Assignment_1 Solution

1. Geometric image formation

a. $p = (30, 20)$

b.

- The behind model projects an inverted image; front model projects an upright image.
- The behind model corresponds better to physical pinhole camera model.
- The other model has same equation except image inversion.

c.

- Focal length gets bigger, the projection becomes bigger.
- Distance gets bigger, the projection becomes smaller.

d.

- 2DH: $(1, 1, 1)$
- Another: $k \times (1, 1, 1)$ and $k \neq 0$

e. 2D point: $(1/2, 1/2)$

f. Point at infinity which represents a direction.

g. In homogeneous coordinates:

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

which is a linear equation, and $u = \frac{\alpha u}{\alpha} = \frac{fx}{z}$, $v = \frac{\alpha v}{\alpha} = \frac{fy}{z}$. Homogeneous coordinates help us postpone the z division until when we want to move back to 2D.

h. Dimension of

- M is 3×4,
- K is 3×3,
- I is 3×3,
- O is 3×1.

i. $p = (1.8, 4.6)$

---

2. Modeling transformations

a. $(3, 4)$

b. $(2, 2)$

c. $(0, \sqrt{2})$

d. $(2, 2 - \sqrt{2})$

e. TR

f. p is scaled by $(3, 2)$ about the origin.

g. p is translated by $(1, 2)$.

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

i. $T(-1,-2)R(-45)$

j. Find (x, y) such that $x + 3y = 0 \Rightarrow$ One solution can be $(3, -1)$ or $k \times (3, -1)$.

k. Vector projection: $(34/29, 85/29) \approx (1.172, 2.931)$ OR scalar projection: $17/\sqrt{29} \approx 3.157$

---

3. General camera model

a. The general projection matrix can help transform objects from different coordinate systems in different situations.

b. $R^T_{4\times4}T_{4\times4}(-t)$ or $\begin{bmatrix} R^T_{3\times3} & -R^T_{3\times3}t_{3\times1} \\ 0 & 1 \end{bmatrix}$

c. $R = [\hat{x}, \hat{y}, \hat{z}]$

d.

- R^*: rotation of the world with respect to the camera.
- T^*: translation of the world with respect to the camera.

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

f.

- K^* contains intrinsic parameters,
- $[R^*|T^*]$ contains extrinsic parameters.

g. Including a 2D skew parameter makes the camera model more accurate.

h.

- The location of the original pixel changed in a non-linear way. The straight lines become curved.
- The camera model more scale away from center.

i.

- Weak-perspective camera: M_∞ is approximation to perspective camera's matrix M where the last raw is $[0, 0, 0, 1]$. The parallel line of object appears to parallel each other.
- Affine camera: M_{affine} is a special case of projective camera and is a computational model. $M_{\text{affine}} = \left[\begin{array}{ccc|c} a & b & c & d \\ e & f & g & h \\ \hline 0 & 0 & 0 & 1 \end{array}\right] = \begin{bmatrix} A & t \\ \mathbf{0}^T & 1 \end{bmatrix}$, where A is an arbitrary rank-2 2×3 matrix and t is an arbitrary vector in $\mathbb{R}^2$. The first two rows are more loose compared with weak-perspective camera.

---

4. Color and photometric image formation

a.

- The surface radiance is light in the scene. (reflected from the surface)
- The image irradiance is light in the image. (received at the image)

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

c. Surface albedo is the reflection coefficient and measures how well the surface reflect the light. It is defined as the ratio of irradiance reflected to the irradiance received by a surface.

d. That’s how human perceive colors.

e. Shades of gray.

f. Mapping can be done using CIE conversion.

g. Y represents the relative luminance, i.e., perceived relative brightness, of the color as perceived by human eye. Y can be used to display the grayscale image.

h. LAB color is approximate to human vision. Colorimetric distances between the individual colors correspond to perceived color differences.