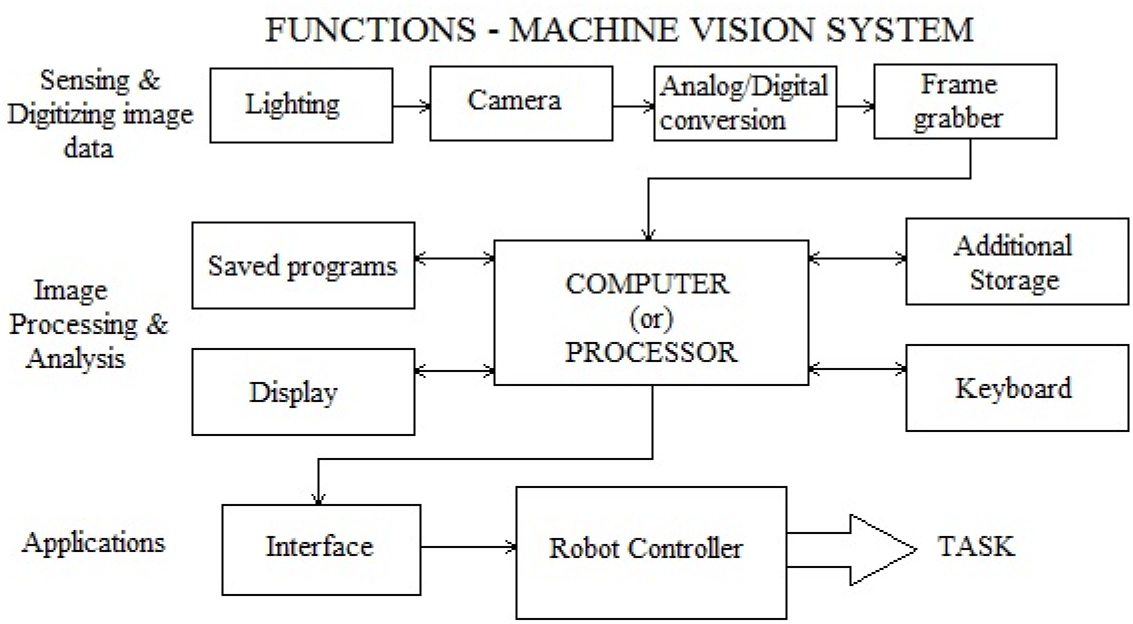
Chapter 2 - Robotic Vision

Robot Vision - Vision of robots requires the ability to identify and accurately determine the positions of all relevant three dimensional objects within the robot work place.

Robot vision may be defined as the process of extracting, characterizing, and interpreting information from images of a three dimensional world.

The robot vision process includes three important tasks, namely:

1. Sensing & Digitizing Image Data
2. Image Processing & Analysis
3. Applications



Sensing & Digitizing Image Data:

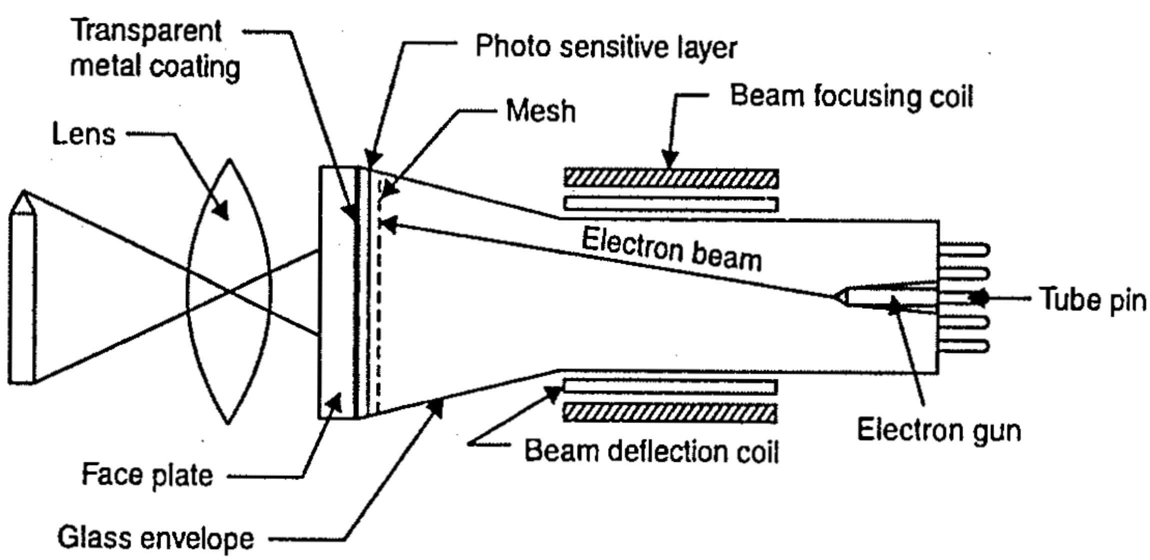
- A camera is used in the sensing and digitizing tasks for viewing the images. It will make use of special lighting methods for gaining better picture contrast.
- These images are changed into the digital form, and it is known as the frame of the vision data.

