Weicong Feng

Professor Zhigang Zhu

CSC 74030-01

03/14/2022

Assignment2

Assignment 2

Due date: Wednesday, 03/15/2022

The work in this assignment is my own. Any outside sources have been properly cited.

1 (Camera Models- 20 points) Prove that the vector from the viewpoint of a pinhole camera to the vanishing point (in the image plane) of a set of 3D parallel lines is parallel to the direction of the parallel lines. Please show the steps of your proof.

Ans: Let $P_0 = (X_0, Y_0, Z_0)^T$ is a fixed point on the line, vector $V = (a, b, c)^T$ represents the direction of the line. For a 3D line $P = P_0 + tV$, where $P = (X, Y, Z)^T$ is any point on the line.

First, we need to find the vanishing point in the image plane:

$$x_{van} = f\frac{(X_0 + ta)}{Z_0 + tc} \overset{t\to\infty}{\Longrightarrow} x_{van} = f\frac{ta}{tc} = \frac{fa}{c}$$

$$y_{van} = f\frac{(Y_0 + tb)}{Z_0 + tc} \overset{t\to\infty}{\Longrightarrow} y_{van} = f\frac{tb}{tc} = \frac{fb}{c}$$

So, the vector from viewpoint of a pinhole camera to the vanishing point is $\begin{bmatrix} \frac{fa}{c} \\ \frac{fb}{c} \end{bmatrix}$.

Then, we can project the direction of the line to the image plane:

$$x_V = f\frac{a}{c}$$

$$y_V = f\frac{b}{c}$$

Thus, it is proved that the vector from the viewpoint of a pinhole camera to the vanishing point of a set of 3D parallel lines is parallel to the direction of the parallel lines.

2. (Camera Models- 20 points) Show that relation between any image point (xim, yim)T of a plane (in the form of (x1,x2,x3)T in projective space) and its corresponding point (Xw, Yw, Zw)T on the plane in 3D space can be represented by a 3×3 matrix. You should start from the general form of the camera model (x1,x2,x3)T = MintMext(Xw, Yw, Zw, 1)T, where M = MintMext is a 3×4 matrix, with the image center (ox, oy), the focal length f, the scaling factors(sx and sy), the rotation matrix R and the translation vector T all unknown. Note that in the course slides and the lecture notes, I used a simplified model of the perspective project by assuming ox and oy are known and sx = sy =1, and only discussed the special cases of planes.. So you cannot directly copy those equations I used. Nor can you simply derive the 3×4 matrix M. Instead you should use the general form of the projective matrix (5 points), and the general form of a plane nx Xw + ny Yw + nz Zw = d (5 points), work on an integration (5 points), to form a 3×3 matrix between a 3D point on the plane and its 2D image projection (5 points).

Ans: using the homogeneous coordinates,

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = M_{int}M_{ext}\begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}, \text{ where } M_{int} = \begin{bmatrix} -f_x & 0 & o_x \\ 0 & -f_y & o_y \\ 0 & 0 & 1 \end{bmatrix}, M_{ext} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix}$$

So, M = $M_{int}M_{ext}$ =

$$\begin{bmatrix} -f_x r_{11} + o_x r_{31} & -f_x r_{12} + o_x r_{32} & -f_x r_{13} + o_x r_{33} & -f_x T_x + o_x T_z \\ -f_y r_{21} + o_y r_{31} & -f_y r_{22} + o_y r_{32} & -f_y r_{23} + o_y r_{33} & -f_y T_y + o_y T_z \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix}$$

Because all points are on a plane, so

$$n_x X_w + n_y Y_w + n_z Z_w = d \Rightarrow Z_w = \frac{d}{n_z} - \frac{n_x}{n_z} X_w - \frac{n_y}{n_z} Y_w$$

Plug above plane equation to the projection equation, we can get:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathrm{MP_w}$$

$$= \begin{bmatrix} -f_x r_{11} + o_x r_{31} & -f_x r_{12} + o_x r_{32} & -f_x r_{13} + o_x r_{33} & -f_x T_x + o_x T_z \\ -f_y r_{21} + o_y r_{31} & -f_y r_{22} + o_y r_{32} & -f_y r_{23} + o_y r_{33} & -f_y T_y + o_y T_z \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix} \begin{pmatrix} \mathrm{X_w} \\ \mathrm{Y_w} \\ \frac{d}{n_z} - \frac{n_x}{n_z} X_w - \frac{n_y}{n_z} Y_w \\ 1 \end{pmatrix}$$

=

$$\begin{bmatrix} (-f_x r_{11} + o_x r_{31})X_w + (-f_x r_{12} + o_x r_{32})Y_w + (f_x r_{13} - o_x r_{33})(\frac{n_x}{n_z}X_w + \frac{n_y}{n_z}Y_w - \frac{d}{n_z}) + (-f_x T_x + o_x T_z) \\ (-f_y r_{21} + o_y r_{31})\mathrm{X_w} + (-f_y r_{22} + o_y r_{32})\mathrm{Y_w} + (f_y r_{23} - o_y r_{33})(\frac{n_x}{n_z}X_w + \frac{n_y}{n_z}Y_w - \frac{d}{n_z}) + (-f_y T_y + o_y T_z) \\ r_{31}X_w + r_{32}Y_w + r_{33}\frac{d}{n_z} - r_{33}\frac{n_x}{n_z}X_w - r_{33}\frac{n_y}{n_z}Y_w + T_z \end{bmatrix}$$

