MIT CSAIL

6.869 Advances in Computer Vision

Fall 2019

Problem Set 1 Solutions

Problem 1.2 *Orthographic projection (2pts)*

This can be shown in several ways. A simple way would be to think about orthographic projection as parallel projection of the world onto a plane. Only that in our setup the plane is tilted by θ degrees about the x-axis (clockwise when the x-axis points towards the observer), and a scaling factor α is involved.

To model this projection, we therefore need to rotate the world coordinates by θ degrees (counter-clockwise) about the X-axis, scale and project in the Z direction, and shift it to the camera origin (x_0, y_0). This is achieved using the following transformations

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \tag{1}$$

which yield the required equations

$$x = \alpha X + x_0 \tag{2}$$

$$y = \alpha\left(\cos(\theta)Y - \sin(\theta)Z\right) + y_0 \tag{3}$$

Problem 1.3 *Constraints (2pts)*

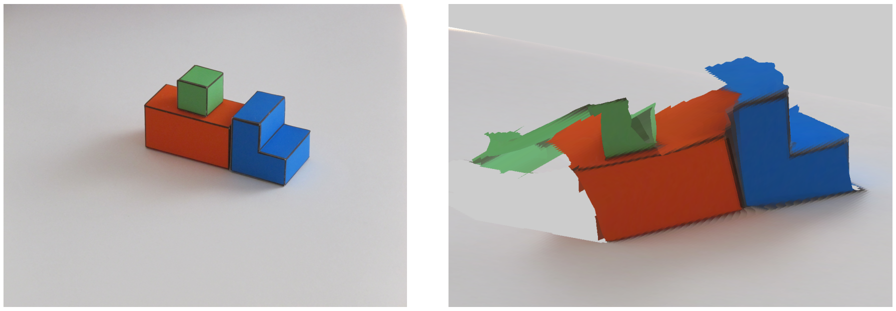
In our simple world we assume edges are either horizontal or vertical. For vertical edges the Z coordinate is constant, and so we can define the constraint $\frac{\partial Z}{\partial y} = 0$ for all points lying on such edges.

The change in the Z coordinate along direction $t = (-n_y, n_x)$ tangent to an edge can be derived from the projection equations as

$$\frac{\partial Z}{\partial t} = \nabla Z \cdot t = -n_y \frac{\partial Z}{\partial x} + n_x \frac{\partial Z}{\partial y} = \frac{-n_x}{\alpha \sin(\theta)} \tag{4}$$

For horizontal edge pixels, we thus set $\frac{\partial Z}{\partial t} = \frac{-n_x}{\alpha \sin(\theta)}$. This equation tells us that the steeper the image edge (recall n_x is the x component of the edge *normal*) – the larger the change in the Z coordinate in the world, which also agrees with our intuition.

For contact edges (edges between objects and the ground plane in our simple world), plugging $Y = 0$ in Equation 3, we get $Z = \frac{y - y_0}{-\alpha \sin(\theta)}$.



(a) Input image (b) Reconstruction

Figure 1: Failure case.

For all other pixels, we use planarity constraints as done in class

$$\frac{\partial^2 Z}{\partial x^2} = 0, \frac{\partial^2 Z}{\partial y^2} = 0, \frac{\partial^2 Z}{\partial y \partial x} = 0 \tag{5}$$

Problem 1.6 *Violating simple world assumptions (2pts)*

Figure 1 shows the 3D reconstruction for one of the images we supplied. The apparent artifacts in this reconstruction are due to (a) occlusion boundaries which are not correctly modeled, producing smooth transitions between the objects, and (b) a misclassified contact edge in the green cube, which pulls the points to the ground plane. Are those problems difficult to fix? How would you improve this result?