UNIT 3: IMAGES AND GRAPHICS

1.1 INTRODUCTION

Image is the spatial representation of an object. It may be 2D or 3D scene or another image. Images may be real or virtual. It can be abstractly thought of as continuous function defining usually a rectangular region of plane. Example:

- Recorded image- photographic, or in digital format
- Computer vision- video image, digital image or picture
- Computer graphics- digital image
- Multimedia- deals about all above formats

1.2 DIGITAL IMAGE REPRESENTATION

A digital image is represented by a matrix of numeric values each representing a quantized **intensity** value. When I is a two-dimensional matrix, then I(r, c) is the intensity value at the position corresponding to row r and column c of the matrix. The points at which an image is sampled are known as picture elements, commonly abbreviated as **pixels**. The pixel values of intensity images are called gray scale levels. The intensity at each pixel is represented by an integer and is determined from the continuous image by averaging over a small neighbourhood around the pixel location. If there are just two intensity values, for example, black, and white, they are represented by the numbers 0 and 1; such images are called **binary-valued images**. If 8-bit integers are used to store each pixel value, the gray levels range from 0 (black) to 255 (white).

1.2.1 Digital Image Format

There are different kinds of image formats in the literature. We shall consider the image format that comes out of an image frame grabber, i.e., the captured image format, and the format when images are stored, i.e., the stored image format.

Captured Image Format

The image format is specified by two main parameters: spatial resolution, which is specified as pixels*pixels (e.g. 640x480) and color encoding, which is specified by bits per pixel. Both parameter values depend on hardware and software for input/output of images.

For example, for image capturing on a SPARCstation, the VideoPix card and its software are used. The spatial resolution is 320 X 240 pixels and the color can be encoded with 1-bit (a binary image format), 8-bit (color or grayscale) or 24-bit (color-RGB).

Stored Image Format

When we store an image, we are storing a two-dimensional array of values, in which each value represents the data associated with a pixel in the image. For a bitmap, this value is a binary digit.

For a color image (pixmap), the value may be a collection of:

- Three numbers representing the intensities of the red, green and blue components of the color at that pixel.
- Three numbers that are indices to tables of the red, green and blue intensities.
- A single number that is an index to a table of color triples.
- An index to any number of other data structures that can represent a color.
- Four or five spectral samples for each other.

In addition, each pixel may have other information associated with it; e.g., three numbers indicating the normal to the surface drawn at that pixel.

Information associated with the image as a whole, e.g., width, height, depth, name of the creator, etc. may also have to be stored.

The image may be compressed before storage for saving storage space. Some current image file formats for storing images include GIF, X11 Bitmap, Sun Rasterfile, PostScript, IRIS, JPEG, TIFF, etc.

1.3 IMAGE AND GRAPHICS FORMATS

Image File Format

There are many file formats used to store bitmaps and vectored drawing. Following is a list of few image file formats:

Format	**Extension**
Microsoft Windows DIB	.bmp .dib .rle
Microsoft Palette	.pal
Autocad format 2D	.dxf
JPEG	.jpg
Windows Meta file	.wmf
Portable network graphic	.png
Compuserve gif	.gif
Apple Macintosh	.pict .pic .pct

Bitmap (BMP):

BMP is a standard format used by Windows to store device-independent and application independent images. The number of bits per pixel (1, 4, 8, 15, 24, 32, or 64) for a given BMP file is specified in a file header. BMP files with 24 bits per pixel are common.

Graphics Interchange Format (GIF)

GIF is a common format for images that appear on webpages. GIFs work well for line drawings, pictures with blocks of solid color, and pictures with sharp boundaries between colors. GIFs are compressed, but no information is lost in the compression process; a decompressed image is exactly the same as the original. One color in a GIF can be designated as transparent, so that the image will have the background color of any Web page that displays it. A sequence of GIF images can be stored in a single file to form an animated GIF. GIFs store at most 8 bits per pixel, so they are limited to 256 colors.

Joint Photographic Experts Group (JPEG)

JPEG is a compression scheme that works well for natural scenes, such as scanned photographs. Some information is lost in the compression process, but often the loss is imperceptible to the human eye. Colour JPEG images store 24 bits per pixel, so they are capable of displaying more than 16 million colors. There is also a grey scale JPEG format that stores 8 bits per pixel. JPEGs do not support transparency or animation. The level of compression in JPEG images is configurable, but higher compression levels (smaller files) result in more loss of information. A 20:1 compression ratio often produces an image that the human eye finds difficult to distinguish from the original. The Figure below shows a BMP image and two JPEG images that were compressed from that BMP image. The first JPEG has a compression ratio of 4:1 and the second JPEG has a compression ratio of about 8:1.

