- In above figure pixel at (x, y) is assigned half the intensity as its ½ area is inside the area boundary.
- Similar adjustments based on percent of area of pixel inside are applied to other pixel.

Transformation

Changing Position, shape, size, or orientation of an object on display is known as transformation.

Basic Transformation

- Basic transformation includes three transformations **Translation**, **Rotation**, and **Scaling**.
- These three transformations are known as basic transformation because with combination of these three transformations we can obtain any transformation.

Translation

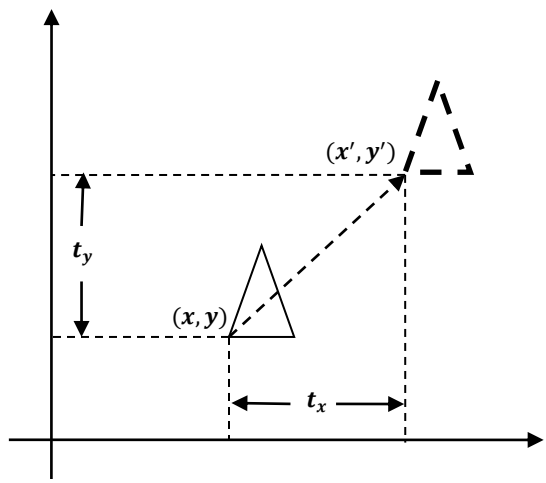


Fig. 3.1: - Translation.

- It is a transformation that used to reposition the object along the straight line path from one coordinate location to another.
- It is rigid body transformation so we need to translate whole object.
- We translate two dimensional point by adding translation distance $\boldsymbol{t_x}$ and $\boldsymbol{t_y}$ to the original coordinate position $(\boldsymbol{x}, \boldsymbol{y})$ to move at new position $(\boldsymbol{x'}, \boldsymbol{y'})$ as:
  $\boldsymbol{x' = x + t_x}$ **&** $\boldsymbol{y' = y + t_y}$
- Translation distance pair $(\boldsymbol{t_x}, \boldsymbol{t_y})$ is called a **Translation Vector** or **Shift Vector**.
- We can represent it into single matrix equation in column vector as;
  $\boldsymbol{P' = P + T}$
  $\begin{bmatrix} \boldsymbol{x'} \\ \boldsymbol{y'} \end{bmatrix} = \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} + \begin{bmatrix} \boldsymbol{t_x} \\ \boldsymbol{t_y} \end{bmatrix}$
- We can also represent it in row vector form as:
  $\boldsymbol{P' = P + T}$
  $[\boldsymbol{x'} \quad \boldsymbol{y'}] = [\boldsymbol{x} \quad \boldsymbol{y}] + [\boldsymbol{t_x} \quad \boldsymbol{t_y}]$
- Since column vector representation is standard mathematical notation and since many graphics package like **GKS** and **PHIGS** uses column vector we will also follow column vector representation.
- **Example**: - Translate the triangle [A (10, 10), B (15, 15), C (20, 10)] 2 unit in x direction and 1 unit in y direction.
  We know that
  $P' = P + T$
  $P' = [P] + \begin{bmatrix} t_x \\ t_y \end{bmatrix}$

For point (10, 10)

$A' = \begin{bmatrix}10\\10\end{bmatrix} + \begin{bmatrix}2\\1\end{bmatrix}$

$A' = \begin{bmatrix}12\\11\end{bmatrix}$

For point (15, 15)

$B' = \begin{bmatrix}15\\15\end{bmatrix} + \begin{bmatrix}2\\1\end{bmatrix}$

$B' = \begin{bmatrix}17\\16\end{bmatrix}$

For point (10, 10)

$C' = \begin{bmatrix}20\\10\end{bmatrix} + \begin{bmatrix}2\\1\end{bmatrix}$

$C' = \begin{bmatrix}22\\11\end{bmatrix}$

- Final coordinates after translation are [A′ (12, 11), B′ (17, 16), C′ (22, 11)].

Rotation

- It is a transformation that used to reposition the object along the circular path in the XY - plane.
- To generate a rotation we specify a rotation angle $\boldsymbol{\theta}$ and the position of the **Rotation Point** (**Pivot Point**) $(\boldsymbol{x_r, y_r})$ about which the object is to be rotated.
- Positive value of rotation angle defines counter clockwise rotation and negative value of rotation angle defines clockwise rotation.
- We first find the equation of rotation when pivot point is at coordinate origin$(\mathbf{0}, \mathbf{0})$.

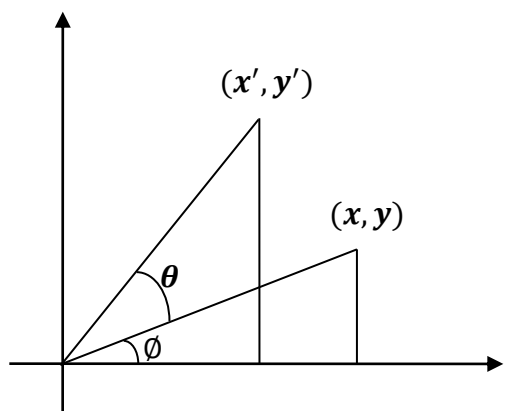


Fig. 3.2: - Rotation.

- From figure we can write.

  $\boldsymbol{x = r\cos\emptyset}$

  $\boldsymbol{y = r\sin\emptyset}$

  and

  $\boldsymbol{x' = r\cos(\theta + \emptyset) = r\cos\emptyset\cos\theta - r\sin\emptyset\sin\theta}$

  $\boldsymbol{y' = r\sin(\emptyset + \theta) = r\cos\emptyset\sin\theta + r\sin\emptyset\cos\theta}$
- Now replace $\boldsymbol{r\cos\emptyset}$ with $\boldsymbol{x}$ and $\boldsymbol{r\sin\emptyset}$ with $\boldsymbol{y}$ in above equation.

  $\boldsymbol{x' = x\cos\theta - y\sin\theta}$

  $\boldsymbol{y' = x\sin\theta + y\cos\theta}$
- We can write it in the form of column vector matrix equation as;

  $\boldsymbol{P' = R \cdot P}$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}$$

- Rotation about arbitrary point is illustrated in below figure.

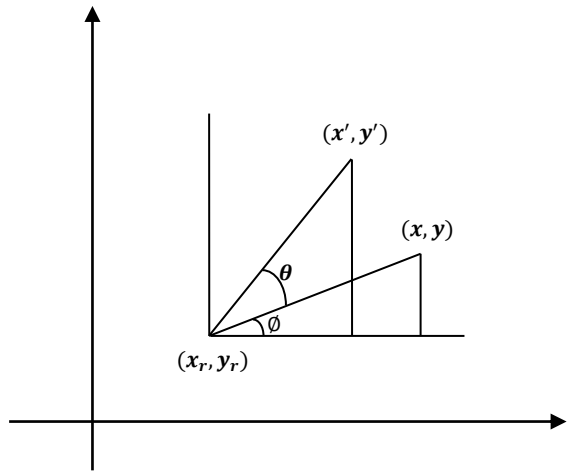


Fig. 3.3: - Rotation about pivot point.

- Transformation equation for rotation of a point about pivot point (x_r, y_r) is:
  $x' = x_r + (x - x_r)\cos\theta - (y - y_r)\sin\theta$
  $y' = y_r + (x - x_r)\sin\theta + (y - y_r)\cos\theta$
- These equations are differing from rotation about origin and its matrix representation is also different.
- Its matrix equation can be obtained by simple method that we will discuss later in this chapter.
- Rotation is also rigid body transformation so we need to rotate each point of object.
- **Example**: - Locate the new position of the triangle [A (5, 4), B (8, 3), C (8, 8)] after its rotation by 90° clockwise about the origin.
  As rotation is clockwise we will take $\theta = -90^{\circ}$.
  $P' = R \cdot P$
  $$P' = \begin{bmatrix} \cos(-90) & -\sin(-90) \\ \sin(-90) & \cos(-90) \end{bmatrix} \begin{bmatrix} 5 & 8 & 8 \\ 4 & 3 & 8 \end{bmatrix}$$
  $$P' = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 5 & 8 & 8 \\ 4 & 3 & 8 \end{bmatrix}$$
  $$P' = \begin{bmatrix} 4 & 3 & 8 \\ -5 & -8 & -8 \end{bmatrix}$$
- Final coordinates after rotation are [A′ (4, -5), B′ (3, -8), C′ (8, -8)].

Scaling

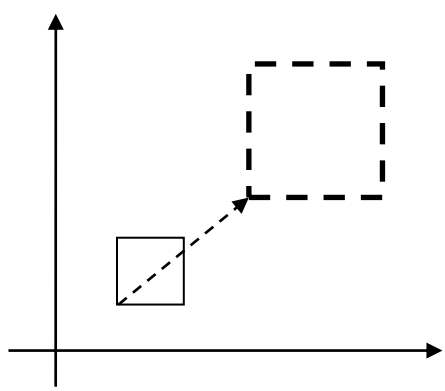


Fig. 3.4: - Scaling.

- It is a transformation that used to alter the size of an object.
- This operation is carried out by multiplying coordinate value $(\boldsymbol{x}, \boldsymbol{y})$ with scaling factor $(\boldsymbol{s_x}, \boldsymbol{s_y})$ respectively.
- So equation for scaling is given by:
  $\boldsymbol{x' = x \cdot s_x}$
  $\boldsymbol{y' = y \cdot s_y}$
- These equation can be represented in column vector matrix equation as:
  $\boldsymbol{P' = S \cdot P}$
  $\begin{bmatrix} \boldsymbol{x'} \\ \boldsymbol{y'} \end{bmatrix} = \begin{bmatrix} \boldsymbol{s_x} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{s_y} \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix}$
- Any positive value can be assigned to$(\boldsymbol{s_x}, \boldsymbol{s_y})$**.**
- Values less than 1 reduce the size while values greater than 1 enlarge the size of object, and object remains unchanged when values of both factor is 1.
- Same values of $\boldsymbol{s_x}$ and $\boldsymbol{s_y}$ will produce **Uniform Scaling**. And different values of $\boldsymbol{s_x}$ and $\boldsymbol{s_y}$ will produce **Differential Scaling**.
- Objects transformed with above equation are both scale and repositioned.
- Scaling factor with value less than 1 will move object closer to origin, while scaling factor with value greater than 1 will move object away from origin.
- We can control the position of object after scaling by keeping one position fixed called **Fix point** $(\boldsymbol{x_f}, \boldsymbol{y_f})$ that point will remain unchanged after the scaling transformation.

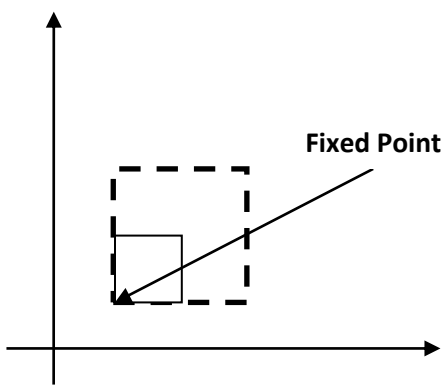


Fig. 3.5: - Fixed point scaling.

- Equation for scaling with fixed point position as $(\boldsymbol{x_f}, \boldsymbol{y_f})$ is:
  $\boldsymbol{x' = x_f + (x - x_f)s_x}$ $\qquad$ $\boldsymbol{y' = y_f + (y - y_f)s_y}$
  $\boldsymbol{x' = x_f + xs_x - x_f s_x}$ $\qquad$ $\boldsymbol{y' = y_f + ys_y - y_f s_y}$
  $\boldsymbol{x' = xs_x + x_f(1 - s_x)}$ $\qquad$ $\boldsymbol{y' = ys_y + y_f(1 - s_y)}$
- Matrix equation for the same will discuss in later section.
- Polygons are scaled by applying scaling at coordinates and redrawing while other body like circle and ellipse will scale using its defining parameters. For example ellipse will scale using its semi major axis, semi minor axis and center point scaling and redrawing at that position.
- **Example**: - Consider square with left-bottom corner at (2, 2) and right-top corner at (6, 6) apply the transformation which makes its size half.
  As we want size half so value of scale factor are $s_x = 0.5, s_y = 0.5$ and Coordinates of square are [A (2, 2), B (6, 2), C (6, 6), D (2, 6)].
  $P' = S \cdot P$

$$P' = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \begin{bmatrix} 2 & 6 & 6 & 2 \\ 2 & 2 & 6 & 6 \end{bmatrix}$$

$$P' = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix} \begin{bmatrix} 2 & 6 & 6 & 2 \\ 2 & 2 & 6 & 6 \end{bmatrix}$$

$$P' = \begin{bmatrix} 1 & 3 & 3 & 1 \\ 1 & 1 & 3 & 3 \end{bmatrix}$$

- Final coordinate after scaling are [A′ (1, 1), B′ (3, 1), C′ (3, 3), D′ (1, 3)].

Matrix Representation and homogeneous coordinates

- Many graphics application involves sequence of geometric transformations.
- For example in design and picture construction application we perform Translation, Rotation, and scaling to fit the picture components into their proper positions.
- For efficient processing we will reformulate transformation sequences.
- We have matrix representation of basic transformation and we can express it in the general matrix form as:

  $\boldsymbol{P' = M_1 \cdot P + M_2}$

  Where $\boldsymbol{P}$ and $\boldsymbol{P'}$ are initial and final point position, $\boldsymbol{M_1}$ contains rotation and scaling terms and $\boldsymbol{M_2}$ contains translation al terms associated with pivot point, fixed point and reposition.
- For efficient utilization we must calculate all sequence of transformation in one step and for that reason we reformulate above equation to eliminate the matrix addition associated with translation terms in matrix $\boldsymbol{M_2}$.
- We can combine that thing by expanding 2X2 matrix representation into 3X3 matrices.
- It will allows us to convert all transformation into matrix multiplication but we need to represent vertex position $(\boldsymbol{x}, \boldsymbol{y})$ with homogeneous coordinate triple $(\boldsymbol{x_h}, \boldsymbol{y_h}, \boldsymbol{h})$ Where $\boldsymbol{x = \frac{x_h}{h}}$, $\boldsymbol{y = \frac{y_h}{h}}$ thus we can also write triple as $(\boldsymbol{h \cdot x}, \boldsymbol{h \cdot y}, \boldsymbol{h})$.
- For two dimensional geometric transformation we can take value of $\boldsymbol{h}$ is any positive number so we can get infinite homogeneous representation for coordinate value $(\boldsymbol{x}, \boldsymbol{y})$.
- But convenient choice is set $\boldsymbol{h = 1}$ as it is multiplicative identity, than $(\boldsymbol{x}, \boldsymbol{y})$ is represented as $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{1})$.
- Expressing coordinates in homogeneous coordinates form allows us to represent all geometric transformation equations as matrix multiplication.
- Let's see each representation with $\boldsymbol{h = 1}$

Translation

$$\boldsymbol{P' = T_{(t_x, t_y)} \cdot P}$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

NOTE: - Inverse of translation matrix is obtain by putting $-\boldsymbol{t_x}$ & $-\boldsymbol{t_y}$ instead of $\boldsymbol{t_x}$ & $\boldsymbol{t_y}$.

Rotation

$$\boldsymbol{P' = R_{(\theta)} \cdot P}$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

NOTE: - Inverse of rotation matrix is obtained by replacing $\boldsymbol{\theta}$ by $-\boldsymbol{\theta}$.

Scaling

$$\boldsymbol{P' = S_{(s_x, s_y)} \cdot P}$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

NOTE: - Inverse of scaling matrix is obtained by replacing s_x & s_y by $\frac{1}{s_x}$ & $\frac{1}{s_y}$ respectively.

Composite Transformation

- We can set up a matrix for any sequence of transformations as a **composite transformation matrix** by calculating the matrix product of individual transformation.
- For column matrix representation of coordinate positions, we form composite transformations by multiplying matrices in order from right to left.

Translations

- Two successive translations are performed as:

$$P' = T(t_{x2}, t_{y2}) \cdot \{T(t_{x1}, t_{y1}) \cdot P\}$$

$$P' = \{T(t_{x2}, t_{y2}) \cdot T(t_{x1}, t_{y1})\} \cdot P$$

$$P' = \begin{bmatrix} 1 & 0 & t_{x2} \\ 0 & 1 & t_{y2} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & t_{x1} \\ 0 & 1 & t_{y1} \\ 0 & 0 & 1 \end{bmatrix} \cdot P$$

$$P' = \begin{bmatrix} 1 & 0 & t_{x1} + t_{x2} \\ 0 & 1 & t_{y1} + t_{y2} \\ 0 & 0 & 1 \end{bmatrix} \cdot P$$

$$P' = T(t_{x1} + t_{x2}, t_{y1} + t_{y2}) \cdot P\}$$

Here P' and P are column vector of final and initial point coordinate respectively.

- This concept can be extended for any number of successive translations.

Example: Obtain the final coordinates after two translations on point $p(2,3)$ with translation vector $(4,\ 3)$ and $(-1,\ 2)$ respectively.

$$P' = T(t_{x1} + t_{x2}, t_{y1} + t_{y2}) \cdot P$$

$$P' = \begin{bmatrix} 1 & 0 & t_{x1} + t_{x2} \\ 0 & 1 & t_{y1} + t_{y2} \\ 0 & 0 & 1 \end{bmatrix} \cdot P = \begin{bmatrix} 1 & 0 & 4 + (-1) \\ 0 & 1 & 3 + 2 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$$

$$P' = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 8 \\ 1 \end{bmatrix}$$

Final Coordinates after translations are $p'(5,\ 8)$.

Rotations

- Two successive Rotations are performed as:

$$P' = R(\theta_2) \cdot \{R(\theta_1) \cdot P\}$$

$$P' = \{R(\theta_2) \cdot R(\theta_1)\} \cdot P$$

$$P' = \begin{bmatrix} \cos\theta_2 & -\sin\theta_2 & 0 \\ \sin\theta_2 & \cos\theta_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 & 0 \\ \sin\theta_1 & \cos\theta_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot P$$

$$P' = \begin{bmatrix} \cos\theta_2\cos\theta_1 - \sin\theta_2\sin\theta_1 & -\sin\theta_1\cos\theta_2 - \sin\theta_2\cos\theta_1 & 0 \\ \sin\theta_1\cos\theta_2 + \sin\theta_2\cos\theta_1 & \cos\theta_2\cos\theta_1 - \sin\theta_2\sin\theta_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot P$$

$$\boldsymbol{P' = \begin{bmatrix} \cos(\theta_1+\theta_2) & -\sin(\theta_1+\theta_2) & 0 \\ \sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2) & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot P}$$

$$\boldsymbol{P' = R(\theta_1+\theta_2) \cdot P}$$

Here $\boldsymbol{P'}$ and $\boldsymbol{P}$ are column vector of final and initial point coordinate respectively.

- This concept can be extended for any number of successive rotations.

Example: Obtain the final coordinates after two rotations on point $p(6,9)$ with rotation angles are 30^o and 60^o respectively.

$$P' = R(\theta_1+\theta_2) \cdot P$$

$$P' = \begin{bmatrix} cos(\theta_1+\theta_2) & -sin(\theta_1+\theta_2) & 0 \\ sin(\theta_1+\theta_2) & cos(\theta_1+\theta_2) & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot P$$

$$P' = \begin{bmatrix} cos(30+60) & -sin(30+60) & 0 \\ sin(30+60) & cos(30+60) & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot P$$

$$P' = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ 9 \\ 1 \end{bmatrix} = \begin{bmatrix} -9 \\ 6 \\ 1 \end{bmatrix}$$

Final Coordinates after rotations are $p'(-9,\ 6)$.

Scaling

- Two successive scaling are performed as:

$$\boldsymbol{P' = S(s_{x2}, s_{y2}) \cdot \{S(s_{x1}, s_{y1}) \cdot P\}}$$

$$\boldsymbol{P' = \{S(s_{x2}, s_{y2}) \cdot S(s_{x1}, s_{y1})\} \cdot P}$$

$$\boldsymbol{P' = \begin{bmatrix} s_{x2} & 0 & 0 \\ 0 & s_{y2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_{x1} & 0 & 0 \\ 0 & s_{y1} & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot P}$$

$$\boldsymbol{P' = \begin{bmatrix} s_{x1} \cdot s_{x2} & 0 & 0 \\ 0 & s_{y1} \cdot s_{y2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot P}$$

$$\boldsymbol{P' = S(s_{x1} \cdot s_{x2}, s_{y1} \cdot s_{y2}) \cdot P}$$

Here $\boldsymbol{P'}$ and $\boldsymbol{P}$ are column vector of final and initial point coordinate respectively.

- This concept can be extended for any number of successive scaling.

Example: Obtain the final coordinates after two scaling on line pq $[p(2,2),\ q(8,\ 8)]$ with scaling factors are $(2,\ 2)$ and $(3,\ 3)$ respectively.

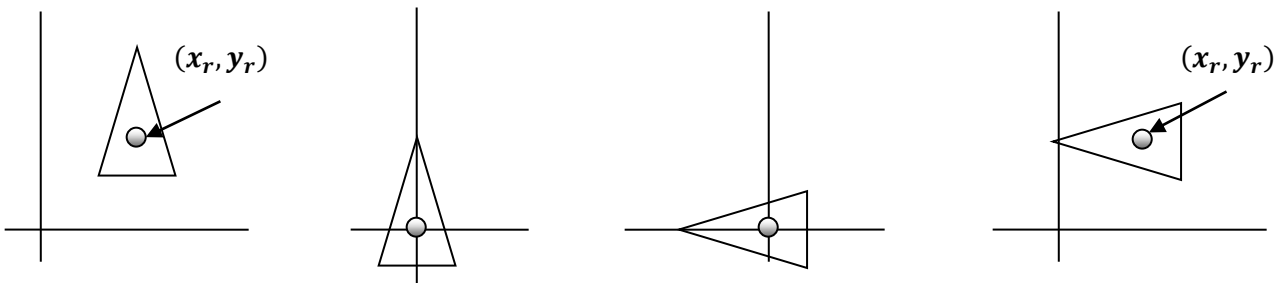
$$P' = S(s_{x1} \cdot s_{x2}, s_{y1} \cdot s_{y2}) \cdot P$$

$$P' = \begin{bmatrix} s_{x1} \cdot s_{x2} & 0 & 0 \\ 0 & s_{y1} \cdot s_{y2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot P = \begin{bmatrix} 2 \cdot 3 & 0 & 0 \\ 0 & 2 \cdot 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot P$$

$$P' = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 8 \\ 2 & 8 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 12 & 48 \\ 12 & 48 \\ 1 & 1 \end{bmatrix}$$

Final Coordinates after rotations are $p'(12,\ 12)$ and $q'(48,\ 48)$.

General Pivot-Point Rotation



(a) Original Position of Object and Pivot Point.

(b) Translation of Object so that Pivot Point (x_r, y_r) is at Origin.

(c) Rotation about Origin.

(d) Translation of Object so that Pivot Point is Return to Position (x_r, y_r) .

Fig. 3.6: - General pivot point rotation.

- For rotating object about arbitrary point called pivot point we need to apply following sequence of transformation.
  1. Translate the object so that the pivot-point coincides with the coordinate origin.
  2. Rotate the object about the coordinate origin with specified angle.
  3. Translate the object so that the pivot-point is returned to its original position (i.e. Inverse of step-1).
- Let's find matrix equation for this

$$\boldsymbol{P' = T(x_r, y_r) \cdot [R(\theta) \cdot \{T(-x_r, -y_r) \cdot P\}]}$$

$$\boldsymbol{P' = \{T(x_r, y_r) \cdot R(\theta) \cdot T(-x_r, -y_r)\} \cdot P}$$

$$\boldsymbol{P' = \begin{bmatrix} 1 & 0 & x_r \\ 0 & 1 & y_r \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -x_r \\ 0 & 1 & -y_r \\ 0 & 0 & 1 \end{bmatrix} \cdot P}$$

$$\boldsymbol{P' = \begin{bmatrix} \cos\theta & -\sin\theta & x_r(1-\cos\theta) + y_r \sin\theta \\ \sin\theta & \cos\theta & y_r(1-\cos\theta) - x_r \sin\theta \\ 0 & 0 & 1 \end{bmatrix} \cdot P}$$

$$\boldsymbol{P' = R(x_r, y_r\theta) \cdot P}$$

Here $\boldsymbol{P'}$ and $\boldsymbol{P}$ are column vector of final and initial point coordinate respectively and (x_r, y_r) are the coordinates of pivot-point.

- **Example**: - Locate the new position of the triangle [A (5, 4), B (8, 3), C (8, 8)] after its rotation by 90° clockwise about the centroid.