- A frame grabber is incorporated for taking digitized image continuously at 30 frames per second. Instead of scene projections, every frame is divided as a matrix. By performing sampling operation on the image, the number of pixels can be identified.
- The pixels are generally described by the elements of the matrix. A pixel is decreased to a value for measuring the intensity of light.
- As a result of this process, the intensity of every pixel is changed into the digital value and stored in the computer's memory.

Imaging devices:

There are a variety of commercial imaging devices available. Camera technologies available include the older black and white and vidicon camera and the newer second generation
solid state cameras. Solid state camera used for robot vision include charge coupled devices
(CCD), charge injection devices (CID) and silicon bipolar sensor cameras. For our use two
devices in this subsection, the vidicon camera and the charge coupled devices.

1. Vidicon camera:



Working Principle of Vidicon Camera

- The metal coating of the faceplate is applied with a positive voltage.
- The photosensitive layer acts like a capacitor with negative charge on the inner surface and positive charge on the opposite side, as electron beam strikes.
- The light striking the photosensitive layer reduces the resistance and the current starts flowing and neutralizes the positive charge.
- As the image is formed on the target layer, the concentration of electrons is high in dark area and low in lighted area.
- The electrons so formed flow through metal layer and through the tube pins.
- Variation in current during the electron beam scanning motion produces a video signal proportional to the intensity of input image.

7.15 ANALOG TO DIGITAL CONVERSION

The imaging device like video camera gives analog signal voltage denoting the two dimensional images of the object. This information about the image has to be stored in the bit memory of the computer only on conversion to digital signal. The digital conversion of the analog signal is an approximation of reality with minimum error, done using a analog to digital (A/D) converter. This process has three staged phases :

(1) Sampling

(2) Quantization

(3) Encoding.

• Sampling

Let the function $f(x, y)$ denote the two-dimensional image pattern on the image device. The geometric co-ordinates x and y of the image plane are digitized to get information by the process known as 'image sampling'. After sampling the digitised function $f(x, y)$ in the spatial co-ordinates is generated which can be easily stored in the computer memory.

Assume

N = number of lines in the face plate of the image device like vidicon camera.

S = sampling capability of A/D converter in sec. (sampling capability is the cycle time or frequency needed to convert the analog signal of a pixel by the A/D converter)

R = scanning rate in second, for complete face plate.

R_d = line change over delay for the electron beam.

Hence the scanning rate per line,

$$R_L = \left(\frac{R}{N} + R_d\right) \quad \text{...(7.4)}$$

Number of pixels that can be processed/line,

$$P_n = \left(\frac{R + NR_d}{N.S}\right) \quad \text{...(7.5)}$$

Generally the scan line change over is given in the percentage of the scan rate for a line.

- **Quantization**

The digitization of the amplitude of the image function $f(x, y)$ depending on the intensity of the pixel is known as Quantization.