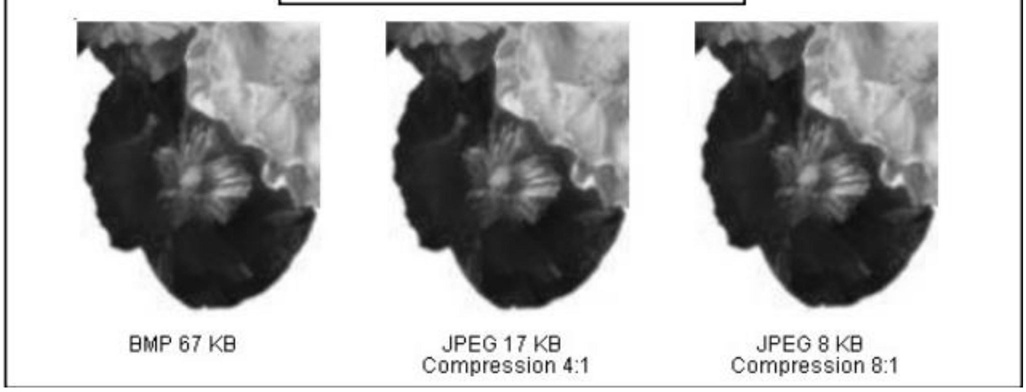


Figure 4.12: A BMP and Two JPEG Images

JPEG compression does not work well for line drawings, blocks of solid color, and sharp boundaries. JPEG is a compression scheme, not a file format. JPEG File Interchange Format (JFIF) is a file format commonly used for storing and transferring images that have been compressed according to the JPEG scheme. JFIF files displayed by Web browsers use the .jpg extension.

Exchangeable Image File (Exif)

Exif is a file format used for photographs captured by digital cameras. An Exif file contains an image that is compressed according to the JPEG specification. An Exif file also contains information about the photograph (date taken, shutter speed, exposure time, and so on) and information Notes about the camera (manufacturer, model, and so on).

Portable Network Graphics (PNG)

The PNG format retains many of the advantages of the GIF format but also provides capabilities beyond those of GIF. Like GIF files, PNG files are compressed with no loss of information. PNG files can store colors with 8, 24, or 48 bits per pixel and gray scales with 1, 2, 4, 8, or 16 bits per pixel. In contrast, GIF files can use only 1, 2, 4, or 8 bits per pixel. A PNG file can also store an alpha value for each pixel, which specifies the degree to which the color of that pixel is blended with the background colour.

Tag Image File Format (TIFF)

TIFF is a flexible and extendable format that is supported by a wide variety of platforms and image-processing applications. TIFF files can store images with an arbitrary number of bits per pixel and can employ a variety of compression algorithms. Several images can be stored in a single, multiple-page TIFF file. Information related to the image (scanner make, host computer, type of compression, orientation, samples per pixel, and so on) can be stored in the file and arranged through the use of tags. The TIFF format can be extended as needed by the approval and addition of new tags.

Graphics Format

Graphic image formats are specified through graphics primitives and their attributes.

- Graphic primitive – line, rectangle, circle, ellipses, specification 2D and 3D objects.
- Graphic attribute – line style, line width, color.

Graphics formats represent a higher level of image representation, i.e., they are not represented by a pixel matrix initially.

- PHIGS (Programmer's Hierarchical Interactive Graphics)
- GKS (Graphical Kernel System)

Vector Drawings

Vector Drawings are completely computer generated. They are otherwise known as **object oriented graphics** as they consist of objects such as shapes. Vectors are used to create graphics such as interface elements (banners, buttons) text, line art and detailed drawings (plans, maps). Essentially they are computer generated drawings. Effects can be added to vector graphics to add realism, however, they need to be converted to bitmaps in order to do this. Vectors don't consist of pixels. Instead, they are made up of co-ordinates, shapes, line, and colour data. Therefore they aren't resolution dependent. It is for this reason that vector graphics can be scaled without losing their quality. Vectors are also easier to edit. In comparison to bitmaps, vectors have nice clean edges.

The following extract from the Adobe site gives a further definition of Vector graphics: "You can freely move or modify vector graphics without losing detail or clarity, because they are resolution-independent-they maintain crisp edges when resized, printed to a PostScript printer, saved in a PDF file, or imported into a vector-based graphics application. As a result, vector graphics are the best choice for artwork, such as logos, that will be used at various sizes and in various output media."