  Pivot point is centroid of the triangle so:

$$x_r = \frac{5+8+8}{3} = 7\,, \qquad y_r = \frac{4+3+8}{3} = 5$$

As rotation is clockwise we will take $\theta = -90^{\circ}$.

$$P' = R_{(x_r, y_r, \theta)} \cdot P$$

$$P' = \begin{bmatrix} \cos\theta & -\sin\theta & x_r(1-\cos\theta) + y_r \sin\theta \\ \sin\theta & \cos\theta & y_r(1-\cos\theta) - x_r \sin\theta \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & 8 & 8 \\ 4 & 3 & 8 \\ 1 & 1 & 1 \end{bmatrix}$$

$$P' = \begin{bmatrix} \cos(-90) & -\sin(-90) & 7(1-\cos(-90)) + 5\sin(-90) \\ \sin(-90) & \cos(-90) & 5(1-\cos(-90)) - 7\sin(-90) \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & 8 & 8 \\ 4 & 3 & 8 \\ 1 & 1 & 1 \end{bmatrix}$$

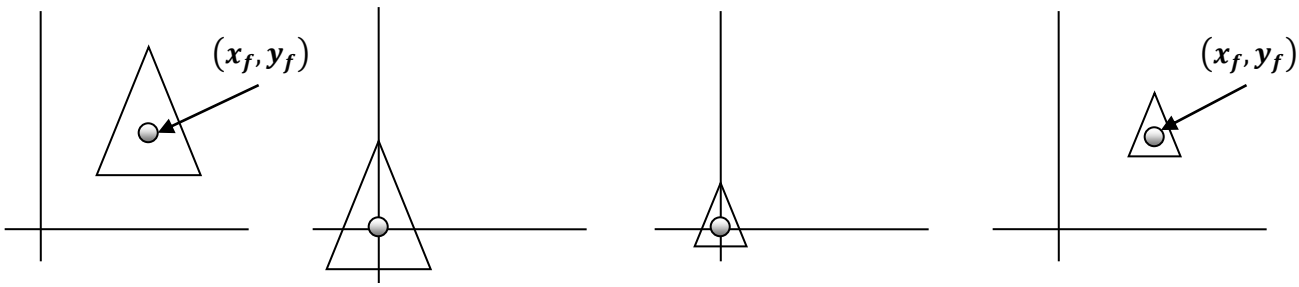
$$P' = \begin{bmatrix} 0 & 1 & 7(1-0) - 5(1) \\ -1 & 0 & 5(1-0) + 7(1) \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & 8 & 8 \\ 4 & 3 & 8 \\ 1 & 1 & 1 \end{bmatrix}$$

$$P' = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 12 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & 8 & 8 \\ 4 & 3 & 8 \\ 1 & 1 & 1 \end{bmatrix}$$

$$P' = \begin{bmatrix} 11 & 13 & 18 \\ 7 & 4 & 4 \\ 1 & 1 & 1 \end{bmatrix}$$

- Final coordinates after rotation are [A′ (11, 7), B′ (13, 4), C′ (18, 4)].

General Fixed-Point Scaling



(a) Original Position of Object and Fixed Point

(b) Translate Object so that Fixed Point (x_f, y_f) is at Origin

(c) Scale Object with Respect to Origin

(d) Translate Object so that Fixed Point is Return to Position (x_f, y_f) .

Fig. 3.7: - General fixed point scaling.

- For scaling object with position of one point called fixed point will remains same, we need to apply following sequence of transformation.
  1. Translate the object so that the fixed-point coincides with the coordinate origin.
  2. Scale the object with respect to the coordinate origin with specified scale factors.
  3. Translate the object so that the fixed-point is returned to its original position (i.e. Inverse of step-1).
- Let's find matrix equation for this

$$\boldsymbol{P' = T(x_f, y_f) \cdot [S(s_x, s_y) \cdot \{T(-x_f, -y_f) \cdot P\}]}$$

$$\boldsymbol{P' = \{T(x_f, y_f) \cdot S(s_x, s_y) \cdot T(-x_f, -y_f)\} \cdot P}$$

$$\boldsymbol{P' = \begin{bmatrix} 1 & 0 & x_f \\ 0 & 1 & y_f \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -x_f \\ 0 & 1 & -y_f \\ 0 & 0 & 1 \end{bmatrix} \cdot P}$$

$$\boldsymbol{P' = \begin{bmatrix} s_x & 0 & x_f(1 - s_x) \\ 0 & s_y & y_f(1 - s_y) \\ 0 & 0 & 1 \end{bmatrix} \cdot P}$$

$$\boldsymbol{P' = S(x_f, y_f, s_x, s_y) \cdot P}$$

Here $\boldsymbol{P'}$ and $\boldsymbol{P}$ are column vector of final and initial point coordinate respectively and (x_f, y_f) are the coordinates of fixed-point.

- **Example**: - Consider square with left-bottom corner at (2, 2) and right-top corner at (6, 6) apply the transformation which makes its size half such that its center remains same.

  Fixed point is center of square so:

$$x_f = 2 + \frac{6-2}{2}, \quad y_f = 2 + \frac{6-2}{2}$$

  As we want size half so value of scale factor are $s_x = 0.5, s_y = 0.5$ and Coordinates of square are [A (2, 2), B (6, 2), C (6, 6), D (2, 6)].

$$P' = S(x_f, y_f, s_x, s_y) \cdot P$$

$$P' = \begin{bmatrix} s_x & 0 & x_f(1 - s_x) \\ 0 & s_y & y_f(1 - s_y) \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 6 & 6 & 2 \\ 2 & 2 & 6 & 6 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$P' = \begin{bmatrix} 0.5 & 0 & 4(1-0.5) \\ 0 & 0.5 & 4(1-0.5) \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 6 & 6 & 2 \\ 2 & 2 & 6 & 6 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$P' = \begin{bmatrix} 0.5 & 0 & 2 \\ 0 & 0.5 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 6 & 6 & 2 \\ 2 & 2 & 6 & 6 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$P' = \begin{bmatrix} 3 & 5 & 5 & 3 \\ 3 & 3 & 5 & 5 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

- Final coordinate after scaling are [A′ (3, 3), B′ (5, 3), C′ (5, 5), D′ (3, 5)]

General Scaling Directions

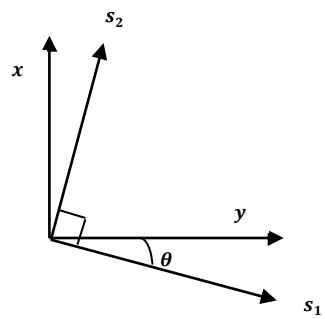


Fig. 3.8: - General scaling direction.

- Parameter $\boldsymbol{s_x}$ and $\boldsymbol{s_y}$ scale the object along $\boldsymbol{x}$ and $\boldsymbol{y}$ directions. We can scale an object in other directions by rotating the object to align the desired scaling directions with the coordinate axes before applying the scaling transformation.
- Suppose we apply scaling factor $\boldsymbol{s_1}$ and $\boldsymbol{s_2}$ in direction shown in figure than we will apply following transformations.
  1. Perform a rotation so that the direction for $\boldsymbol{s_1}$ and $\boldsymbol{s_2}$ coincide with $\boldsymbol{x}$ and $\boldsymbol{y}$ axes.
  2. Scale the object with specified scale factors.
  3. Perform opposite rotation to return points to their original orientations. (i.e. Inverse of step-1).
- Let's find matrix equation for this

$$\boldsymbol{P' = R^{-1}(\theta) \cdot [S(s_1, s_2) \cdot \{R(\theta) \cdot P\}]}$$

$$\boldsymbol{P' = \{R^{-1}(\theta) \cdot S(s_1, s_2) \cdot R(\theta)\} \cdot P}$$

$$\boldsymbol{P' = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot P}$$

$$\boldsymbol{P' = \begin{bmatrix} s_1\cos^2\theta + s_2\sin^2\theta & (s_2 - s_1)\cos\theta\sin\theta & 0 \\ (s_2 - s_1)\cos\theta\sin\theta & s_1\sin^2\theta + s_2\cos^2\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot P}$$

Here $\boldsymbol{P'}$ and $\boldsymbol{P}$ are column vector of final and initial point coordinate respectively and $\boldsymbol{\theta}$ is the angle between actual scaling direction and our standard coordinate axes.

Other Transformation

- Some package provides few additional transformations which are useful in certain applications. Two such transformations are reflection and shear.

Reflection

- A reflection is a transformation that produces a mirror image of an object.

- The mirror image for a two –dimensional reflection is generated relative to an **axis of reflection** by rotating the object 180° about the reflection axis.
- Reflection gives image based on position of axis of reflection. Transformation matrix for few positions are discussed here.

Transformation matrix for reflection about the line $\boldsymbol{y = 0}$, ***the x axis***.

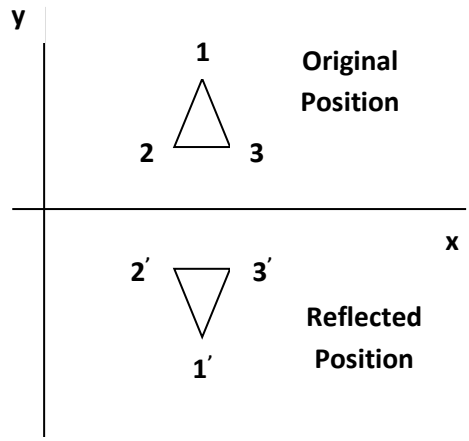


Fig. 3.9: - Reflection about x - axis.

- This transformation keeps x values are same, but flips (Change the sign) y values of coordinate positions.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Transformation matrix for reflection about the line $\boldsymbol{x = 0}$, ***the y axis***.

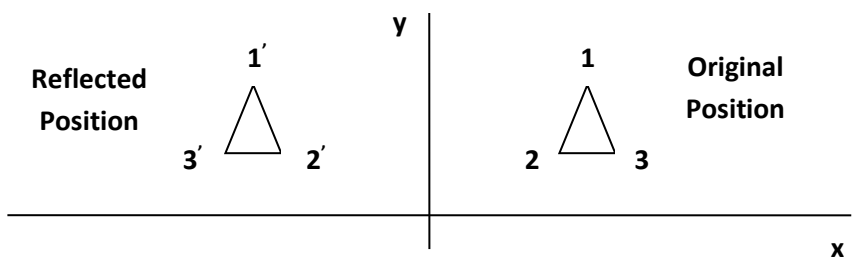


Fig. 3.10: - Reflection about y - axis.

- This transformation keeps y values are same, but flips (Change the sign) x values of coordinate positions.

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Transformation matrix for reflection about the ***Origin***.

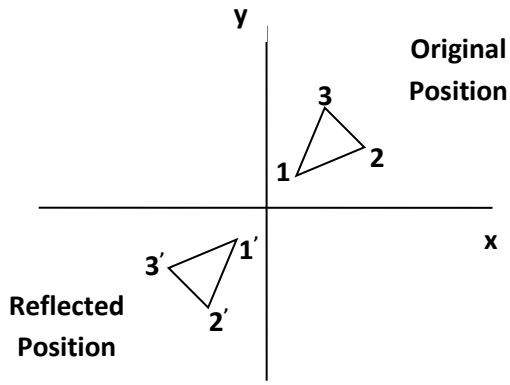


Fig. 3.11: - Reflection about origin.

- This transformation flips (Change the sign) x and y both values of coordinate positions.

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Transformation matrix for reflection about the line $\boldsymbol{x = y}$.

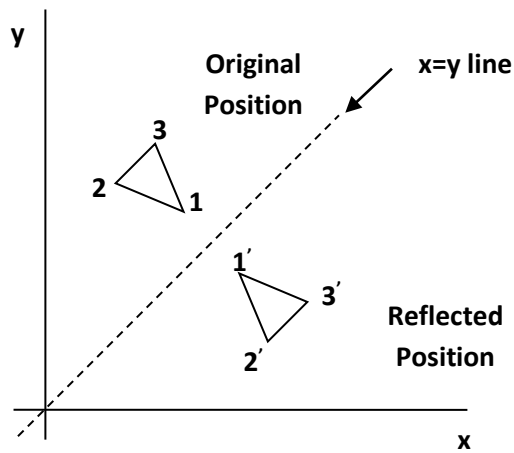


Fig. 3.12: - Reflection about x=y line.

- This transformation interchange x and y values of coordinate positions.

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Transformation matrix for reflection about the line $\boldsymbol{x = -y}$.

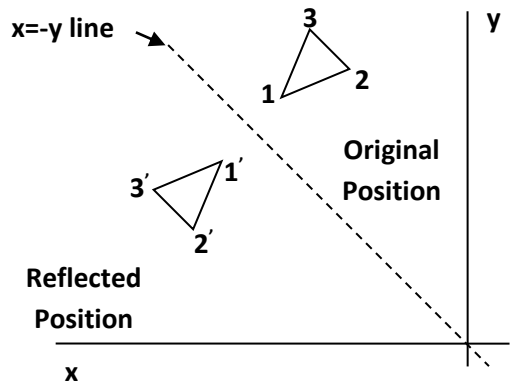


Fig. 3.12: - Reflection about x=-y line.

- This transformation interchange x and y values of coordinate positions.

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- **Example**: - Find the coordinates after reflection of the triangle [A (10, 10), B (15, 15), C (20, 10)] about x axis.

$$P' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 10 & 15 & 20 \\ 10 & 15 & 10 \\ 1 & 1 & 1 \end{bmatrix}$$

$$P' = \begin{bmatrix} 10 & 15 & 20 \\ -10 & -15 & -10 \\ 1 & 1 & 1 \end{bmatrix}$$

- Final coordinate after reflection are [A′ (10, -10), B′ (15, -15), C′ (20, -10)]

Shear

- A transformation that distorts the shape of an object such that the transformed shape appears as if the object were composed of internal layers that had been caused to slide over each other is called **shear**.
- Two common shearing transformations are those that shift coordinate x values and those that shift y values.

Shear in $\boldsymbol{x - direction}$.

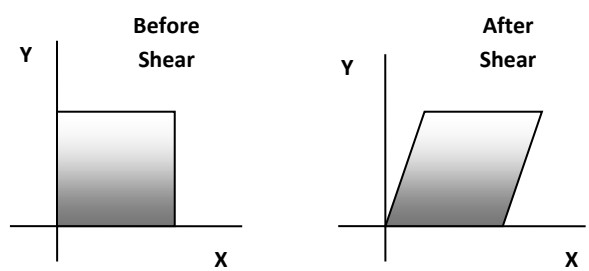


Fig. 3.13: - Shear in x-direction.

- Shear relative to $x - axis$ that is $y = 0$ line can be produced by following equation:
  $\boldsymbol{x' = x + sh_x \cdot y}$, $\boldsymbol{y' = y}$
- Transformation matrix for that is:
  $$\begin{bmatrix} \mathbf{1} & \boldsymbol{sh_x} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}$$
  Here $\boldsymbol{sh_x}$ is shear parameter. We can assign any real value to $\boldsymbol{sh_x}$.
- We can generate $x - direction$ shear relative to other reference line $y = y_{ref}$ with following equation:
  $\boldsymbol{x' = x + sh_x \cdot (y - y_{ref})}$, $\boldsymbol{y' = y}$
- Transformation matrix for that is:
  $$\begin{bmatrix} \mathbf{1} & \boldsymbol{sh_x} & \boldsymbol{-sh_x \cdot y_{ref}} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}$$
- **Example**: - Shear the unit square in x direction with shear parameter ½ relative to line $y = -1$.
  Here $y_{ref} = -1$ and $sh_x = 0.5$
  Coordinates of unit square are [A (0, 0), B (1, 0), C (1, 1), D (0, 1)].
  $$P' = \begin{bmatrix} 1 & sh_x & -sh_x \cdot y_{ref} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$
  $$P' = \begin{bmatrix} 1 & 0.5 & -0.5 \cdot (-1) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$
  $$P' = \begin{bmatrix} 1 & 0.5 & 0.5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$
  $$P' = \begin{bmatrix} 0.5 & 1.5 & 2 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$
- Final coordinate after shear are [A′ (0.5, 0), B′ (1.5, 0), C′ (2, 1), D′ (1, 1)]

Shear in $\boldsymbol{y - direction}$.

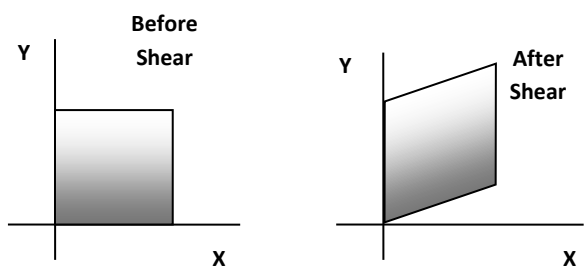


Fig. 3.14: - Shear in y-direction.

- Shear relative to $y - axis$ that is $x = 0$ line can be produced by following equation:
  $\boldsymbol{x' = x}$, $\boldsymbol{y' = y + sh_y \cdot x}$
- Transformation matrix for that is:
  $$\begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{sh_y} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}$$
  Here $\boldsymbol{sh_y}$ is shear parameter. We can assign any real value to $\boldsymbol{sh_y}$.

- We can generate $y - direction$ shear relative to other reference line $x = x_{ref}$ with following equation:
  $\boldsymbol{x' = x, \qquad y' = y + sh_y \cdot (x - x_{ref})}$
- Transformation matrix for that is:
  $$\begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{sh_y} & \mathbf{1} & \boldsymbol{-sh_y \cdot x_{ref}} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}$$
- **Example**: - Shear the unit square in y direction with shear parameter ½ relative to line $x = -1$.
  Here $x_{ref} = -1$ and $sh_y = 0.5$
  Coordinates of unit square are [A (0, 0), B (1, 0), C (1, 1), D (0, 1)].
  $$P' = \begin{bmatrix} 1 & 0 & 0 \\ sh_y & 1 & -sh_y \cdot x_{ref} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$
  $$P' = \begin{bmatrix} 1 & 0 & 0 \\ 0.5 & 1 & -0.5 \cdot (-1) \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$
  $$P' = \begin{bmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0.5 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$
  $$P' = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0.5 & 1 & 2 & 1.5 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$
- Final coordinate after shear are [A′ (0, 0.5), B′ (1, 1), C′ (1, 2), D′ (0, 1.5)]

The Viewing Pipeline

- **Window:** Area selected in world-coordinate for display is called window. It defines what is to be viewed.
- **Viewport:** Area on a display device in which window image is display (mapped) is called viewport. It defines where to display.
- In many case window and viewport are rectangle, also other shape may be used as window and viewport.
- In general finding device coordinates of viewport from word coordinates of window is called as **viewing transformation.**
- Sometimes we consider this viewing transformation as window-to-viewport transformation but in general it involves more steps.

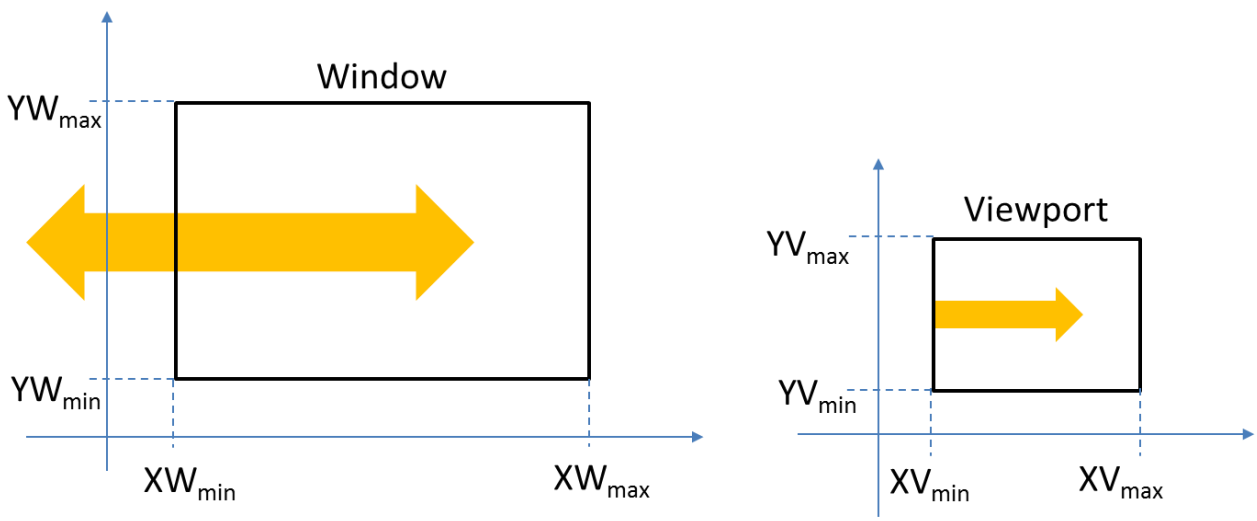


Fig. 3.1: - A viewing transformation using standard rectangles for the window and viewport.

- Now we see steps involved in viewing pipeline.

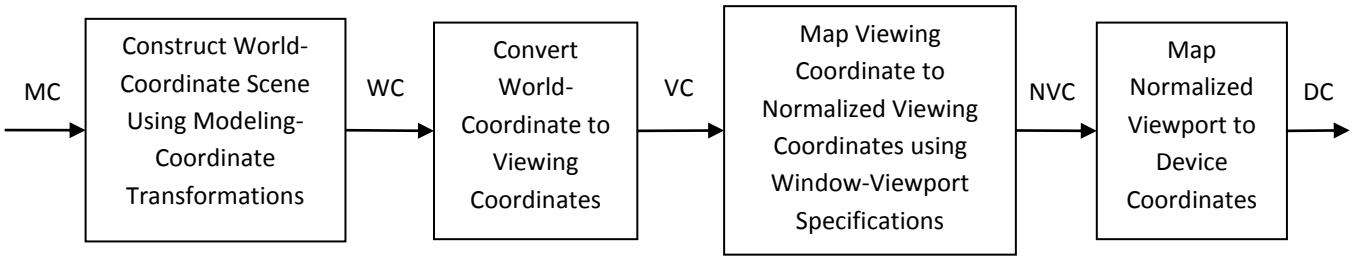


Fig. 3.2: - 2D viewing pipeline.

- As shown in figure above first of all we construct world coordinate scene using modeling coordinate transformation.
- After this we convert viewing coordinates from world coordinates using window to viewport transformation.
- Then we map viewing coordinate to normalized viewing coordinate in which we obtain values in between 0 to 1.
- At last we convert normalized viewing coordinate to device coordinate using device driver software which provide device specification.
- Finally device coordinate is used to display image on display screen.
- By changing the viewport position on screen we can see image at different place on the screen.
- By changing the size of the window and viewport we can obtain zoom in and zoom out effect as per requirement.
- Fixed size viewport and small size window gives zoom in effect, and fixed size viewport and larger window gives zoom out effect.
- View ports are generally defines with the unit square so that graphics package are more device independent which we call as normalized viewing coordinate.

Viewing Coordinate Reference Frame

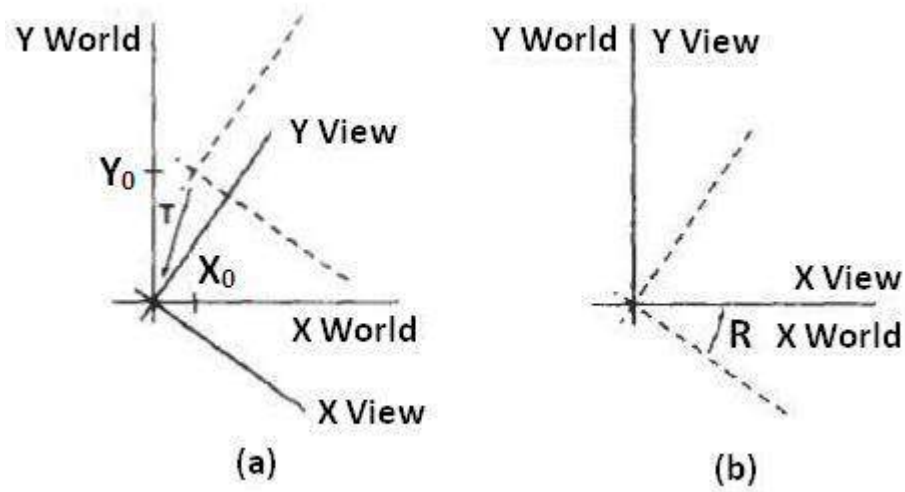


Fig. 3.3: - A viewing-coordinate frame is moved into coincidence with the world frame in two steps: (a) translate the viewing origin to the world origin, and then (b) rotate to align the axes of the two systems.

- We can obtain reference frame in any direction and at any position.
- For handling such condition first of all we translate reference frame origin to standard reference frame origin and then we rotate it to align it to standard axis.
- In this way we can adjust window in any reference frame.
- this is illustrate by following transformation matrix:

$$\mathbf{M_{wc,vc} = RT}$$

- Where T is translation matrix and R is rotation matrix.

Window-To-Viewport Coordinate Transformation

- Mapping of window coordinate to viewport is called window to viewport transformation.
- We do this using transformation that maintains relative position of window coordinate into viewport.
- That means center coordinates in window must be remains at center position in viewport.
- We find relative position by equation as follow:

$$\frac{\mathbf{x_v - x_{vmin}}}{\mathbf{x_{vmax} - x_{vmin}}} = \frac{\mathbf{x_w - x_{wmin}}}{\mathbf{x_{wmax} - x_{wmin}}}$$

$$\frac{\mathbf{y_v - y_{vmin}}}{\mathbf{y_{vmax} - y_{vmin}}} = \frac{\mathbf{y_w - y_{wmin}}}{\mathbf{y_{wmax} - y_{wmin}}}$$

- Solving by making viewport position as subject we obtain:

$$\mathbf{x_v = x_{vmin} + (x_w - x_{wmin})s_x}$$
$$\mathbf{y_v = y_{vmin} + (y_w - y_{wmin})s_y}$$

- Where scaling factor are :

$$\mathbf{s_x = \frac{x_{vmax} - x_{vmin}}{x_{wmax} - x_{wmin}}}$$
$$\mathbf{s_y = \frac{y_{vmax} - y_{vmin}}{y_{wmax} - y_{wmin}}}$$

- We can also map window to viewport with the set of transformation, which include following sequence of transformations:
  1. Perform a scaling transformation using a fixed-point position of (x_{Wmin},y_{wmin}) that scales the window area to the size of the viewport.
  2. Translate the scaled window area to the position of the viewport.
- For maintaining relative proportions we take ($s_x = s_y$). in case if both are not equal then we get stretched or contracted in either the x or y direction when displayed on the output device.
- Characters are handle in two different way one way is simply maintain relative position like other primitive and other is to maintain standard character size even though viewport size is enlarged or reduce.
- Number of display device can be used in application and for each we can use different window-to-viewport transformation. This mapping is called the **workstation transformation.**

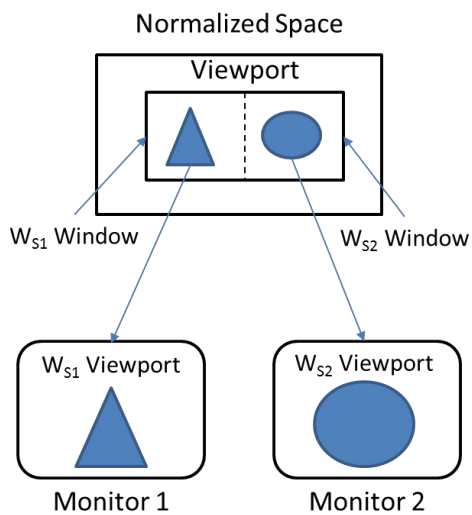


Fig. 3.4: - workstation transformation.