The number of quantization level,

$$Q = 2^n \quad ...(7.6)$$

where the, n is number of memory bits in the A/D converter.

The quantization level spacing is given as

$$L = \frac{F_r}{Q} \quad ...(7.7)$$

where F_r = full scale range of the camera in volts.

From equations (7.6) and (7.7)

$$L = \frac{F_r}{2^n} \quad ...(7.8)$$

The digital approximation of the analog signal gives the error in quantization as

Quantization error, $$e_q = \pm \frac{1}{2}(L)$$

or $$e_q = \pm \frac{1}{2}\left(\frac{F_r}{2^n}\right) \quad ...(7.9)$$

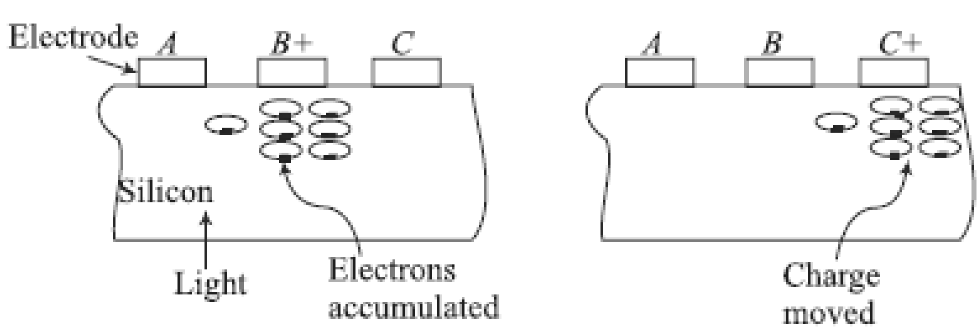
- **Encoding**

Depending on the image created on the faceplate of the camera, the intensity of the different pixel would be different. The conversion into the digital code of the amplitude levels follows the process of quantization. The digitized amplitude code is represented by the binary sequence of digits, which is known as encoding. The spaced quantization levels show the difference in intensities and the amplitude levels. All zeros in the binary sequence of digits represent dark (black) intensity level. All ones in the bit memory is the representative bright (white) intensity pixel. In between the two with combination of zeros and ones shows the gray color.

2. CHARGE-COUPLED DEVICE (CCD):

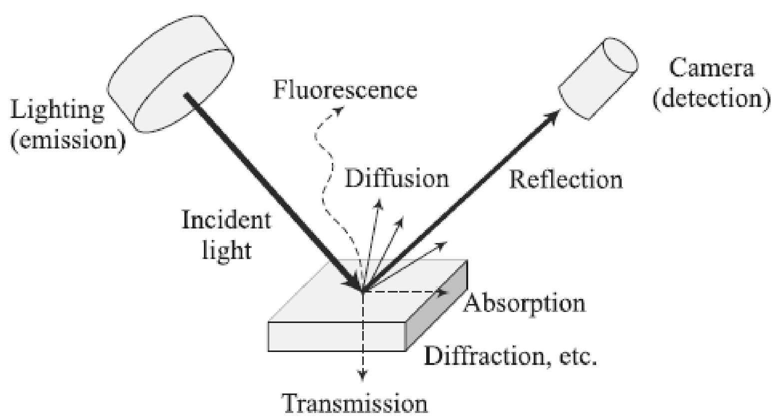
- A charge-coupled device (CCD) is a device for the movement of electrical charge, usually from within the device to an area where the charge can be manipulated, for example conversion into a digital value. This is achieved by "shifting" the signals between stages within the device one at a time.
- CCDs move charge between capacitive bins in the device, with the shift allowing for the transfer of charge between bins. The CCD is a major piece of technology in digital imaging. In a CCD image sensor, pixels are represented by p-doped MOS capacitors.

