Assignment 1: CS-512 - Computer Vision

Fall 2018

Diego Martin Crespo

A20432558

Question 1: Geometric Image Formation

(a)

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{1}{z}\begin{bmatrix} f & 0 \\ 0 & f \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{1}\begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}\begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 30 \\ 20 \end{bmatrix}$$

(b) The pinhole camera model, where the image plane is behind the center of projection, corresponds better to the physical pinhole camera model because in a camera the image is produced behind the center of projection. The equation is the same for both models but with the addition of a minus sign (from one to another).

(c) Given the image formation equation, (a):
When the focal length (f) gets bigger, the projection gets bigger.
When the distance to the objects (z) gets bigger, the projection gets smaller.

(d) Knowing the vector $(u, v) = (1, 1)$ and that the 2DH vector has the form $(\tilde{u}, \tilde{v}, \tilde{w})$, where $u = \frac{\tilde{u}}{\tilde{w}}$ and $v = \frac{\tilde{v}}{\tilde{w}}$.
Therefore, for instance $(1, 1, 1)$ or $(2, 2, 2)$ 2DH vectors can correspond to $(1, 1)$ 2D vector.

(e) Having $(1, 1, 2)$ 2DH point;

$u = 1/2$

$v = 1/2$

2D point is $(1/2, 1/2) = (0.5, 0.5)$

(f) It represents a vector (is a point in the infinite).

(g) Homogenious coordinates able you to work linearly. You simply postpone the conversion to basic coordinates at the last moment (when you divide by z).

(h) M is 3x4
K is 3x3
J is 3x3
O is 3x1

(i) $$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = M \underbrace{\begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \end{bmatrix}}_{P} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 1 & 2 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 18 \\ 46 \\ 10 \end{bmatrix}$$

$$\text{2D point} = \left\{ \begin{aligned} u &= \frac{U}{W} = \frac{18}{10} = 1.8 \\ v &= \frac{V}{W} = \frac{46}{10} = 4.6 \end{aligned} \right\} = (1.8, 4.6)$$

Question 2: Modeling Transformations

(a) $$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}, \quad (3, 4)$$

(b) $$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \quad (2, 2)$$

(c) $$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos 45 & -\sin 45 & 0 \\ \sin 45 & \cos 45 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \sqrt{2} \\ 1 \end{bmatrix}, \quad (0, \sqrt{2})$$

(d) $R_p(\theta) = T(p) \cdot R(\theta)\, T(-p)$

$$R_p(\theta) = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos 45 & -\sin 45 & 0 \\ \sin 45 & \cos 45 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 2 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 2-2\sqrt{2} \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = R_p(\theta) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2-\sqrt{2} \\ 1 \end{bmatrix}, \quad (2,\ 2-\sqrt{2})$$

(e) $P' = T \cdot R \cdot P$

(f) $M = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, it is scaled by $(3,2)$

(g) $M = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$, it is translated by $(1,2)$

(h) $M' = M^{-1} = \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(i) $M' = M^{-1} = [R(45)\, T(1,2)]^{-1} = T^{-1}(1,2)\, R^{-1}(45) = T(-1,-2)\, R(-45)$

(j) $(1,3) \cdot (u, v) = 0 \rightarrow (3,-1)$

(k) $\bar{u} = (1,3) \quad \bar{v} = (2,5)$

$$\text{proj}\, \bar{u} = \frac{\bar{u} \cdot \bar{v}}{|\bar{v}|^2} \cdot \bar{v} = \frac{1.2 + 3.5}{(\sqrt{2^2+5^2})^2} \cdot (2,5) = \frac{17}{29}(2,5) = (1.17, 1.93)$$

pg 2

<u>Question 3</u> : General Camera Model

(a) For not need to know where the camera is to refer in the world.

(b) $p^{(c)} = R^{-1} \cdot T^{-1} p^{(w)}$

(c) $R = \begin{bmatrix} \hat{x}_1 & \hat{y}_1 & \hat{z}_1 \\ \hat{x}_2 & \hat{y}_2 & \hat{z}_2 \\ \hat{x}_3 & \hat{y}_3 & \hat{z}_3 \end{bmatrix}$

(d) R^* is the rotation matrix from word to camera $R^* = R^T$
T^* is the translation matrix from word to camera $T^* = -R^T T$

(e) $M_{i \leftarrow c} = \begin{bmatrix} K_u & 0 & u_0 \\ 0 & K_v & u_1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} K_u & 0 & 512 \\ 0 & K_v & 512 \\ 0 & 0 & 1 \end{bmatrix}$

(f) → K^* contains the intrinsic parameters:
- focal length (mm)
- scale (pixel/mm)
- translation of optical center (mm)

→ R^* and T^* contains the extrinsic parameters: (d):
- rotation
- translation

(g) It is used to adjust the desviation angle introduced when the camera captures the image. It is a small desviation, but it is better to work with higher precision.

(h) It introduces a new scale to the model. The dificulty is that it is not linear.

(i) A weak prespective camera makes all lines parallel, there is no vanishing point. It is a good approximation when the depth variation is small compared to distance from object.

Question 4: Color and photometric image formation

(a) → Surface variance: power of light per surface are reflected from surface

→ Image irradiance: power of light per surface are received at each field

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

(c) The albedo is the ratio of irradiance reflected to the irradiance received by a surface.

(d) The main purpose of RGB color model is for sensing, representation and display of images in electronic systems. Before the electronic age, the RGB color already had a solid theory based on human perception of colors.

(e) It is the grey scale

(f) the RGB scale was created by comparing the real colours to the resulting mix of the RGB values. It was done empirically

(g) It contains the black (1) and the white (0) information of each pixel.

(h) Its advantage is that in LAB color space euclidean distance is representative of the perception of distance. That is not true RGB.