Computer Vision Assignment 3

Yating Jing

November 20, 2015

1 Scene Recognition with Bag of Words

6/12 images are classified correctly, the accuracy is 0.5.

2 Stereo Matching and Reconstruction

a.

Setting limit to the maximum disparity yields different results:

If searching the whole scan line:

Figure 1: Maximum disparity = image width

Meanwhile, if limiting the maximum disparity to 50 and 25, then the maps look like:

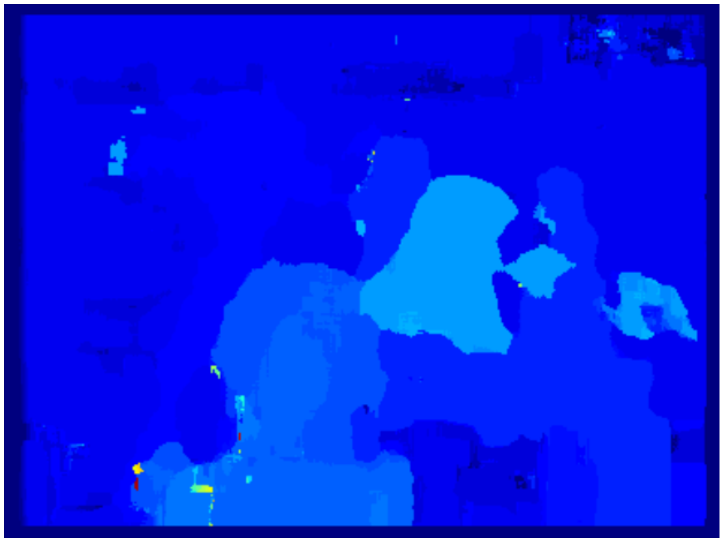

Figure 2: Maximum disparity = 50

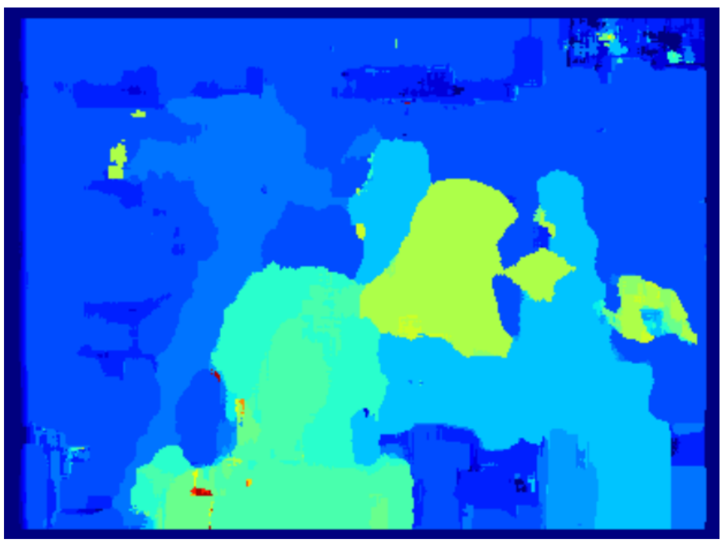

Figure 3: Maximum disparity = 25

b.

For points corresponding to 0 disparity (such as the margins of the depth image), x, y, z coordinates are set to sys.maxint.

3 Projection Matrix

a.

Given calibration matrix K, the focal length in pixel units is 1072 and the image center of the camera $(o_x, o_y) = (500, 390)$.

b.

R represents the orientation of the camera in the world coordinate frame.

$\sin(0.7854) \approx \cos(0.7854) \approx 0.7071$, 0.7854 is approximately 45 degrees, this means that the camera is rotated by about 45 degrees around the world X camera axis.

$\mathbf{t} = -RC$ is the position of the world origin in camera coordinates. According to the definition of P, vector $C = [25, 50, 0]^T$ is the camera center in the world coordinate frame.

c.

The initial rotation matrix is:

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

Now rotate the camera by 10 degrees around world X camera axis, which means that the camera is now rotated by 55 degrees in total around X, $55 * \pi/180 \approx 0.9599$. the new R then becomes:

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(0.9599) & -\sin(0.9599) \\ 0 & \sin(0.9599) & \cos(0.9599) \end{bmatrix}$$

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5736 & -0.8191 \\ 0 & 0.8191 & 0.5736 \end{bmatrix}$$

d.

The initial translation vector $\mathbf{t} = [t_x, t_y, t_z]^T$ is:

$$\mathbf{t} = -RC = \begin{bmatrix} -25 \\ -35.355 \\ -35.355 \end{bmatrix}$$

Now further translate the camera by 10 units along the world positive X axis, which is to add 10 to t_x, then the translation vector becomes:

$$\mathbf{t}' = \begin{bmatrix} -15 \\ -35.355 \\ -35.355 \end{bmatrix}$$

Thus the new vector C becomes:

$$C = \begin{bmatrix} 15 \\ 50 \\ 0 \end{bmatrix}$$

4 Fundamental Matrix and Epipolar Lines

a.

As shown in the figure below, there are quite a few outliers.

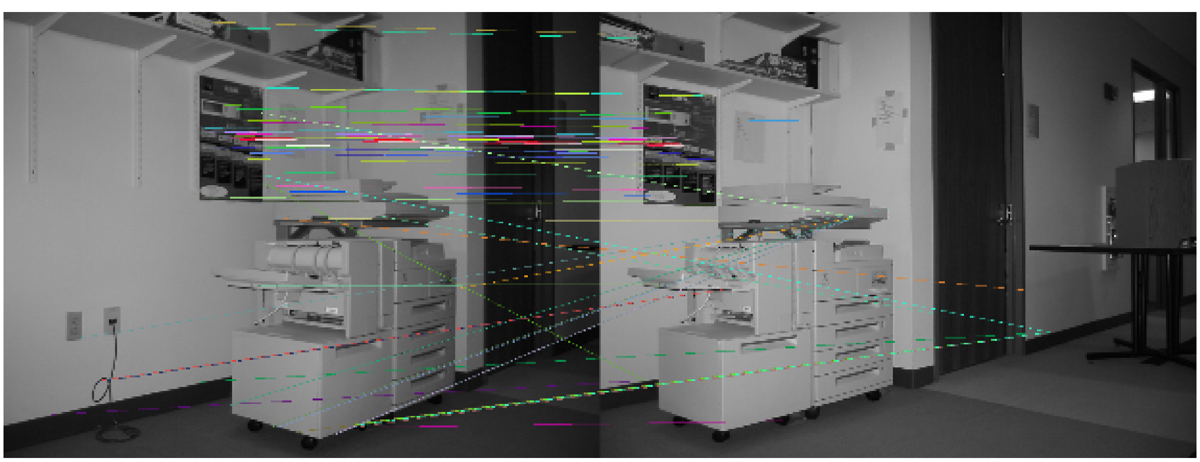

Figure 4: Correspondence

b.

Based on Appendix A of Lecture 15, F is the eigenvector corresponding to the smallest eigenvalue of $A^T A$.

$$F = \begin{bmatrix} -2.04971649e-06 & 3.15469059e-06 & -4.00873984e-06 \\ 2.16092972e-06 & -2.48832456e-03 & 6.56772352e-04 \\ -1.94349197e-03 & -1.24176658e-03 & 9.99994029e-01 \end{bmatrix}$$

References

[1] Dissecting the Camera Matrix: https://ksimek.github.io/2012/08/14/decompose/