1. Introduction to Computer Graphics and Display Systems

1.1 Introduction

Computer graphics is a field of computer science that is concerned with digitally synthesizing and manipulating visual contents. It is also sometimes called the science and technology of creating, storing, displaying, and manipulating images and objects. Computer graphics broadly studies the manipulation of visual and geometric information using computational techniques.

The study of "real" world and its geometry requires us to build a mathematical model. After representing the world mathematically, we need to transform (continuous) mathematical representations into finite or discrete representation in order to make them accessible to computers, usually to display them on a monitor or output on a printer in some cases. Computer graphics, as an academic discipline, focuses on the mathematical and computational foundations of image generation and processing rather than purely aesthetic issues. It includes almost everything on computers that is not text or sound. All computers used today use one or the other aspect some graphics, and users expect to interact with their computers through icons and pictures rather than just by typing. The term computer graphics deals with:

- The representation and manipulation of pictorial data by a computer
- The various technologies used to create and manipulate such pictorial data
- The images so produced
- The study of methods for digitally synthesizing and manipulating visual content

The potential of designing, modeling, and visualizing objects play an important role not only in computer graphics, but also in the generation of realistic models and representations of objects based on measurement data. The combination of real data and images with techniques from computer graphics has gained importance today than ever before. The chapter introduces to the basics of the graphics field by understanding what an image is, and how it is different from an object?

1.2 Image and objects

Computer graphics provides methods to generate images using a computer. The word "image" should be understood in a more abstract sense here. An image can be a representation of a realistic scene or an object from the real world. It can take different forms. An image may be an actual picture display, a page stored in video memory, or a source code obtained from a program. An image is a primarily symbolic visual representation other than a text. In computer graphics and image analysis, it is a two-dimensional (2D) array of data with values at each element of the array related to intensity or a color.

Objects are real entities defined in three-dimensional (3D) world coordinates. A careful observation reveals that an image is a 2D description of a 3D object. In fact, what we perceive is an image of a real-world object. Computer graphics deals with both 2D and 3D descriptions of an object that performs transformations generating visual displays. These transformations may be directly performed on object coordinates or image coordinates. We often refer it to as an image plane and an object space for simplicity. Object manipulation gives a better flexibility for generalization at the cost of complexity of operations. Models and algorithms for this task combine knowledge from different areas of mathematics and computer science. We explore them shortly in the subsequent chapters.

Computer graphics, at its basic level, discusses the representation of objects and images, algorithms and procedures for generating and manipulating them.

1.3 Image representation

Computer graphics is all about digital models for handling 3D geometric objects, as well as 2D images. Representations of shapes and images may be approximations of the real world or synthetic constructs that exist only in the computer. The primary goals of computer graphics initiatives are the generation of reasonable and informative images and computation in a short amount of time with little storage requirements.

In computer graphics, pictures (as graphical subjects) are represented as a collection of discrete picture element called pixels. A pixel is the smallest addressable screen element. The process of determining the appropriate pixels for representing pictures or object is called rasterization, and the process of representing continuous pictures as graphical object (as a collection of discrete pixel) is called a scan conversion.

1.4 Raster and vector-based graphics

There are two kinds of computer graphics: raster (composed of pixels) and vector (composed of paths). Raster images are more commonly called bitmap images. The raster and vector graphic systems use completely different internal representation, different image entry methods and differ in handling as well as manipulation techniques.

Bitmap images are pixel-based. The raster system represents an image as a 2D array where the value in the array defines color or gray scale, for the corresponding image. The only primitive element of raster systems is pixel. In other words, location and color information about the image is stored in individual pixels within a grid. In bitmap images, each pixel has an assigned color. The information stored in a bitmap image regarding pixel location and color forms the image. Bitmap images are edited at the pixel level that allows the color of any one pixel to change. For raster systems, image manipulation is simply a matter of addressing memory, together with an image independent method of display and large number of raster operations. Additional attributes of bitmap images include:

- Bitmap images are usually created and edited in photo or paint programs such as Adobe® Photoshop®.
- Bitmap images are mapped to a grid or an array of pixels.
- The size of an image is based on the image's resolution.
- Bitmap images are not easily scalable.
- Bitmap images are used for photorealistic images and, therefore, may involve complex color variations.