- As shown in figure two different displays devices are used and we map different window-to-viewport on each one.

Clipping Operations

- Generally, any procedure that identifies those portions of a picture that are either inside or outside of a specified region of space is referred to as a **clipping algorithm**, or simply **clipping**. The region against which an object is to clip is called a **clip window.**
- Clip window can be general polygon or it can be curved boundary.

Application of Clipping

- It can be used for displaying particular part of the picture on display screen.
- Identifying visible surface in 3D views.
- Antialiasing.
- Creating objects using solid-modeling procedures.
- Displaying multiple windows on same screen.
- Drawing and painting.

Point Clipping

- In point clipping we eliminate those points which are outside the clipping window and draw points which are inside the clipping window.
- Here we consider clipping window is rectangular boundary with edge (x_{wmin},x_{wmax},y_{wmin},y_{wmax}).
- So for finding wether given point is inside or outside the clipping window we use following inequality:

  $\boldsymbol{x_{wmin} \leq x \leq x_{wamx}}$

  $\boldsymbol{y_{wmin} \leq y \leq y_{wamx}}$
- If above both inequality is satisfied then the point is inside otherwise the point is outside the clipping window.

Line Clipping

- Line clipping involves several possible cases.
  1. Completely inside the clipping window.
  2. Completely outside the clipping window.
  3. Partially inside and partially outside the clipping window.

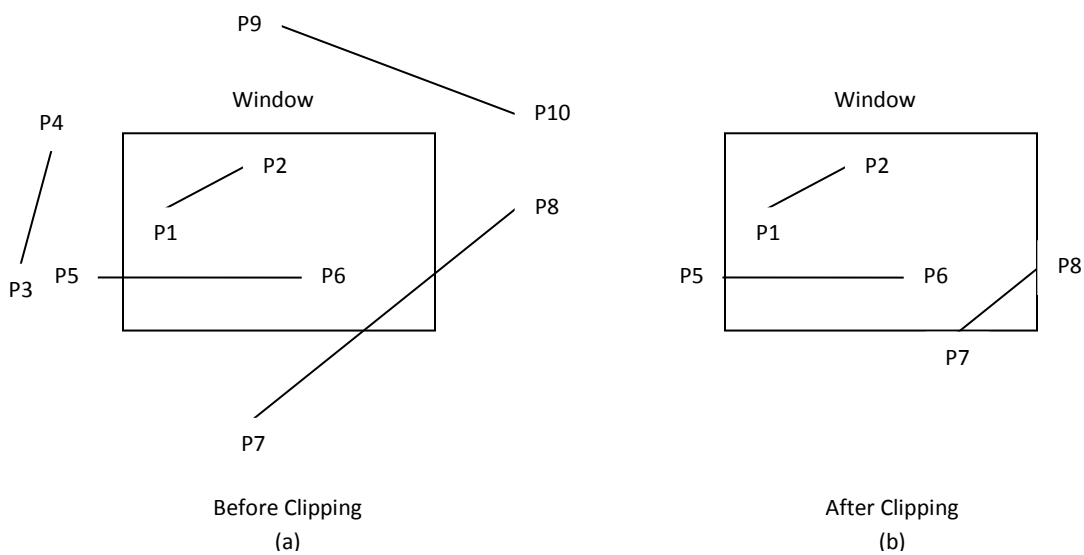


Fig. 3.5: - Line clipping against a rectangular window.

- Line which is completely inside is display completely. Line which is completely outside is eliminated from display. And for partially inside line we need to calculate intersection with window boundary and find which part is inside the clipping boundary and which part is eliminated.
- For line clipping several scientists tried different methods to solve this clipping procedure. Some of them are discuss below.

Cohen-Sutherland Line Clipping

- This is one of the oldest and most popular line-clipping procedures.

Region and Region Code

- In this we divide whole space into nine region and assign 4 bit code to each endpoint of line depending on the position where the line endpoint is located.

1001	1000	1010
0001	0000	0010
0101	0100	0110

Fig. 3.6: - Workstation transformation.

- Figure 3.6 shows code for line end point which is fall within particular area.
- Code is deriving by setting particular bit according to position of area.
  Set bit 1: For left side of clipping window.
  Set bit 2: For right side of clipping window.
  Set bit 3: For below clipping window.
  Set bit 4: For above clipping window.
- All bits as mention above are set means 1 and other are 0.

Algorithm

Step-1:

Assign region code to both endpoint of a line depending on the position where the line endpoint is located.

Step-2:

If both endpoint have code '0000'

Then line is completely inside.

Otherwise

Perform logical ending between this two codes.

If result of logical ending is non-zero

Line is completely outside the clipping window.

Otherwise

Calculate the intersection point with the boundary one by one.

Divide the line into two parts from intersection point.

Recursively call algorithm for both line segments.

Step-3:

Draw line segment which are completely inside and eliminate other line segment which found completely outside.

Intersection points calculation with clipping window boundary

- For intersection calculation we use line equation "$y = mx + b$".
- 'x' is constant for left and right boundary which is:
  - for left "$x = x_{wmin}$"
  - for right "$x = x_{wmax}$"
- So we calculate y coordinate of intersection for this boundary by putting values of x depending on boundary is left or right in below equation.

  $\boldsymbol{y = y_1 + m(x - x_1)}$
- 'y' coordinate is constant for top and bottom boundary which is:
  - for top "$y = y_{wmax}$"
  - for bottom "$y = y_{wmin}$"
- So we calculate x coordinate of intersection for this boundary by putting values of y depending on boundary is top or bottom in below equation.

  $\boldsymbol{x = x_1 + \frac{y - y_1}{m}}$

Liang-Barsky Line Clipping

- Line clipping approach is given by the Liang and Barsky is faster than cohen-sutherland line clipping. Which is based on analysis of the parametric equation of the line which are as below.

  $\boldsymbol{x = x_1 + u\Delta x}$

  $\boldsymbol{y = y_1 + u\Delta y}$

  Where $0 \leq u \leq 1$,$\Delta x = x_2 - x_1$ and $\Delta y = y_2 - y_1$.

Algorithm

1. Read two end points of line $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$
2. Read two corner vertices, left top and right bottom of window: (x_{wmin}, y_{wmax}) and (x_{wmax}, y_{wmin})
3. Calculate values of parameters p_k and q_k for $k = 1, 2, 3, 4$ such that,

   $p_1 = -\Delta x, \quad q_1 = x_1 - x_{wmin}$

   $p_2 = \Delta x, \quad q_2 = x_{wmax} - x_1$

   $p_3 = -\Delta y, \quad q_3 = y_1 - y_{wmin}$

   $p_4 = \Delta y, \quad q_4 = y_{wmax} - y_1$
4. If $p_k = 0$ for any value of $k = 1, 2, 3, 4$ then,

   Line is parallel to k^{th} boundary.

   If corresponding $q_k < 0$ then,

   Line is completely outside the boundary. Therefore, discard line segment and Go to Step 10.

   Otherwise

   Check line is horizontal or vertical and accordingly check line end points with corresponding boundaries.

   If line endpoints lie within the bounded area

   Then use them to draw line.

   Otherwise

Use boundary coordinates to draw line. And go to Step 8.

5. For $k = 1,2,3,4$ calculate r_k for nonzero values of p_k and q_k as follows:
   $r_k = \frac{q_k}{p_k}$, $for\ k = 1,2,3,4$
6. Find $u_1\ and\ u_2$ as given below:
   $u_1 = \max\{0, r_k | where\ k\ takes\ all\ values\ for\ which\ p_k < 0\}$
   $u_2 = \min\{1, r_k | where\ k\ takes\ all\ values\ for\ which\ p_k > 0\}$
7. If $u_1 \leq u_2$ then
   Calculate endpoints of clipped line:
   $x_1' = x_1 + u_1\Delta x$
   $y_1' = y_1 + u_1\Delta y$
   $x_2' = x_1 + u_2\Delta x$
   $y_2' = y_1 + u_2\Delta y$
   Draw line (x_1', y_1', x_2', y_2');
8. Stop.

Advantages

1. More efficient.
2. Only requires one division to update u_1 and u_2.
3. Window intersections of line are calculated just once.

Nicholl-Lee-Nicholl Line Clipping

- By creating more regions around the clip window the NLN algorithm avoids multiple clipping of an individual line segment.
- In Cohen-Sutherlan line clipping sometimes multiple calculation of intersection point of a line is done before actual window boundary intersection or line is completely rejected.
- These multiple intersection calculation is avoided in NLN line clipping procedure.
- NLN line clipping perform the fewer comparisons and divisions so it is more efficient.
- But NLN line clipping cannot be extended for three dimensions while Cohen-Sutherland and Liang-Barsky algorithm can be easily extended for three dimensions.
- For given line we find first point falls in which region out of nine region shown in figure below but three region shown in figure by putting point are only considered and if point falls in other region than we transfer that point in one of the three region.

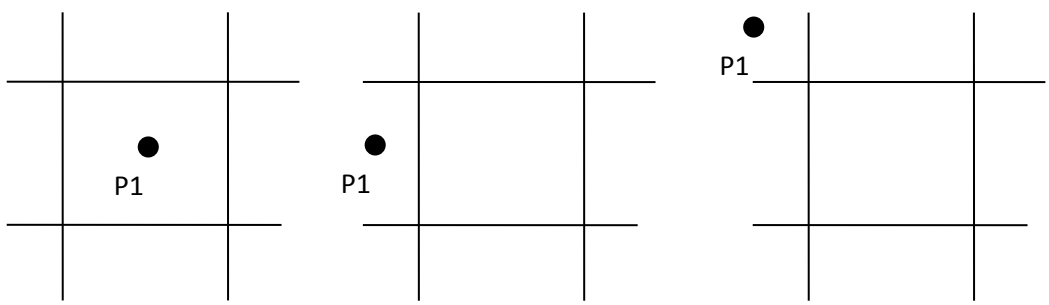


Fig. 3.7: - Three possible position for a line endpoint p1 in the NLN line-clipping algorithm.

- We can also extend this procedure for all nine regions.
- Now for p1 is inside the window we divide whole area in following region:

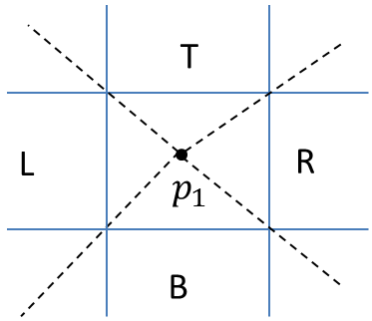


Fig. 3.8: - Clipping region when p1 is inside the window.

- Now for p1 is in edge region we divide whole area in following region:

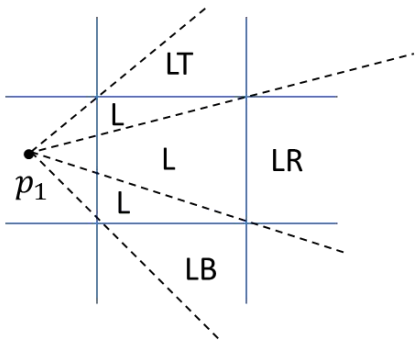


Fig. 3.9: - Clipping region when p1 is in edge region.

- Now for p1 is in corner region we divide whole area in following region:

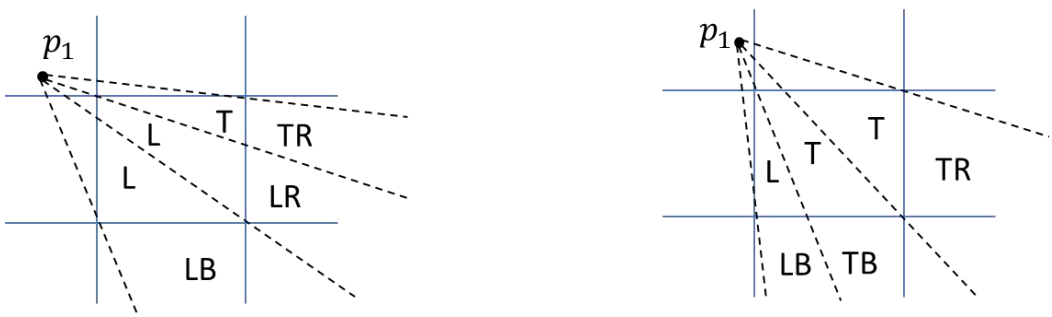


Fig. 3.10: - Two possible sets of clipping region when p1 is in corner region.

- Regions are name in such a way that name in which region p2 falls is gives the window edge which intersects the line.
- For example region LT says that line need to clip at left and top boundary.
- For finding that in which region line $\boldsymbol{p_1p_2}$ falls we compare the slope of the line to the slope of the boundaries:

  $\boldsymbol{slope\ \overline{p_1p_{B1}} < slope\ \overline{p_1p_2} < slope\ \overline{p_1p_{B2}}}$

  Where $\boldsymbol{\overline{p_1p_{B1}}}$ and $\boldsymbol{\overline{p_1p_{B2}}}$ are boundary lines.
- For example p1 is in edge region and for checking whether p2 is in region LT we use following equation.

  $\boldsymbol{slope\ \overline{p_1p_{TR}} < slope\ \overline{p_1p_2} < slope\ \overline{p_1p_{TL}}}$

  $$\frac{y_T - y_1}{x_R - x_1} < \frac{y_2 - y_1}{x_2 - x_1} < \frac{y_T - y_1}{x_L - x_1}$$
- After checking slope condition we need to check weather it crossing zero, one or two edges.
- This can be done by comparing coordinates of p_2 with coordinates of window boundary.
- For left and right boundary we compare x coordinates and for top and bottom boundary we compare y coordinates.
- If line is not fall in any defined region than clip entire line.

- Otherwise calculate intersection.
- After finding region we calculate intersection point using parametric equation which are:
- $\boldsymbol{x = x_1 + (x_2 - x_1)u}$
- $\boldsymbol{y = y_1 + (y_2 - y_1)u}$
- For left or right boundary $x = x_l \text{ or } x_r$ respectively, with $u = (x_{l/r} - x1)/(x2 - x1)$, so that y can be obtain from parametric equation as below:
- $\boldsymbol{y = y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x_L - x_1)}$
- Keep the portion which is inside and clip the rest.

Polygon Clipping

- For polygon clipping we need to modify the line clipping procedure because in line clipping we need to consider about only line segment while in polygon clipping we need to consider the area and the new boundary of the polygon after clipping.

Sutherland-Hodgeman Polygon Clipping

- For correctly clip a polygon we process the polygon boundary as a whole against each window edge.
- This is done by whole polygon vertices against each clip rectangle boundary one by one.
- Beginning with the initial set of polygon vertices we first clip against the left boundary and produce new sequence of vertices.
- Then that new set of vertices is clipped against the right boundary clipper, a bottom boundary clipper and a top boundary clipper, as shown in figure below.

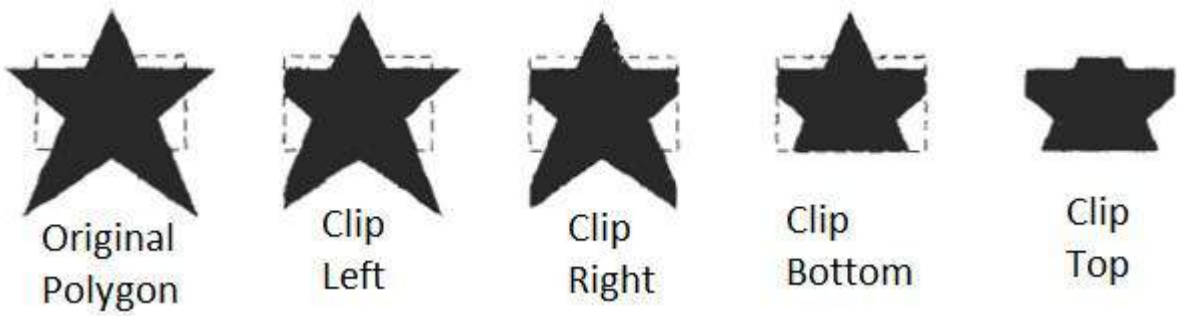


Fig. 3.11: - Clipping a polygon against successive window boundaries.

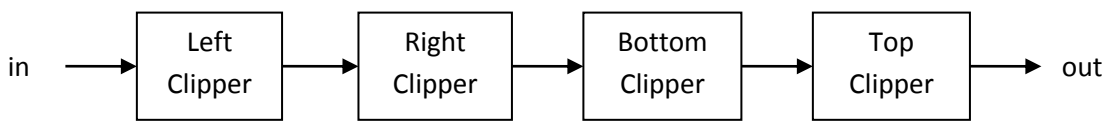


Fig. 3.12: - Processing the vertices of the polygon through boundary clipper.

- There are four possible cases when processing vertices in sequence around the perimeter of a polygon.

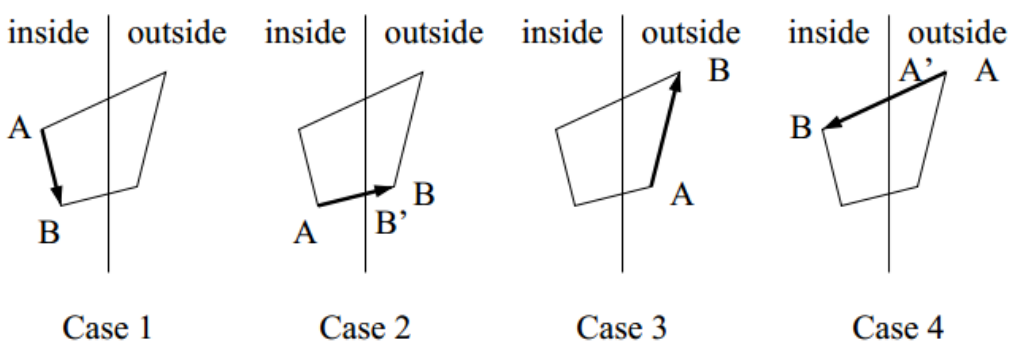


Fig. 3.13: - Clipping a polygon against successive window boundaries.

- As shown in case 1: if both vertices are inside the window we add only second vertices to output list.
- In case 2: if first vertices is inside the boundary and second vertices is outside the boundary only the edge intersection with the window boundary is added to the output vertex list.
- In case 3: if both vertices are outside the window boundary nothing is added to window boundary.
- In case 4: first vertex is outside and second vertex is inside the boundary, then adds both intersection point with window boundary, and second vertex to the output list.
- When polygon clipping is done against one boundary then we clip against next window boundary.
- We illustrate this method by simple example.

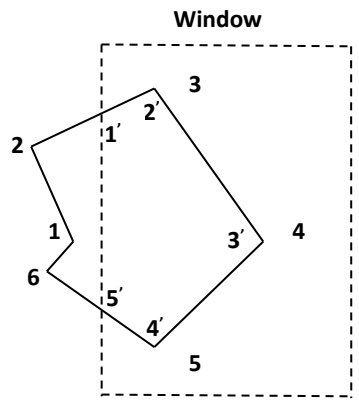
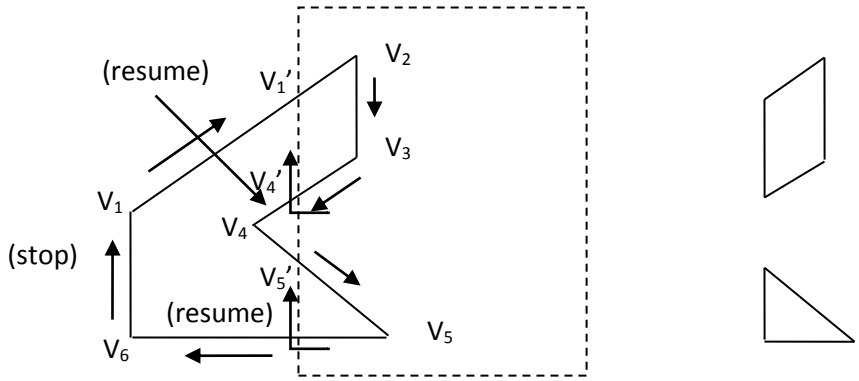


Fig. 3.14: - Clipping a polygon against left window boundaries.

- As shown in figure above we clip against left boundary vertices 1 and 2 are found to be on the outside of the boundary. Then we move to vertex 3, which is inside, we calculate the intersection and add both intersection point and vertex 3 to output list.
- Then we move to vertex 4 in which vertex 3 and 4 both are inside so we add vertex 4 to output list, similarly from 4 to 5 we add 5 to output list, then from 5 to 6 we move inside to outside so we add intersection pint to output list and finally 6 to 1 both vertex are outside the window so we does not add anything.
- Convex polygons are correctly clipped by the Sutherland-Hodgeman algorithm but concave polygons may be displayed with extraneous lines.
- For overcome this problem we have one possible solution is to divide polygon into numbers of small convex polygon and then process one by one.
- Another approach is to use Weiler-Atherton algorithm.

Weiler-Atherton Polygon Clipping

- In this algorithm vertex processing procedure for window boundary is modified so that concave polygon also clip correctly.
- This can be applied for arbitrary polygon clipping regions as it is developed for visible surface identification.
- Main idea of this algorithm is instead of always proceeding around the polygon edges as vertices are processed we sometimes need to follow the window boundaries.
- Other procedure is similar to Sutherland-Hodgeman algorithm.
- For clockwise processing of polygon vertices we use the following rules:
  - For an outside to inside pair of vertices, follow the polygon boundary.
  - For an inside to outside pair of vertices, follow the window boundary in a clockwise direction.
- We illustrate it with example:



(a) (b)

Fig. 3.14: - Clipping a concave polygon (a) with the Weiler-Atherton algorithm generates the two se

- As shown in figure we start from v1 and move clockwise towards v2 and add intersection point and next point to output list by following polygon boundary, then from v2 to v3 we add v3 to output list.
- From v3 to v4 we calculate intersection point and add to output list and follow window boundary.
- Similarly from v4 to v5 we add intersection point and next point and follow the polygon boundary, next we move v5 to v6 and add intersection point and follow the window boundary, and finally v6 to v1 is outside so no need to add anything.
- This way we get two separate polygon section after clipping.

Three Dimensional Display Methods

Parallel Projection

- This method generates view from solid object by projecting parallel lines onto the display plane.
- By changing viewing position we can get different views of 3D object onto 2D display screen.

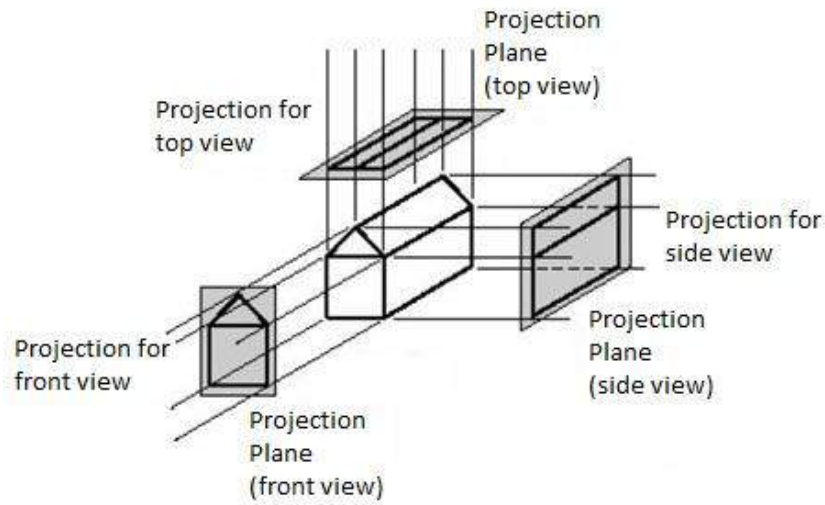


Fig. 4.1: - different views object by changing viewing plane position.

- Above figure shows different views of objects.
- This technique is used in Engineering & Architecture drawing to represent an object with a set of views that maintain relative properties of the object e.g.:- orthographic projection.

Perspective projection

- This method generating view of 3D object by projecting point on the display plane along converging paths.

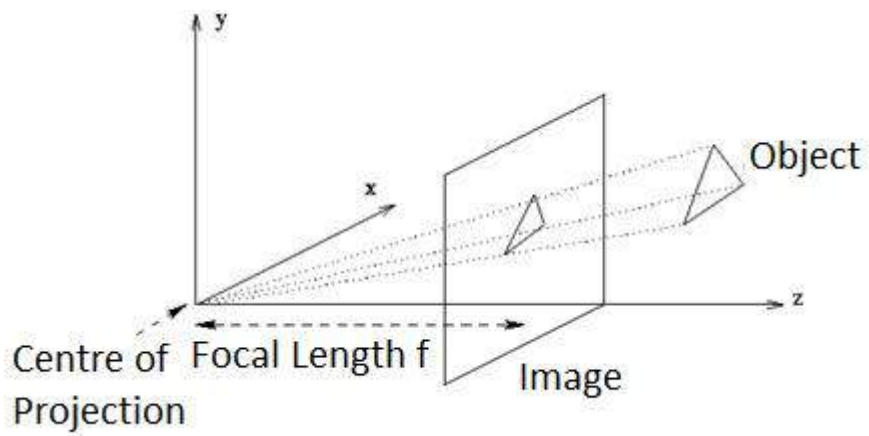


Fig. 4.2: - perspective projection

- This will display object smaller when it is away from the view plane and of nearly same size when closer to view plane.
- It will produce more realistic view as it is the way our eye is forming image.

Depth cueing

- Many times depth information is important so that we can identify for a particular viewing direction which are the front surfaces and which are the back surfaces of display object.