Vector file formats create smaller file sizes than bitmaps generally. Vectors still aren't well supported in the world wide web. To include a vector in a website it is still best to rasterize it, for instance, convert it to a bitmap. This can however, result in the edges of the vector graphic loosing definition and become pixelated.

1.4 Image Synthesis, analysis and Transmission

1.4.1 Computer Image Processing

Image processing is a method to perform some operations on an image, in order to get an enhanced image or to extract some useful information from it. Computer image processing comprises of **image synthesis** (generation) and **image analysis** (recognition). It is a type of signal processing in which input is an image and output may be image or characteristics/features associated with that Image, processing basically includes the following three steps:

- Importing the image via image acquisition tools;
- Analyzing and manipulating the image;
- Output in which result can be altered image or report that is based on image analysis.

There are two types of methods used for image processing namely, analogue and digital image processing. Analogue image processing can be used for the hard copies like printouts and photographs. Digital image processing techniques help in manipulation of the digital images by using computers. The three general phases that all types of data have to undergo while using digital technique are pre-processing, enhancement, and display, information extraction. image.

1.4.2 Dynamics in Graphics

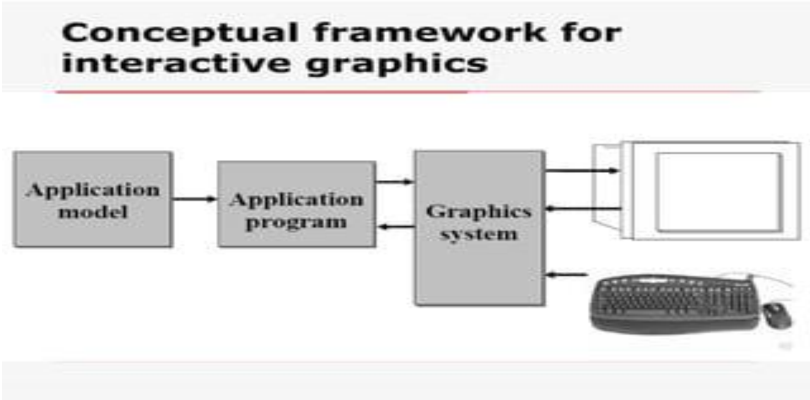
Dynamic graphics for data, means simulating motion or movement using the computer. It may also be thought of as multiple plots linked by time. Two main examples of dynamic graphics are animations, and tours. An animation, very generally defined, may be produced for time-indexed data by showing the plots in time order, for example as generated by an optimization algorithm. With dynamic simulation, you can create many impressive effects such as explosion, flood, storm, tornado, ocean, etc., for animations and computer games.

Motion Dynamic: With motion dynamic, objects can be moved and enabled with respect to a stationary observer.

Update Dynamic: Update dynamic is the actual change of the shape, color, or other properties of the objects being viewed.

1.4.3 The Framework of Interactive Graphics Systems

In interactive Computer Graphics user have some controls over the picture, i.e., the user can make any change in the produced image. Interactive Computer Graphics require two-way communication between the computer and the user. A User can see the image and make any change by sending his command with an input device. The framework of interactive graphics systems have following three components:



- **Application model:**
  The application model represents the data or objects to be pictured on the screen; it is stored in an application database. The model typically stores descriptions of primitives that define the shape of components of the object, object attributes and connectivity relationships that describe how the components fit together. The model is application-specific and is created independently of any particular display system.
- **Application program:**
  Therefore, the application program must convert a description of the portion of the model to whatever procedure calls or commands the graphics system uses to create an image. This conversion process has two phases. The application program traverses the application database that stores the model to extract the portions to be viewed, using some selection or query system.
- **Graphics system:**
  Second, the extracted geometry is put in a format that can be sent to the graphics system. The application program handles user input. It produces views by sending to the third component, the graphics system, a series of graphics output commands that contain both a detailed geometric description of what is to be viewed and the attributes describing how the objects should appear. The graphics system is responsible for actually producing the picture from the detailed descriptions and for passing the user's input to the application program for processing.

1.4.4 Graphics input/ output hardware

Graphics Hardware – Input: Current input technology provides us with the ubiquitous mouse, the data tablet and the transparent, touch-sensitive panel mounted on the screen. Even fancier input devices that supply, in addition to (x,y) screen location, 3D and higher-dimensional input values, are becoming common, such as **track-balls**, **space balls** or the **data glove**.

Track-balls can be made to sense rotation about the vertical axis in addition to that about the two horizontal axes. However, there is no direct relationship between hand movements with the device and the corresponding movement in 3D space.

A **space-ball** is a rigid sphere containing strain gauges. The user pushes or pulls the sphere in any direction, providing 3D translation and orientation. In this case, the directions of movement correspond to the user's attempts to move the rigid sphere, although the hand does not actually move.

The **data glove** records hand position and orientation as well as finger movements. It is a glove covered with small, lightweight sensors. Each sensor is a short length of fiber-optic cable, with a Light-Emitting Diode (LED) at one end and a photo- transistor at the other. Wearing the data glove, a user can grasp objects, move and rotate them and then release them.

Graphics Hardware –Output: Current output technology uses **raster displays**, which store display primitives in a refresh buffer in terms of their component pixels. The architecture of a raster display is shown in figure below. In some raster displays, there is a hardware display controller that receives and interprets sequences of output commands. In simpler, more common systems, such as those in personal computers, the display controller exists only as a software component of the graphics library package, and the refresh buffer is no more than a piece of the CPU's memory that can be read by the image display subsystem that produces the actual image on the screen.

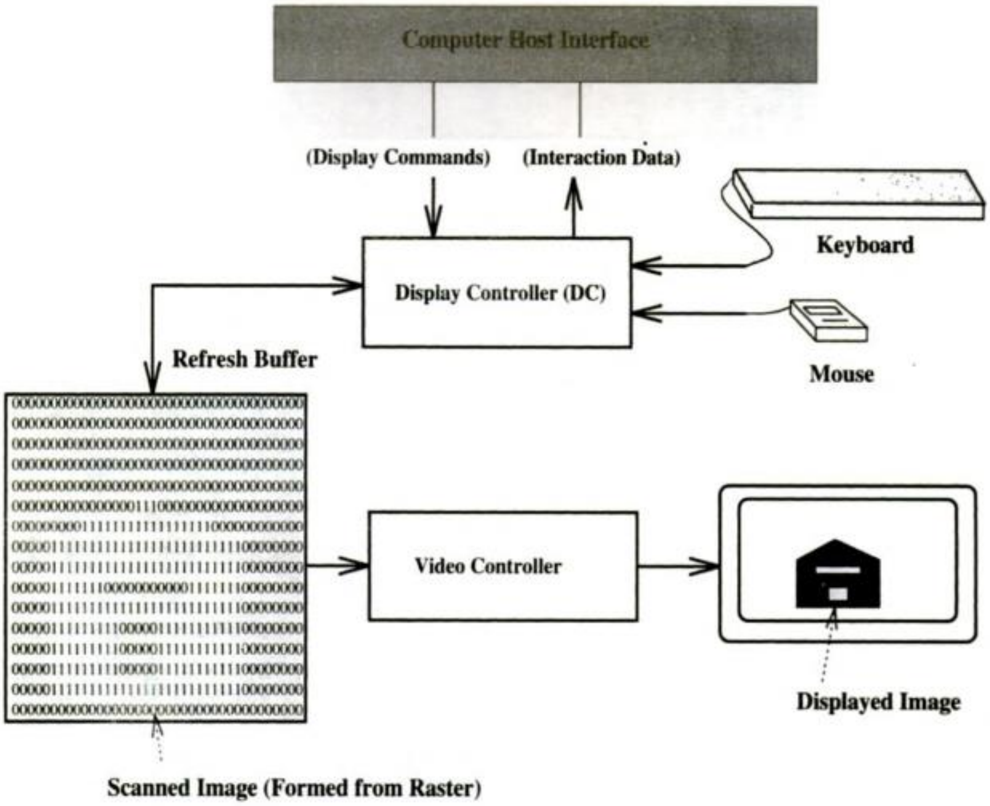


Figure 1: Architecture of Raster Display

The complete image on a raster display is formed from the raster, which is a set of horizontal raster lines, each a row of individual pixels; the raster is thus stored as a matrix of pixels representing the entire screen area. The entire image is scanned out sequentially by the video controller. The raster scan is shown in figure below.

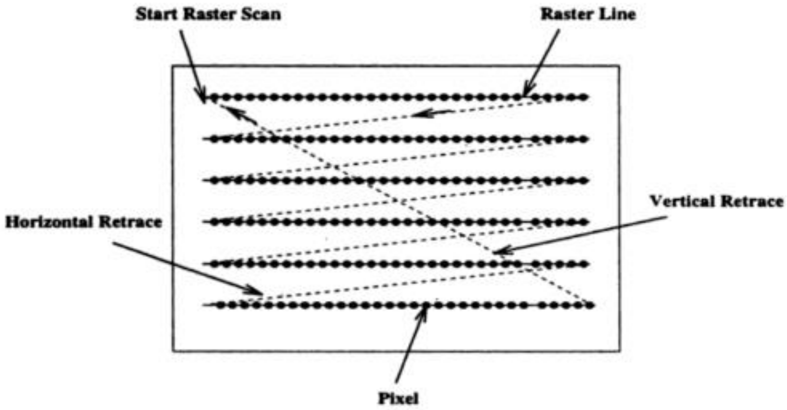


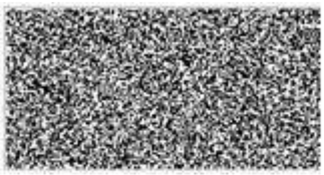
Figure 2: Raster Scan

At each pixel, the beam's intensity is set to reflect the pixel's intensity; in color systems, three beams are controlled - one for each primary color (red, green, blue) as specified by the three color components of each pixel's value.

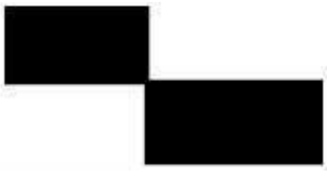
Raster graphics systems have other characteristics. To avoid flickering of the image, a 60 Hz or higher refresh rate is used today; an entire image of 1024 lines of 1024 pixels each must be stored explicitly and a bitmap or pixmap is generated. Raster graphics can display areas filled with solid colors or patterns, i.e., realistic images of 3D objects.

1.4.5 Dithering

Dithering is the process by which we create illusions of the color that are not present actually. It is done by the random arrangement of pixels. For example. Consider this image.



This is an image with only black and white pixels in it. Its pixels are arranged in an order to form another image that is shown below. Note at the arrangement of pixels has been changed, but not the quantity of pixels.



The growth of raster graphics has made color and grayscale an integral part of contemporary computer graphics. The color of an object depends not only on the object itself, but also on the light source illuminating it, on the color of the surrounding area and on the human visual system.

What we see on a black-and-white television set or display monitor is achromatic light. Achromatic light is determined by the attribute quality of light. Quality of light is determined by the intensity and luminance parameters. For example, if we have hardcopy devices or displays which are only bi-levelled, which means they produce just two intensity levels, then we would like to expand the range of available intensity.

The solution lies in our eye's capability for spatial integration. If we view a very small area from a sufficiently large viewing distance, our eyes average fine detail within the small area and record only the overall intensity of the area. This phenomenon is exploited in the technique called **half toning**, or **clustered-dot ordered dithering** (half toning approximation). Each small resolution unit is imprinted with a circle of black ink whose area is proportional to the blackness 1- I (I=intensity) of the area in the original photograph.

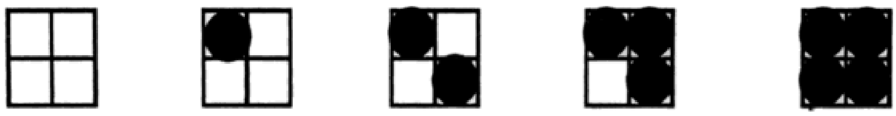


*Figure 3:five intensity levels approximated with two 2*2 dither patterns*

Graphics output devices can approximate the variable- area circles of halftone reproduction. For example, a 2 x 2 pixel area of a bi-level display can be used to produce five different intensity levels at the cost of halving the spatial resolution along each axis. The patterns, shown in above Figure, can be filled by 2 x 2 areas, with the number of 'on' pixels proportional to the desired intensity. The patterns can be represented by the dither matrix. This technique is used on devices which are not able to display individual dots (e.g., laser printers). This means that these devices are poor at reproducing isolated 'on' pixels (the black dots in figure above). All pixels that are 'on' for a particular intensity must be adjacent to other 'on' pixels.

1.4.6 Image Analysis

Image analysis is concerned with techniques for extracting descriptions from images that are necessary for higher-level scene analysis methods. By itself, knowledge of the position and value of any particular pixel almost conveys no information related to the recognition of an object, the description of an object's shape, its position or orientation, the measurement of any distance on the object or whether the object is defective. Hence, image analysis techniques include computation of perceived brightness and color, partial or complete recovery of three-dimensional data in the scene, location of discontinuities corresponding to objects in the scene and characterization of the properties of uniform regions in the image.

Image analysis is important in many areas: aerial surveillance photographs, Scan television images of the moon or of planets gathered from space probes, television images taken from an industrial robot's visual sensor, X-ray images and computerized axial tomography (CAT) scans. Subareas of image processing include image enhancement, pattern detection and recognition and scene analysis and computer vision.

