Machine vision is inclusion of all industrial and non industrial applications in which, there is combination of hardware and software that provide operational guidance to devices in execution of their functions that are related to capture & processing of images.

industrial machine vision implies

- low cost
- acceptance accuracy.
- high robustness
- high reliability
- high mechanical
- temperature stability.

cost of govnt/military applications is high

machine vision systems rely on digital sensors protected inside industrial cameras with specialized optics to acquire images.

Machine vision excels at quantitative measurement of a structured scene because of its speed, accuracy and repeatibility.

- eliminates maintenance time & costs associated with wear and tear on mechanichal components.

~~Explore the~~

PAGE NO.
DATE

Part location is the critical first step in MV
four major categories of machine vision applications
categories are
- guidance
- identification
- gauging
- inspection

} GIGI

guidance
- It is used to report the location & orientation of a part in 2D or 3D space.
- greater speed and accuracy than manual positioning in tasks.
- Sometimes guidance requires geometric pattern matching.

identification
- identification & recognition reads barcodes (1-D) data matrix codes (2-D), direct part marks (DPM) and characters printed on parts labels, packages.

OCR
- optical character recognition
- reads alphanumeric characters.

OVR
- optical verification recognition.
- confirms presence of character string.

- DPM applications mark a code or character string directly on part.
- technique is used for error-proofing, enabling effecient containment strategies, monitoring process control & quality control metrics, quantifying problematic areas in a plant such as bottlenecks.

Gauging

- A MV system for gauging calculates the distance between two or more points or geometric locations on an object and determines whether these measurement meets specifications.

inspection

- detects defects, contaminants, functional flaws and other irregularities in manufactured products.

Imaging basics - Calculating lens focal length

Q. We want to image an object that is 400 mm from the front of the lens to the object and desire field of view of 90 mm.

a camera uses 1/1.2" format which measures 10.67 mm x 8 mm

Ans: FOV = 90 mm
Working distance = 400 mm

Sensor size : ~~[illegible]~~ 10.67 mm - we will calculate for a 90 mm horizontal FOV, in turn use the horizontal sensor dimension.

$$\text{Focal length} = \frac{\text{Sensor size} \times WD}{FOV}$$

$$= \frac{10.67 \times 400}{90}$$

$$= 47.4\,mm$$

Note: Lenses are only available off the shelf in various focal length (ie 25mm, 35mm, 50mm); so this calculate is theoretical and may need an iteration to adjust working distance.

Alternatively if you have small or large FOV the closest FL lens to your calculation may be suitable.

Constraints while selecting lens:

- Lenses have minimum working distances.
- Lenses need to be paired with appropriate sensor.
  for eg:) if you have 1/2" sensor, you need to ensure you are using a 1/2" format lens or larger.
- Lens should have enough resolution, to resolve pixels on your camera.

Image sensing & aquistion.

Digitization = sampling + quantization

to convert analog image to digital image you need digitization.

Sampling deals with coordinates.
Quantization deals with amplitude.

sampling is on x-axis
Quantization is on y-axis

More the samples, better is the image quality.
Quantization removes fluctuations, and rounds off the values of samples. (i.e int form)

Types of sampling

uniform
- complete image is equally sampled.

non uniform
- only picks up few samples.

Sampling
- There are finite no of pixels. (sampling - spatial resolution)
- amplitude of pixel is represented by a finite no of bits (quantization - gray scale resolution)

0 black
255 white

(-) 256 X 256 → 65000 pixels → picture quality is always unacceptable

(-) 640 X 480 → · low end on most real cameras
→ ideal for emailing pictures
or posting pictures on website.

(-) 1216 X 912 - megapixel - 1109000 pixels
- good for printing pictures

(-) 1600 X 1200 - 2×10^6 pixels
- high res
- can print 4X5 inch

(-) 2240 X 1680 - found on 4 megapixel
- allows larger printed photos with good quality.
- 16 X 20 inches

(-) 4064 X 2704 - 11.1 megapixels
- 13.5 X 9 inch image

Digital cameras
- CCD (charge coupled devices)
- CMOS (complementary metal oxide semiconductor)

CCD and CMOS contains million of photosites (pixels). So these photosites convert the incoming light into charge or electrons.

light → charge conversion → charge accumulation

↓

Transfer for further processing

↓

Charge is converted to voltage.

↓

amplified used amplifiers

These are the common steps you see in a sensors but this sequence might vary in CCD & CMOS.

