3D Computer Vision HW3

Lachin

November 2023

1

Intrinsic Camera Properties

These properties are inherent to the camera and its lens:

- **Focal Length:** The distance between the lens and the image sensor when the subject is in focus. Determines field of view and magnification.
- **Sensor Size:** Physical dimensions of the camera's image sensor.
- **Aperture Size:** The opening in the lens through which light passes, expressed as f-numbers (f/2.8, f/5.6, etc.). Larger apertures allow more light and decrease depth of field.
- **Lens Distortion:** Includes barrel, pincushion, and mustache distortions, affecting the straightness of lines.
- **Pixel Size:** Size of individual pixels on the sensor, affecting resolution and light sensitivity.

Extrinsic Camera Properties

These properties relate to the camera's position and orientation in space:

- **Position:** The physical location of the camera in space.
- **Orientation:** The direction in which the camera is pointed, described in terms of pitch, roll, and yaw angles.
- **Field of View:** The extent of the observable world seen at any given moment. Influenced by both intrinsic and extrinsic factors.

Changes in Parameters

Moving the Position of the Camera in the Scene: Changes Extrinsic Parameters, Affects camera's position and orientation.

Changing the Optic (Lens) of the Camera: Changes Intrinsic Parameters, Affects focal length, aperture size, lens distortion, and potentially field of view.

2

For converting a point from the world coordinate system to the camera coordinate system, we use the following relation:

$$\mathbf{P}_{\text{camera}} = R \cdot \mathbf{P}_{\text{world}} + \mathbf{t}$$

where R is the rotation matrix, $\mathbf{t}$ is the translation vector, $\mathbf{P}_{\text{world}}$ is the point in world coordinates, and $\mathbf{P}_{\text{camera}}$ is the point in camera coordinates.

Point $[25, 40, 50]$ without Rotation

Given:

$$R = \begin{bmatrix} 0.1 & 0.5 & 0.3 \\ 0.6 & 0.1 & 0.2 \\ 0.4 & 0.5 & 0.8 \end{bmatrix}, \quad \mathbf{t} = \begin{bmatrix} 10 \\ 20 \\ 30 \end{bmatrix}, \quad \mathbf{P}_{\text{world}} = \begin{bmatrix} 25 \\ 40 \\ 50 \end{bmatrix}$$

Calculating $\mathbf{P}_{\text{camera}}$:

$$\mathbf{P}_{\text{camera}} = R \cdot \mathbf{P}_{\text{world}} + \mathbf{t}$$

$$\mathbf{P}_{\text{camera}} = \begin{bmatrix} 0.1 \cdot 25 + 0.5 \cdot 40 + 0.3 \cdot 50 \\ 0.6 \cdot 25 + 0.1 \cdot 40 + 0.2 \cdot 50 \\ 0.4 \cdot 25 + 0.5 \cdot 40 + 0.8 \cdot 50 \end{bmatrix} + \begin{bmatrix} 10 \\ 20 \\ 30 \end{bmatrix}$$

$$\mathbf{P}_{\text{camera}} == \begin{bmatrix} 47.5 \\ 49 \\ 100 \end{bmatrix}$$

Point $[50, 80, 100]$ with Rotation

For rotations around the x-axis by 45° and the y-axis by 45°, we define rotation matrices R_x and R_y respectively:

$$R_x(\theta_x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_x) & -\sin(\theta_x) \\ 0 & \sin(\theta_x) & \cos(\theta_x) \end{bmatrix}$$

$$R_y(\theta_y) = \begin{bmatrix} \cos(\theta_y) & 0 & \sin(\theta_y) \\ 0 & 1 & 0 \\ -\sin(\theta_y) & 0 & \cos(\theta_y) \end{bmatrix}$$

where $\theta_x = \theta_y = 45°$ or $\frac{\pi}{4}$ radians.

The combined rotation matrix R_{new} is computed as:

$$R_{\text{new}} = R_y(\theta_y) \cdot R_x(\theta_x) \cdot R$$

Substituting $\cos(45°) = \sin(45°) = \frac{\sqrt{2}}{2}$ we get:

$$R_{\text{new}} = \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \\ -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \cdot R$$

Using this updated rotation matrix, we calculate the camera coordinates for the second point:

$$\mathbf{P}_{\text{camera2}} = R_{\text{new}} \cdot \mathbf{P}_{\text{world2}} + \mathbf{t}$$

where $\mathbf{P}_{\text{world2}} = \begin{bmatrix} 50 \\ 80 \\ 100 \end{bmatrix}$.

The calculated camera coordinates for the second point after rotation are approximately $[162.03, -37.98, 75.97]$.

3

Consider a line L in 3-D space defined by two distinct points $P_1(x_1, y_1, z_1)$ and $P_2(x_2, y_2, z_2)$. Any point P on this line can be represented as a linear combination of P_1 and P_2:

$$P = (1 - t)P_1 + tP_2$$

$$P = (x_1 + t(x_2 - x_1), y_1 + t(y_2 - y_1), z_1 + t(z_2 - z_1))$$

where t is a scalar that varies along the line L.

Using the pinhole camera model, a 3-D point (X, Y, Z) is projected onto a 2-D image plane (x, y) by the projection equations:

$$x = f\frac{X}{Z}$$
$$y = f\frac{Y}{Z}$$

where f is the focal length of the camera.

For any point P on line L, its projection p onto the image plane is:

$$p = \left(f\frac{x_1 + t(x_2 - x_1)}{z_1 + t(z_2 - z_1)}, f\frac{y_1 + t(y_2 - y_1)}{z_1 + t(z_2 - z_1)} \right)$$

As t varies, the projected points p will form a set of points in the image plane. We can observe that:

$$\frac{x_1 + t(x_2 - x_1)}{z_1 + t(z_2 - z_1)}$$
$$\frac{y_1 + t(y_2 - y_1)}{z_1 + t(z_2 - z_1)}$$

are linear functions of t. Thus, as t changes, the ratios change linearly, defining a line in the 2-D image plane.

4

Given the essential matrix H, we want to find the possible rotation matrices R and translation vectors T. The matrix H is given by:

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

The decomposition of the essential matrix is performed using Singular Value Decomposition (SVD), which decomposes H into $U\Sigma V^T$.

The SVD of H is computed as follows:

$$H = U\Sigma V^T$$

With H defined as above, we find U, Σ, and V^T. Using these, we can compute two possible rotation matrices R_1 and R_2, and two possible translation vectors T_1 and T_2, as follows:

$$R_1 = UWV^T, \quad R_2 = UW^TV^T$$

$$T_1 = u_3, \quad T_2 = -u_3$$

where W is a matrix used to enforce the property that a rotation matrix has a determinant of 1, defined as:

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

and u_3 is the third column of U.

The calculated rotation matrices and translation vectors are:

Rotation matrices:

$$R_1 = \begin{bmatrix} 0.7071 & 0 & 0.7071 \\ 0 & 1 & 0 \\ -0.7071 & 0 & 0.7071 \end{bmatrix}$$

$$R_2 = \begin{bmatrix} 0.7071 & 0 & 0.7071 \\ 0 & -1 & 0 \\ 0.7071 & 0 & -0.7071 \end{bmatrix}$$

Translation vectors: $T_1 = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}$ and $T_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$