Assignment 1 Answers

1 a) $f = 10$

$P = (3,2,1)$, be a world point

Coordinates of P when projecting onto image is given by

$$U = \frac{-x}{z} f \quad , \quad V = \frac{-y}{z} f.$$

$$U = -\frac{3}{1} \times 10 \, , \quad V = -\frac{2}{1} \times 10 \quad \Rightarrow (U, V) = (-30, 20)$$

c) projection gets bigger when the focal length gets bigger i.e, projection is directly proportional to focal length.

projection gets smaller as the distance to the object gets bigger i.e, projection is inversly proportional to the distance of the object.

d) 2D point : $(1,1)$

2DH coordinates of same point is $(1,1,1)$ or $(2,2,2)$

e) 2DH point : $(1,1,2)$

corresponding 2D point is : $\left(\frac{1}{2}, \frac{1}{2}\right)$

f) 2DH point $(1,1,0)$ is a point of infinity i.e, giving the direction of a vector in 2DH coordinates, whose length is 1 in both x & y in 2D coordinates.

g) postponing the division by z until when required i.e, until the point where I want to convert the points into corresponding 2D/3D coordinates makes it possible to write the non-linear projection equation as a linear equation in homogenous coordinates.

h) dimensions of

M : 3×4 $\quad$ I : 3×3

K : 3×3 $\quad$ O : 3×1

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad K = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$O = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

i) $m = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 1 & 2 & 1 & 2 \end{bmatrix}$ $\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \underset{3\times4}{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 1 & 2 & 1 & 2 \end{bmatrix}} \underset{3DH}{\begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \end{bmatrix}}$

2DH

$P = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

$$= \begin{bmatrix} 1\times1 + 2\times2 + 3\times3 + 4\times1 \\ 5\times1 + 6\times2 + 7\times3 + 8\times1 \\ 1\times1 + 2\times2 + 1\times3 + 2\times1 \end{bmatrix} = \begin{bmatrix} 18 \\ 46 \\ 10 \end{bmatrix}$$

$$\Rightarrow x = \frac{u}{w}, \quad y = \frac{v}{w}$$

$$x = \frac{18}{10}, \quad y = \frac{46}{10}$$

$$\boxed{(x, y) = (1.8, 4.6)}$$

b) In the pinhole camera model where the image plane is behind the center of projection the image is inverted. And this a real-world model.

$$u = -\frac{x}{z} f, \quad v = -\frac{y}{z} f$$

In the pinhole camera model where the image plane is in front of the center of projection is the image is not inverted and this is a non-real model.

$$u = \frac{x}{z} f, \quad v = \frac{y}{z} f.$$

2) a) $P = (1, 1)$ $\Rightarrow$ $x = 1,\ y = 1$

translate p by $(2, 3)$ $\quad t_x = 2,\ t_y = 3$

translation equation :

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \times 2 + 0 + 2 \times 1 \\ 0 + 1 \times 3 + 3 \times 1 \\ 0 + 0 + 1 \times 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \\ 1 \end{bmatrix}$$

$(1, 1)$, translated by $(2, 3)$ is $(4, 6)$

b) $P = (1, 1)$, Scale p by $(2, 2)$

Scaling equation.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

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

$(1, 1)$, Scaled by $(2, 2)$ is $(2, 2)$

c) $P = (1, 1)$, rotate by 45°

rotation equation :

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

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

$$= \begin{bmatrix} 0.525 & -0.85 & 0 \\ 0.85 & 0.525 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.525 - 0.85 + 0 \\ 0.85 + 0.525 + 0 \\ 0 + 0 + 1 \end{bmatrix}$$

$$= \begin{bmatrix} -0.3259 \\ 1.375 \\ 1 \end{bmatrix}$$

$(1, 1)$, rotated by 45° is $(-0.3259, 1.375)$

d) $P = (1,1)$, rotate by $45°$ about point $(2,2)$

this can be ~~and~~ achieved by

1. translate $(1,1)$ by $(2,2)$
2. rotate the translated point by $45°$
3. translate the rotated point by $(-1,-1)$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

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

$\cos 45° = 0.525$
$\sin 45° = 0.85$

$(1,1)$, translated point is $(3,3)$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos 45° & -\sin 45° & 0 \\ \sin 45° & \cos 45° & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3\times0.525 - 3\times0.85 + 0 \\ 3\times0.85 + 3\times0.525 + 0 \\ 0+0+1 \end{bmatrix}$$

$$= \begin{bmatrix} -0.975 \\ 4.125 \\ 1 \end{bmatrix}$$

$(3,3)$, rotated by $45°$ is $(-0.975, 4.125)$

Now translate $(-0.975, 4.125)$ back i.e, $T^{-1}(1,1)$ or $T(-1,-1)$
(i.e inverse translation)