=

$$\begin{bmatrix} \frac{n_x}{n_z}(f_x r_{13} - o_x r_{33}) - (f_x r_{11} - o_x r_{31}) & \frac{n_y}{n_z}(f_x r_{13} - o_x r_{33}) - (f_x r_{12} - o_x r_{32}) & -\frac{d}{n_z}(f_x r_{13} - o_x r_{33}) - (f_x T_x - o_x T_z) \\ \frac{n_x}{n_z}(f_y r_{23} - o_y r_{33}) - (f_y r_{21} - o_y r_{31}) & \frac{n_y}{n_z}(f_y r_{23} - o_y r_{33}) - (f_y r_{22} - o_y r_{32}) & -\frac{d}{n_z}(f_y r_{23} - o_y r_{33}) - (f_y T_y - o_y T_z) \\ r_{31} - \frac{n_x}{n_z}r_{33} & r_{32} - \frac{n_y}{n_z}r_{33} & T_z + \frac{d}{n_z}r_{33} \end{bmatrix} \begin{pmatrix} \mathrm{X_w} \\ \mathrm{Y_w} \\ 1 \end{pmatrix}$$

Hence, when $\mathrm{n_z} \neq 0$, the projective matrix M =

$$\begin{bmatrix} \frac{n_x}{n_z}(f_x r_{13} - o_x r_{33}) - (f_x r_{11} - o_x r_{31}) & \frac{n_y}{n_z}(f_x r_{13} - o_x r_{33}) - (f_x r_{12} - o_x r_{32}) & -\frac{d}{n_z}(f_x r_{13} - o_x r_{33}) - (f_x T_x - o_x T_z) \\ \frac{n_x}{n_z}(f_y r_{23} - o_y r_{33}) - (f_y r_{21} - o_y r_{31}) & \frac{n_y}{n_z}(f_y r_{23} - o_y r_{33}) - (f_y r_{22} - o_y r_{32}) & -\frac{d}{n_z}(f_y r_{23} - o_y r_{33}) - (f_y T_y - o_y T_z) \\ r_{31} - \frac{n_x}{n_z}r_{33} & r_{32} - \frac{n_y}{n_z}r_{33} & T_z + \frac{d}{n_z}r_{33} \end{bmatrix}$$

.

In the case of $\mathrm{n_z} = 0$, we can get $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathrm{MP_w} =$

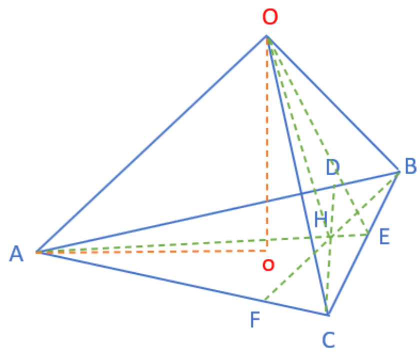
$$\begin{bmatrix} -f_x r_{11} + o_x r_{31} & -f_x r_{12} + o_x r_{32} & -f_x T_x + o_x T_z \\ -f_y r_{21} + o_y r_{31} & -f_y r_{22} + o_y r_{32} & -f_y T_y + o_y T_z \\ r_{31} & r_{32} & T_z \end{bmatrix} \begin{pmatrix} \mathrm{X_w} \\ \mathrm{Y_w} \\ 1 \end{pmatrix}$$

So, when $\mathrm{n_z} = 0$, the projective matrix M =

$$\begin{bmatrix} -f_x r_{11} + o_x r_{31} & -f_x r_{12} + o_x r_{32} & -f_x T_x + o_x T_z \\ -f_y r_{21} + o_y r_{31} & -f_y r_{22} + o_y r_{32} & -f_y T_y + o_y T_z \\ r_{31} & r_{32} & T_z \end{bmatrix}$$

3. (Calibration- 20 points) Prove the Orthocenter Theorem by geometric arguments: Let T be the triangle on the image plane defined by the three vanishing points of three mutually orthogonal sets of parallel lines in space. Then the image center is the orthocenter of the triangle T (i.e., the common intersection of the three altitudes.

(1) Basic proof: use the result of Question 1, assuming the aspect ratio of the camera is 1. Note that you are asked to prove the Orthcenter Theorem, not just the orthcenter of a triangle (7 points)



Ans:

A, B, C are the three vanishing points, which are the three vertices on triangle T. The three altitudes of T, AE, BF, CD insert on point H. O is the viewpoint and the image center is o, so Oo perpendicular to the image plane, ABC.

