CS 6384: Computer Vision Homework 1

Yapeng Tian
The University of Texas at Dallas

Problem 1

(3 points)

Suppose a pinhole camera has a camera intrinsic matrix K. Let the camera extrinsics be a 3D rotation R, and a 3D translation $\mathbf{t}$. Given a pixel $(x, y)^T$ in an image, assume the depth of the pixel is d, where depth is the distance between the 3D point of pixel and the camera center. Compute the coordinates of the 3D point in the **world coordinate system**.

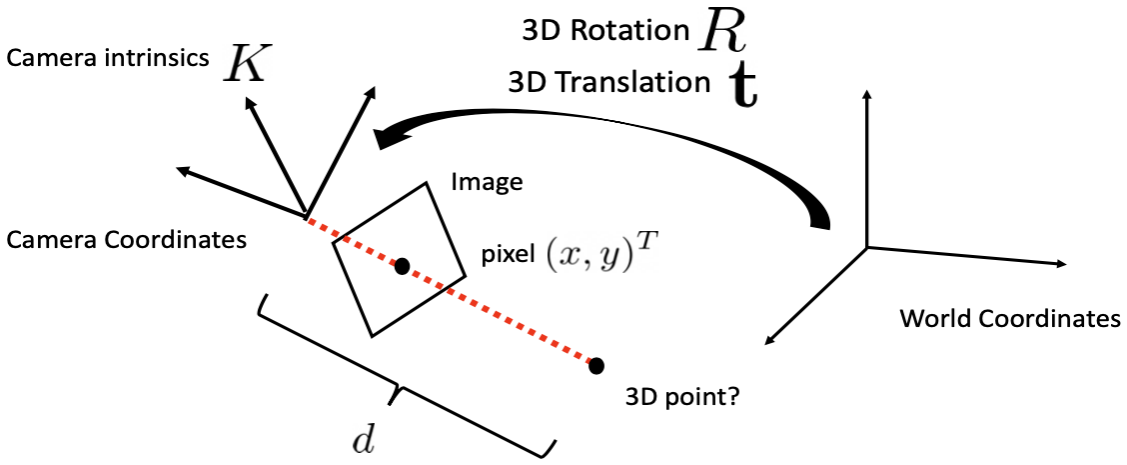

Figure 1: Backprojection of a pixel.

Solution

Given

- Camera intrinsic matrix, K
- Camera extrinsics
  - 3D Rotation R
  - 3D Translation t
- depth d
- and pixel $(x, y)^T$

To calculate pixel coordinates from world coordinates is straight forward.

$$X_{cam} = K [R | t] X_{world}$$

$$x_{pixel} = \begin{bmatrix} f \frac{x_{cam}}{d} + p_x \\ f \frac{y_{cam}}{d} + p_y \end{bmatrix}$$

So we invert this process to calculate world coordinates.

To get X_{cam} we do the following

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f \frac{x_{cam}}{d} + p_x \\ f \frac{y_{cam}}{d} + p_y \end{bmatrix}$$

$$X_{cam} = \begin{bmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \end{bmatrix} = \begin{bmatrix} d\,\frac{x - p_x}{f} \\ d\,\frac{y - p_y}{f} \\ d \end{bmatrix}$$

where p_x and p_y are the principal point offsets.

From before

$$X_{cam} = K\,[R \mid t]\,X_{world}$$

$$K^{-1} X_{cam} = [R \mid t]\,X_{world}$$

$$K^{-1} X_{cam} = R\,X_{world} + t$$

$$R^{-1} K^{-1} X_{cam} = X_{world} + R^{-1} t$$

$$\underline{X_{world} = R^{-1} K^{-1} X_{cam} - R^{-1} t}$$

Given the formula for X_{cam} from the pixel $(x, y)^T$ and the formula for the world coordinates from the camera coordinates you can solve for world coordinates from pixel coordinates.