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -0.975 \\ 4.125 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.975+0-1 \\ 0+4.125-1 \\ 0+0+1 \end{bmatrix} = \begin{bmatrix} -1.975 \\ 3.125 \\ 1 \end{bmatrix}$$

So, $(1,1)$, rotated by $45°$ about point $(2,2)$ is $(-1.975, 3.125)$

e) Combined matrix to be applied on the object is TR

f) $M = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ will scale a point by $(3,2)$

g) $M = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$ will translate a point by $(1,2)$

3 a) The need for a general projection matrix that uses different coordinate systems for camera & image is to transform the object from 3D to 2D.

b) $M_{c \leftarrow w} = T^{-1} R^{-1} \Rightarrow \left[\begin{array}{c|c} R^T & -T \\ \hline 0 & 1 \end{array}\right]$

c) $R = [\hat{x}\ \hat{y}\ \hat{z}]$ where $\hat{x}, \hat{y}$ & $\hat{z}$ are column vectors.

d) R^* & T^* are rotation and translation of world w.r.t camera

e) $M_{i \leftarrow c} = T(U_0, V_0)\ S(K_U, K_V)$

$$= \begin{bmatrix} 1 & 0 & U_0 \\ 0 & 1 & V_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} K_U & 0 & 0 \\ 0 & K_V & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} K_U & 0 & U_0 \\ 0 & K_V & V_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} K_U & 0 & 512 \\ 0 & K_V & 512 \\ 0 & 0 & 1 \end{bmatrix}$$

f) K^* contains intrinsic parameters of camera

$[R^* | T^*]$ contains extrinsic parameters of camera.

g) The reason for adding a 2D skew parameter in the camera model is to make the model more accurate.

h) A scaling parameter $\lambda = 1 + K_1 d^2 + K_2 d^4$ is introduced into the camera model when taking into account the radial lens distortion. This introduces a complication of non-linearity into the model as this parameter, λ, increases as the distance from center increases.

i) Weak perspective camera is a simplified version, which gives a good approximation when there is no much perspective i.e, the depth variation is small compared to the distance of the object from the camera.

Affine camera uses arbitrary coefficients and this model behaves worse than the weak-perspective camera.

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

the transformation matrix that will reverse the effects of transformation by m is m^{-1}

$$M^{-1} = \begin{bmatrix} 0.3333 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

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

inverse of this transformation is

$$m^{-1} = \left(R(45^\circ)\, T(1,2)\right)^{-1}$$

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

$$= T(-1,-2),\ R(-45^\circ)$$

j) Vector perpendicular to $(1,3)$ by (x,y)

$\cancel{x} + 3y = 0$

if $x = 1$, $y = -1/3$

$(1, -1/3)$ is perpendicular to $(1,3)$

k) projection of vector $(1,3)$ onto direction defined by $(2,5)$

Unit vector in the direction of $(2,5) = \left(\frac{2}{\sqrt{2^2+5^2}}, \frac{5}{\sqrt{2^2+5^2}}\right)$

$$= \left(\frac{2}{\sqrt{29}}, \frac{5}{\sqrt{29}}\right)$$

projection $(1,3)$ onto $\left(\frac{2}{\sqrt{29}}, \frac{5}{\sqrt{29}}\right)$

$$= (1 \cdot 3) \cdot \left(\frac{2}{\sqrt{29}}, \frac{5}{\sqrt{29}}\right) = 1 \times \frac{2}{\sqrt{29}} + 3 \times \frac{5}{\sqrt{29}}$$

$$= \frac{2}{\sqrt{29}} + \frac{15}{\sqrt{29}} = \frac{17}{\sqrt{29}}$$

4) a) Scene radiance is the power of light per surface area reflected from the surface, denoted as $L(p)$

Image radiance is the power of light per surface area received at each pixel, denoted as $E(p)$

b) $$\boxed{E(p) = L(p)\,\frac{\pi}{4}\left(\frac{d}{f}\right)(\cos\alpha)^4} \leftarrow \text{fundamental equation of photometric image formation}$$

where f: focal length of camera

d; diameter of lens

α; angle between optical axis & surface normal

$$E(p) \propto L(p)$$
$$\propto d$$
$$\propto \frac{1}{f}$$

c) albedo of the surface is the fraction of the light that the surface reflects.

d) RGB Color model is used to represent colors because human vision/eye have only sensors for Red, green & blue

e) Color along $(0,0,0)$ & $(1,1,1)$ is grey

f) RGB colors are mapped to real-world colors by controlling the amount of R G & B.

g) luminance component Y is used to represent G (green)

h) In LAB color space, the euclidean distance is more representative of perception distance. i.e., when colors are similar in LAB space, they have lesser euclidean distance.