- Simple method to do this is depth cueing in which assign higher intensity to closer object & lower intensity to the far objects.
- Depth cuing is applied by choosing maximum and minimum intensity values and a range of distance over which the intensities are to vary.
- Another application is to modeling effect of atmosphere.

Visible line and surface Identification

- In this method we first identify visible lines or surfaces by some method.
- Then display visible lines with highlighting or with some different color.
- Other way is to display hidden lines with dashed lines or simply not display hidden lines.
- But not drawing hidden lines will loss the some information.
- Similar method we can apply for the surface also by displaying shaded surface or color surface.
- Some visible surface algorithm establishes visibility pixel by pixel across the view plane. Other determines visibility of object surface as a whole.

Surface Rendering

- More realistic image is produce by setting surface intensity according to light reflect from that surface & the characteristics of that surface.
- It will give more intensity to the shiny surface and less to dull surface.
- It also applies high intensity where light is more & less where light falls is less.

Exploded and Cutaway views

- Many times internal structure of the object is need to store. For ex., in machine drawing internal assembly is important.
- For displaying such views it will remove (cutaway) upper cover of body so that internal part's can be visible.

Three dimensional stereoscopic views

- This method display using computer generated scenes.
- It may display object by three dimensional views.
- The graphics monitor which are display three dimensional scenes are devised using a technique that reflects a CRT image from a vibrating flexible mirror.

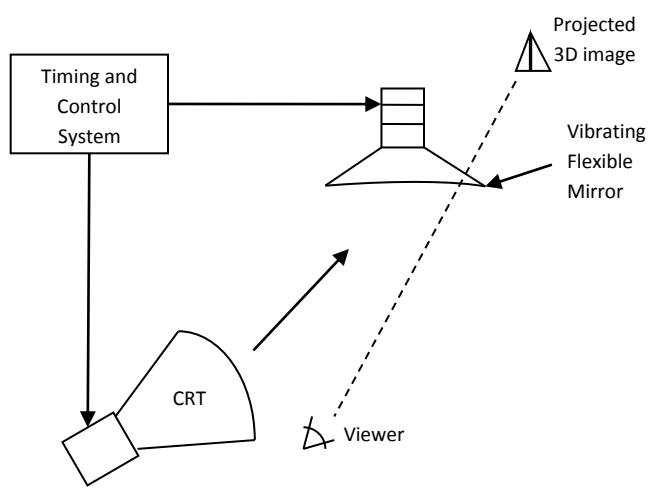


Fig. 4.3: - 3D display system uses a vibrating mirror.

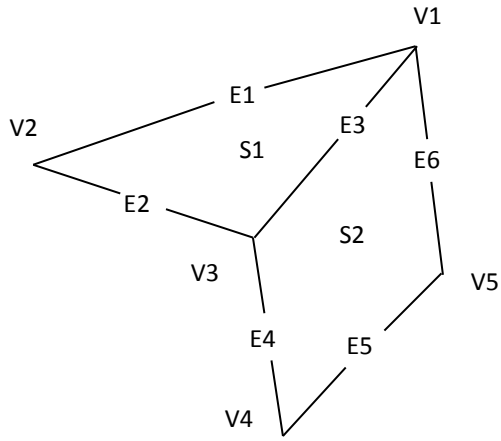
- Vibrating mirror changes its focal length due to vibration which is synchronized with the display of an object on CRT.
- The each point on the object is reflected from the mirror into spatial position corresponding to distance of that point from a viewing position.
- Very good example of this system is GENISCO SPACE GRAPH system, which use vibrating mirror to project 3D objects into a 25 cm by 25 cm by 25 cm volume. This system is also capable to show 2D cross section at different depth.
- Another way is stereoscopic views.
- Stereoscopic views does not produce three dimensional images, but it produce 3D effects by presenting different view to each eye of an observer so that it appears to have depth.
- To obtain this we first need to obtain two views of object generated from viewing direction corresponding to each eye.
- We can contract the two views as computer generated scenes with different viewing positions or we can use stereo camera pair to photograph some object or scene.
- When we see simultaneously both the view as left view with left eye and right view with right eye then two views is merge and produce image which appears to have depth.
- One way to produce stereoscopic effect is to display each of the two views with raster system on alternate refresh cycles.
- The screen is viewed through glasses with each lance design such a way that it act as a rapidly alternating shutter that is synchronized to block out one of the views.

Polygon Surfaces

- A polygonal surface can be thought of as a surface composed of polygonal faces.
- The most commonly used boundary representation for a three dimensional object is a set of polygon surfaces that enclose the object interior

Polygon Tables

- Representation of vertex coordinates, edges and other property of polygon into table form is called polygon table.
- Polygon data tables can be organized into two groups: geometric table and attributes table.
- Geometric table contains vertex coordinate and the other parameter which specify geometry of polygon.
- Attributes table stores other information like Color, transparency etc.
- Convenient way to represent geometric table into three different table namely vertex table, edge table, and polygon table.



Vertex Table	
V1:	X1, Y1, Z1
V2:	X2, Y2, Z2
V3:	X3, Y3, Z3
V4:	X4, Y4, Z4
V5:	X5, Y5, Z5

Edge Table	
E1:	V1, V2
E2:	V2, V3
E3:	V3, V1
E4:	V3, V4
E5:	V4, V5
E6:	V5, V1

Polygon Surface Table	
S1:	E1, E2, E3
S2:	E3, E4, E5, E6

Fig. 4.4: - Geometric Data Table representation.

- Vertex table stores each vertex included into polygon.
- Edge table stores each edge with the two endpoint vertex pointers back to vertex table.
- Polygon table stores each surface of polygon with edge pointer for the each edge of the surface.
- This three table representation stores each vertex one time only and similarly each edge is also one time. So it will avoid representation of storing common vertex and edge so it is memory efficient method.
- Another method to represent with two table vertex table and polygon table but it is inefficient as it will store common edge multiple times.
- Since tables have many entries for large number of polygon we need to check for consistency as it may be possible that errors may occurs during input.
- For dealing with that we add extra information into tables. For example figure below shows edge table of above example with information of the surface in which particular edge is present.

E1:	V1, V2, S1
E2:	V2, V3, S1
E3:	V3, V1, S1, S2
E4:	V3, V4, S2
E5:	V4, V5, S2
E6:	V5, V1, S2

Fig. 4.5: - Edge table of above example with extra information as surface pointer.

- Now if any surface entry in polygon table will find edge in edge table it will verify whether this edge is of particular surface's edge or not if not it will detect errors and may be correct if sufficient information is added.

Plane Equations

- For producing display of 3D objects we must process the input data representation for the object through several procedures.
- For this processing we sometimes need to find orientation and it can be obtained by vertex coordinate values and the equation of polygon plane.
- Equation of plane is given as
  $Ax + By + Cz + D = 0$
- Where (x, y, z) is any point on the plane and A, B, C, D are constants by solving three plane equation for three non collinear points. And solve simultaneous equation for ratio A/D, B/D, and C/D as follows
  $\frac{A}{D}x_1 + \frac{B}{D}y_1 + \frac{C}{D}z_1 = -1$
  $\frac{A}{D}x_2 + \frac{B}{D}y_2 + \frac{C}{D}z_2 = -1$
  $\frac{A}{D}x_3 + \frac{B}{D}y_3 + \frac{C}{D}z_3 = -1$
- Solving by determinant
  $A = \begin{vmatrix} 1 & y_1 & z_1 \\ 1 & y_2 & z_2 \\ 1 & y_3 & z_3 \end{vmatrix}$ $B = \begin{vmatrix} x_1 & 1 & z_1 \\ x_2 & 1 & z_2 \\ x_3 & 1 & z_3 \end{vmatrix}$ $C = \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}$ $D = -\begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$
- By expanding a determinant we get
  $A = y_1(z_2 - z_3) + y_2(z_3 - z_1) + y_3(z_1 - z_2)$
  $B = z_1(x_2 - x_3) + z_2(x_3 - x_1) + z_3(x_1 - x_2)$
  $C = x_1(y_2 - y_3) + x_2(y_3 - y_1) + x_3(y_1 - y_2)$
  $D = -x_1(y_2z_3 - y_3z_2) - x_2(y_3z_1 - y_1z_3) - x_3(y_1z_2 - y_2z_1)$
- This values of A, B, C, D are then store in polygon data structure with other polygon data.
- Orientation of plane is described with normal vector to the plane.

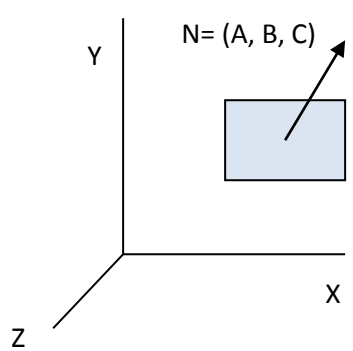


Fig. 4.6: - the vector N normal to the surface.

- Here N= (A,B,C) where A, B, C are the plane coefficient.
- When we are dealing with the polygon surfaces that enclose object interior we define the side of the faces towards object interior is as inside face and outward side as outside face.
- We can calculate normal vector N for any particular surface by cross product of two vectors in counter clockwise direction in right handed system then.
  $N = (v2 - v1)X(v3 - v1)$
- Now N gives values A, B, C for that plane and D can be obtained by putting these values in plane equation for one of the vertices and solving for D.
- Using plane equation in vector form we can obtain D as
  $N \cdot P = -D$
- Plane equation is also used to find position of any point compare to plane surface as follows

If $Ax + By + Cz + D \neq 0$ the point (x,y,z) is not on that plane.
If $Ax + By + Cz + D < 0$ the point (x,y,z) is inside the surface.
If $Ax + By + Cz + D > 0$ the point (x,y,z) is outside the surface.

- These equation are valid for right handed system provides plane parameter A, B, C, and D were calculated using vertices selected in a counter clockwise order when viewing the surface in an outside to inside direction.

Polygon Meshes

Fig. 4.7: -A triangle strip formed with 11 triangle connecting 13 vertices

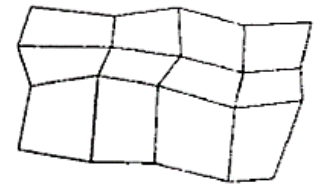


Fig. 4.8: -A quadrilateral mesh containing 12 quadrilaterals constructed from a 5 by 4 input vertex array

- Polygon mesh is a collection of edges, vertices and faces that defines the shape of the polyhedral object in 3D computer graphics and solid modeling.
- An edge can be shared by two or more polygons and vertex is shared by at least two edges.
- Polygon mesh is represented in following ways
  - Explicit representation
  - Pointer to vertex list
  - Pointer to edge list

Explicit Representation

- In explicit representation each polygon stores all the vertices in order in the memory as,
  $P = \left((x_1, y_1, z_1), (x_2, y_2, z_2)\right), \dots, \left((x_m, y_m, z_m), (x_n, y_n, z_n)\right)\Big)$
- It process fast but requires more memory for storing.

Pointer to Vertex list

- In this method each vertex stores in vertex list and then polygon contains pointer to the required vertex.
  $V = ((x_1, y_1, z_1), (x_2, y_2, z_2), \dots, (x_n, y_n, z_n))$
- And now polygon of vertices 3, 4, 5 is represented as $P = ((v_3, v_4), (v_4, v_5), (v_5, v_3))$.
- It is considerably space saving but common edges is difficult to find.

Pointer to Edge List

- In this polygon have pointers to the edge list and edge list have pointer to vertex list for each edge two vertex pointer is required which points in vertex list.
  $V = ((x_1, y_1, z_1), (x_2, y_2, z_2), \dots, (x_n, y_n, z_n))$
  $E = ((v_1, v_2), (v_2, v_3), \dots, (v_n, v_m))$
  $P = (E_1, E_2, E_n)$
- This approach is more memory efficient and easy to find common edges.

Spline Representations

- Spline is flexible strip used to produce a smooth curve through a designated set of points.
- Several small weights are attached to spline to hold in particular position.
- Spline curve is a curve drawn with this method.

- The term spline curve now referred to any composite curve formed with polynomial sections satisfying specified continuity condition at the boundary of the pieces.
- A spline surface can be described with two sets of orthogonal spline curves.

Interpolation and approximation splines

- We specify spline curve by giving a set of coordinate positions called control points. This indicates the general shape of the curve.
- **Interpolation Spline**: - When curve section passes through each control point, the curve is said to interpolate the set of control points and that spline is known as Interpolation Spline.

Fig. 4.9: -interpolation spline.

Fig. 4.10: -Approximation spline.

- **Approximation Spline**: - When curve section follows general control point path without necessarily passing through any control point, the resulting curve is said to approximate the set of control points and that curve is known as Approximation Spline.
- Spline curve can be modified by selecting different control point position.
- We can apply transformation on the curve according to need like translation scaling etc.
- The convex polygon boundary that encloses a set of control points is called **convex hull**.

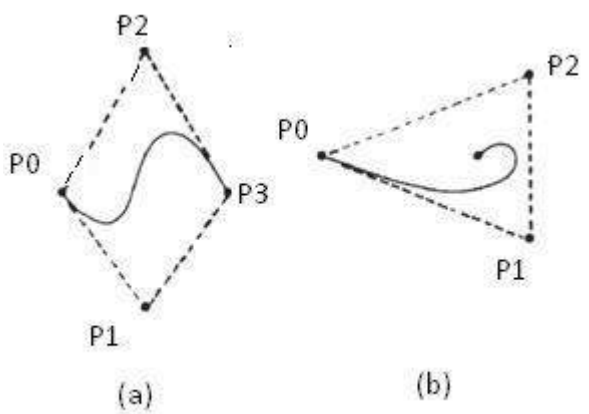


Fig. 4.11: -convex hull shapes for two sets of control points.

- A poly line connecting the sequence of control points for an approximation spline is usually displayed to remind a designer of the control point ordering. This set of connected line segment is often referred as control graph of the curve.
- Control graph is also referred as control polygon or characteristic polygon.

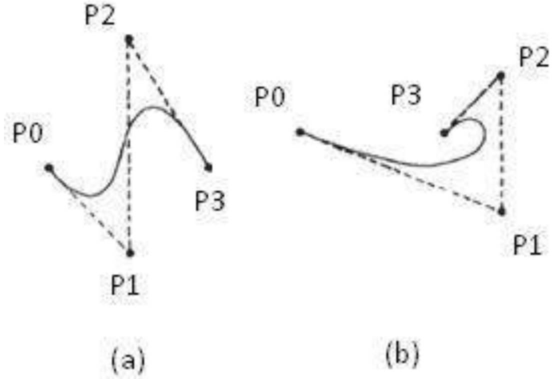


Fig. 4.12: -Control-graph shapes for two different sets of control points.

Parametric continuity condition

- For smooth transition from one curve section on to next curve section we put various continuity conditions at connection points.
- Let parametric coordinate functions as
  $x = x(u),\ y = y(u),\ z = z(u) \qquad\qquad \because u_1 \ll u \ll u_2$
- Then **zero order parametric continuity (c^0)** means simply curves meets i.e. last point of first curve section & first points of second curve section are same.
- **First order parametric continuity (c^1)** means first parametric derivatives are same for both curve section at intersection points.
- **Second order parametric continuity (c^2)** means both the first & second parametric derivative of two curve section are same at intersection.
- Higher order parametric continuity is can be obtain similarly.

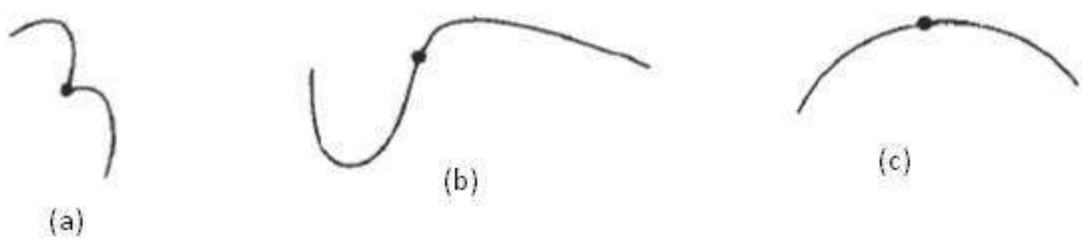


Fig. 4.13: - Piecewise construction of a curve by joining two curve segments uses different orders of continuity: (a) zero-order continuity only, (b) first-order continuity, and (c) second-order continuity.

- First order continuity is often sufficient for general application but some graphics package like cad requires second order continuity for accuracy.

Geometric continuity condition

- Another method for joining two successive curve sections is to specify condition for geometric continuity.
- **Zero order geometric continuity (g^0)** is same as parametric zero order continuity that two curve section meets.
- **First order geometric continuity (g^1)** means that the parametric first derivatives are proportional at the intersection of two successive sections but does not necessary Its magnitude will be equal.
- **Second order geometric continuity (g^2)** means that the both parametric first & second derivatives are proportional at the intersection of two successive sections but does not necessarily magnitude will be equal.

Cubic Spline Interpolation Methods

- Cubic splines are mostly used for representing path of moving object or existing object shape or drawing.
- Sometimes it also used for design the object shapes.
- Cubic spline gives reasonable computation on as compared to higher order spline and more stable compare to lower order polynomial spline. So it is often used for modeling curve shape.
- Cubic interpolation splines obtained by fitting the input points with piecewise cubic polynomial curve that passes through every control point.

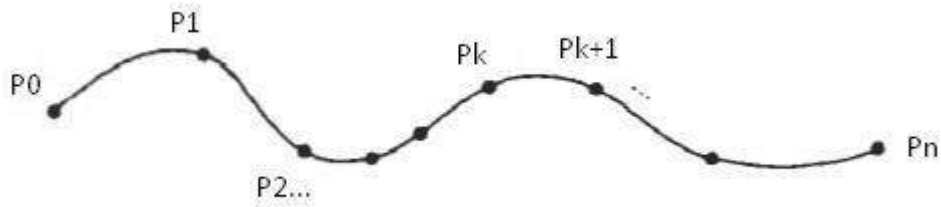


Fig. 4.14: -A piecewise continuous cubic-spline interpolation of n+1 control points.

$p_k = (x_k, y_k, z_k)$ Where, k=0, 1, 2, 3 ..., n

- Parametric cubic polynomial for this curve is given by

$x(u) = a_x u^3 + b_x u^2 + c_x u + d_x$

$y(u) = a_y u^3 + b_y u^2 + c_y u + d_y$

$z(u) = a_z u^3 + b_z u^2 + c_z u + d_z$

$where(\ 0 \leq u \leq 1)$

- For above equation we need to determine for constant a, b, c and d the polynomial representation for each of n curve section.
- This is obtained by settling proper boundary condition at the joints.
- Now we will see common method for settling this condition.

Natural Cubic Splines

- Natural cubic spline is a mathematical representation of the original drafting spline.
- We consider that curve is in c^2 continuity means first and second parametric derivatives of adjacent curve section are same at control point.
- For the ''n+1'' control point we have n curve section and 4n polynomial constants to find.
- For all interior control points we have four boundary conditions. The two curve section on either side of control point must have same first & second order derivative at the control points and each curve passes through that control points.
- We get other two condition as p_0 (first control points) starting & p_n(last control point) is end point of the curve.
- We still required two conditions for obtaining coefficient values.
- One approach is to setup second derivative at p_0 & p_n to be 0. Another approach is to add one extra dummy point at each end. I.e. we add p_{-1} & p_{n+1} then all original control points are interior and we get 4n boundary condition.
- Although it is mathematical model it has major disadvantage is with change in the control point entire curve is changed.
- So it is not allowed for local control and we cannot modify part of the curve.

Hermit Interpolation

- It is named after French mathematician Charles hermit

- It is an interpolating piecewise cubic polynomial with specified tangent at each control points.
- It is adjusted locally because each curve section is depends on it's end points only.
- Parametric cubic point function for any curve section is then given by:

  $p(0) = p_k$
  $p(1) = p_{k+1}$
  $p'(0) = dp_k$
  $p''(1) = dp_{k+1}$

  Where dp_k & dp_{k+1} are values of parametric derivatives at point p_k & p_{k+1} respectively.
- Vector equation of cubic spline is:

  $p(u) = au^3 + bu^2 + cu + d$
- Where x component of p is
- $x(u) = a_x u^3 + b_x u^2 + c_x u + d_x$ and similarly y & z components
- Matrix form of above equation is

  $$P(u) = [u^3\ u^2\ u\ 1]\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$
- Now derivatives of p(u) is $p'(u)=3au^2+2bu+c+0$
- Matrix form of p'(u) is

  $$P'(u) = [3u^2\ 2u\ 1\ 0]\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$
- Now substitute end point value of u as 0 & 1 in above equation & combine all four parametric equations in matrix form:

  $$\begin{bmatrix} p_k \\ p_{k+1} \\ dp_k \\ dp_{k+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 3 & 2 & 1 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$
- Now solving it for polynomial co efficient

  $$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} p_k \\ p_{k+1} \\ dp_k \\ dp_{k+1} \end{bmatrix}$$

  $$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = M_H \begin{bmatrix} p_k \\ p_{k+1} \\ dp_k \\ dp_{k+1} \end{bmatrix}$$
- Now Put value of above equation in equation of $p(u)$

  $$p(u) = [u^3\ u^2\ u\ 1]\begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} p_k \\ p_{k+1} \\ dp_k \\ dp_{k+1} \end{bmatrix}$$

  $$p(u) = [2u^3 - 3u^2 + 1 \quad -2u^3 + 3u^2 \quad u^3 - 2u^2 + u \quad u^3 - u^2]\begin{bmatrix} p_k \\ p_{k+1} \\ dp_k \\ dp_{k+1} \end{bmatrix}$$

  $p(u) = p_k(2u^3 - 3u^2 + 1) + p_{k+1}(-2u^3 + 3u^2) + dp_k(u^3 - 2u^2 + u) + dp_{k+1}(u^3 - u^2)$
  $p(u) = p_k H_0(u) + p_{k+1} H_1(u) + dp_k H_2(u) + dp_{k+1} H_3(u)$

  Where $H_k(u)$ for k=0 , 1 , 2 , 3 are referred to as blending functions because that blend the boundary constraint values for curve section.

- Shape of the four hermit blending function is given below.

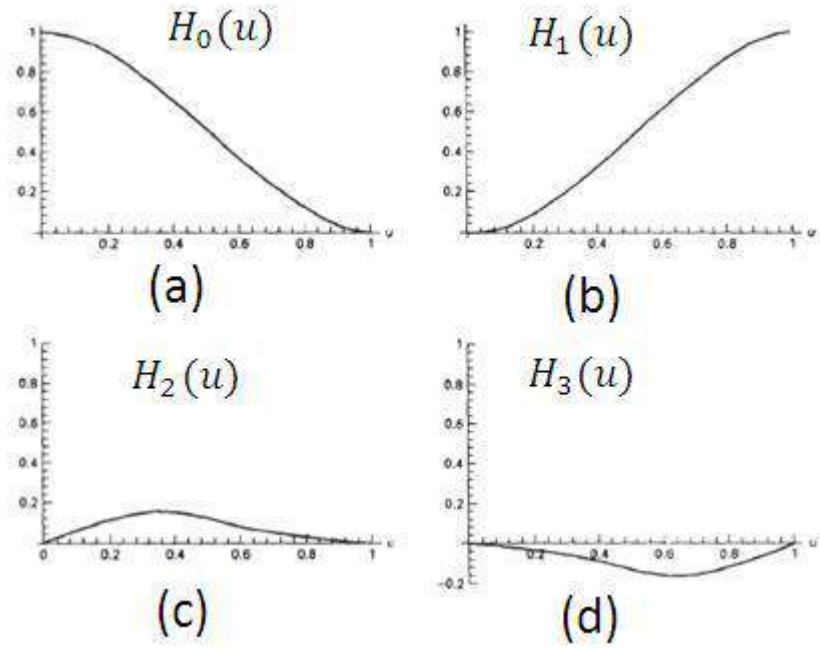


Fig. 4.15: -the hermit blending functions.

- Hermit curves are used in digitizing application where we input the approximate curve slope means DP_k & DP_{k+1}.
- But in application where this input is difficult to approximate at that place we cannot use hermit curve.

Cardinal Splines

- As like hermit spline cardinal splines also interpolating piecewise cubics with specified endpoint tangents at the boundary of each section.
- But in this spline we need not have to input the values of endpoint tangents.
- In cardinal spline values of slope at control point is calculated from two immediate neighbor control points.
- It's spline section is completely specified by the 4-control points.

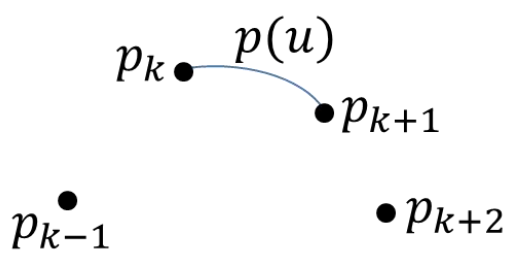


Fig. 4.16: -parametric point function p(u) for a cardinal spline section between control points p_k and p_{k+1}.

- The middle two are two endpoints of curve section and other two are used to calculate slope of endpoints.
- Now parametric equation for cardinal spline is:

$$p(0) = p_k$$
$$p(1) = p_{k+1}$$
$$p'(0) = \frac{1}{2}(1-t)(p_{k+1} - p_{k-1})$$
$$p'(1) = \frac{1}{2}(1-t)(p_{k+2} - p_k)$$

Where parameter t is called **tension** parameter since it controls how loosely or tightly the cardinal spline fit the control points.

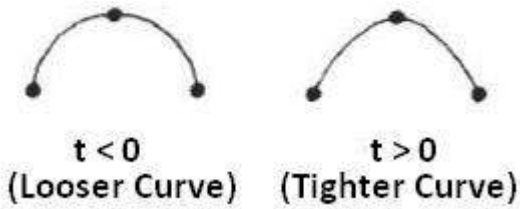


Fig. 4.17: -Effect of the tension parameter on the shape of a cardinal spline section.

- When t = 0 this class of curve is referred to as **catmull-rom spline** or **overhauser splines.**
- Using similar method like hermit we can obtain:

$$p(u) = [u^3 \quad u^2 \quad u \quad 1] \cdot M_c \cdot \begin{bmatrix} p_{k-1} \\ p_k \\ p_{k+1} \\ p_{k+2} \end{bmatrix}$$

- Where the cardinal matrix is

$$M_c = \begin{bmatrix} -s & 2-s & s-2 & s \\ 2s & s-3 & 3-2s & -s \\ -s & 0 & s & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

- With $s = (1-t)/2$
- Put value of M_c in equation of p(u)

$$p(u) = [u^3 \quad u^2 \quad u \quad 1] \cdot \begin{bmatrix} -s & 2-s & s-2 & s \\ 2s & s-3 & 3-2s & -s \\ -s & 0 & s & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} p_{k-1} \\ p_k \\ p_{k+1} \\ p_{k+2} \end{bmatrix}$$

$$p(u) = [-su^3 + 2su^2 - su \quad (2-s)u^3 + (s-3)u^2 + 1 \quad (s-2)u^3 + (3-s)u^2 + su \quad su^3 - su^2] \cdot \begin{bmatrix} p_{k-1} \\ p_k \\ p_{k+1} \\ p_{k+2} \end{bmatrix}$$

$$p(u) = p_{k-1}(-su^3 + 2su^2 - su) + p_k\big((2-s)u^3 + (s-3)u^2 + 1\big) + p_{k+1}\big((s-2)u^3 + (3-s)u^2 + su\big) + p_{k+2}(su^3 - su^2)$$

$$p(u) = p_{k-1}CAR_0(u) + p_kCAR_1(u) + p_{k+1}CAR_2(u) + p_{k+2}CAR_3(u)$$

Where polynomial $CAR_k(u)\ for\ k = 0,1,2,3$ are the cardinals blending functions.

- Figure below shows this blending function shape for t = 0.

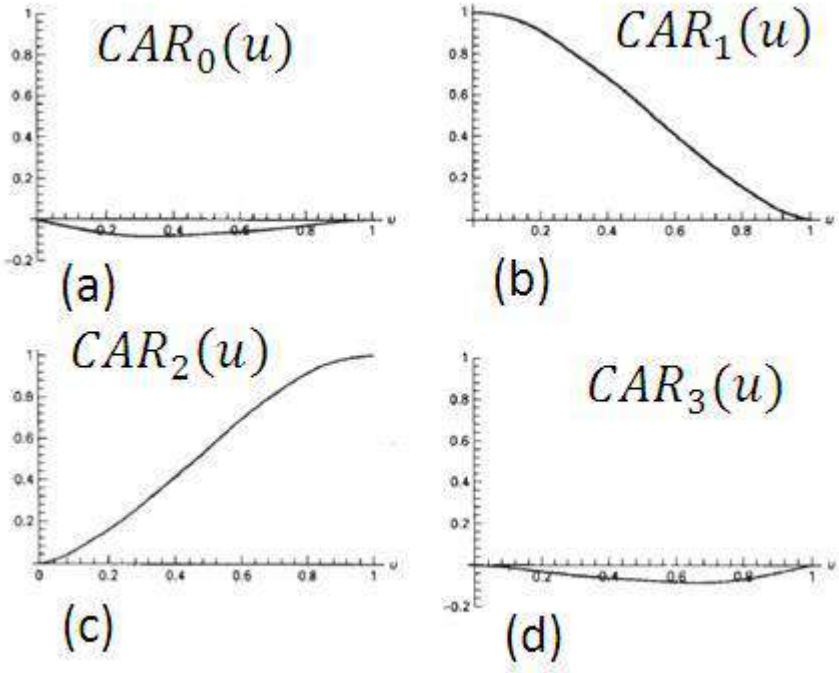


Fig. 4.18: -The cardinal blending function for t=0 and s=0.5.

Kochanek-Bartels spline

- It is extension of cardinal spline
- Two additional parameters are introduced into the constraint equation for defining kochanek-Bartels spline to provide more flexibility in adjusting the shape of curve section.
- For this parametric equations are as follows:

  $p(0) = p_k$

  $p(1) = p_{k+1}$

  $p'(0) = \frac{1}{2}(1-t)[(1+b)(1-c)(p_k - p_{k-1}) + (1-b)(1+c)(p_{k+1} - p_k)]$

  $p'(1) = \frac{1}{2}(1-t)[(1+b)(1+c)(p_{k+1} - p_k) + (1-b)(1-c)(p_{k+2} - p_{k+1})]$

  Where 't' is tension parameter same as used in cardinal spline.
- B is **bias parameter** and C is the **continuity parameter.**
- In this spline parametric derivatives may not be continuous across section boundaries.
- Bias B is used to adjust the amount that the curve bends at each end of section.

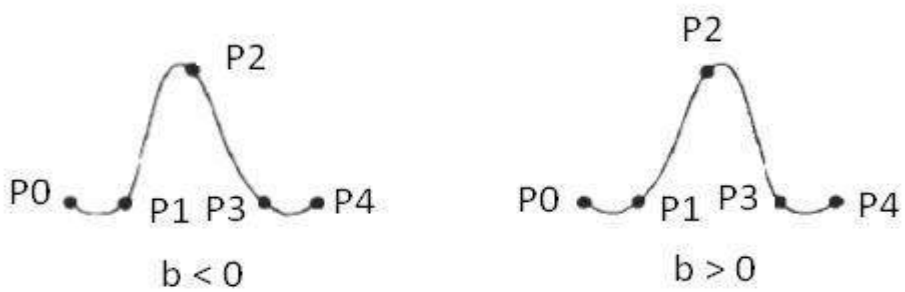


Fig. 4.19: -Effect of bias parameter on the shape of a Kochanek-Bartels spline section.

- Parameter c is used to controls continuity of the tangent vectors across the boundaries of section. If C is nonzero there is discontinuity in the slope of the curve across section boundaries.

- It is used in animation paths in particular abrupt change in motion which is simulated with nonzero values for parameter C.

Bezier Curves and Surfaces

- It is developed by French engineer Pierre Bezier for the Renault automobile bodies.
- It has number of properties and easy to implement so it is widely available in various CAD and graphics package.

Bezier Curves

- Bezier curve section can be fitted to any number of control points.
- Number of control points and their relative position gives degree of the Bezier polynomials.
- With the interpolation spline Bezier curve can be specified with boundary condition or blending function.
- Most convenient method is to specify Bezier curve with blending function.
- Consider we are given n+1 control point position from p_0 to p_n where $p_k = (x_k, y_k, z_k)$.
- This is blended to gives position vector p(u) which gives path of the approximate Bezier curve is:

  $$p(u) = \sum_{k=0}^{n} p_k BEZ_{k,n}(u) \qquad 0 \le u \le 1$$

  Where $BEZ_{k,n}(u) = C(n,k)u^k(1-u)^{n-k}$

  And $C(n,k) = {}^{n!}/_{k!\,(n-k)!}$
- We can also solve Bezier blending function by recursion as follow:

  $BEZ_{k,n}(u) = (1-u)BEZ_{k,n-1}(u) + uBEZ_{k-1,n-1}(u) \qquad n > k \ge 1$

  Here $BEZ_{k,k}(u) = u^k$ and $BEZ_{0,k}(u) = (1-u)^k$
- Parametric equation from vector equation can be obtain as follows.

  $$x(u) = \sum_{k=0}^{n} x_k BEZ_{k,n}(u)$$

  $$y(u) = \sum_{k=0}^{n} y_k BEZ_{k,n}(u)$$

  $$z(u) = \sum_{k=0}^{n} z_k BEZ_{k,n}(u)$$
- Bezier curve is a polynomial of degree one less than the number of control points.
- Below figure shows some possible curve shapes by selecting various control point.

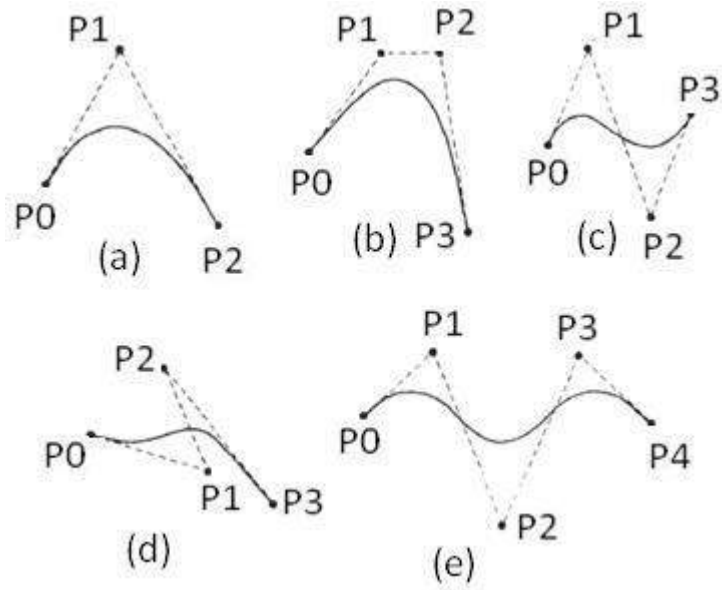


Fig. 4.20: -Example of 2D Bezier curves generated by different number of control points.

- Efficient method for determining coordinate positions along a Bezier curve can be set up using recursive calculation
- For example successive binomial coefficients can be calculated as

$$C(n,k) = \frac{n-k+1}{k} C(n,k-1) \qquad n \geq k$$

Properties of Bezier curves

- It always passes through first control point i.e. p(0) = p_0
- It always passes through last control point i.e. p(1) = p_n
- Parametric first order derivatives of a Bezier curve at the endpoints can be obtain from control point coordinates as:

  $p'(0) = -np_0 + np_1$

  $p'(1) = -np_{n-1} + np_n$
- Parametric second order derivatives of endpoints are also obtained by control point coordinates as:

  $p''(0) = n(n-1)[(p_2 - p_1) - (p_1 - p_0)]$

  $p''(1) = n(n-1)[(p_{n-2} - p_{n-1}) - (p_{n-1} - p_n)]$
- Bezier curve always lies within the convex hull of the control points.
- Bezier blending function is always positive.
- Sum of all Bezier blending function is always 1.

$$\sum_{k=0}^{n} BEZ_{k,n}(u) = 1$$

- So any curve position is simply the weighted sum of the control point positions.
- Bezier curve smoothly follows the control points without erratic oscillations.

Design Technique Using Bezier Curves

- For obtaining closed Bezier curve we specify first and last control point at same position.

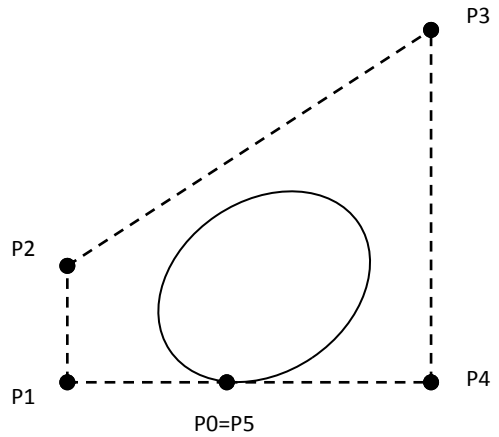


Fig. 4.21: -A closed Bezier Curve generated by specifying the first and last control points at the same location.

- If we specify multiple control point at same position it will get more weight and curve is pull towards that position.

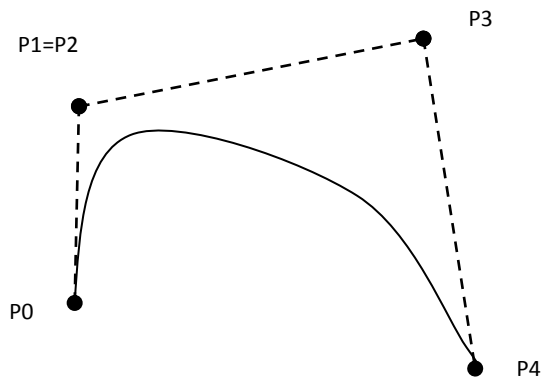


Fig. 4.22: -A Bezier curve can be made to pass closer to a given coordinate position by assigning multiple control point at that position.

- Bezier curve can be fitted for any number of control points but it requires higher order polynomial calculation.
- Complicated Bezier curve can be generated by dividing whole curve into several lower order polynomial curves. So we can get better control over the shape of small region.
- Since Bezier curve passes through first and last control point it is easy to join two curve sections with zero order parametric continuity (C^0).
- For first order continuity we put end point of first curve and start point of second curve at same position and last two points of first curve and first two point of second curve is collinear. And second control point of second curve is at position
  $p_n + (p_n - p_{n-1})$
- So that control points are equally spaced.

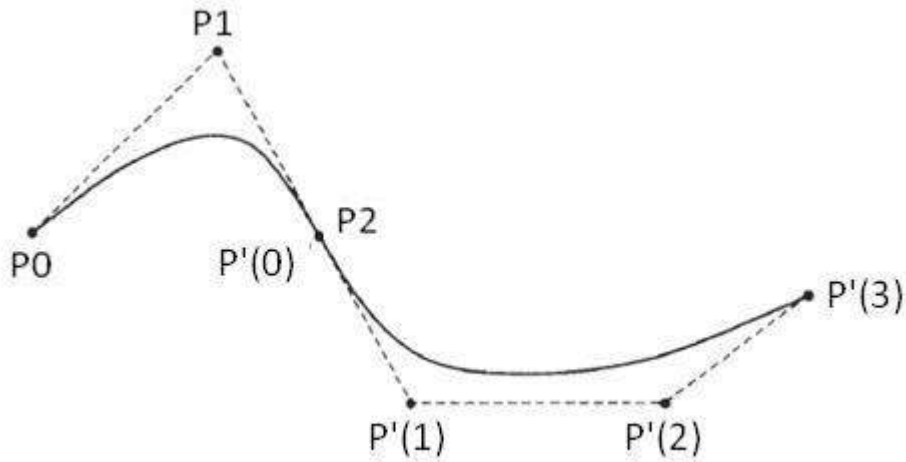


Fig. 4.23: -Zero and first order continuous curve by putting control point at proper place.

- Similarly for second order continuity the third control point of second curve in terms of position of the last three control points of first curve section as
  $p_{n-2} + 4(p_n - p_{n-1})$
- C^2 continuity can be unnecessary restrictive especially for cubic curve we left only one control point for adjust the shape of the curve.

Cubic Bezier Curves

- Many graphics package provides only cubic spline function because this gives reasonable design flexibility in average calculation.
- Cubic Bezier curves are generated using 4 control points.
- 4 blending function obtained by substituting n=3
  $BEZ_{0,3}(u) = (1-u)^3$
  $BEZ_{1,3}(u) = 3u(1-u)^2$
  $BEZ_{2,3}(u) = 3u^2(1-u)$
  $BEZ_{3,3}(u) = u^3$
- Plots of this Bezier blending function are shown in figure below

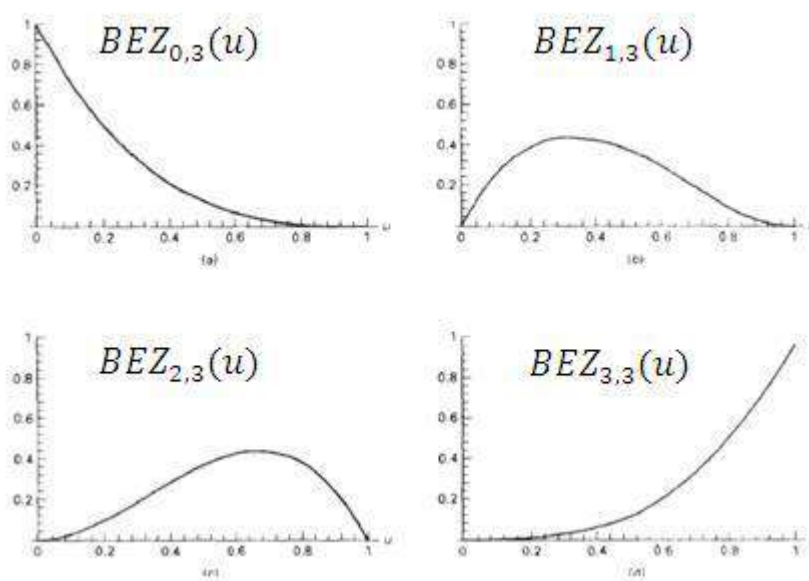


Fig. 4.24: -Four Bezier blending function for cubic curve.

- The form of blending functions determines how control points affect the shape of the curve for values of parameter u over the range from 0 to 1.
  At u = 0 $BEZ_{0,3}(u)$ is only nonzero blending function with values 1.
  At u = 1 $BEZ_{3,3}(u)$ is only nonzero blending function with values 1.
- So the cubic Bezier curve is always pass through p_0 and p_3.
- Other blending function is affecting the shape of the curve in intermediate values of parameter u.
- $BEZ_{1,3}(u)$ is maximum at $u = {}^{1}/_{3}$and $BEZ_{2,3}(u)$ is maximum at $u = {}^{2}/_{3}$
- Blending function is always nonzero over the entire range of u so it is not allowed for local control of the curve shape.
- At end point positions parametric first order derivatives are :
  $p'(0) = 3(p_1 - p_0)$
  $p'(1) = 3(p_3 - p_2)$
- And second order parametric derivatives are.
  $p''(0) = 6(p_0 - 2p_1 + p_2)$
  $p''(1) = 6(p_1 - 2p_2 + p_3)$
- This expression can be used to construct piecewise curve with C^1 and C^2 continuity.
- Now we represent polynomial expression for blending function in matrix form:

$$p(u) = [u^3 \quad u^2 \quad u \quad 1] \cdot M_{BEZ} \cdot \begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{bmatrix}$$

$$M_{BEZ} = \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

- We can add additional parameter like tension and bias as we did with the interpolating spline.

Bezier Surfaces

- Two sets of orthogonal Bezier curves can be used to design an object surface by an input mesh of control points.
- By taking Cartesian product of Bezier blending function we obtain parametric vector function as:

$$p(u,v) = \sum_{j=0}^{m} \sum_{k=0}^{n} p_{j,k} BEZ_{j,m}(v) BEZ_{k,n}(u)$$

- $p_{j,k}$ Specifying the location of the (m+1) by (n+1) control points.
- Figure below shows Bezier surfaces plot, control points are connected by dashed line and curve is represented by solid lines.

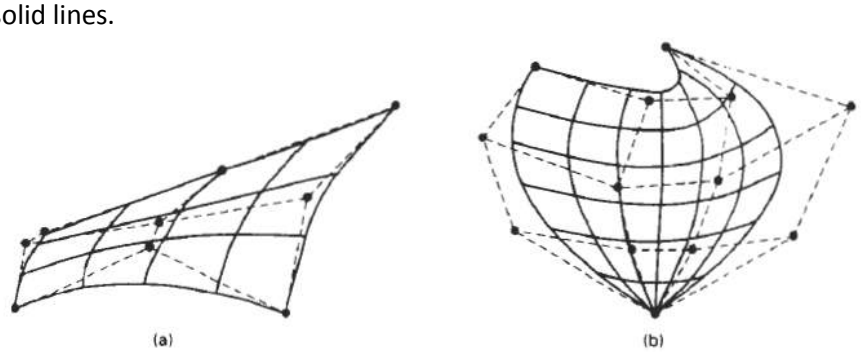


Fig. 4.25: -Bezier surfaces constructed for (a) m=3, n=3, and (b) m=4, n=4. Dashed line connects the control points.

- Each curve of constant u is plotted by varying v over interval 0 to 1. And similarly we can plot for constant v.
- Bezier surfaces have same properties as Bezier curve, so it can be used in interactive design application.
- For each surface patch we first select mesh of control point XY and then select elevation in Z direction.
- We can put two or more surfaces together and form required surfaces using method similar to curve section joining with continuity C^0, C^1, and C^2 as per need.

B-Spline Curves and Surfaces

- B-Spline is most widely used approximation spline.
- It has two advantage over Bezier spline
  1. Degree of a B-Spline polynomial can be set independently of the number of control points (with certain limitation).
  2. B-Spline allows local control.
- Disadvantage of B-Spline curve is more complex then Bezier spline

B-Spline Curves

- General expression for B-Spline curve in terms of blending function is given by:

  $$p(u) = \sum_{k=0}^{n} p_k B_{k,d}(u) \qquad u_{min} \le u \le u_{max}, 2 \le d \le n+1$$

  Where p_k is input set of control points.
- The range of parameter u is now depends on how we choose the B-Spline parameters.
- B-Spline blending function $B_{k,d}$ are polynomials of degree d-1 , where d can be any value in between 2 to n+1.
- We can set d=1 but then curve is only point plot.
- By defining blending function for subintervals of whole range we can achieve local control.
- Blending function of B-Spline is solved by Cox-deBoor recursion formulas as follows.

  $$B_{k,1}(u) = \begin{cases} 1 & if\ u_k \le u \le u_{k+1} \\ 0 & otherwise \end{cases}$$

  $$B_{k,d}(u) = \frac{u - u_k}{u_{k+d-1} - u_k} B_{k,d-1}(u) + \frac{u_{k+d} - u}{u_{k+d} - u_{k+1}} B_{k+1,d-1}(u)$$
- The selected set of subinterval endpoints u_j is reffered to as a **knot vector**.
- We can set any value as a subinterval end point but it must follow $u_j \le u_{j+1}$
- Values of u_{min} and u_{max} depends on number of control points, degree d, and knot vector.
- Figure below shows local control

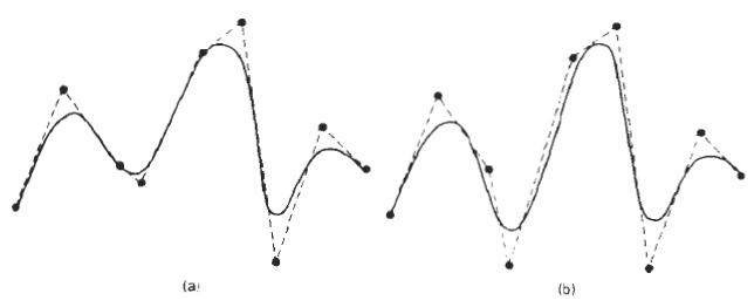


Fig. 4.26: -Local modification of B-Spline curve.

- B-Spline allows adding or removing control points in the curve without changing the degree of curve.
- B-Spline curve lies within the convex hull of at most d+1 control points so that B-Spline is tightly bound to input positions.

- For any u in between u_{d-1} to u_{n+1}, sum of all blending function is 1 i.e. $\sum_{k=0}^{n} B_{k,d}(u) = 1$
- There are three general classification for knot vectors:
  - Uniform
  - Open uniform
  - Non uniform

Properties of B-Spline Curves

- It has degree d-1 and continuity C^{d-2} over range of u.
- For n+1 control point we have n+1 blending function.
- Each blending function $B_{k,d}(u)$ is defined over d subintervals of the total range of u, starting at knot value u_k.
- The range of u is divided into n+d subintervals by the n+d+1 values specified in the knot vector.
- With knot values labeled as $\{u_0, u_1, \ldots, u_{n+d}\}$ the resulting B-Spline curve is defined only in interval from knot values u_{d-1} up to knot values u_{n+1}
- Each spline section is influenced by d control points.
- Any one control point can affect at most d curve section.

Uniform Periodic B-Spline

- When spacing between knot values is constant, the resulting curve is called a uniform B-Spline.
- For example $\{0.0, 0.1, 0.2, \ldots, 1.0\}$ or $\{0,1,2,3,4,5,6,7\}$
- Uniform B-Spline have periodic blending function. So for given values of n and d all blending function has same shape. And each successive blending function is simply a shifted version of previous function.
  $B_{k,d}(u) = B_{k+1,d}(u + \Delta u) = B_{k+2,d}(u + 2\Delta u)$
  Where Δu is interval between adjacent knot vectors.

Cubic Periodic B-Spline

- It commonly used in many graphics packages.
- It is particularly useful for generating closed curve.
- If any three consecutive control points are identical the curve passes through that coordinate position.
- Here for cubic curve d = 4 and n = 3 knot vector spans d+n+1 =4+3+1=8 so it is {0,1,2,3,4,5,6,7}
- Now boundary conditions for cubic B-Spline curve is obtain from equation.

$$p(u) = \sum_{k=0}^{n} p_k B_{k,d}(u) \qquad u_{min} \le u \le u_{max}, 2 \le d \le n+1$$

- That are

$$p(0) = \frac{1}{6}(p_0 + 4p_1 + p_2)$$

$$p(1) = \frac{1}{6}(p_1 + 4p_2 + p_3)$$

$$p'(0) = \frac{1}{2}(p_2 - p_0)$$

$$p'(1) = \frac{1}{2}(p_3 - p_1)$$

- Matrix formulation for a cubic periodic B-Splines with the four control points can then be written as

$$p(u) = [u^3 \quad u^2 \quad u \quad 1] \cdot M_B \cdot \begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{bmatrix}$$

Where

$$M_B = \frac{1}{6}\begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 0 & 3 & 0 \\ 1 & 4 & 1 & 0 \end{bmatrix}$$

- We can also modify the B-Spline equation to include a tension parameter t.
- The periodic cubic B-Spline with tension matrix then has the form:

$$M_{Bt} = \frac{1}{6}\begin{bmatrix} -t & 12-9t & 9t-12 & t \\ 3t & 12t-18 & 18-15t & 0 \\ -3t & 0 & 3t & 0 \\ t & 6-2t & t & 0 \end{bmatrix}$$

When t = 1 $M_{Bt} = M_B$

- We can obtain cubic B-Spline blending function for parametric range from 0 to 1 by converting matrix representation into polynomial form for t = 1 we have

$$B_{0,3}(u) = \frac{1}{6}(1-u)^3$$

$$B_{1,3}(u) = \frac{1}{6}(3u^3 - 6u^2 + 4)$$

$$B_{2,3}(u) = \frac{1}{6}(-3u^3 + 3u^2 + 3u + 1)$$

$$B_{3,3}(u) = \frac{1}{6}u^3$$

Open Uniform B-Splines

- This class is cross between uniform B-Spline and non uniform B-Splines.
- Sometimes it is treated as a special type of uniform B-Spline, and sometimes as non uniform B-Spline
- For open uniform B-Spline (open B-Spline) the knot spacing is uniform except at the ends where knot values are repeated d times.
- For example {0,0,1,2,3,3} for d=2 and n=3, and {0,0,0,0,1,2,2,2,2} for d=4 and n=4.
- For any values of parameter d and n we can generate an open uniform knot vector with integer values using the calculations as follow:

$$u_j = \begin{cases} 0, & for\ 0 \le j < d \\ j - d + 1 & for\ d \le j \le n \\ n - d + 2 & for\ j > n \end{cases}$$

Where $0 \le j \le n + d$

- Open uniform B-Spline is similar to Bezier spline if we take d=n+1 it will reduce to Bezier spline as all knot values are either 0 or 1.
- For example cubic open uniform B-Spline with d=4 have knot vector {0,0,0,0,1,1,1,1}.
- Open uniform B-Spline curve passes through first and last control points.
- Also slope at each end is parallel to line joining two adjacent control points at that end.
- So geometric condition for matching curve sections are same as for Bezier curves.
- For closed curve we specify first and last control point at the same position.

Non Uniform B-Spline

- For this class of spline we can specify any values and interval for knot vector.
- For example {0,1,2,3,3,4}, and {0,0,1,2,2,3,4}
- It will give more flexible shape of curves. Each blending function have different shape when plots and different intervals.
- By increasing knot multiplicity we produce variation in curve shape and also introduce discontinuities.

- Multiple knot value also reduces continuity by 1 for each repeat of particular value.
- We can solve non uniform B-Spline using similar method as we used in uniform B-Spline.
- For set of n+1 control point we set degree d and knot values.
- Then using the recurrence relations we can obtain blending function or evaluate curve position directly for display of the curve.

B-Spline Surfaces

- B-Spline surface formation is also similar to Bezier splines orthogonal set of curves are used and for connecting two surface we use same method which is used in Bezier surfaces.
- Vector equation of B-Spline surface is given by cartesion product of B-Spline blending functions:

$$p(u,v) = \sum_{k1=0}^{n1} \sum_{k2=0}^{n2} p_{k1,k2} B_{k1,d1}(u) B_{k2,d2}(v)$$

- Where $p_{k1,k2}$ specify control point position.
- It has same properties as B-Spline curve.

3D Translation

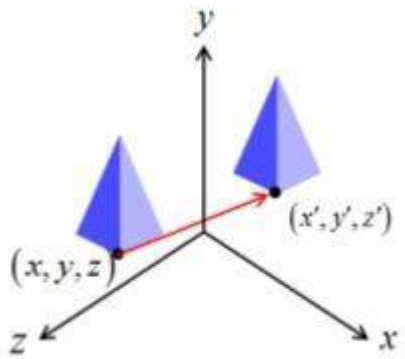


Fig. 5.1: - 3D Translation.

- Similar to 2D translation, which used 3x3 matrices, 3D translation use 4X4 matrices (X, Y, Z, h).
- In 3D translation point (X, Y, Z) is to be translated by amount tx, ty and tz to location (X', Y', Z').

  $\boldsymbol{x' = x + tx}$

  $\boldsymbol{y' = y + ty}$

  $\boldsymbol{z' = z + tz}$
- Let's see matrix equation

  $\boldsymbol{P' = T \cdot P}$

  $$\begin{bmatrix} \boldsymbol{x'} \\ \boldsymbol{y'} \\ \boldsymbol{z'} \\ \boldsymbol{1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{1} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{t_x} \\ \boldsymbol{0} & \boldsymbol{1} & \boldsymbol{0} & \boldsymbol{t_y} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{1} & \boldsymbol{t_z} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{1} \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \\ \boldsymbol{z} \\ \boldsymbol{1} \end{bmatrix}$$
- Example : - Translate the given point P (10,10,10) into 3D space with translation factor T (10,20,5).

  $P' = T \cdot P$

  $$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

  $$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 10 \\ 0 & 1 & 0 & 20 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 10 \\ 10 \\ 10 \\ 1 \end{bmatrix}$$

  $$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 20 \\ 30 \\ 15 \\ 1 \end{bmatrix}$$

  Final coordinate after translation is P' (20, 30, 15).

Rotation

- For 3D rotation we need to pick an axis to rotate about.
- The most common choices are the X-axis, the Y-axis, and the Z-axis

Coordinate-Axes Rotations

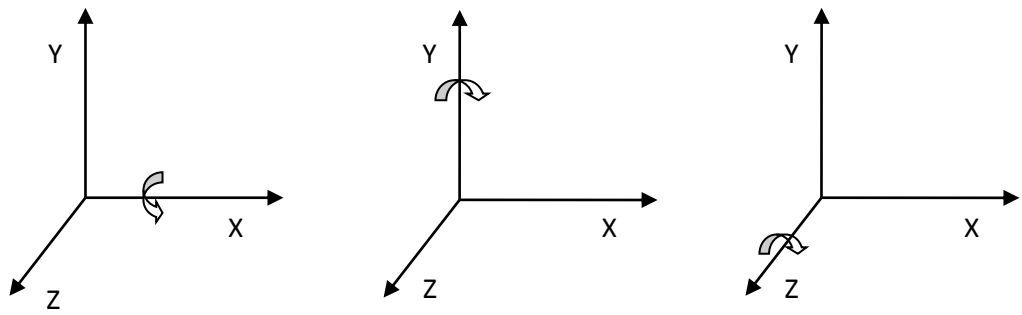


Fig. 5.2: - 3D Rotations.

Z-Axis Rotation

- Two dimension rotation equations can be easily convert into 3D Z-axis rotation equations.
- Rotation about z axis we leave z coordinate unchanged.

  $x' = x\cos\theta - y\sin\theta$

  $y' = x\sin\theta + y\cos\theta$

  $z' = z$

  Where Parameter θ specify rotation angle.
- Matrix equation is written as:

  $P' = R_z(\theta) \cdot P$

  $$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} . \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

X-Axis Rotation

- Transformation equation for x-axis is obtain from equation of z-axis rotation by replacing cyclically as shown here

  $x \rightarrow y \rightarrow z \rightarrow x$
- Rotation about x axis we leave x coordinate unchanged.

  $y' = y\cos\theta - z\sin\theta$

  $z' = y\sin\theta + z\cos\theta$

  $x' = x$

  Where Parameter θ specify rotation angle.
- Matrix equation is written as:

  $P' = R_x(\theta) \cdot P$

  $$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} . \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Y-Axis Rotation

- Transformation equation for y-axis is obtain from equation of x-axis rotation by replacing cyclically as shown here

  $x \rightarrow y \rightarrow z \rightarrow x$

- Rotation about y axis we leave y coordinate unchanged.

  $\boldsymbol{z' = z\cos\theta - x\sin\theta}$

  $\boldsymbol{x' = z\sin\theta + x\cos\theta}$

  $\boldsymbol{y' = y}$

  Where Parameter $\boldsymbol{\theta}$ specify rotation angle.
- Matrix equation is written as:

  $\boldsymbol{P' = R_y(\theta)\cdot P}$

  $$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$
- Example: - Rotate the point P(5,5,5) 90° about Z axis.

  $P' = R_z(\theta)\cdot P$

  $$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

  $$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos 90 & -\sin 90 & 0 & 0 \\ \sin 90 & \cos 90 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

  $$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 5 \\ 5 \\ 1 \end{bmatrix}$$

  $$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} -5 \\ 5 \\ 5 \\ 1 \end{bmatrix}$$

  Final coordinate after rotation is P' (-5, 5, 5).

General 3D Rotations when rotation axis is parallel to one of the standard axis

- Three steps require to complete such rotation
  1. Translate the object so that the rotation axis coincides with the parallel coordinate axis.
  2. Perform the specified rotation about that axis.
  3. Translate the object so that the rotation axis is moved back to its original position.
- This can be represented in equation form as:

  $\boldsymbol{P' = T^{-1}\cdot R(\theta)\cdot T\cdot P}$

General 3D Rotations when rotation axis is inclined in arbitrary direction

- When object is to be rotated about an axis that is not parallel to one of the coordinate axes, we need rotations to align the axis with a selected coordinate axis and to bring the axis back to its original orientation.
- Five steps require to complete such rotation.
  1. Translate the object so that the rotation axis passes through the coordinate origin.
  2. Rotate the object so that the axis of rotation coincides with one of the coordinate axes.
  3. Perform the specified rotation about that coordinate axis.
  4. Apply inverse rotations to bring the rotation axis back to its original orientation.
  5. Apply the inverse translation to bring the rotation axis back to its original position.
- We can transform rotation axis onto any of the three coordinate axes. The Z-axis is a reasonable choice.

- We are given line in the form of two end points P1 (x1,y1,z1), and P2 (x2,y2,z2).
- We will see procedure step by step.

1) **Translate the object so that the rotation axis passes through the coordinate origin.**

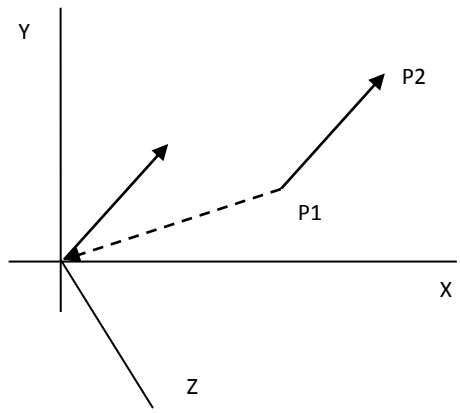


Fig. 5.3: - Translation of vector V.

- For translation of step one we will bring first end point at origin and transformation matrix for the same is as below

$$\boldsymbol{T} = \begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & -\boldsymbol{x_1} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & -\boldsymbol{y_1} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & -\boldsymbol{z_1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}$$

2) **Rotate the object so that the axis of rotation coincides with one of the coordinate axes.**

- This task can be completed by two rotations first rotation about x-axis and second rotation about y-axis.
- But here we do not know rotation angle so we will use dot product and vector product.
- Lets write rotation axis in vector form.

  $\boldsymbol{V = P_2 - P_1 = (x_2 - x_1, y_2 - y_1, z_2 - z_1)}$
- Unit vector along rotation axis is obtained by dividing vector by its magnitude.

$$\boldsymbol{u} = \frac{\boldsymbol{V}}{|\boldsymbol{V}|} = \left(\frac{\boldsymbol{x_2 - x_1}}{|\boldsymbol{V}|}, \frac{\boldsymbol{y_2 - y_1}}{|\boldsymbol{V}|}, \frac{\boldsymbol{z_2 - z_1}}{|\boldsymbol{V}|}\right) = (\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$$

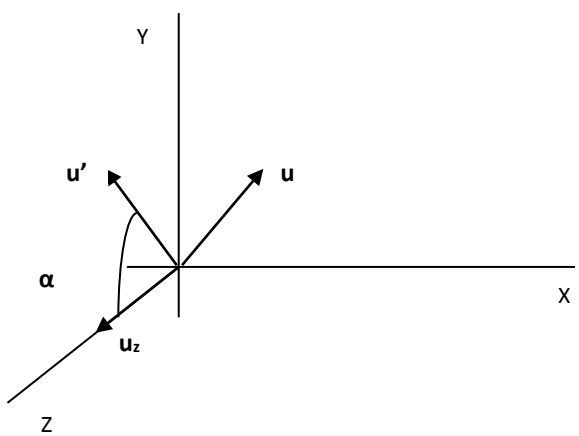


Fig. 5.4: - Projection of u on YZ-Plane.

- Now we need cosine and sin value of angle between unit vector '**u**' and XZ plane and for that we will take projection of u on YZ-plane say '**u''**' and then find dot product and cross product of '**u''**' and '$\mathbf{u_z}$' .
- Coordinate of '**u''**' is (0,b,c) as we will take projection on YZ-plane x value is zero.

  $\boldsymbol{u' \cdot u_z = |u'||u_z| \cos \alpha}$

$$\cos\alpha = \frac{u' \cdot u_z}{|u'||u_z|} = \frac{(0,b,c)(0,0,1)}{\sqrt{b^2+c^2}} = \frac{c}{d} \quad where\ d = \sqrt{b^2+c^2}$$

And

$$u' \times u_z = u_x|u'||u_z| \sin\alpha = u_x \cdot b$$

$$u_x|u'||u_z| \sin\alpha = u_x \cdot b$$

Comparing magnitude

$$|u'||u_z| \sin\alpha = b$$

$$\sqrt{b^2+c^2} \cdot (1) \sin\alpha = b$$

$$d \sin\alpha = b$$

$$\sin\alpha = \frac{b}{d}$$

- Now we have $\sin\alpha$ and $\cos\alpha$ so we will write matrix for rotation about X-axis.

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha & 0 \\ 0 & \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{c}{d} & -\frac{b}{d} & 0 \\ 0 & \frac{b}{d} & \frac{c}{d} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- After performing above rotation '**u**' will rotated into '**u'''**' in XZ-plane with coordinates (a, 0, √(b^2+c^2)). As we know rotation about x axis will leave x coordinate unchanged, '**u'''**' is in XZ=plane so y coordinate is zero, and z component is same as magnitude of '**u''**'.
- Now rotate '**u'''**' about Y-axis so that it coincides with Z-axis.

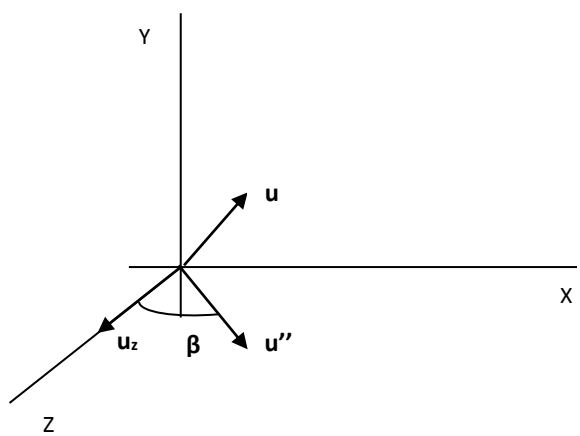


Fig. 5.5: - Rotation of u about X-axis.

- For that we repeat above procedure between '**u'''**' and '$\mathbf{u_z}$' to find matrix for rotation about Y-axis.

$$u'' \cdot u_z = |u''||u_z| \cos\beta$$

$$\cos\beta = \frac{u' \cdot u_z}{|u'||u_z|} = \frac{\left(a,0,\sqrt{b^2+c^2}\right)(0,0,1)}{1} = \sqrt{b^2+c^2} = d \quad where\ d = \sqrt{b^2+c^2}$$

And

$$u'' \times u_z = u_y|u''||u_z| \sin\beta = u_y \cdot (-a)$$

$$u_y|u''||u_z| \sin\beta = u_y \cdot (-a)$$

Comparing magnitude

$$|u''||u_z| \sin\beta = (-a)$$

$(1)\ \sin\beta = -a$

$\sin\beta = -a$

- Now we have $\sin\beta$ and $\cos\beta$ so we will write matrix for rotation about Y-axis.

$$R_y(\beta) = \begin{bmatrix} \cos\beta & 0 & \sin\beta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$R_y(\beta) = \begin{bmatrix} d & 0 & -a & 0 \\ 0 & 1 & 0 & 0 \\ a & 0 & d & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Now by combining both rotation we can coincides rotation axis with Z-axis

3) **Perform the specified rotation about that coordinate axis.**

- As we know we align rotation axis with Z axis so now matrix for rotation about z axis

$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

4) **Apply inverse rotations to bring the rotation axis back to its original orientation.**

- This step is inverse of step number 2.

5) **Apply the inverse translation to bring the rotation axis back to its original position.**

6) This step is inverse of step number 1.

So finally sequence of transformation for general 3D rotation is

$$P' = T^{-1} \cdot {R_x}^{-1}(\alpha) \cdot {R_y}^{-1}(\beta) \cdot R_z(\theta) \cdot R_y(\beta) \cdot R_x(\alpha) \cdot T \cdot P$$

Scaling

- It is used to resize the object in 3D space.
- We can apply uniform as well as non uniform scaling by selecting proper scaling factor.
- Scaling in 3D is similar to scaling in 2D. Only one extra coordinate need to consider into it.

Coordinate Axes Scaling

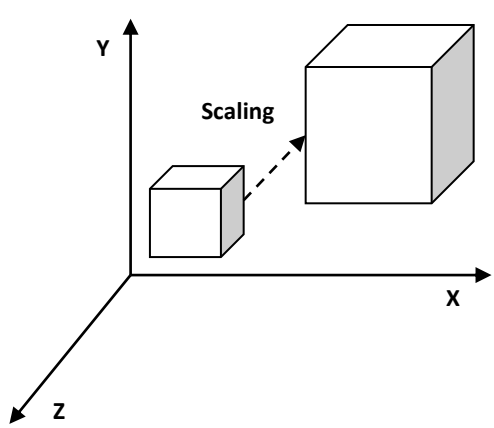


Fig. 5.6: - 3D Scaling.

- Simple coordinate axis scaling can be performed as below.

$P' = S \cdot P$

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

- Example: - Scale the line AB with coordinates (10,20,10) and (20,30,30) respectively with scale factor S(3,2,4).

  $P' = S \cdot P$

  $$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

  $$\begin{bmatrix} A_x' & B_x' \\ A_y' & B_y' \\ A_z' & B_z' \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 10 & 20 \\ 20 & 30 \\ 10 & 30 \\ 1 & 1 \end{bmatrix}$$

  $$\begin{bmatrix} A_x' & B_x' \\ A_y' & B_y' \\ A_z' & B_z' \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 30 & 60 \\ 40 & 60 \\ 40 & 120 \\ 1 & 1 \end{bmatrix}$$

  Final coordinates after scaling are A' (30, 40, 40) and B' (60, 60, 120).

Fixed Point Scaling

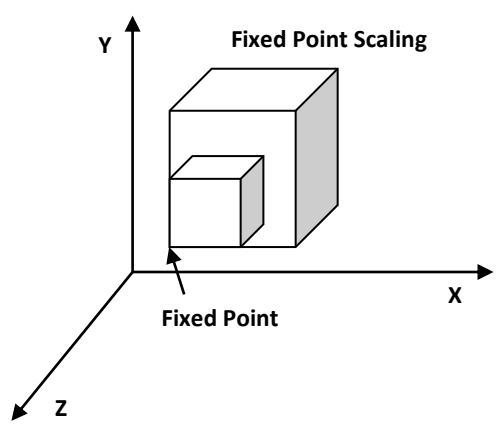


Fig. 5.7: - 3D Fixed point scaling.

- Fixed point scaling is used when we require scaling of object but particular point must be at its original position.
- Fixed point scaling matrix can be obtained in three step procedure.
  1. Translate the fixed point to the origin.
  2. Scale the object relative to the coordinate origin using coordinate axes scaling.
  3. Translate the fixed point back to its original position.
- Let's see its equation.

  $\boldsymbol{P' = T(x_f, y_f, z_f) \cdot S(s_x, s_y, s_z) \cdot T(-x_f, -y_f, -z_f) \cdot P}$

  $$\boldsymbol{P' = \begin{bmatrix} 1 & 0 & 0 & x_f \\ 0 & 1 & 0 & y_f \\ 0 & 0 & 1 & z_f \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & -x_f \\ 0 & 1 & 0 & -y_f \\ 0 & 0 & 1 & -z_f \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot P}$$

$$P' = \begin{bmatrix} s_x & 0 & 0 & (1-s_x)x_f \\ 0 & s_y & 0 & (1-s_y)y_f \\ 0 & 0 & s_z & (1-s_z)z_f \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot P$$

Other Transformations

Reflections

- Reflection means mirror image produced when mirror is placed at require position.
- When mirror is placed in XY-plane we obtain coordinates of image by just changing the sign of z coordinate.
- Transformation matrix for reflection about XY-plane is given below.

$$RF_z = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Similarly Transformation matrix for reflection about YZ-plane is.

$$RF_x = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Similarly Transformation matrix for reflection about XZ-plane is.

$$RF_y = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Shears

- Shearing transformation can be used to modify object shapes.
- They are also useful in 3D viewing for obtaining general projection transformations.
- Here we use shear parameter **'a'** and **'b'**
- Shear matrix for Z-axis is given below

$$SH_z = \begin{bmatrix} 1 & 0 & a & 0 \\ 0 & 1 & b & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Similarly Shear matrix for X-axis is.

$$SH_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ b & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Similarly Shear matrix for X-axis is.

$$SH_y = \begin{bmatrix} 1 & a & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & b & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Viewing Pipeline

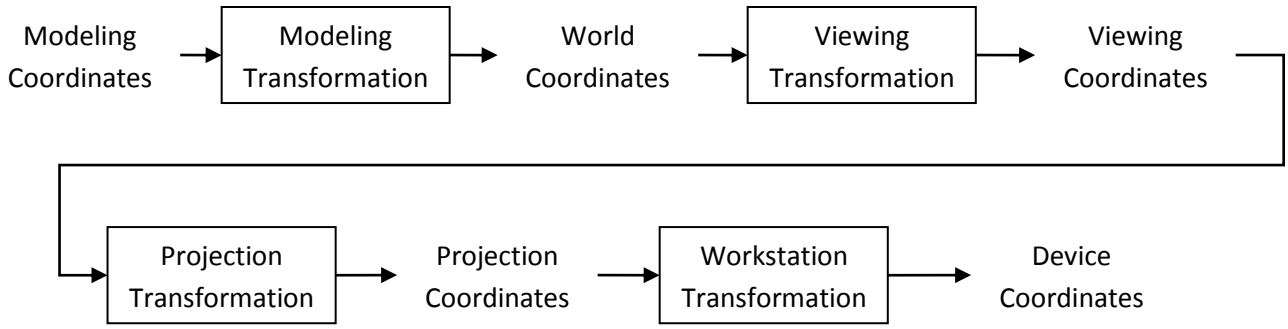


Fig. 5.8: - General 3D viewing pipeline.

- Steps involved in 3D pipeline are similar to the process of taking a photograph.
- As shown in figure that initially we have modeling coordinate of any object which we want to display on the screen.
- By applying modeling transformation we convert modeling coordinates to world coordinates which gives which part or portion is to be display.
- Then by applying viewing transformation we obtain the viewing coordinate which is fitted in viewing coordinate reference frame.
- Then in case of three dimensional objects we have three dimensions coordinate but we need to display that object on two dimensional screens so we apply projection transformation on it which gives projection coordinate.
- Finally projection coordinate is converted into device coordinate by applying workstation transformation which gives coordinates which is specific to particular device.

Viewing Co-ordinates.

- Generating a view of an object is similar to photographing the object.
- We can take photograph from any side with any angle & orientation of camera.
- Similarly we can specify viewing coordinate in ordinary direction.

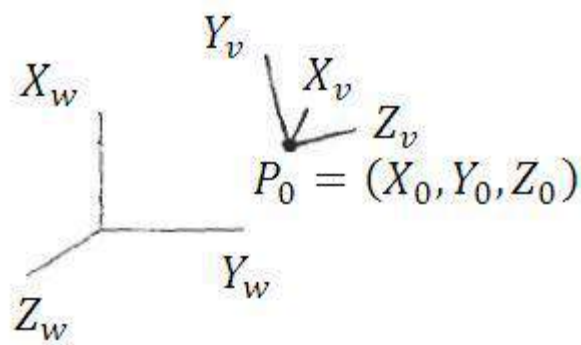


Fig. 5.9: -A right handed viewing coordinate system, with axes Xv, Yv, and Zv, relative to a world-coordinate scene.

Specifying the view plan

- We decide view for a scene by first establishing viewing coordinate system, also referred as view reference coordinate system.
- Then projection plane is setup in perpendicular direction to Zv axis.

- Then projections positions in the scene are transferred to viewing coordinate then viewing coordinate are projected onto the view plane.
- The origin of our viewing coordinate system is called view reference point.
- View reference point is often chosen to be close to or on the surface as same object scene. We can also choose other point also.
- Next we select positive direction for the viewing Zv axis and the orientation of the view plane by specifying the view plane normal vector N.
- Finally we choose the up direction for the view by specifying a vector V called the view up vector. Which specify orientation of camera.
- View up vector is generally selected perpendicular to normal vector but we can select any angle between V & N.
- By fixing view reference point and changing direction of normal vector N we get different views of same object this is illustrated by figure below.

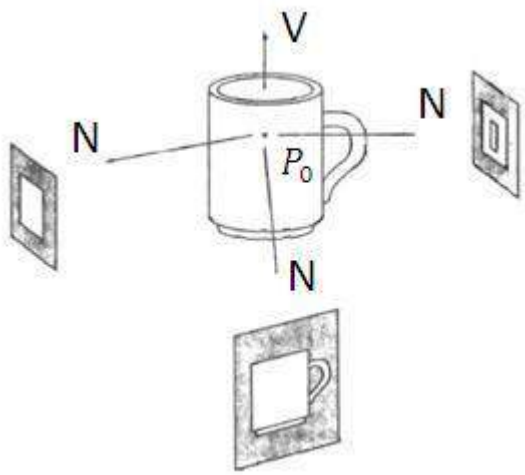


Fig. 5.10: -Viewing scene from different direction with a fixed view-reference point.

Transformation from world to viewing coordinates

- Before taking projection of view plane object description is need to transfer from world to viewing coordinate.
- It is same as transformation that superimposes viewing coordinate system to world coordinate system.
- It requires following basic transformation.
- 1) Translate view reference point to the origin of the world coordinate system.
- 2) Apply rotation to align

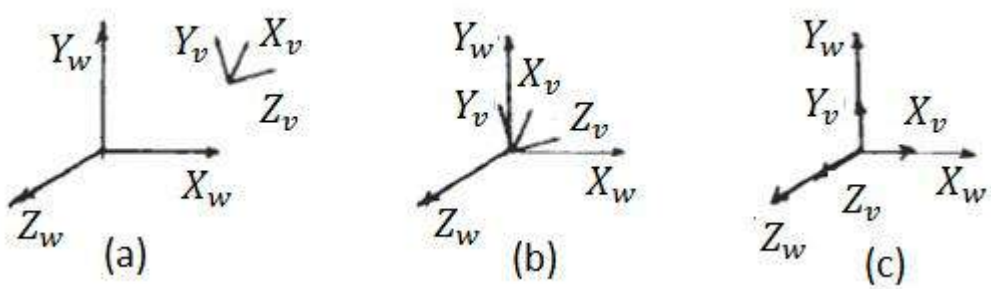


Fig. 5.11: - Aligning a viewing system with the world-coordinate axes using a sequence of translate-rotate transformations.

- As shown in figure the steps of transformation

- Consider view reference point in world coordinate system is at position (x_0, y_0, z_0)than for align view reference point to world origin we perform translation with matrix:
- $T = \begin{bmatrix} 1 & 0 & 0 & -x_0 \\ 0 & 1 & 0 & -y_0 \\ 0 & 0 & 1 & -z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
- Now we require rotation sequence up-to three coordinate axis rotations depending upon direction we choose for N.
- In general case N is at arbitrary direction then we can align it with word coordinate axes by rotation sequence $Rz \cdot Ry \cdot Rx$.
- Another method for generating the rotation transformation matrix is to calculate unit *uvn* vectors and from the composite rotation matrix directly.
- Here

  $n = \frac{N}{|N|} = (n_1, n_2, n_3)$

  $u = \frac{V \times N}{|V \times N|} = (u_1, u_2, u_3)$

  $v = n \times u = (v_1, v_2, v_3)$
- This method also automatically adjusts the direction for *u* so that *v* is perpendicular to *n*.
- Than composite rotation matrix for the viewing transformation is then:

  $R = \begin{bmatrix} u_1 & u_2 & u_3 & 0 \\ v_1 & v_2 & v_3 & 0 \\ n_1 & n_2 & n_3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
- This aligns *u* to Xw axis, *v* to Yw axis and *n* to Zw axis.
- Finally composite matrix for world to viewing coordinate transformation is given by:
- $M_{wc,vc} = R \cdot T$
- This transformation is applied to object's coordinate to transfer them to the viewing reference frame.

Projections

- Once world-coordinate descriptions of the objects in a scene are converted to viewing coordinates, we can project the three-dimensional objects onto the two-dimensional view plane.
- Process of converting three-dimensional coordinates into two-dimensional scene is known as **projection**.
- There are two projection methods namely.
  1. Parallel Projection.
  2. Perspective Projection.
- Lets discuss each one.

Parallel Projections

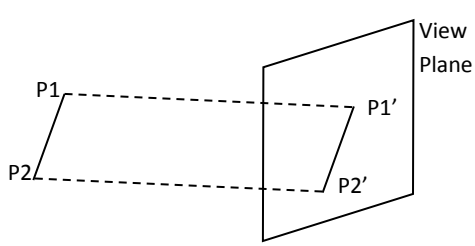


Fig. 5.12: - Parallel projection.

- In a parallel projection, coordinate positions are transformed to the view plane along parallel lines, as shown in the, example of above Figure.
- We can specify a parallel projection with a projection vector that defines the direction for the projection lines.
- It is further divide into two types.
  1. Orthographic parallel projection.
  2. Oblique parallel projection.

Orthographic parallel projection

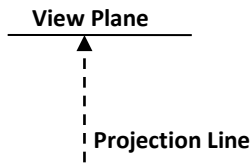


Fig. 5.13: - Orthographic parallel projection.

- When the projection lines are perpendicular to the view plane, we have an orthographic parallel projection.
- Orthographic projections are most often used to produce the front, side, and top views of an object, as shown in Fig.

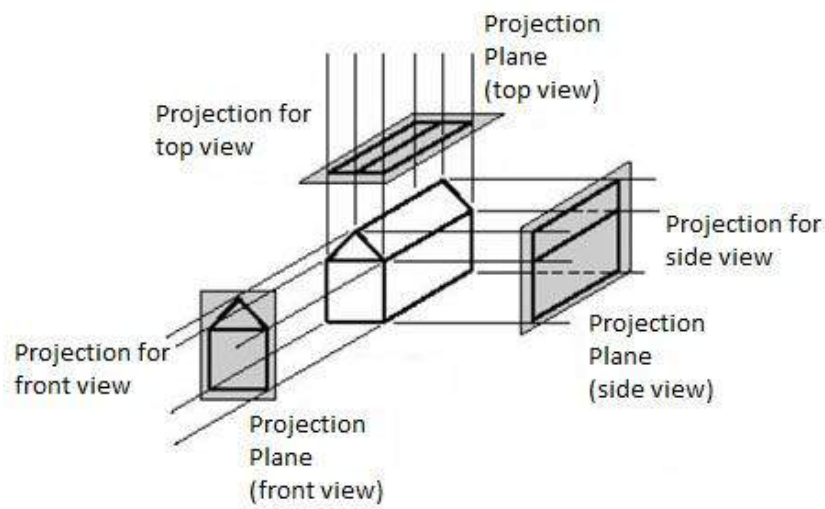


Fig. 5.14: - Orthographic parallel projection.

- Engineering and architectural drawings commonly use orthographic projections, because lengths and angles are accurately depicted and can be measure from the drawings.
- We can also form orthographic projections that display more than one face of an object. Such view are called **axonometric orthographic projections**. Very good example of it is **Isometric** projection.
- Transformation equations for an orthographic parallel projection are straight forward.
- If the view plane is placed at position z_{vp} along the z_v axis, then any point (x, y, z) in viewing coordinates is transformed to projection coordinates as

  $x_p = x, \qquad y_p = y$

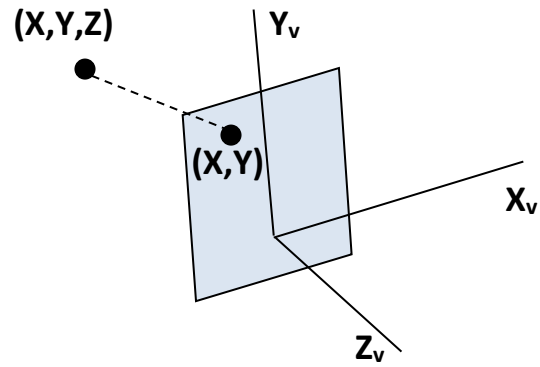


Fig. 5.15: - Orthographic parallel projection.

- Where the original z-coordinate value is preserved for the depth information needed in depth cueing and visible-surface determination procedures.

Oblique parallel projection.

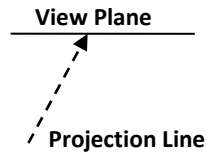


Fig. 5.16: - Oblique parallel projection.

- An oblique projection is obtained by projecting points along parallel lines that are not perpendicular to the projection plane.
- Coordinate of oblique parallel projection can be obtained as below.

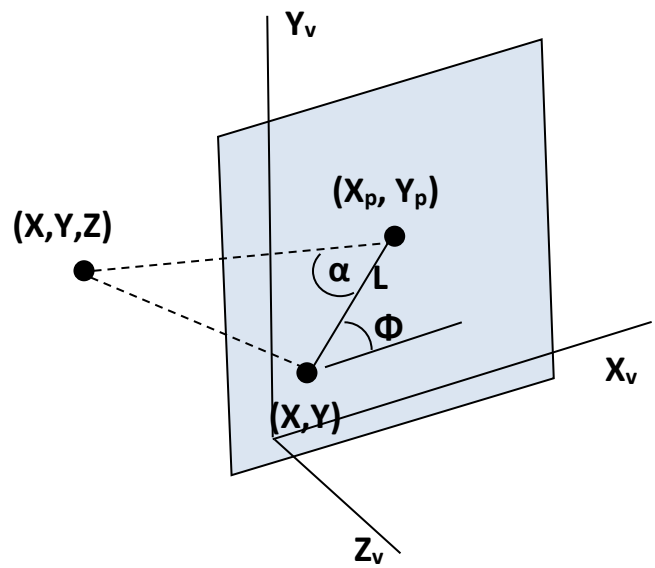


Fig. 5.17: - Oblique parallel projection.

- As shown in the figure (X,Y,Z) is a point of which we are taking oblique projection (Xp,Yp) on the view plane and point (X,Y) on view plane is orthographic projection of (X,Y,Z).
- Now from figure using trigonometric rules we can write

$$x_p = x + L\cos\emptyset$$
$$y_p = y + L\sin\emptyset$$

- Length L depends on the angle α and the z coordinate of the point to be projected:

$$\tan\alpha = \frac{Z}{L}$$
$$L = \frac{Z}{\tan\alpha}$$
$$L = ZL_1, \quad Where\ L_1 = \frac{1}{\tan\alpha}$$

- Now put the value of L in projection equation.

$$x_p = x + ZL_1\cos\emptyset$$
$$y_p = y + ZL_1\sin\emptyset$$

- Now we will write transformation matrix for this equation.

$$M_{parallel} = \begin{bmatrix} 1 & 0 & L_1\cos\emptyset & 0 \\ 0 & 1 & L_1\sin\emptyset & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- This equation can be used for any parallel projection. For orthographic projection L_1=0 and so whole term which is multiply with z component is zero.
- When value of $\tan\alpha = 1$ projection is known as **Cavalier projection**.
- When value of $\tan\alpha = 2$ projection is known as **Cabinet projection**.

Perspective Projection

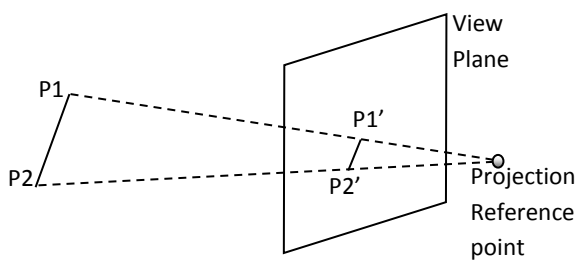


Fig. 5.18: - Perspective projection.

- In perspective projection object positions are transformed to the view plane along lines that converge to a point called the **projection reference point** (or **center of projection** or **vanishing point**).

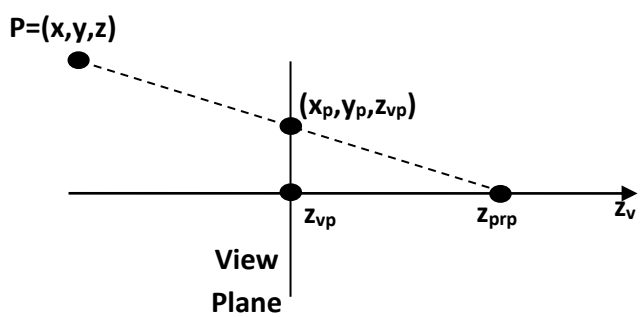


Fig. 5.19: - Perspective projection.

- Suppose we set the projection reference point at position z_{prp} along the z_v axis, and we place the view plane at z_{vp} as shown in Figure above. We can write equations describing coordinate positions along this perspective projection line in parametric form as

$$x' = x - xu$$
$$y' = y - yu$$
$$z' = z - (z - z_{prp})u$$

- Here parameter u takes the value from 0 to 1, which is depends on the position of object, view plane, and projection reference point.
- For obtaining value of u we will put $z'=z_{vp}$ and solve equation of z'.

$$z' = z - (z - z_{prp})u$$
$$z_{vp} = z - (z - z_{prp})u$$
$$u = \frac{z_{vp} - z}{z_{prp} - z}$$

- Now substituting value of u in equation of x' and y' we will obtain.

$$x_p = x\left(\frac{z_{prp} - z_{vp}}{z_{prp} - z}\right) = x\left(\frac{d_p}{z_{prp} - z}\right)$$
$$y_p = y\left(\frac{z_{prp} - z_{vp}}{z_{prp} - z}\right) = y\left(\frac{d_p}{z_{prp} - z}\right), \qquad Where\ d_p = z_{prp} - z_{vp}$$

- Using three dimensional homogeneous-coordinate representations, we can write the perspective projection transformation matrix form as.

$$\begin{bmatrix} x_h \\ y_h \\ z_h \\ h \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -z_{vp}/d_p & z_{vp}(z_{prp}/d_p) \\ 0 & 0 & -1/d_p & z_{prp}/d_p \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

- In this representation, the homogeneous factor is.

$$h = \frac{z_{prp} - z}{d_p}\ and$$
$$x_p = x_h/h \ \ and\ y_p = y_h/h$$

- There are number of special cases for the perspective transformation equations.
- If view plane is taken to be uv plane, then $z_{vp} = 0$ and the projection coordinates are.

$$x_p = x\left(\frac{z_{prp}}{z_{prp} - z}\right) = x\left(\frac{1}{1 - z/z_{prp}}\right)$$
$$y_p = y\left(\frac{z_{prp}}{z_{prp} - z}\right) = y\left(\frac{1}{1 - z/z_{prp}}\right)$$

- If we take projection reference point at origin than $z_{prp} = 0$ and the projection coordinates are.

$$x_p = x\left(\frac{z_{vp}}{z}\right) = x\left(\frac{1}{z/z_{vp}}\right)$$
$$y_p = y\left(\frac{z_{vp}}{z}\right) = y\left(\frac{1}{z/z_{vp}}\right)$$

- The vanishing point for any set of lines that are parallel to one of the principal axes of an object is referred to as a principal vanishing point
- We control the number of principal vanishing points (one, two, or three) with the orientation of the projection plane, and perspective projections are accordingly classified as one-point, two-point, or three-point projections.
- The number of principal vanishing points in a projection is determined by the number of principal axes intersecting the view plane.

View Volumes and General Projection Transformations

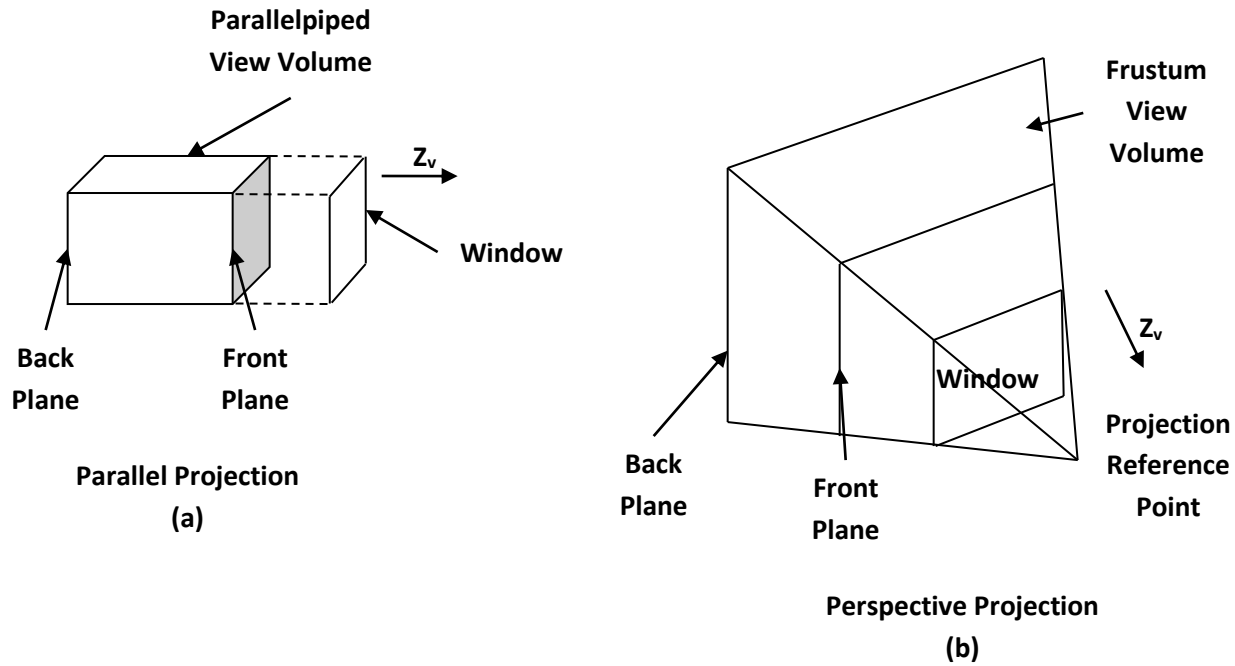


Fig. 5.20: - View volume of parallel and perspective projection.

- Based on view window we can generate different image of the same scene.
- Volume which is appears on the display is known as view volume.
- Given the specification of the view window, we can set up a view volume using the window boundaries.
- Only those objects within the view volume will appear in the generated display on an output device; all others are clipped from the display.
- The size of the view volume depends on the size of the window, while the shape of the view volume depends on the type of projection to be used to generate the display.
- A finite view volume is obtained by limiting the extent of the volume in the z_v direction.
- This is done by specifying positions for one or two additional boundary planes. These z_v-boundary planes are referred to as the **front plane** and **back plane**, or the **near plane** and the **far plane**, of the viewing volume.
- Orthographic parallel projections are not affected by view-plane positioning, because the projection lines are perpendicular to the view plane regardless of its location.
- Oblique projections may be affected by view-plane positioning, depending on how the projection direction is to be specified.

General Parallel-Projection Transformation

- Here we will obtain transformation matrix for parallel projection which is applicable to both orthographic as well as oblique projection.

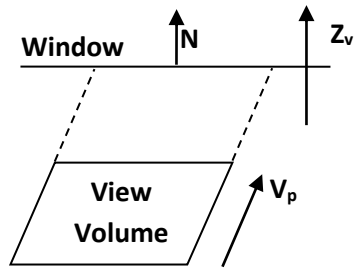


Fig. 5.21: - General parallel projection.

- As shown on figure parallel projection is specified with a projection vector from the projection reference point to the view window.
- Now we will apply shear transformation so that view volume will convert into regular parallelepiped and projection vector will become parallel to normal vector N.

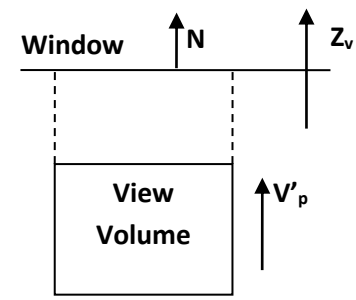


Fig. 5.22: - Shear operation in General parallel projection.

- Let's consider projection vector $\boldsymbol{V_p = (p_x, p_y, p_z)}$.
- We need to determine the elements of a shear matrix that will align the projection vector $\boldsymbol{V_p}$ with the view plane normal vector **N**. This transformation can be expressed as

  $\boldsymbol{V_p{}' = M_{parallel} \cdot V_p}$

  $$\boldsymbol{V_p{}' = \begin{bmatrix} 0 \\ 0 \\ p_z \\ 1 \end{bmatrix}}$$

- where $\boldsymbol{M_{parallel}}$ is equivalent to the parallel projection matrix and represents a z-axis shear of the form

  $$\boldsymbol{M_{parallel} = \begin{bmatrix} 1 & 0 & a & 0 \\ 0 & 1 & b & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}$$

- Now from above equation we can write

  $$\boldsymbol{\begin{bmatrix} 0 \\ 0 \\ p_z \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & a & 0 \\ 0 & 1 & b & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} p_x \\ p_y \\ p_z \\ 1 \end{bmatrix}}$$

- From matrix we can write.

  $\boldsymbol{0 = p_x + ap_z}$

  $\boldsymbol{0 = p_y + bp_z}$

  So

  $$\boldsymbol{a = \frac{-p_x}{p_z}, \quad b = \frac{-p_y}{p_z}}$$

- Thus, we have the general parallel-projection matrix in terms of the elements of the projection vector as

$$M_{parallel} = \begin{bmatrix} 1 & 0 & \frac{-p_x}{p_z} & 0 \\ 0 & 1 & \frac{-p_y}{p_z} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- For an orthographic parallel projection, $p_x = p_y = 0$, and is the identity matrix.

General Perspective-Projection Transformations

- The projection reference point can be located at any position in the viewing system, except on the view plane or between the front and back clipping planes.

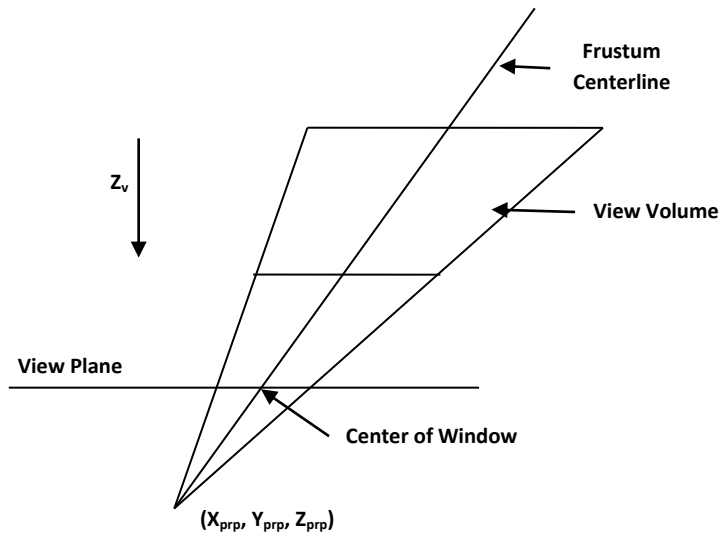


Fig. 5.23: - General perspective projection.

- We can obtain the general perspective-projection transformation with the following two operations:
  1. Shear the view volume so that the center line of the frustum is perpendicular to the view plane.
  2. Scale the view volume with a scaling factor that depends on 1/z .
- A shear operation to align a general perspective view volume with the projection window is shown in Figure.

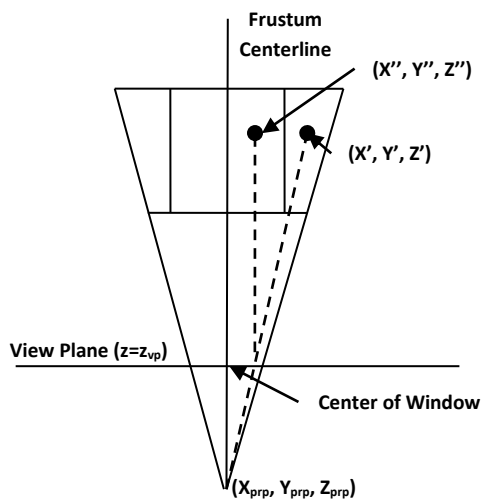


Fig. 5.24: - Shear and scaling operation in general perspective projection.

- With the projection reference point at a general position (X_{prp}, Y_{prp}, Z_{prp}) the transformation involves a combination z-axis shear and a translation:

$$\boldsymbol{M_{shear}} = \begin{bmatrix} \mathbf{1} & \mathbf{0} & \boldsymbol{a} & -\boldsymbol{az_{prp}} \\ \mathbf{0} & \mathbf{1} & \boldsymbol{b} & -\boldsymbol{bz_{prp}} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}$$

Where the shear parameters are

$$\boldsymbol{a} = -\frac{\boldsymbol{x_{prp}} - (\boldsymbol{xw_{min}} + \boldsymbol{xw_{max}})/\mathbf{2}}{\boldsymbol{z_{prp}}}$$

$$\boldsymbol{b} = -\frac{\boldsymbol{y_{prp}} - \frac{\boldsymbol{yw_{min}} + \boldsymbol{yw_{max}}}{\mathbf{2}}}{\boldsymbol{z_{prp}}}$$

- Points within the view volume are transformed by this operation as

$$\boldsymbol{x'} = \boldsymbol{x} + \boldsymbol{a}(\boldsymbol{z} - \boldsymbol{z_{prp}})$$

$$\boldsymbol{y'} = \boldsymbol{y} + \boldsymbol{b}(\boldsymbol{z} - \boldsymbol{z_{prp}})$$

$$\boldsymbol{z'} = \boldsymbol{z}$$

- After shear we apply scaling operation.

$$\boldsymbol{x''} = \boldsymbol{x'}\left(\frac{\boldsymbol{z_{prp}} - \boldsymbol{z_{vp}}}{\boldsymbol{z_{prp}} - \boldsymbol{z}}\right) + \boldsymbol{x_{prp}}\left(\frac{\boldsymbol{z_{vp}} - \boldsymbol{z}}{\boldsymbol{z_{prp}} - \boldsymbol{z}}\right)$$

$$\boldsymbol{y''} = \boldsymbol{y'}\left(\frac{\boldsymbol{z_{prp}} - \boldsymbol{z_{vp}}}{\boldsymbol{z_{prp}} - \boldsymbol{z}}\right) + \boldsymbol{y_{prp}}\left(\frac{\boldsymbol{z_{vp}} - \boldsymbol{z}}{\boldsymbol{z_{prp}} - \boldsymbol{z}}\right)$$

- Homogeneous matrix for this transformation is:

$$\boldsymbol{M_{scale}} = \begin{bmatrix} \mathbf{1} & \mathbf{0} & \frac{-\boldsymbol{x_{prp}}}{\boldsymbol{z_{prp}} - \boldsymbol{z_{vp}}} & \frac{\boldsymbol{x_{prp}}\boldsymbol{z_{vp}}}{\boldsymbol{z_{prp}} - \boldsymbol{z_{vp}}} \\ \mathbf{0} & \mathbf{1} & \frac{-\boldsymbol{y_{prp}}}{\boldsymbol{z_{prp}} - \boldsymbol{z_{vp}}} & \frac{\boldsymbol{y_{prp}}\boldsymbol{z_{vp}}}{\boldsymbol{z_{prp}} - \boldsymbol{z_{vp}}} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{-\mathbf{1}}{\boldsymbol{z_{prp}} - \boldsymbol{z_{vp}}} & \frac{\boldsymbol{z_{prp}}}{\boldsymbol{z_{prp}} - \boldsymbol{z_{vp}}} \end{bmatrix}$$

- Therefore the general perspective-projection transformation is obtained by equation:

$$\boldsymbol{M_{perspective}} = \boldsymbol{M_{scale}} \cdot \boldsymbol{M_{shear}}$$

Classification of Visible-Surface Detection Algorithms

- It is broadly divided into two parts
  - Object-Space methods
  - Image-Space methods
- Object space method compares objects and parts of objects to each other within the scene definition to determine which surface is visible.
- In image space algorithm visibility is decided point by point at each pixel position on the projection plane.

