CS 4391: Introduction to 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**.

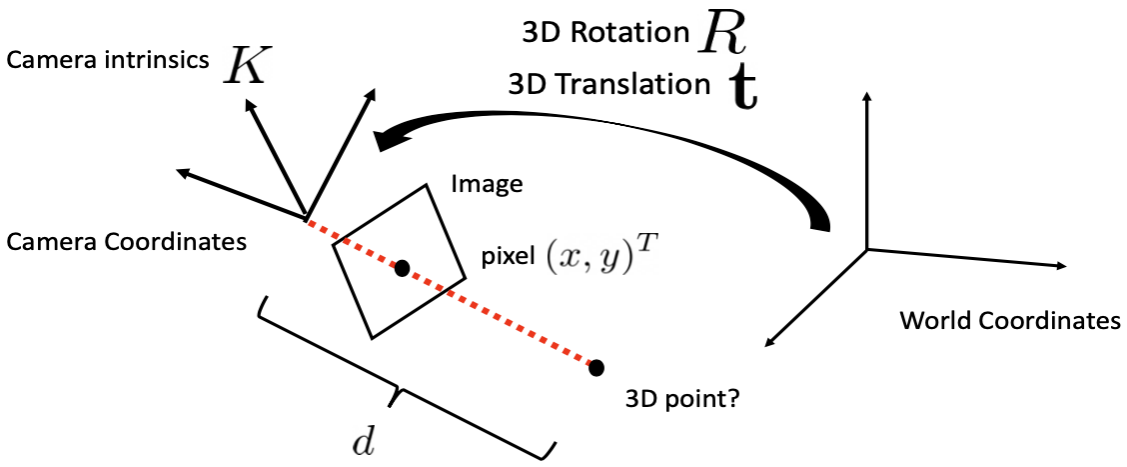

Figure 1: Backprojection of a pixel.

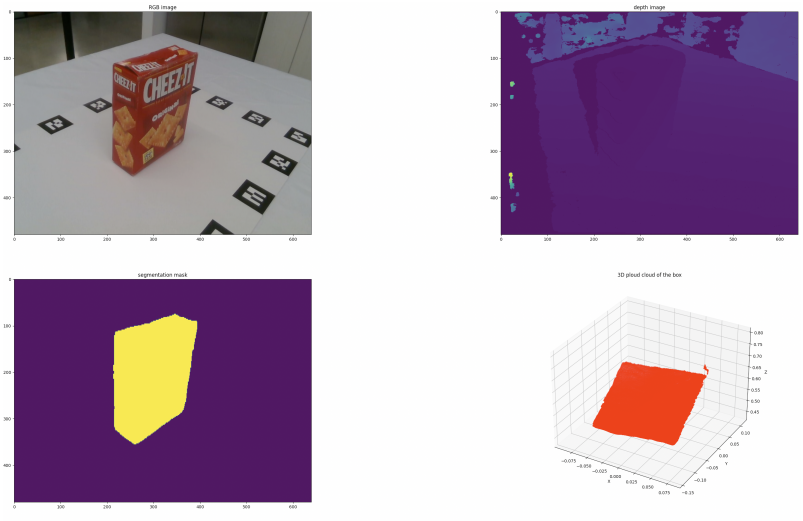

Figure 2: Backprojection.

Problem 2

(3 points)

Download the homework1_programming.zip file from eLearning, Course Homepage, Assignments, Homework 1. Finish the following programming problem and submit your script to eLearning. You can zip all the data and files for submission. Our TA will run your script to verify it.

Install the Python packages needed by

- pip install -r requirement.txt

Implement the backproject() function in backproject.py. The function takes a depth image with height H and width W, and a camera intrinsics matrix as input, and outputs a point cloud with shape $H \times W \times 3$ generated from the depth image.

After your implementation, run the backproject.py in Python to verify it. Figure 2 shows an example of running the script. It shows the 3D points on the cracker box.

Here are some useful resources:

- Python basics `https://pythonbasics.org/`
- Numpy `https://numpy.org/doc/stable/user/basics.html`
- OpenCV `https://docs.opencv.org/4.x/d6/d00/tutorial_py_root.html`