- These capacitors are biased above the threshold for inversion when image acquisition begins, allowing the conversion of incoming photons into electron charges at the semiconductor-oxide interface; the CCD is then used to read out these charges.
- In a CCD for capturing images, there is a photoactive region (an epitaxial layer of silicon), and a transmission region made out of a shift register (the CCD, properly speaking).
- An image is projected through a lens onto the capacitor array (the photoactive region),causing each capacitor to accumulate an electric charge proportional to the light intensity at that location.
- A one-dimensional array, used in line-scan cameras, captures a single slice of the image, whereas a two-dimensional array, used in video and still cameras, captures a two dimensional picture corresponding to the scene projected onto the focal plane of the sensor.
- Once the array has been exposed to the image, a control circuit causes each capacitor to transfer its contents to its neighbor (operating as a shift register). The last capacitor in the array dumps its charge into a charge amplifier, which converts the charge into a voltage. By repeating this process, the controlling circuit converts the entire contents of the array in the semiconductor to a sequence of voltages.
- In a digital device, these voltages are then sampled, digitized, and usually stored in memory; in an analog device (such as an analog video camera), they are processed into a continuous analog signal (e.g. by feeding the output of the charge amplifier into a low-pass filter), which is then processed and fed out to other circuits for transmission, recording, or other processing.



Lighting Techniques

- It involves radiometry (measurement of the flow and transfer of radiant energy), general illumination models, and surface having both diffuse and specular reflection components.
- Different points on the objects in front of the imaging system will have different intensity values on the image, depending on the amount of incident radiance, how they are illuminated, how they reflect light, how the reflected light is collected by a lens system, and how the sensor camera responds to the incoming light.
- Figure below shows the basic reflection phenomenon. Hence, proper illumination of the scene is important. It also affects the complexity level of the image-processing algorithm required.
- The lighting techniques must avoid reflections and shadow unless they are designed for the purpose of image processing. The main task of lighting is to create contrast between the object features to be detected. Typical lighting techniques are explained below.



\

Light reflections

Direct Incident Lighting This simple lighting technique can be used for non reflective materials which strongly scatter the light due to their matte, porous, fibrous, non-glossy surface. Ideally, a ring light is chosen for smaller illuminated fields that can be arranged around the lens. Shadows are avoided to the greatest extent due to the absolutely vertical illumination. Halogen lamps and large fluorescence illumination can be used too.

<u>Diffuse Incident Lighting</u> Diffused light is necessary for many applications, e.g., to test reflective, polished, glossy, or metallic objects. It is particularly difficult if these surfaces are not glossy, perfectly flat, but individually shaped, wrinkled, curved, or cylindrical. To create diffused lighting, one may use incident light with diffusers, coaxial illumination, i.e. Light is coupled into the axis of the camera by means of a beam splitter or half-mirror, or the dome-shaped illumination where light is diffused by means of a diffused coated dome in which the camera looks through an opening in the dome onto the workpiece.

<u>Lateral Lighting</u> Light from the side can be radiated at a relatively wide or narrow angle. The influence on the camera image can be significant. In an extreme case, the
image information can almost be inverted.

<u>Dark Field Lighting</u> At first sight, images captured using dark field illumination seem unusual to the viewer. The light shines at a shallow angle. According to the principle of angle of incidence equals the angle of reflection, all the light is directed away from the camera. The field of view, therefore, remains dark. Inclined edges, scratches, imprints, slots, and elevations interfere with the beam of light. At these anomalies, the light is reflected towards the camera. Hence, these defects appear bright in the camera image.

<u>Backlighting</u> Transmitted light illumination is the first choice of lighting when it is necessary to measure parts as accurately as possible. The lighting is arranged on the opposite side of the camera, the component itself is put in the light beam.

Image Processing & Analysis:

- In this function, the image interpretation and data reduction processes are done. The threshold of an image frame is developed as a binary image for reducing the data.
- The data reduction will help in converting the frame from raw image data to the feature value data. The feature value data can be calculated via computer

programming. This is performed by matching the image descriptors like size and appearance with the previously stored data on the computer.

- The image processing and analysis function will be made more effective by training the machine vision system regularly. There are several data collected in the training process like length of perimeter, outer & inner diameter, area, and so on.
- Here, the camera will be very helpful to identify the match between the computer models and new objects of feature value data.

Applications:

Some of the important applications of the machine vision system in the robots are:

- Inspection
- Orientation
- Part Identification
- Location

- **Image Data Reduction**

The purpose of image data reduction is to reduce the volume of data either by ellimination of some or part processing, leading to the following sub-techniques.

(*a*) Digital conversion

(*b*) Windowing.

* *Digital conversion* is characterized by reduction in number of gray levels. For a 8-bit register each pixel would have $2^8 = 256$ gray levels. When fewer bits are used to represent pixel intensity the digital conversion is reduced, to suit the requirements.

The data reduction is effected in the following manner generalized as

Total number of bits on the face plate,

$$T_1 = N_r . N_c (2)^n$$

where N_r = number of lines or rows

N_c = number of points per line

2^n = total gray levels.

Binary bit conversion for totally black and white intensities,

$$T_2 = N_c . N_r . (2)$$

Reduction in data volume $= (T_1 - T_2)$

$= 2N_c N_r (2^{n-1} - 1)$

* *Windowing* is processing a portion of the stored digital image. The portion of focus extracted for image processing is the window. A rectangular window is selected as to highlight the component of interest on the screen. The pixels of the faceplate within the window are processed and analized by the computer.

Segmentation:

Segmentation is a general term which is applies to the various methods of data reduction. In segmentation , the objective is to group areas of an image having similar characteristics or features into distinct entities representing the parts of the image.

Example: Boundaries (Edges) or regions (areas) represents two natural segments of an image.

There are many ways to segment an image

1. Thresholding
2. Region growing
3. Edge detection

1. Thresholding

- *Thresholding of image* is one of the principal object detection techniques when the binary image data to be processed is high in volume. Mathematically the threshold function can be defined as

$$T = T[s(x, y), p(x, y), f(x, y)]$$

where $s(x, y)$ = spatial co-ordinate of a point on the screen

$p(x, y)$ = local property at that point.

$f(x, y)$ = the intensity function of the image at the point of analysis.

The comparison of the threshold with that of the intensity decides the type of point on the plate.

when $f(x, y) > T$ $s(x, y)$ is the object point

$f(x, y) < T$ $s(x, y)$ is the background point.

The value of the image threshold,

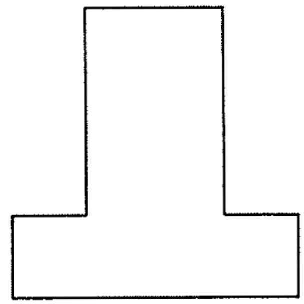
$$g(x, y) = \begin{cases} 1 \text{ if } & f(x, y) > T \\ 0 \text{ if } & f(x, y) \le T \end{cases}$$

Dependability of T	Threshold type
on $f(x, y)$	global threshold
on $f(x, y)$ and $p(x, y)$	local threshold
on $s(x, y)$	dynamic threshold

2. Region growing

- *Region growing* is the processing technique where grid elements processing similar attributes are grouped to form a region. The grid elements are collection of pixels which discretize the object image and the background formed on the faceplate. The properties and the spatial geometric co-ordinates of a region decide, on the process of merging or if to be left independent as a separate entity. The region growing procedure can be better understood by assigning 0s for the background and 1s for the object regions. This is depicted in the Fig. below

0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	1	1	0	0	0	0
0	0	0	0	1	1	1	0	0	0	0
0	0	0	0	1	1	1	0	0	0	0
0	0	0	0	1	1	1	0	0	0	0
0	0	0	0	1	1	1	0	0	0	0
0	0	1	1	1	1	1	1	1.	0	0
0	0	1	1	1	1	1	1	1	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0