System integration

CCD is very old technology.
It is not ~~feasible to~~ possible to integrate peripheral components like timers, A/D converters in main sensor.

CMOS, its fabrication ~~process~~ procedure is similar to those ~~peripheral components~~ of integrated circuit.
- It is possible to integrate peripheral componouts into the single chip.
- It is possible to have camera on chip or system on chip.
- CMOS is quite compact.

Power consumption

- CCD requires different power supplies
- typical voltage is 7V to 10V.
- overall power consumption of CCD is very high.

- CMOS requires single power supply
- Typical voltage is 3.3 to 5V
- Overall power consumption is very less.

Proccessing speed

CCD:
- speed is less as charge converting to voltage will ~~be lesser compared to~~ be one by one.
- we can use multiple shift registers to increase speed, but hardware need increases

PAGE NO
DATE

- In CMOS, charge to voltage conversion is carried out in same pixel, so higher speed.
- we can increase speed in CMOS by multiple column select lines. ~~are~~

Noise & sensitivity

- CMOS' sensitivity is less compared to CCD due to ~~due to~~ which CCD has high dynamic range
- In CMOS' not only that amplifiers used in each pixel is not identical, you will see non uniform amplification, and that will act as additional noise.

image distortion

- In CCD there is effect of blooming, if ~~photopixels~~ photosites are exposed to light for more longer time.
- In CMOS there is distortion called rolling shutter, as it reads pixels line by line, so if objects are moving faster it can cause rolling shutter issue.

not an issue in CCD as ~~it~~ all pixels are read or exposed at the same time. (i.e global shutter)

What is MV?

- automatic extraction of information from digital images.

• faster
• consistent
• works for longer period of time

Machine vision common uses

- measurement
- counting
- decoding
- location

Components of MV system

- camera : is responsible for taking light information through use of CCD / CMOS sensor and converting that to digital information which is pixels.

- Optics : lens res & camera resolution are matching.
  - ~~picking~~ : when you pair a lens with camera make sure that you are using compatible mounts. (C or CS mounts)

but before the light falls on camera, it needs to be focused on sensor and that's job of optics.

- illumination - imp aspect of your vision system
  - here again the choice of right type of light is very critical.
  - choose the right light which gives you best consistency & best contrast.

These three components are most imp part of your vision system.

Why use MV?

- to save money & save profitibability
  - increase yield
  - track & trace
  - reduce ~~effects~~ defects
  - comply with regulations.
- they never get tired.
- faster.
- consistentency & repeatable

low light → dark pixels
bright light → bright pixels

Key parts of vision system

- lighting
- lens
- sensor
- vision processing
- communication

- lighting is important part as it illuminates part to be inspected allowing its features to standout so camera can clearly see them.

- lens : captures image present it to sensor in form of light.

sensor : converts this light into digital image, which is sent to processor for analysis

vission processing : review the image and extract required information then system will run whatever measurements or other processes it has been instructed to run.

communication : the resulting data is communicated out to the world in useful manner.
. It is communicated by either discrete I/O signal or data sent over a serial connection to a device that is logging info or using it.

vision processing consists of algos

i.e. acquire image → pre processing → analysis → geometry & tolerance → results.
(to optimize the image and ensure all necessary features stand out)

for detection of minimum features of an object in machine vision we atleast need an object resolution of atleast 10 pixels / mm

e.g.) we want 5 pixels / 10 um then what is camera resolution.

$$\frac{5 \text{ pix}}{10 \text{ um}} = \frac{5 \text{ pix}}{10 \times 10^{-3} \times \text{mm}} = 500 \text{ pix/mm}$$

500 pix/mm → object resolution

given FOV = 70 mm

Hence we would need 500 × 70 = 35000 pixels for height

short exposure : ~~sharp~~ image
long " : blur image

if ratio is 3 : 4 then

$$\frac{3}{4} = \frac{35000}{?}$$

? = 46668 pixels

Camera resolution = 46668 × 35000 pixels
≃ 47 000 × 35000
= 1645 mega pixels

↓

it is very high, it is not possible, we need to increase object resolution.

$$F.L = \frac{\text{sensor size} \times WD}{FOV}$$

generally we assume WD to be in approx 2 times the FOV.

now 1 inch = 25.4 mm

we can take 2/3 inch sensor

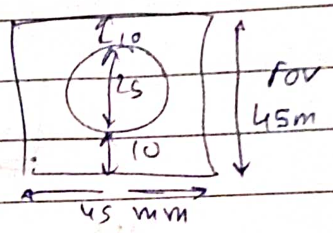
$$\text{sensor size} = \frac{2 \times 25.4}{3} = 50.8 \text{ mm}$$

FOV = 90 mm
WD = 190 mm

(Image segmentation)

indirect lighting works well on glossy surface.