According to the conclusion of question 1, OA, OB, OC parallel to 3 mutually orthogonal sets of parallel lines, so they perpendicular to each other. It means OA perpendicular to plane OBC, then OA perpendicular to BC. On the other hand, AE is an altitude on the BC, AE⊥BC. Hence, BC perpendicular to plane OAHE. Thus, BC⊥OH.

Same reason, it can be proved that AC⊥OH and AB⊥OH. So, we can say OH perpendicular to plane ABC, and OH⊥AE.

As definition, the o is image center, Oo is on the axis and perpendicular to image plane ABC. Both Oo and OH go through the same point O and perpendicular to same plane ABC, so o and H are the same point. In short, the image center is the orthocenter of the triangle T.

(2) If you do not know the focal length of the camera, can you still find the image center using the Orthocenter Theorem? Explain why or why not (3 points). Can you also estimate the focal length after you find the image center? If yes, how, and if not, why (5 points)

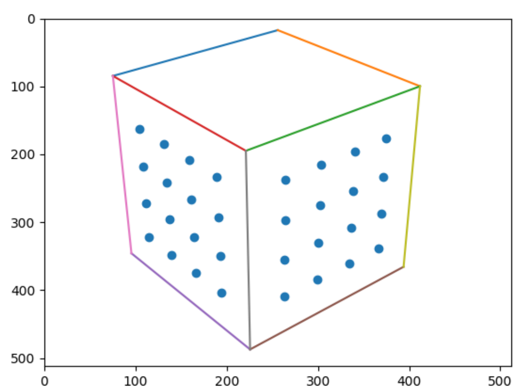
Ans: By Orthocenter Theorem, the image center and orthocenter coincide. So, the image center has nothing to do with the focal length. After finding the location of the image center, it is easy to calculate the position of viewpoint of O. For instance, because the triangle of OoA is similar to OoE, so f/oE = oA/f. oE and oA can be calculated by the position of image center and A, B, C.

(3) If you do not know the aspect ratio and the focal length of the camera, can you still find the image center using the Orthocenter Theorem? Explain why or why not. (5 points)

Ans: No. The focal length is nothing about the position of image center, however, different aspect ratio would obtain a scaled image, so the position of image center will change.

4. Calibration Programming Exercises (40 points): Implement the direct parameter calibration method in order to (1) learn how to use SVD to solve systems of linear equations; (2) understand the physical constraints of the camera parameters; and (3) understand important issues related to calibration, such as calibration pattern design, point localization accuracy and robustness of the algorithms. Since calibrating a real camera involves lots of work in calibration pattern design, image processing and error controls as well as solving the equations, we will use simulated data to understand the algorithms. As a by-product we will also learn how to generate 2D images from 3D models using a "virtual" pinhole camera.

A.Calibration pattern "design". Generate data of a "virtual" 3D cube similar to the one shown in here of the lecture notes in camera calibration. For example, you can hypothesize a 1x1x1 m3 cube and pick up coordinates of 3-D points on one corner of each black square in your world coordinate system. Make sure that the number of your 3-D points is sufficient for the following calibration procedures. In order to show the correctness of your data, draw your cube (with the control points marked) using Matlab (or whatever language you are using). I have provided a piece of starting code in Matlab for you to use. (5 points)



B. "Virtual" camera and images. Design a "virtual" camera with known intrinsic parameters including focal length f, image center (ox, oy) and pixel size (sx, sy). As an example, you can assume that the focal length is f = 16 mm, the image frame size is 512*512 (pixels) with an image center (ox,oy) = (256, 256), and the size of the image sensor inside your camera is 8.8 mm *6.6 mm (so the pixel size is (sx,sy) = (8.8/512, 6.6/512)). Capture an image of your "virtual" calibration cube with your virtual camera with a given pose (rotation R and translation T). For example, you can take the picture of the cube 4 meters away and with a tilt angle of 30 degree. Use three rotation angles alpha, beta, gamma to generate the rotation matrix R (refer to the lecture notes in camera model – please double check the equation since it might have typos in signs). You may need to try different poses in order to have a suitable image of your calibration target. (5 points)

Ans: The virtual camera parameters setting is shown as below:

```
focal length:  0.016
image center:  256 , 256
pixel size:  1.71875e-05 , 1.2890625e-05
rotation angles:
alpha:  -119.99999999999999
beta:  0.0
gamma:  40.0
Rotation Matrix:  [[ 0.76604444 -0.64278761  0.        ]
 [-0.3213938  -0.38302222  0.8660254 ]
 [-0.5566704  -0.66341395 -0.5       ]]
```

C. Direction calibration method: Estimate the intrinsic (fx, fy, aspect ratio a, image center (ox,oy)) and extrinsic (R, T and further alpha, beta, gamma) parameters. Use SVD to solve the homogeneous linear system and the least square problem, and to enforce the orthogonality constraint on the estimate of R.

C(i). Use the accurately simulated data (both 3D world coordinates and 2D image coordinates) to the algorithms and compare the results with the “ground truth” data (which are given in step (a) and step (b)). