Vismaya Veeramanju Kalyem
CV512 FALL-18
A20423189

1 Geometric image formation

a) $f = 10$

$p = (3, 2, 1)$

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{1}{z} \begin{bmatrix} f & 0 \\ 0 & f \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= 1 \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 30 \\ 20 \end{bmatrix}$$

b) Image plane is behind COP

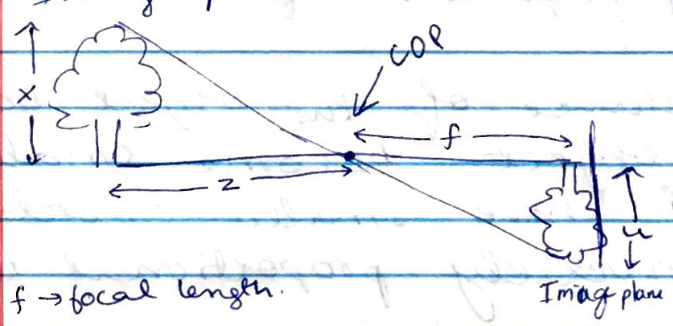
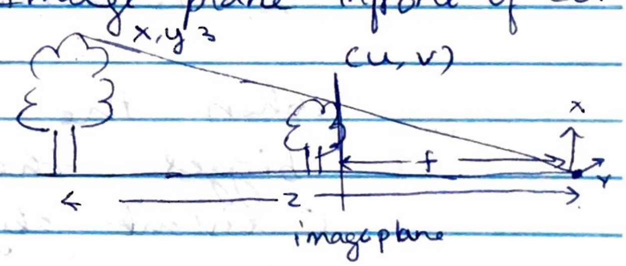


Image plane infront of COP



1. Inverted image $(-f)$
2. projection Equation

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{1}{z} \begin{bmatrix} -f & 0 \\ 0 & -f \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

3. This model is used in real camera where instead of pinhole a len is placed.

1. Upright image $(+f)$
2. $$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{1}{z} \begin{bmatrix} f & 0 \\ 0 & f \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$
3. The model for understanding and easy of using.

conti

* The first model where the image plane is behind the center of ~~o~~ projection corresponds to a ~~better~~ physical pinhole camera model.
* The other model is justified as when we take a image we dont see a inverted image as the software does the rotating before displaying it to us. Hence the model is justified and it is also easy to understand & use.

c) When the focal length becomes bigger the object becomes bigger (Everything becomes magnified)

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{1}{z} \begin{bmatrix} f & 0 \\ 0 & f \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

u, v is directly proportional f.

When the distance of the object becomes bigger the object become smaller (distant object have smaller co-ordinates)
u, v is inversely proportional to z

d)

2D	2DH
(x, y)	(x, y, w)
$(1, 1)$	$(1, 1, 1)$

$\downarrow$

2D point

$(1, 1) \longleftarrow (3, 3, 3)$

other 2DH point

$\left(\frac{4}{4}, \frac{4}{4}\right) = (1, 1) \longleftarrow (4, 4, 4)$

e) 2DH

$(1, 1, 2)$

2D point

$\left(\frac{1}{2}, \frac{1}{2}\right)$

f) Meaning of 2DH point $(1,1,0)$ which points to $\left(\frac{1}{0}, \frac{1}{0}\right)$ a point at infinity.

It can used to ~~represnt~~ represent direction. 0 meaning same direction.

g) ~~F~~ Non-linear projection equation can be made to linear equation in homogeneous coordinates by postponing the division by Z untill the ~~the~~ end

(ie. perspective division done at end).