Q If we want 10 pix/mm object resolution, resolution of a side becomes:



45 x 10 = 450 pixels

rso = 640 x 480

sensor size = 2/3 x 25.4

FOV = 45 mm

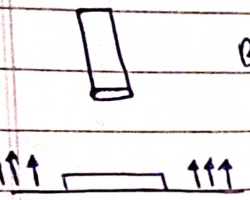
~~WD = 50 mm or 40 mm~~

FL = 18.7mm

nearest is 16mm so we can select that for our applican

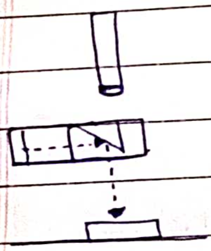
Lighting

- MV systems create images by analyzing the reflected light from object, not by analyzing object itself.



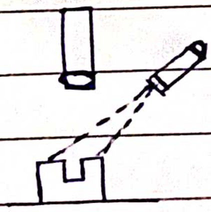
Back lighting : enhances an object's outline for applications that need only external or edge measurement.

- It helps detect shapes & makes dimensional measurements.



axial diffuse lighting

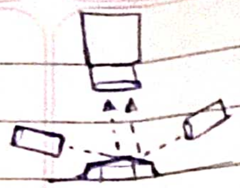
- axial diffuse lighting couples light into optical path from the side.
A semitransparent mirror illuminated from side casts light downwards on the part.
- the part reflects light back through the semi-transparent mirror resulting in ~~very~~ evenly illuminated and homogeneous looking image.



structured light

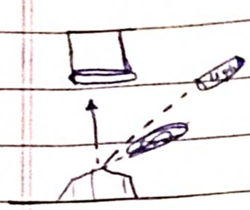
- projection of light pattern at a known angle onto a object.
- useful for contrast-independent surface inspections, acquiring dimensional info and calculating volume

dark field illumination



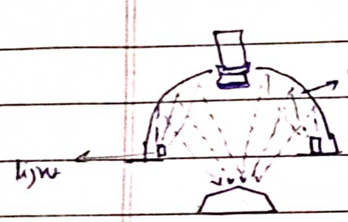
- dark field illumination generally preferred for low contrast applications.
- specular light is reflected away from camera and diffused light from surface texture & elevation changes are reflected into camera.

Bright field illumination



- for high contrast applications.
- given object is not shiny otherwise can create hotspots. to prevent that we need more diffused light source to provide even illumination.

Diffused dome lighting (indirect light)



- ~~for shiny glossy objects~~
- most uniform illumination of features of interest, and can mask irregularities that are not of interest & may be confusing to the scene.

Strobe lighting : for high speed applications
- to freeze moving objects for examination
- use a strobe light also helps to prevent blurring.

Types of MV system

1D

- analyzes a digital signal one line at a time instead of looking at a whole picture at once.
- commonly detects and classifies defects on materials manufactured in a continuous process, such as paper metal, plastic & other non woven sheet or roll goods.

2D

- can be done by 2D snapshots.
- or can be done by vision-line scan-builds a 2D image line by line.

Area scan vs line scan

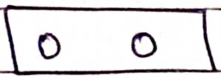
- Line scan systems have specific advantages over area scan.

for eg: inspecting round or cylindrical parts may require multiple area scans to cover entire surface.
However rotating the part in front of a single line scan camera captures entire surface by unwrapping image.

- Line scan fits more easily in tight space.
- Line scan provides higher res than traditional cameras.
- Line scan suitable for products which are in continuous motion.

Algorithms

① Stereo vision

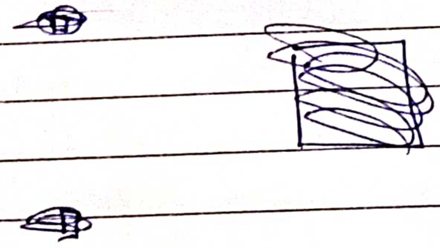


A person with one eye cannot measure depth.