Image enhancement deals with improving image quality by eliminating noise (extraneous or missing pixels) or by enhancing contrast. Pattern detection and recognition deal with detecting and clarifying standard patterns and finding distortions from these patterns. A particularly important example is Optical Character Recognition (OCR) technology, which allows for the economical bulk input of pages of typeset, typewritten or even hand-printed characters.

Scene Analysis and computer vision deal with recognizing and reconstructing 3D models of a scene from several 2D images. An example is an industrial robot sensing the relative sizes, shapes, positions and colors of objects.

1.4.7 Image Recognition

To fully recognize an object in an image means knowing that there is an agreement between the sensory projection and the observed image. Agreement between the observed spatial configuration and the expected sensory projection requires the following capabilities:

- Infer explicitly or implicitly an object's position and orientation from the spatial configuration.
- Confirm that the inference is correct.

To infer an object's (e.g. a cup) position, orientation and category or class from the spatial configuration of gray levels requires the capability to infer which pixels are part of the object. Further, from among those pixels that are part of the object, it requires the capability to distinguish observed object features, such as special markings, lines, curves, surfaces or boundaries (e.g. edges of the cup).

Analytic inference of object shape, position and orientation depends on matching the distinguishing image features (in 2D, a point, line segment or region) with corresponding object features (in 3D, a point, line segment, arc segment, or a curved or planar surface). The kind of object, background, imaging sensor and viewpoint of the sensor determine whether the recognition problem is easy or difficult.

Computer recognition and inspection of objects is, in general, a complex procedure, requiring a variety of steps that successively transform the iconic data into recognition information. A recognition methodology must pay substantial attention to each of the following six steps: image formatting, conditioning, labelling, grouping, extracting and matching.

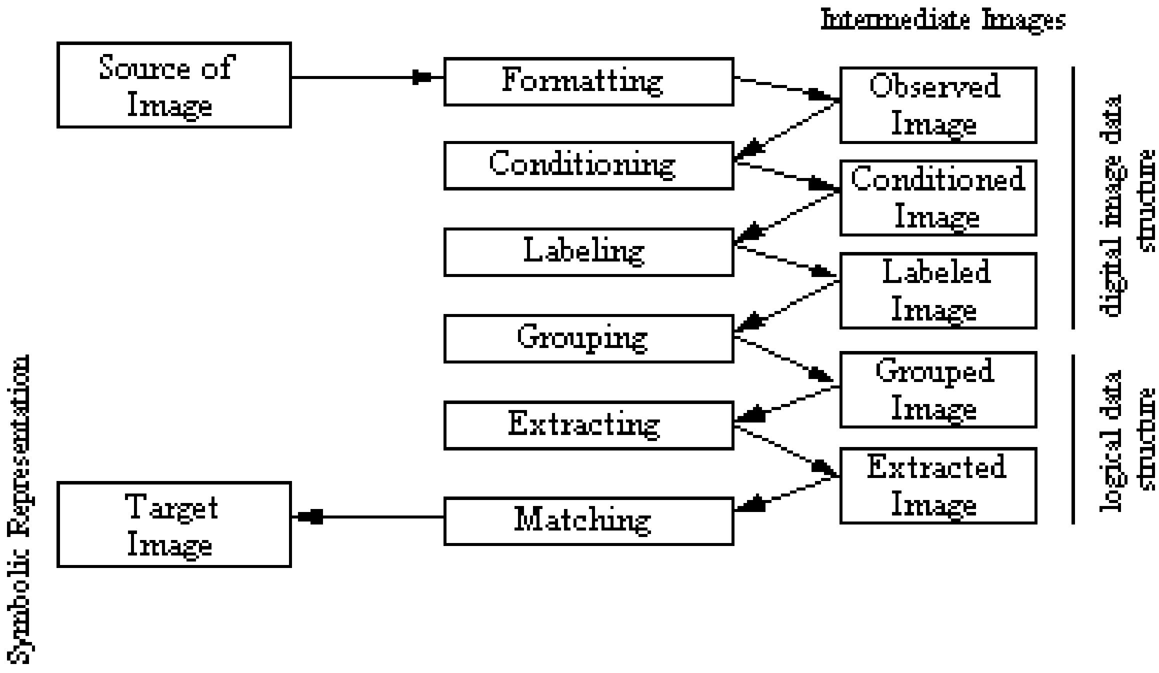


Figure 4: Image Recognition Steps

1. Image Formatting

Image formatting means capturing an image from a camera and bringing it into a digital form. It means that we will have a digital representation of an image in the form of pixels.

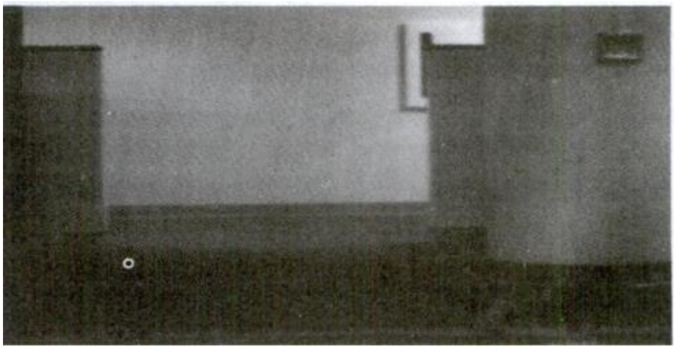


Figure 5: Observed image

Conditioning, labelling, grouping, extracting and matching constitute a canonical decomposition of the image recognition problem, each step preparing and transforming the data to facilitate the next step. As these steps work on any level in the unit transformation process, they prepare the data for the unit transformation, identify the next higher-level unit and interpret it. The five transformation steps, in more detail, are:

2. Conditioning:

Conditioning is based on a model that suggests the observed image is composed of an informative pattern modified by uninteresting variations that typically add to or multiply the informative pattern. Conditioning estimates the informative pattern on the basis of the observed image. Thus, conditioning suppresses noise, which can be thought of as random un-patterned variations affecting all measurements. Conditioning can also perform background normalization by suppressing uninteresting systematic or patterned variations. Conditioning is typically applied uniformly and is context-independent.

3. Labelling:

Labelling is based on a model that suggests the informative pattern has structure as a spatial arrangement of events, each spatial event being a set of connected pixels. Labelling determines in what kinds of spatial events each pixel participates.

An example of a labelling operation is edge detection. Edge detection is an important part of the recognition process. Edge detection techniques find local discontinuities in some image attribute, such as intensity or color (e.g. detection of cup edges). These discontinuities are of interest because they are likely to occur at the boundaries of objects.

An edge is said to occur at a point in the image if some image attribute changes in value discontinuously at that point. Examples are intensity edges. An ideal edge, in one dimension, may be viewed as a step change in intensity; for example, a step between high-valued and low-valued pixels. If the step is detected, the neighbouring high-valued and low-valued pixels are labelled as part of an edge.

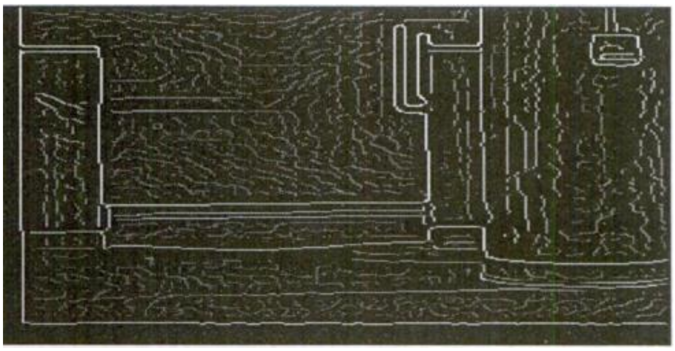


Figure 6: Edge detection of the image

Edge detection recognizes many edges, but not all of them are significant. Therefore, another labelling operation must occur after edge detection, namely thresholding. Thresholding specifies which edges should be accepted and which should not; the thresholding operation filters only the significant edges from the image and labels them. Other edges are removed. Other kinds of labelling operations include corner finding and identification of pixels that participate in various shape primitives.

4. Grouping:

The labelling operation labels the kinds of primitive spatial events in which the pixel participates. The grouping operation identifies the events by collecting together or identifying maximal connected sets of pixels participating in the same kind of event. When the reader recalls the intensity edge detection viewed as a step change in intensity, the edges are labelled as step edges, and the grouping operation constitutes the step edge linking.

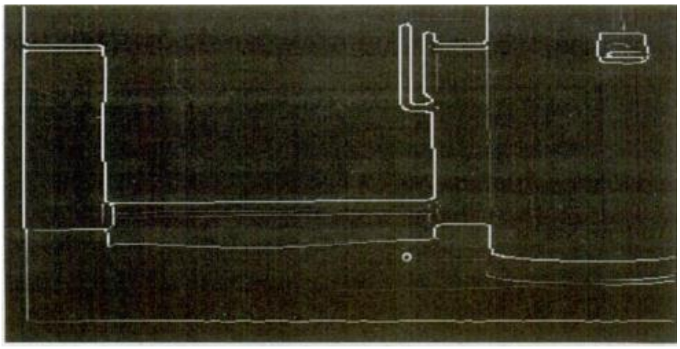


Figure 7: Thresholding the image

A grouping operation, where edges are grouped into lines, is called line-fitting. Again the grouping operation line-fitting is performed on the image shown in Figure below:

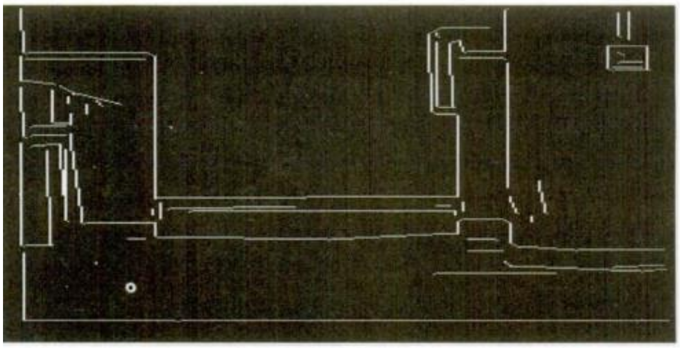


Figure 8: Line-fitting of the image

The grouping operation involves a change of logical data structure. The observed image, the conditioned image and the labelled image are all digital image data structures. Depending on the implementation, the grouping operation can produce either an image data structure in which each pixel is given an index associated with the spatial event to which it belongs or a data structure that is a collection of sets. Each set corresponds to a spatial event and contains the pairs of positions (row, column) that participate in the event. In either case, a change occurs in the logical data structure.

The entities of interest prior to grouping are pixels; the entities of interest after grouping are sets of pixels.

5. Extraction

Extracting the grouping operation determines the new set of entities, but they are left naked in the sense that the only thing they poses is their identity. The extracting operation computes for each group of pixels a list of properties. Example properties might include its centroid, area, orientation, spatial moments, gray tone moments, spatial-gray tone moments, circumscribing circle, inscribing circle, and so on.

Other properties might depend on whether the group is considered a region or an arc. If the group is a region, the number of holes might be a useful property. If the group is an arc, average curvature might be a useful property. Extraction can also measure topological or spatial relationships between two or more groupings. For example, an extracting operation may make explicit that two groupings touch, or are spatially close, or that one grouping is above another.

6. Matching:

After the completion of the extracting operation, the events occurring on the image have been identified and measured, but the events in and of themselves have no meaning. The meaning of the observed spatial events emerges when a perceptual organization has occurred such that a specific set of spatial events in the observed spatial organization clearly constitutes an imaged instance of some previously known object, such as a chair or the letter A. Once an object or set of object parts has been recognized, measurements (such as the distance between two parts, the angle between two lines or the area of an object part) can be made and related to the allowed tolerance, as may be the case in an inspection scenario. It is the matching operation that determines the interpretation of some related set of image events, associating these events with some given three-dimensional object or two-dimensional shape. There are a wide variety of matching operations. The classic example is tem- plate matching, which compares the examined pattern with stored models (templates) of known patterns and chooses the best match.

1.4.8 Image Transmission:

Image transmission takes into account transmission of digital images through computer networks. There are several requirements on the networks when images are transmitted: (1) The network must accommodate bursty data transport because image transmission is bursty (The burst is caused by the large size of the image.); (2) Image transmission requires reliable transport; (3) Time-dependence is not a dominant characteristic of the image in contrast to audio/video transmission.

Image size depends on the image representation format used for transmission. There are several possibilities:

- Raw image data transmission:

In this case, the image is generated through a video digitizer and transmitted in its digital format. The size can be computed in the following manner:

Size = spatial resolution x pixel quantization

For example, the transmission of an image with a resolution of 640 x 480 pixels and pixel quantization of8 bits per pixel requires transmission of 307,200 bytes through the network.

- Compressed image data transmission

In this case, the image is generated through a video digitizer and compressed before transmission. Methods such as JPEG or MPEG, are used to downsize the image. The reduction of image size depends on the compression method and compression rate.

- Symbolic image data transmission

In this case, the image is represented through symbolic data representation as image primitives (e.g., 2D or 3D geometric representation), attributes and other control information. This image representation method is used in computer graphics. Image size is equal to the structure size, which carries the transmitted symbolic information of the image.