$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

2DH 3DH

$u = \frac{U}{W} = \frac{fx}{z}$ $v = \frac{V}{W} = \frac{fy}{z}$

h) $M = K[I \mid 0]$

M dimensions ~~3x4~~ 3×4

K 3×3

I 3×3

0 3×1

i)
$$\underset{2DH}{\begin{bmatrix} U \\ V \\ W \end{bmatrix}} = \underbrace{\left[\begin{array}{ccc|c} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]}_{\text{projection matrix}} \underset{3DH}{\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}}$$

$$= \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 1 & 2 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 + 4 + 9 + 4 \\ 5 + 12 + 21 + 8 \\ 1 + 4 + 3 + 2 \end{bmatrix}$$

$$= \begin{bmatrix} 18 \\ 46 \\ 10 \end{bmatrix}$$

$$\frac{U}{W} = \frac{18}{10} = 1.8$$

$$\frac{V}{W} = \frac{46}{10} = 4.6$$

$(u, v) =$ 2D point $(1.8, 4.6)$

2. Modeling transformation

a)
$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & tx \\ 0 & 1 & ty \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

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

$$= \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}$$

2DH - Point
(3, 4, 1)

b)
$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} Sx & 0 & 0 \\ 0 & Sy & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}.$$

2DH Point

(2, 2, 1)

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

$= \begin{bmatrix} \cos 45 & -\sin 45 & 0 \\ \sin 45 & \cos 45 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

$= \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

$= \begin{bmatrix} 0 \\ \sqrt{2} \\ 1 \end{bmatrix}$

d) translate (2,2) to origin

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

rotate by 45°

$\begin{bmatrix} \cos 45 & -\sin 45 & 0 \\ \sin 45 & \cos 45 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -\sqrt{2} \\ 1 \end{bmatrix}$

Translate it back to (2,2)

$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -\sqrt{2} \\ 1 \end{bmatrix} = \cancel{\begin{bmatrix} -\sqrt{2}+2 \\ 2 \\ 1 \end{bmatrix}}, \begin{bmatrix} 2 \\ -\sqrt{2}+2 \\ 1 \end{bmatrix}$

e) $P' = MP$

$= T((RP))$

$= TRP$

f) $\begin{bmatrix} 3, & 0, & 0 \\ 0, & 2, & 0 \\ 0, & 0, & 1 \end{bmatrix}$

The point is scaled 3 times in x-axis & 2 times Y axis.

g) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$

The point is translated 1 times in x-axis and 2 times y-axis.

h) $\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \text{Inverse} \Rightarrow \frac{1}{6} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{bmatrix}$

The matrix that will reverse the effect of the above is

$\begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

i $M = R(45)T(1,2)$

$M^{-1} = (R(45)T(1,2))^{-1}$

$= (T(1,2))^{-1}(R(45))^{-1}$

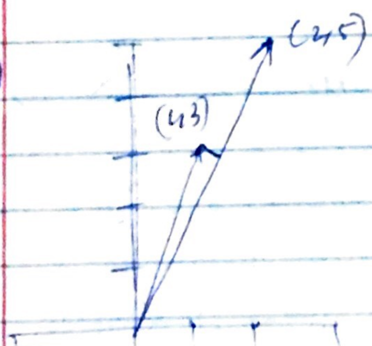
j) $\theta = 90$ point $\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$$

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

$$= \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}$$

k)



$$\text{proj of}_b(a) = \frac{a \cdot b}{|b|^2} \cdot b$$

$$= \frac{17}{29} \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

$$= 0.586 \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

3 a) We need a general projection matrix that uses different coordinate systems for camera and image as we want the projection to consider there is 3D co-ordinates system not jus the camera co-ordinates.

b) project $P-T$ onto $\hat{x}_c, \hat{y}_c, \hat{z}_c$

$$x' = (P-T) \cdot \hat{x}_c = \hat{x}_c^T (P-T)$$
$$y' = (P-T) \cdot \hat{y}_c = \hat{y}_c^T (P-T)$$
$$z' = (P-T) \cdot \hat{z}_c = \hat{z}_c^T (P-T)$$