Figure 1.4(a) shows a small blue arrow image at its actual size and Figure 1.4(b) the same image magnified. The upper-left quadrant of Figure 1.4(b) shows the individual pixels of the image. Figure 1.4(c) shows magnified view of the left eye illustrating the pixel composition of the left eye. Note that each pixel in the magnified image defines the location and color information in the array.

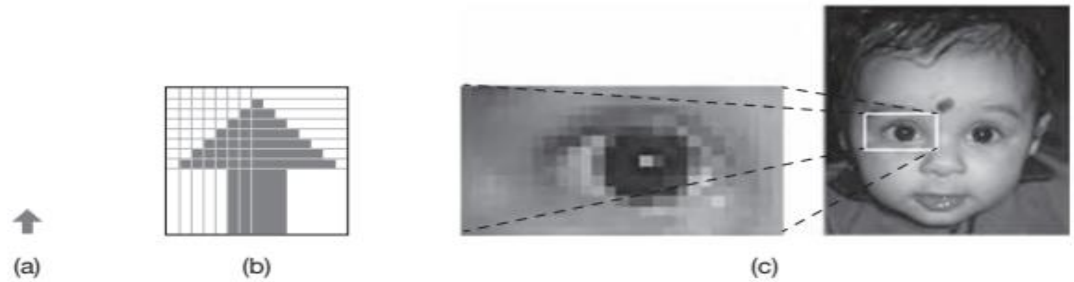

Figure 1.4 (a) An arrow image in its actual size; (b) a magnified arrow image with pixel grid; (c) bitmap representation of eye.

A vector drawing system stores a graphic object as a set of primitives; therefore all vector images are mathematically-based images. All lines, shapes, etc., (also called objects) of a vectorbased image are independent of one another. The primitives that make up the object are arranged in a user-defined order, that when output in that order produces correct display. Vector systems use some kid of data entry (generally either keyboard or mouse) for the graphics primitives. The internal computational model of vector systems is different from raster systems. The vector system relies on quick transformation and fast drawing of the primitives.

Additional attributes of vector-based images include:

• Vector-based images are usually created and edited in software like CoralDRAW® and Adobe Illustrator®.

• These images have smooth edges and create curves or shapes.

• Vector-based images are good for precise illustrations, but are not as good as bitmaps for photorealistic images.

• Vector-based images are easily scalable due to their use of mathematics formulas.

• The long undo/redo list is possible to maintain due to well defined mathematical primitive which is another convenience during image manipulation.

Figure 1.5(a) shows an image of a rose (in actual size). Figure 1.5(b) shows one leaf of the leaves magnified, along with the paths and handles used within vector images.

Let us consider another example of a bitmap and a vector graphic image of the letter A (Figure 1.6). It is clear that the larger we display a bitmap, the jagged it appears, while a vector image remains smooth in any size.

Note that the vector representation is converted into a bitmap for the computer to display it. You can easily convert vector computer aided design (CAD) data to raster representation but not vice versa. The cost of vector operation is roughly proportional to the number of primitives involved, whereas the costs of raster operations are proportional to the area changed.

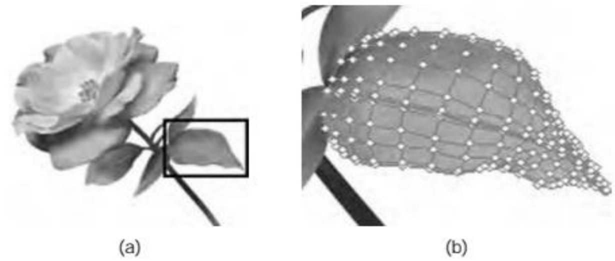

Figure 1.5 (a) A rose; (b) vector description of rose leaf.

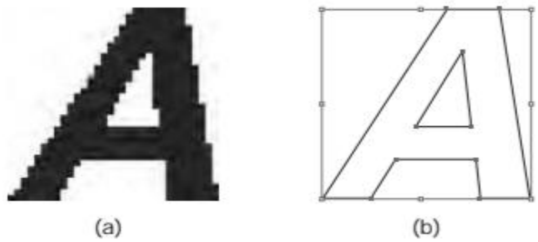

Figure 1.6 (a) A bitmap image; (b) a vector image.

1.6 Applications of computer graphics

Computer graphics finds applications in our lives starting from the most basic needs of objects representation to animation and visual effects. The following applications of computer graphics attempt to solve many of our day-to-day problems:

• Digital art: A digital art, most commonly, refers to an art created on a computer in a digital form. A digital art can be purely computer-generated, such as fractals and an algorithmic art or taken from another source like a scanned photograph or image drawn using vector graphics software.

• Special effects: Illusions used in film, television, and entertainment industries to simulate the imagined events in a story are traditionally called special effects (also known as SFX or SPFX). In modern films, special effects are usually used to alter previously-filmed elements by adding, removing, or enhancing objects within the scene.

• Visual effects: Visual effects (or VFX, in short) are the term given in which images or film frames are created and manipulated for film and video. Visual effects usually involve the integration of live-action footage with computer-generated imagery or other elements (such as pyrotechnics or model work) in order to create environments or scenarios which look realistic. But visual effects may be dangerous, costly, or simply impossible to capture on a film.

• Video games: Video games find a good place in the gaming and animation industry.

• Computer-aided design: Some of the other areas of applications of computer graphics include computer-aided design (CAD). This is used by civil, mechanical, and electronic engineers to build a model of building, bridges, printed circuit board (PCB), etc.

• Medical imaging: It can be used to train new doctors on understanding and experimenting surgery. It is used in molecular biology to study the biological model in genes.

In addition, there are many other applications for computer graphics that are not listed here. We might be interested in automating a complex picture that we could make by hand for convenience, or we might be interested in creating an image that is abstract or random in nature. In fact, there are many applications of computer graphics that fall somewhere in extremes presented above.

1.7 Display devices

Today computers are integral part of our lives, as they are effectively used in computing, control, and entertainment. Roles of display units for any of these tasks are equally important. Over the decades, computer graphics has made noticeable advances in the display and visualization technology.

A display device is a device for presentation of information, such as image or a text, for visual or tactile reception, acquired, stored, or transmitted in various forms. When the input information is supplied as an electrical signal, the display is called electronic display. Electronic displays are available for presentation of visual and tactile information. We explore here some electronic displays used for graphics output. Common types of display devices include analog and digital electronic displays, projectors segment displays, and many more.

1.8 Cathode ray tubes

The most talked about and pioneer in the category of analog electronic displays are cathode ray tubes (CRT). A CRT is a specialized vacuum tube in which images are produced when an electron beam strikes a phosphorescent surface. Most desktop computer displays make use of CRTs. The CRT in a computer display is similar to the picture tube in a television receiver. A CRT consists of several basic components as illustrated in Figure 1.7.

Heat is supplied to the cathode by passing current through a heater element. The cathode is a cylindrical metallic structure that is rich in electrons. On heating, electrons are released from cathode surface. The control grid is the next element that follows the cathode. It almost covers the cathode leaving small opening for electrons to come out. Intensity of the electron beam is controlled by setting voltage levels on the control grid. A high negative voltage applied to the control grid shuts off the beam by repelling electrons and stopping them from passing through the small hole at the end of the control grid structure. A smaller negative voltage on the control grid simply decreases the number of electrons passing through the cathode. Thus, we can control the brightness of display by varying the voltage on the control grid. Positively charged anodes, in the sequence accelerating anodes, accelerate the electrons towards phosphor screen.

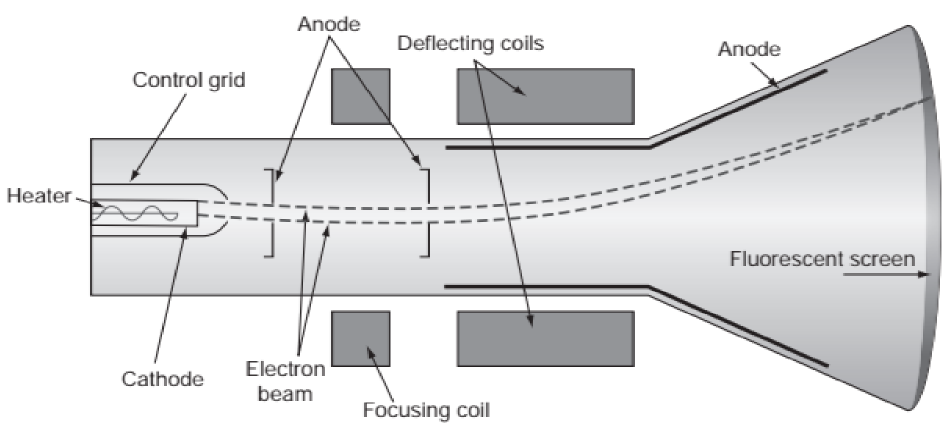

Figure 1.7 A cathode ray tube.

Focusing and deflection coils are together needed to force the electron beam to converge into a small spot, as it strikes the screen, otherwise the electrons would repel each other and the beam would spread out as it approaches the screen. Deflecting coils produce an extremely lowfrequency electromagnetic field that allows for constant adjustment of the direction of the electron beam. There are two sets of deflecting coils: horizontal and vertical. Electrostatic focusing is commonly used in television and computer graphics monitors. Finally, when the accelerating electron beam collides on the phosphor coating, a part of kinetic energy is converted into light and heat. When the electrons in the beam collide with the phosphor coating, they are stopped and their kinetic energy is absorbed by the phosphor, resulting in screen display. Note that the amount of light emitted by the phosphor coating depends on the number of electrons striking the screen and can be controlled by varying the voltage on the control grid.

Graphical displays for early computers used vector monitors, a type of CRT, similar to an oscilloscope that uses magnetic (rather than electrostatic) deflection. Magnetic deflection allows the construction of much shorter tubes for a given viewable image size. Here, the beam traces straight lines between arbitrary points, repeatedly, refreshing the display as quickly as possible. Vector displays for computers did not noticeably suffer from the display artifacts of aliasing and pixilation, but were limited in that they could display only a shape's outline. Advanced vector systems could provide a limited amount of shading, and only a limited amount of crudely-drawn text. The number of shapes and textual characters drawn was severely limited, because the speed of refresh was roughly inversely proportional to the number of vectors needed to be drawn. Some newer vector monitors are capable of displaying multiple colors, using either a typical tricolor CRT or two phosphor layers (so-called penetration color).

CRTs produce crisp, vibrant images. But they do have a serious drawback: They are bulky. In order to increase the screen width in a CRT set, you also have to increase the length of the tube (to give the scanning electron gun room to reach all parts of the screen). It has been claimed that the electromagnetic fields emitted by CRT monitors constitute a health hazard and can affect the functioning of living cells, and people are moving towards other forms of digital displays replacing them slowly.

1.8.1 Raster-scan display

A raster-scan display is the most common method of drawing images on a CRT screen. In this method, horizontal and vertical deflection signals are generated to move a beam all over the screen in a pattern for displaying any image (Figure 1.8). The electron beam is swept across the screen one row at a time from top to bottom.

The electron beam sweeps back and forth from left to right across the screen. The beam is on, while it moves from left to right. The beam is off, when it moves back from right to left. This phenomenon is called the horizontal retrace, as shown by dotted lines in the figure.

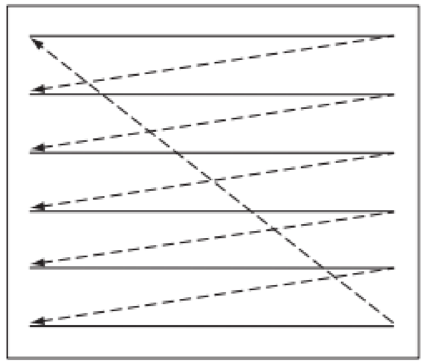

Figure 1.8 A raster-scan mechanism.

As soon as the beam reaches the bottom of the screen, it is turned off and is rapidly retraced back to the top to start again. This is called the vertical retrace. Raster-scan displays maintain the steady image on the screen by repeating scanning of the same image. This process is known as refreshing of screen.

Typically, a graphics display consists of three components: frame buffer, display controller, and a TV screen or monitor. The frame buffer stores an image as a matrix of intensity values. It is an interface between what are usually relatively slow graphics computations and the high data rate video image display. In a typical personal computer, the frame buffer is located on the graphics card that manages the video subsystem of the computer. It basically used to be not much more than some extra memory. Stored intensity values are then retrieved from the refresh buffer and displayed on the screen one row at a time. Each intensity value is represented by bit zero (0) or one (1) in the frame buffer.

The video or display controller has direct access to memory locations in the frame buffer. It is responsible for retrieving data from the frame buffer and passing it to the display device. It reads each successive bytes of data from frame buffer and converts this 0's and 1's in one line into a corresponding video signals, and this line is called a scan line. If the intensity is one (1) then the controller sends a signal to display a dot in the corresponding position on the screen. If the intensity is zero (0) then no dot is displayed. This is the simplest way in which a black and white image is displayed. The display-controller repeats the read, convert, and fill operations at least 60 times per seconds that maintains a study picture on the screen (refresh rate).

1.8.2 Random-scan display

A CRT, as a random-scan display unit, has an electron beam directed only to the parts of the screen where a picture is to be drawn. Random-scan monitors draw a picture one line at a time. These are also referred to as vector displays (or stroke-writing or calligraphic displays). The component of a picture (lines and curves) can be drawn and refreshed by a random-scan system in any specified order. A pen plotter operates in a similar way and is an example of a randomscan, hard-copy device.

The refresh rate, on a random-scan system, depends on the number of primitives like lines to be displayed. A picture definition is stored as a set of line-drawing commands in an area of memory called a refresh display file (or a refresh buffer). To display a specified picture, the system cycles through the set of commands in the display file, drawing each component line one by one. After all line-drawing commands have been processed, the system cycles back to the first line command in the list and repeat the procedure of scan, display, and retrace. Random-scan displays are designed to draw all the component lines of a picture 30 to 60 times each second. High-quality vector systems are capable of handling approximately 100,000 short lines at this refresh rate. It is important to note that the faster refreshing of the set of lines could burn out the phosphor. Therefore, when a small set of lines are to be displayed, each refresh cycle is delayed to avoid greater refresh rates, typically 60 frames per second.

Random-scan systems are designed for line-drawing applications; hence cannot display realistic shaded scenes. Vector displays generally have a higher resolution than raster systems, as picture definition is stored as a set of line-drawing instructions instead of a set of intensity values for all screen points. These vector displays produce smooth line drawings, because the CRT beam directly follows the line path. A raster system, in contrast, produces jagged lines that are plotted as discrete point sets. Table 1.1 presents the differences between a raster-scan display and a random-scan display.

Table 1.1 Difference between a raster-scan display and a random-scan display

Raster-scan display	Random-scan display
It draws the image by scanning one row at a time.	It draws the image by directing the electron beam directly to the part of the screen, where the image is to be drawn.
They generally have resolution limited to pixel size.	They have higher resolution than the raster scan system.
Lines are jiggered, and curves are less smooth.	Line plots are straight, and curves are smooth.
They are more suited to geometric area drawing applications, e.g. monitors, TV.	They are more suited to line drawing application, e.g. CRO and pen plotter.