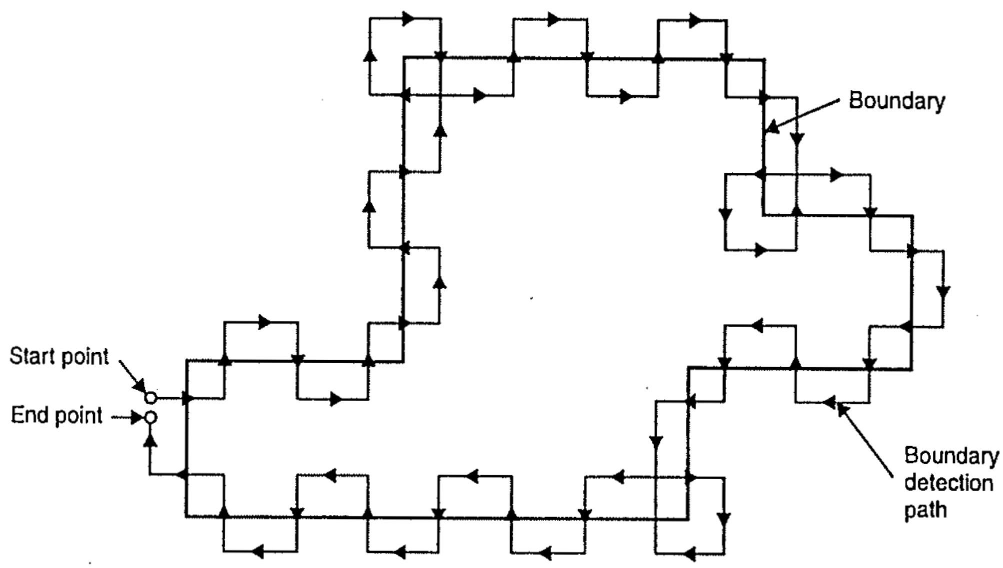
Object image

The region growing procedure can be briefed as follows :

- A pixel on the object is identified and assigned the value 1.
- The adjacent pixel are tracked for match in the attributes. The matching pixel is assigned 1 and non-matching pixel with 0.
- The terms are repeated till the complete screen is covered resulting in growth and identification of region.

3.Edge detection

At boundary the pixels on the faceplate different intensity levels which are stored in the computer in the binary form. This is the distinguishing feature of the object images. The features of similar region, at the edges show demarcation representing change over of the attributes. The edge detection is based on follow-the-edge procedure as shown in the Fig. below The procedure is to scan the pixel within the region, for which turn left and step or otherwise turnright and step from a starting point outside the boundary. This is continued till the end point meets the starting point.



Object Recognition

The next step in image data processing is to identify the object the image represents. The object recognition techniques used in industry today may be classified into two major categories

a) Template matching techniques

b) Structural techniques

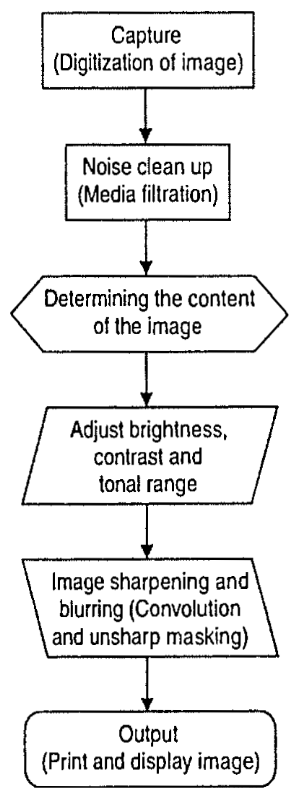
Template matching techniques are a subset of the more general statistical pattern recognition techniques that serve to classify objects in an image into predetermined categories. The basic problem in template matching is to match the object into a stored pattern feature set defined as a model template. These techniques are applicable if there is no requirement for a large number of model templates. When the match is found, allowing for a certain statistical variations in the comparison process. Then the object has been properly classified.

Structural techniques of pattern recognition consider relationships between features or edges of an object. For examples, if an image of an object can be divided into four straight lines (the lines are called primitives)connected at their endpoints and the connected lines are at right angles, then the object is rectangle .

This kind of technique is known as syntactic pattern recognition is the most widely used structural techniques.
Structural techniques differ from decision theoretic techniques in that the later deals with a pattern on a quantitative basis and ignores for the most interrelationship among object primitives.