$\downarrow$

camera co-ordinate

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} - & \hat{x}_c^T & - \\ - & \hat{y}_c^T & - \\ - & \hat{z}_c^T & - \end{bmatrix} (P-T)$$

The transformation between world and camera is obtained by aligning the camera with world.

c) $$R^T \hat{x}_c = \begin{bmatrix} \hat{x}_c^T & \hat{x}_c \\ \hat{y}_c^T & \hat{x}_c \\ \hat{z}_c^T & \hat{x}_c \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = x_w$$

$$R^T \hat{y}_c = \begin{bmatrix} \hat{x}_c^T & \hat{y}_c \\ \hat{y}_c^T & \hat{y}_c \\ \hat{z}_c^T & \hat{y}_c \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = y_w$$

$$R^T \hat{z}_c = \begin{bmatrix} \hat{x}_c^T & \hat{z}_c \\ \hat{y}_c^T & \hat{z}_c \\ \hat{z}_c^T & \hat{z}_c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = z_w.$$

d) $M = \begin{bmatrix} R^* & T^* \\ 0 & 1 \end{bmatrix}$

T^* and R^* are translation of world wrt camera.

$$M_{c \leftarrow w} = \tilde{R}^{-1} \tilde{T}^{-1}$$

$$\begin{bmatrix} R & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} I & T \\ 0 & 1 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} R^T & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I & -T \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} R^T & -R^T T \\ 0 & 1 \end{bmatrix}$$

$$= \left[\begin{array}{c|c} R^* & T^* \\ \hline 0 & 1 \end{array} \right]$$

e) $(u_0 \; v_0) = (512, 512)$

Transformation metrix $\rightarrow$ camera to image

$$M_{c \leftarrow i} = \begin{bmatrix} 1/k_u & 0 & 0 \\ 0 & 1/k_v & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$m_{i \leftarrow c} = (M_{c \leftarrow i})^{-1}$$

$$= \begin{bmatrix} 1 & & -u_0 \\ & 1 & -v_0 \\ & & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1/k_u & & \\ & 1/k_v & \\ & & 1 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} 1 & & u_0 \\ & 1 & v_0 \\ & & 1 \end{bmatrix} \begin{bmatrix} k_u & & \\ & k_v & \\ & & 1 \end{bmatrix}$$

$$\begin{bmatrix} k_u & 0 & u_0 \\ 0 & k_v & V_0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} k_u & 0 & 512 \\ 0 & k_v & 512 \\ 0 & 0 & 1 \end{bmatrix}$$

f) $M = \underbrace{K^*}_{\text{Intrinsic paramter.}} \underbrace{[R^* \mid T^*]}_{\text{extrinsic parameter.}}$

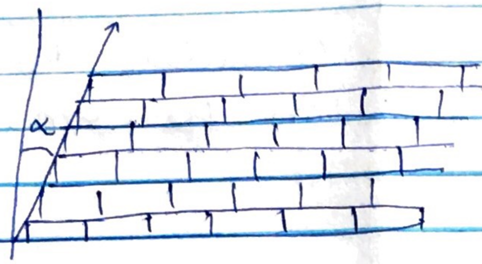
Extrinsic parametuss depend on where you position and move the camera.

Intrinsic parameter.

$$K^* = \begin{bmatrix} f k_u & 0 & u_0 \\ 0 & f k_v & V_0 \\ 0 & 0 & 1 \end{bmatrix}$$

focal length and no. of pixel / mm depends on the camera which are fixed.

g) Reason for adding a skew.



When Image is scanned from sensor, there is a small shift leading to a skewed image. ~~There is~~ Only if you want a very precise camera model you need to add this.

$$M_{skew} = \begin{bmatrix} 1 & \tan\alpha & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

~~Radial lens distortion~~

~~What happens etc~~

h)

$$\begin{bmatrix} 1/\lambda & & \\ & 1/\lambda & \\ & & 1 \end{bmatrix}$$

$$\lambda = 1 + k_1 d + k_2 d^2$$

Linear distortion co-efficient (k_1)　　quadratic distortion co-effient (k_2)

d = distance from center.

as λ is dependent on distance from the center taking into account radial lens distortion adds computational expenses. Once the parameter are determined, the image can be warped to correct distortions

i) Weak perspective camera.

In such camera we dont see much perspective distortion on distant objects.

The sides are parallel.

Since the last row is always [0, 0, 0 1],
there is no perspective division.

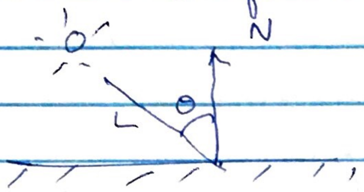
In affine model rather than being projected orthogonally to the plane, they are projected parallel to the line of sight to the object center. ~~[illegible]~~

$$\tilde{x} = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ 0 & 0 & 0 & 1 \end{bmatrix} \tilde{p}$$

4 a)

Surface Radiance

Considers the reflectance on the surface.

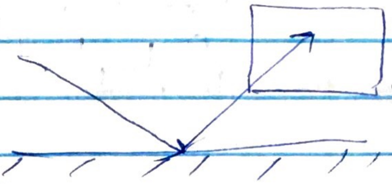


$$I_{ref} = I \rho \cos\theta$$
$$= I \rho \, (N \cdot C)$$

Depends on the albed constant & Intensity of light.

Image irradiance

Considers the reflectance on the camera sensors



$$\underset{\uparrow}{E(p)} = \underset{\uparrow}{L(p)} \frac{\pi}{4} \left(\frac{d}{f}\right)^2 (\cos\alpha)^4$$

light at the image — light at surface.

Depends on the focal length, diameter of lens ~~distance~~ an α (angle between principle axis and surface normal.

b) $$E(p) = L(p) \frac{\pi}{4} \left(\frac{d}{f}\right)^2 (\cos\alpha)^4$$

$E(p)$ → light at image.

$L(p)$ → light at surface.

f → focal length.

d → diameter of lens

$(\cos\alpha)^4$ ← angle b/w principal axis and surface normal.

c) ρ surface albedo $\in [0, 1]$ (reflection co-efficient

It is a intrinsic parameter of the surface that say how good the reflectance of the surface is

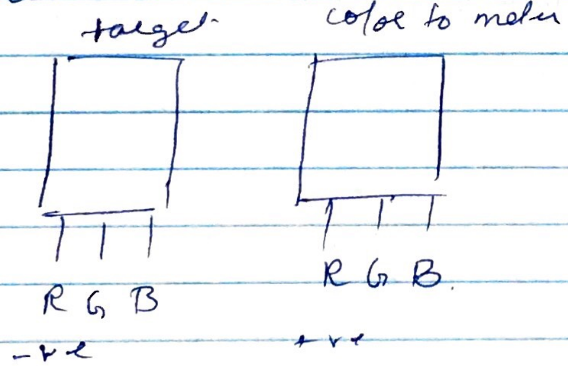
0 being a poor reflector

1 very good reflector.

d) The Eye works in the same way.
It consists of cells that are sensitive Red green and blue, having different combination of these RGB. our brain knows how to make a color. Hence the reason for choosing RGB color model to represent colors.

e) We can see the different shades of gray along the line (0,0,0) & (1,1,1) where (0,0,0) Black being the darkest shade & (1,1,1) white being the lightest shade of gray.

f) Human target where used to map RGB colors to real world colors.



By increasing the nobs of RGB they could match the values to the target and find the values of RGB. They also introduced negative values on the target to help in matching..

g) Y is called the luminance
It gives the intensity of the color
but not the color itself.
Example in Black and white TV only
the luminance is showed.

h) Can be used for shade detection
- Can be used to recognize faces as
the skin tone is similar in shades
- Can be used in detecting the
yellowness in the tooth. (dental surgery)