Back-Face Detection

- Back-Face Detection is simple and fast object –space method.
- It identifies back faces of polygon based on the inside-outside tests.
- A point (x, y, z) is inside if Ax + By + Cz + d < 0 where A, B, C, and D are constants and this equation is nothing but equation of polygon surface.
- We can simplify test by taking normal vector N= (A, B, C) of polygon surface and vector V in viewing direction from eye as shown in figure

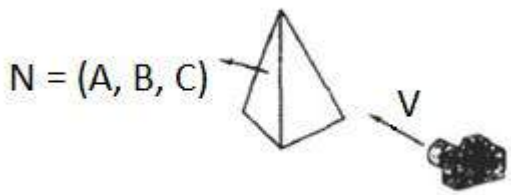


Fig. 6.1:- vector V in the viewing direction and back-face normal vector N of a polyhedron.

- Then we check condition if $V \cdot N > 0$ then polygon is back face.
- If we convert object description in projection coordinates and our viewing direction is parallel to z_v then v= (0,0,v_z) and
  $V \cdot N = V_z C$.
- So now we only need to check sign of C.
- In right handed viewing system V is along negative z_v axis. And in that case
  If C<0 the polygon is backface.
- Also we cannot see any face for which C=0.
- So in general for right handed system
  If $C \leq 0$ polygon is back face.
- Similar method can be used for left handed system.
- In left handed system V is along the positive Z direction and polygon is back face if $C \geq 0$.
- For a single convex polyhedron such as the pyramid by examining parameter C for the different plane we identify back faces.
- So far the scene contains only non overlapping convex polyhedral, back face method works properly.
- For other object such as concave polyhedron as shown in figure below we need to do more tests for determining back face.

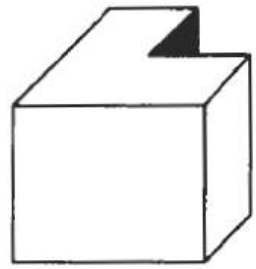


Fig. 6.2:-view of a concave polyhedron with one face partially hidden by other faces.

Depth Buffer Method/ Z Buffer Method

Algorithm

- Initialize the depth buffer and refresh buffer so that for all buffer positions***(x,*** y),
  - $depth(x, y) = 0, \qquad refresh(x, y) = I_{backgnd}$
- For each position on each polygon surface, compare depth values to previously stored values in the depth buffer to determine visibility.
- Calculate the depth z for each *(x,* y) position on the polygon.
- If z > depth(x, y), then set
  - $depth(x, y) = z, \qquad refresh(x, y) = I_{surf}(x, y)$
- Where $I_{backgnd}$ **is** the value for the background intensity, and $I_{surf}(x, y)$ is the projected intensity value for the surface at pixel position (x,y). After all surfaces have been processed, the depth buffer contains depth values for the visible surfaces and the refresh buffer contains the corresponding intensity values for those surfaces.

- It is image space approach.
- It compares surface depth at each pixel position on the projection plane.
- It is also referred to as z-buffer method since generally depth is measured in z-direction.
- Each surface of the scene is process separately one point at a time across the surface.

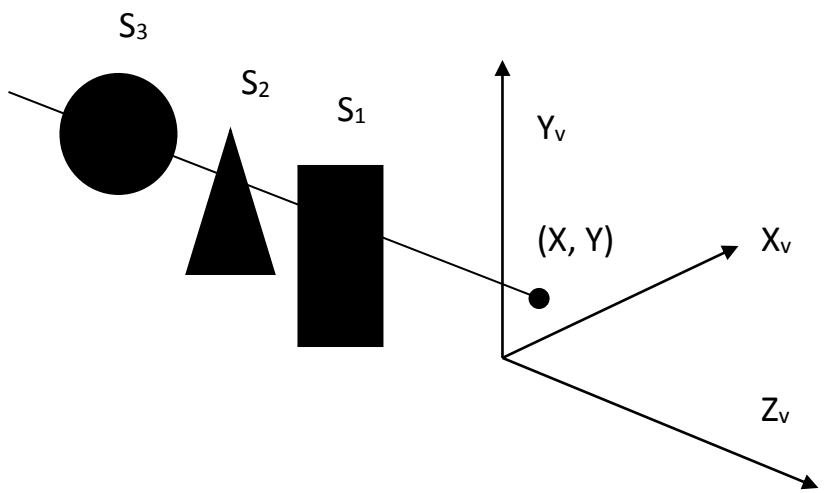


Fig. 6.3:- At view plane position (x, y), surface s_1 has smallest depth from the view plane and so is visible at that position.

- We are starting with pixel position of view plane and for particular surface of object.
- If we take orthographic projection of any point (x,y,z) of the surface on the view plane we get two dimension coordinate (x,y) for that point to display.
- Here we are taking (x.y) position on plan and find particular surface is at how much depth.
- We can implement depth buffer algorithm in normalized coordinates so that z values range from 0 at the back clipping plane to zmax at the front clipping plane.

- Zmax value can be 1 for unit cube or the largest value.
- Here two buffers are required. A depth buffer to store depth value of each (x,y) position and refresh buffer to store corresponding intensity values.
- Initially depth buffer value is 0 and refresh buffer value is intensity of background.
- Each surface of polygon is then process one by one scanline at a time.
- Calculate the z values at each (x,y) pixel position.
- If calculated depth value is greater than the value stored in depth buffer it is replaced with new calculated values and store intensity of that point into refresh buffer at (x,y) position.
- Depth values are calculated from plane equation $Ax + By + Cz + D = 0$ as:

$$z = \frac{-Ax - By - D}{C}$$

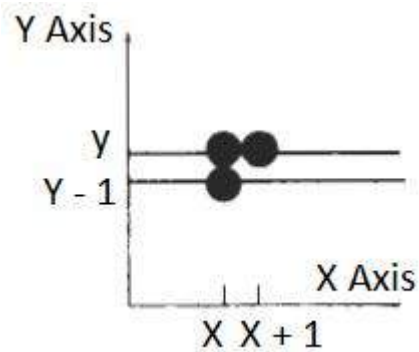


Fig. 6.4:-From position (x,y) on a scan line, the next position across the line has coordinates (x+1,y), and the position immediately below on the next line has coordinates (x,y-1).

- For horizontal line next pixel's z values can be calculated by putting x'=x+1 in above equation.

$$z' = \frac{-A(x+1) - By - D}{C}$$

$$z' = z - \frac{A}{C}$$

- Similarly for vertical line pixel below the current pixel has y'=y-1 so it's z values can be calculated as follows.

$$z' = \frac{-Ax - B(y-1) - D}{C}$$

$$z' = z + \frac{B}{C}$$

- If we are moving along polygon boundary then it will improve performance by eliminating extra calculation.
- For this if we move top to bottom along polygon boundary we get x'=x-1/m and y'=y-1, so z value is obtain as follows.

$$z' = \frac{-A(x - 1/m) - B(y-1) - D}{C}$$

$$z' = z + \frac{A/_m + B}{C}$$

- Alternately we can use midpoint method to find the z values.

Light source

- When we see any object we see reflected light from that object. Total reflected light is the sum of contribution from all sources and reflected light from other object that falls on the object.
- So that the surface which is not directly exposed to light may also visible if nearby object is illuminated.

- The simplest model for light source is **point source.** Rays from the source then follows radial diverging paths from the source position.

Fig. 6.5:- Diverging ray paths from a point light source.

- This light source model is reasonable approximation for source whose size is small compared to the size of object or may be at sufficient distance so that we can see it as point source. For example sun can be taken as point source on earth.
- A nearby source such as the long fluorescent light is more accurately modelled as a **distributed light source.**
- In this case the illumination effects cannot be approximated with point source because the area of the source is not small compare to the size of object.
- When light is falls on the surface the part of the light is reflected and part of the light is absorbed. Amount of reflected and absorbed light is depends on the property of the object surface. For example shiny surface reflect more light while dull surface reflect less light.

Basic Illumination Models/ Shading Model/ Lighting Model

- These models give simple and fast method for calculating the intensities of light for various reflections.

Ambient Light

- This is a simple way to model combination of light reflection from various surfaces to produce a uniform illumination called **ambient light,** or **background light.**
- Ambient light has no directional properties. The amount of ambient light incident on all the surfaces and object are constant in all direction.
- If consider that ambient light of intensity I_a and each surface is illuminate with I_a intensity then resulting reflected light is constant for all the surfaces.

Diffuse Reflection

- When some intensity of light is falls on object surface and that surface reflect light in all the direction in equal amount then the resulting reflection is called **diffuse reflection.**
- Ambient light reflection is approximation of global diffuse lighting effects.
- Diffuse reflections are constant over each surface independent of our viewing direction.
- Amount of reflected light is depend on the parameter K_d, the **diffuse reflection coefficient** or **diffuse reflectivity.**
- K_d is assign value in between 0 and 1 depending on reflecting property. Shiny surface reflect more light so K_d is assign larger value while dull surface assign small value.
- If surface is exposed to only ambient light we calculate ambient diffuse reflection as:
  $I_{ambdiff} = K_d I_a$
  Where I_a the ambient light is falls on the surface.
- Practically most of times each object is illuminated by one light source so now we discuss diffuse reflection intensity for point source.

- We assume that the diffuse reflection from source are scattered with equal intensity in all directions, independent of the viewing direction such a surface are sometimes referred as **ideal diffuse reflector** or **lambertian reflector.**
- This is modelled by **lambert's cosine law.** this law states that the radiant energy from any small surface area dA in any direction $\emptyset_n$ relative to surface normal is proportional to $cos\emptyset_n$.

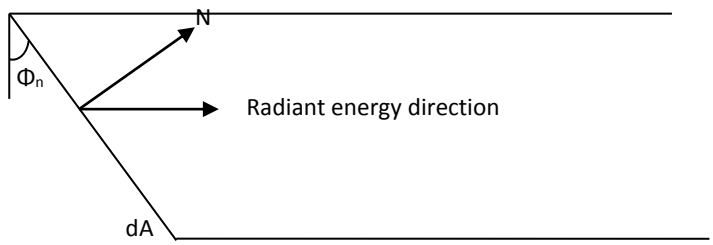


Fig. 6.6:- Radiant energy from a surface area dA in direction Φ_n relative to the surface normal direction.

- As shown reflected light intensity is does not depends on viewing direction so for lambertian reflection, the intensity of light is same in all viewing direction.
- Even though there is equal light distribution in all direction from perfect reflector the brightness of a surface does depend on the orientation of the surface relative to light source.
- As the angle between surface normal and incidence light direction increases light falls on the surface is decreases

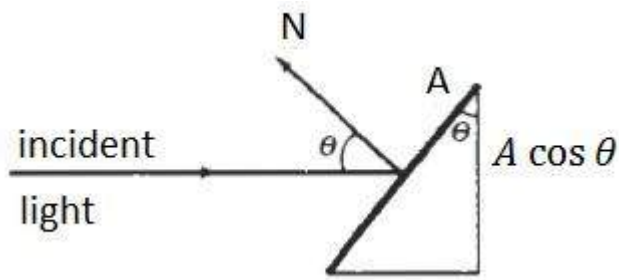


Fig. 6.7:- An illuminated area projected perpendicular to the path of the incoming light rays.

- If we denote the **angle of incidence** between the incoming light and surface normal asθ, then the projected area of a surface patch perpendicular to the light direction is proportional to $cos\theta$.
- If I_l is the intensity of the point light source, then the diffuse reflection equation for a point on the surface can be written as
  $I_{l,diff} = K_d I_l cos\theta$
- Surface is illuminated by a point source only if the angle of incidence is in the range 0^0 to 90^0 other than this value of θ light source is behind the surface.

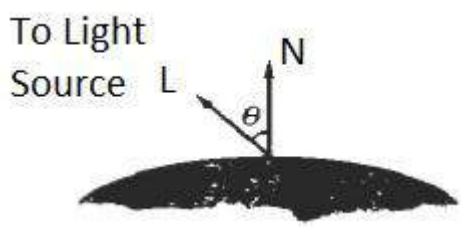


Fig. 6.8:-Angle of incidence θ between the unit light-source direction vector L and the unit surface normal N.

- As shown in figure **N** is the unit normal vector to surface and **L** is unit vector in direction of light source then we can take dot product of this to is:
  $N \cdot L = \cos\theta$
  And

$I_{l,diff} = K_d I_l (N \cdot L)$

- Now in practical ambient light and light source both are present and so total diffuse reflection is given by:

  $I_{diff} = K_a I_a + K_d I_l (N \cdot L)$
- Here for ambient reflection coefficient K_a is used in many graphics package so here we use K_a instead of K_d.

Specular Reflection and the Phong Model.

- When we look at an illuminated shiny surface, such as polished metal we see a highlight, or bright spot, at certain viewing directions. This phenomenon is called **specular reflection,** is the result of total, or near total reflection of the incident light in a concentrated region around the **specular reflection angle.**

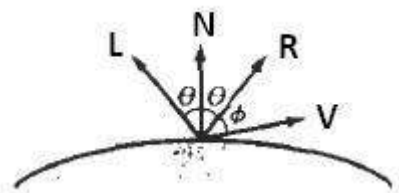


Fig. 6.9:-Specular reflection angle equals angle of incidence θ.

- Figure shows specular reflection direction at a point on the illuminated surface. The specular reflection angle equals the angle of the incident light.
- Here we use R as unit vector in direction of reflection L is unit vector point towards light vector N is unit normal vector and V is unit vector in viewing direction.
- Objects other than ideal reflectors exhibits specular reflection over a finite range of viewing positions around vector R. Shiny surface have a narrow specular reflection range and dull surface have wide specular reflection range.
- **By phong specular reflection model** or simply **phong model** sets the intensity of specular reflection proportional to $cos^{ns}\emptyset$. Angle ∅ varies in between 0^0 to 90^0.
- Values assigned to **specular reflection parameter** ns is determined by the type of surface that we want to display. A shiny surface assigned ns values large nearly 100 and dull surface assigned small nearly 1.
- Intensity of specular reflection depends on the material properties of the surface and the angle of incidence as well as **specular reflection coefficient, $w(\theta)$ for each surfaces.**
- Then specular reflection is given by:

  $I_{spec} = w(\theta) I_l cos^{ns}\emptyset$

  Where I_l is the intensity of light source and ∅ is angle between viewing direction V and specular reflection direction R.
- Since ∅ is angle between two unit vector V and R we can put $cos\emptyset = V \cdot R$.
- And also for many surfaces $w(\theta)$ is constant so we take specular reflection constant as K_s so equation becomes.

  $I_{spec} = K_s I_l (V \cdot R)^{ns}$
- Vector r is calculated in terms of vector L and N as shown in figure

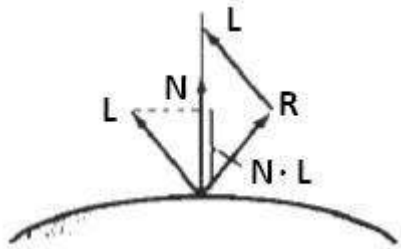


Fig. 6.10:- Calculation of vector R by considering projection onto the direction of the normal vector N.

$R + L = (2N \cdot L)N$

$R = (2N \cdot L)N - L$

- Somewhat simplified phong model is to calculate between half way vectors H and use product of H and N instead of V and R.
- Here H is calculated as follow:

$$H = \frac{L + V}{|L + V|}$$

Combined Diffuse and Specular Reflections With Multiple Light Sources

- For a single point light source we can combined both diffuse and specular reflection by adding intensity due to both reflection as follows:

  $I = I_{diff} + I_{spec}$

  $I = K_a I_a + K_d I_l (N \cdot L) + K_s I_l (N \cdot H)^{ns}$
- And for multiple source we can extend this equation as follow:

$$I = K_a I_a + \sum_{i=1}^{n} I_l [K_d (N \cdot L) + K_s (N \cdot H)^{ns}]$$

Properties of Light

- Light is an electromagnetic wave. Visible light is have narrow band in electromagnetic spectrum nearly 400nm to 700nm light is visible and other bands not visible by human eye.

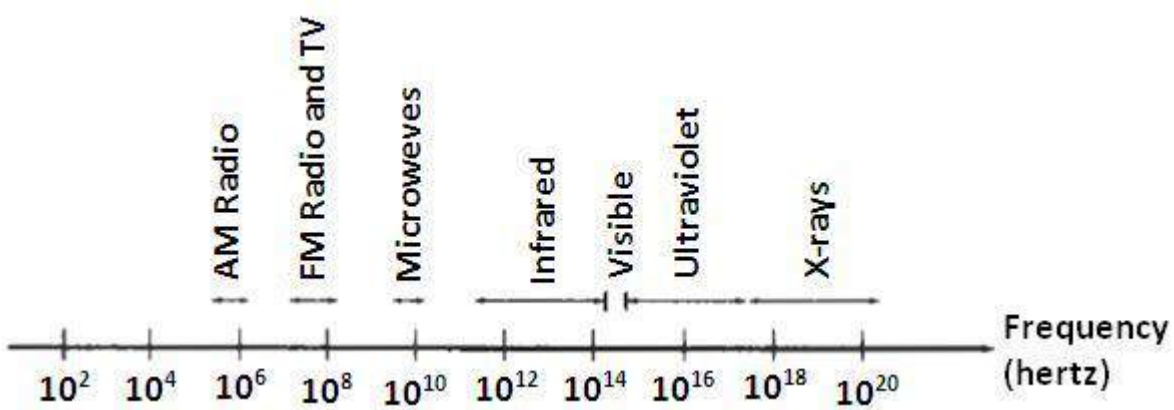


Fig. 6.11:- Electromagnetic spectrum.

- Electromagnetic spectrum shown in figure shows other waves are present in spectrum like microwave infrared etc.
- Frequency value from 4.3 X 10^14 hertz (red) to 7.5 X 10^14 (violet) is visible renge.
- We can specify different color by frequency f or by wavelength λ of the wave.
- We can find relation between f and λ as follows:

  $c = \lambda f$
- Frequency is constant for all the material but speed of the light and wavelength are material dependent.

- For producing white light source emits all visible frequency light.
- Reflected light have some frequency and some are absorbed by the light. This frequency reflected back is decide the color we see and this frequency is called as **dominant frequency (hue)** and corresponding reflected wavelength is called **dominant wavelength.**
- Other property are **purity** and **brightness.** Brightness is perceived intensity of light. Intensity is the radiant energy emitted per unit time, per unit solid angle and per unit projected area of the source.
- **Purity** or **saturation** of the light describes how washed out or how "pure" the color of the light appears.
- Dominant frequency and purity both collectively refers as **chromaticity.**
- If two color source combined to produce white light they are called **complementary color** of each other. For example red and cyan are complementary color.
- Typical color models that are uses to describe combination of light in terms of dominant frequency use three colors to obtain reasonable wide range of colors, called the **color gamut** for that model.
- Two or three colors are used to obtain other colors in the range are called **primary colors.**

XYZ Color Model

- The set of CIE primaries is generally referred to as XYZ or (X, Y, Z) color model.

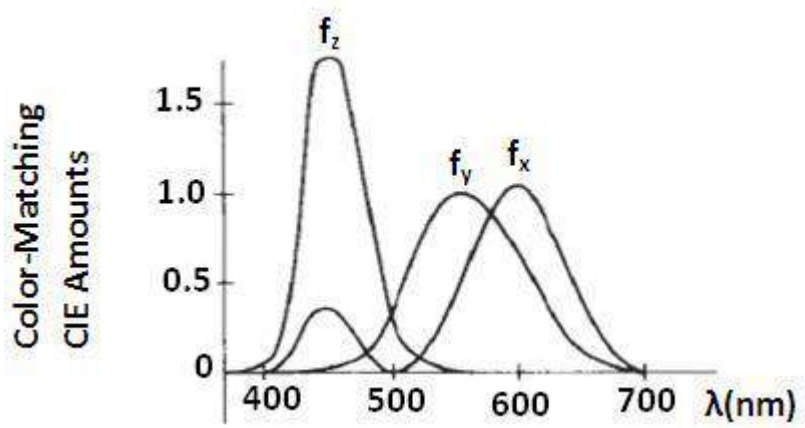


Fig. 6.12:- Amount of CIE primaries needed to display spectral colors.

- X, Y, Z represents vectors in a three dimensional, additive color space.
- Any color C_λ is a combination of three primary colors as
  $C_\lambda = XX + YY + ZZ$
  Where X, Y, Z, is the amount of standard primary need to combine for obtaining color C_λ.
- If we normalize it then.
  $x = \frac{X}{X+Y+Z}$ $y = \frac{Y}{X+Y+Z}$ $z = \frac{Z}{X+Y+Z}$
  With $x + y + z = 1$
- Now we can represent any color with x,y only as z we can find z=1-x-y.
- X, and y are called chromaticity values because they depends only on hue and purity.
- Now if we specify colors with only x, and y values we cannot find amount X, Y, and Z.
- So we specify color with x, y, and Y and rest CIE amount is calculated as:
  $X = \frac{x}{y}Y$ $Z = \frac{z}{y}Y$
  Where z=1-x-y

RGB Color Model

- Based on tristimulus theory of vision our eye perceives color through stimulate one of three visual pigments in the cones of the retina.

- These visual pigments have peak sensitivity at red, green and blue color.
- So combining these three colors we can obtain wide range of color this concept is used in RGB color model.

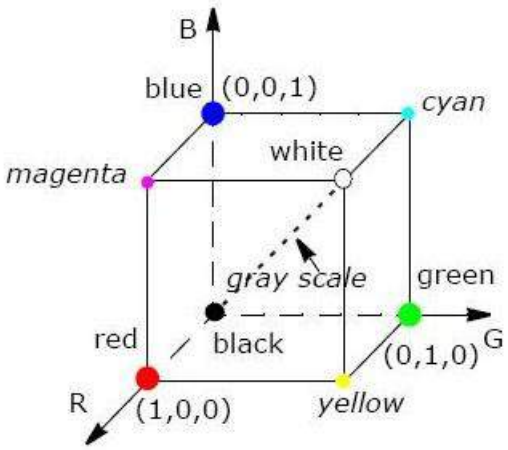


Fig. 6.13:- The RGB color model.

- As shown in figure this model is represented as unit cube.
- Origin represent black color and vertex (1,1,1) is white.
- Vertex of the cube on the axis represents primary color R, G, and B.
- In XYZ color model any color intensity is obtained by addition of primary color.
  $C_\lambda = RR + GG + BB$
- Where R, G, and B is amount of corresponding primary color
- Since it is bounded in between unit cube it's values is very in between 0 to 1 and represented as triplets (R,G,B). For example magenta color is represented with (1,0,1).
- Shades of gray are represented along the main diagonal of cube from black to white vertex.
- For half way gray scale we use triplets (0.5,0.5,0.5).

YIQ Color Model

- As we know RGB monitors requires separates signals for red, green, and blue component of an image but television monitors uses single composite signals.
- For this composite signal NTSC use YIQ color model.
- Here parameter Y is represented as luminance (brightness) while chromaticity information (hue and purity) is specified into I and Q parameter.
- Combination of all red, green, and blue intensities are chosen for Y so black and white television monitors only use signal Y for brightness.
- So largest bandwidth (about 4 MHz) is assigned to Y information signal.
- Parameter I contain orange-cyan hue information that provides the flash-tone shading, and occupies a bandwidth approximately 1.5 MHz.
- Parameter Q carries green-magenta hue information in a bandwidth of about 0.6 MHz.
- An RGB signal can be converted to a television signal using encoder which converts RGB to YIQ values.
- This conversion by transformation is given by:
  $$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.528 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$
- Similarly reverse of this is performed by decoder and by transformation using inverse of above matrix as.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.000 & 0.956 & 0.620 \\ 1.000 & -0.272 & -0.647 \\ 1.000 & -1.108 & 1.705 \end{bmatrix} \begin{bmatrix} Y \\ I \\ Q \end{bmatrix}$$

CMY Color Model

- A color model CMY is used for hardcopy devices as we produce picture by coating a paper with color pigments, we see the color by reflected light a subtractive process.
- When white light is reflected from cyan colored ink the reflected light must have no red component that is red light is absorbed or subtracted by the ink.
- Similarly magenta is subtracting green component.
- Unit cube for CMY model is shown in figure below.

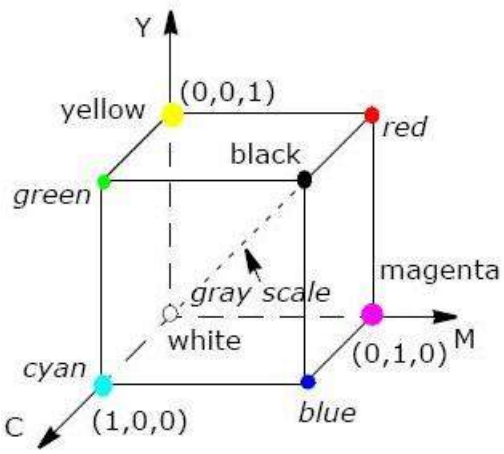


Fig. 6.14:- The CMY color model.

- Point (1,1,1) represents black because all components are subtracts and origin represents white light.
- Gray can be produce among main diagonal by using all three color in equal amount.
- Printing process often use CMY model generates a color points with a collection of four ink dots, one for each primary color C, M, and Y and one dot is black.
- Conversion of RGB to CMY is done by:

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

- And similarly reverse is done by:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} C \\ M \\ Y \end{bmatrix}$$