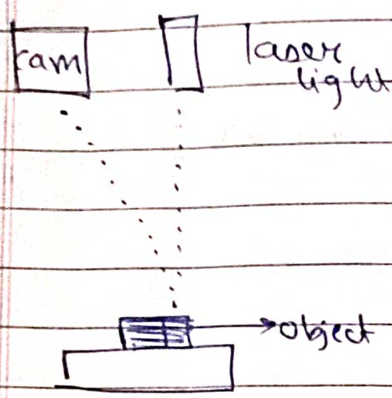
It can be used to find height
- applications in both industrial & commercial scale.
- also used in autonomous car
- ~~for finding rough distance between two objects two objects~~
  - It is used for measuring height & depth but not good for [illegible] distan

LIDAR - can only give information of a point in a line

for navigation stereo + LIDAR is used.

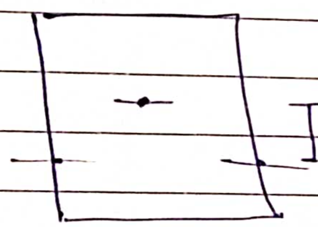


② 3D laser triangulation



cam, laser, and obj are arranged in triangular fashion.

image looks



- works well in units of milimeters or micrometers
- **celant** (care liquid) issue can be solved here.
  transparent

③ Dot projection / fringe pattern projection
- Apple phones projects lot of dots on face, to open the face locks
- It is projecting know pattern on unknow patter and acquire the image.
- It uses infrared lights, so it can be operated at night also.

69 - It's also a kind of triangular fashion, either right angle, isoceles and scalar.

DL|

256 → 80

PAGE NO.

DATE:

- to identify tissues or defects.

100 mm ka object hai

1 line 0.1 mm ka hai

how many lines to capture

$$\frac{1}{x} = \frac{0.1}{100\cancel{0}}$$

$$x = 1000 \text{ lines.}$$

A digital image is an image $f(x,y)$ that has been discretized in spatial coordinate & brightness.

$F(x,y) = I(x,y)\, R(x,y)$

m-adjacency : Two pixels p and q with the values from set 'V' are m-adjacent if

(i) q is in $N_4(p)$

OR

(ii) q is in $N_D(p)$ & the set $N_4(p) \cap N_4(q)$ have no pixels whose values are from 'V'.

e.g.) $V = \{1\}$ set

0_a	1_b	1_c
0_d	1_e	0_f
0_g	0_h	1_i

- b & c are m-adjacent, 1st property.
- b & e are m-adjacent, 1st property.
- e & i are m-adjacent, 2nd property. (as 0_f, 0_h not in set)
- c & e are not m-adjacent, as (1_b ~~& 0_f finds a intersection~~ is in set)

adjacent : p is adjacent to q if they are connected

path : a path from (x,y) to (x_n, y_n) where (x_i, y_i) is adjacent to (x_{i-1}, y_{i-1})

PAGE NO.
DATE:

Connected component

- if $p, q \in S$, p is connected to q in S; if there is a path from p to q consisting entirely of pixels in S.

Euclidean distance

$p(x, y)$, $q(s, t)$

$$[(x-s)^2 + (y-t)^2]^{1/2}$$

City block

$|x-s| + |y-t|$

Chessboard

$\max(|x-s|, |y-t|)$

Intensity transformation

$s = T(r)$

Digital ⇒ values of T typically stored in one-dimensional array + mapping from r to s implemented via table lookups.

Image negatives

$s = (L-1) - r$ ~~where L-1~~

basically in an image $L = 256$

- basically we need it because sometimes in image processing objects are white and background is black, we have the opposite, we need to therefore invert the colors of our image.

image enhancement in spatial domain

Point Processing
- image negative
- log transformation
- powerlaw transformation
- contrast stretching
- bit plane slicing
- histogram equalization
- histogram specification

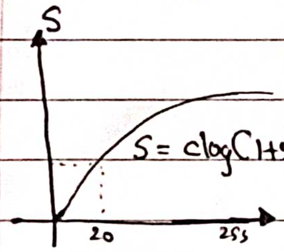
neighbourhood
- smoothening spatial filter
- sharpening spatial
- median
- max-min
- unsharp masking.

log transformations

→ $s = c\log(1+r)$

$r \geq 0$, c = constant

used for expanding dark pixels in an image while compressing the higher level values.



Powerlaw transfor or gamma law

$s = cr^{\gamma}$

also

$s = c(r+\varepsilon)^{\gamma}$

↓ offset

we can do expansion or compression of intensity levels

maps a narrow range of dark input values to wider range.