Training the vision system:
The process of vision system training is to program the vision system with known objects. The system stores these objects in the form of extracted feature values which can be subsequently compared against the corresponding features values from images of unknown objects. Physical parameters such as camera placemet, aperture setting, part position and lighting are the critical conditions that should be simulated as closely as possible during the training vision.

COMPONENTS OF DIGITAL IMAGE PROCESSING



Robot applications:

1. The object can be controlled in both position and appearance
2. Either position or appearance of the object can be controlled but not both.
3. The third level of difficulty requires advanced vision capabilities.
4. Large scale industrial manufacture
5. Short iron unique object manufacture.
6. Inspection of pre manufactured objects.
7. Visual stock control and management systems(counting , barcode reading , store interfaces for digital systems)
8. Control of automated guided vehicles(AGV's)
9. Quality control and refinement of food prodects.
10. Retail automation.
11. Machine vision systems are widely used in semiconductor fabrication. Inspect silicon wafers, processor chips and subcomponents such as resistors and capacitors.
12. In the automotive industry machine vision systems are used to guide industrial robots.

These levels depend on whether the object to be viewed is controlled in position and or appearance. Controlling the position of an object in a manufacturing environment usually requires precise fixturing. Controlling the appearance of an object is accomplished by lighting techniques.

Robot applications of machine vision fall into the three categories

a) Inspection
b) Identification
c) Visual servoing and navigation

Inspection:

The first category is one in which the primary function is the inspection process. This is carried out by the machine vision system and the robot is used in a secondary role to support the applications

The objectives of the machine vision inspection include checking for

Gross surface defects
Discovery of flaws in labeling verification of the presence of components in assembly
Measuring for dimensional accuracy
Checking for presence of holds and other feature in a part.
When these kinds of inspection operation are performed manually, there is a tendency for human
error and also time required in manual inspection operation requires sampling basis. With
machine vision these procedures are carried out automatically using hundred percent inspections
and usually in much less time.

Identification:
The second category identification is concerned with applications in which the purpose of the machine vision system is to recognize and classify an object rather than to inspect it.
Inspection implies the part must be either accepted or rejected. Identification implies that the part involves recognition process in which the part itself or its position and / or orientation is determined. This is usually followed by a subsequent decision and action taken by the robot.
Identification applications of machine vision include
Part sorting
Palletizing
Depalletizing
Picking parts

Visual servoing and navigation:
In the third application category, visual servoing and navigational control the purpose of the vision system is to direct the actions of the robot based on its visual input. The generic example of the robot visual servoing is where the machine vision system is used to control the trajectory of the robots end effector toward an object in the workspace. Industrial example of this applications include part positioning, retrieving parts moving along the conveyor retrieving and

reorienting parts moving along a conveyor, assembly, bin picking and tracking in continuous arc welding.
An example of navigational control would be in automatic robot path planning and collision avoidance using visual data. The bin picking application is an interesting and complex application of machine vision in robotics which involves both identification and servoing. Bin picking involves the use of a robot to grasp and retrieve randomly oriented parts will be overlapping each other.
The vision system must first recognize a target part and its orientation in the container and then it must direct the end effecter to a position to permit grasping and pickup. Solution of the bin picking problem owes much to the pioneering work in vision research at the University of Rhode Island. There are two commercially available bin picking systems one offered by object recognition systems inc called the i-bot 1 system and the other by general electric co called bin vision. Tracking in continuous arc welding is another example of the visual servoing and navigation in robotic vision systems.

EXPECTED QUESTIONS:

- State methods of Analog to Digital conversion and explain any one in detail
- State various image devices used in robot lighting techniques.
- Explain region growing in image processing.
- Explain object recognition techniques.
- Explain feature extraction method of image processing
- Explain the working of Charge Coupled Device.
- Explain robot vision systems with the help of block diagrams.
- Define Robot Vision.
- Explain Vidicon camera with diagram.
- Explain segmentation and explain their types.