$\gamma\uparrow$ images darker $\gamma\downarrow$ images lighter

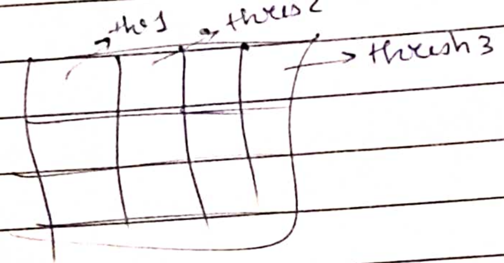
(Image segmentation)

Threshold

- local
  - ~~[illegible]~~ n. black
  - Background correction
- global
  - manual
  - Otsu — depends
  - automatic
    - clustering
    - entropy metric
    - moments.

global op → same for entire image.

local →



for applying otsu in local subimages we apply otsu in subimage.

global → you choose when you have uniformity

local → when you have non uniformity in image

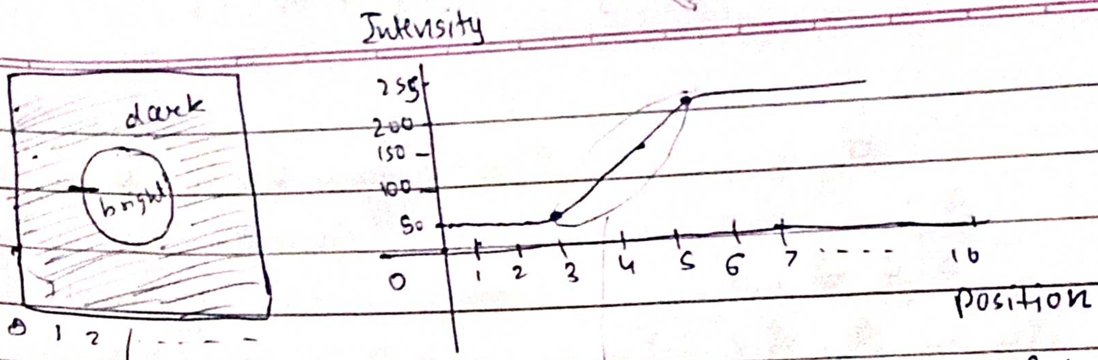
if you have shadow, but is uniformly ~~[illegible]~~ ~~[illegible]~~ distributed in image then global threshold can be used.

filters & spatial (coordinates, intensity)

↳ smoothing:

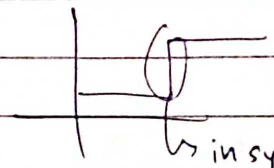
↳ sharpening:

also called local processing.



ramp edge (normal image)

as we are going from background to image.

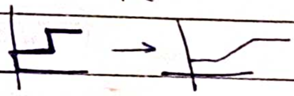


in synthetic or vector images or raster images

images generated by mobile devices

images generated by software applica

so we want to get ~~increase~~ smooth img ~~the step~~ i.e



Sharpening

image ~~xxx~~

$$\begin{bmatrix} 30 & 31 & 35 \\ 40 & 151 & 38 \\ 34 & 33 & 31 \end{bmatrix} \times \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

kernel/mask

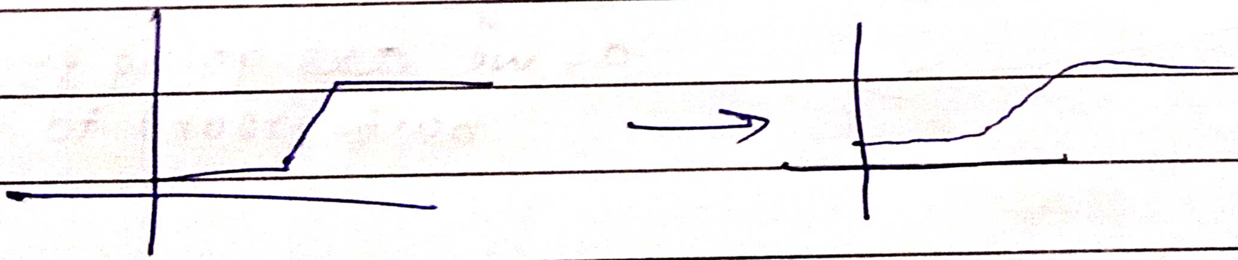
centre pixel $= \overline{30 + 31 + 35 + 40 + 151 + 38 + 34 + 33 + 31}$

$$g(x,y) = \overset{\text{img}}{f(x,y)} * \overset{\text{kernel}}{h(s,t)}$$

(~~3×3 5×5~~) (always odd)

e can see vertical lines because its size greater than kernel size.

the bigger noise you want to smooth, the bigger the kernel should be



(10 Marks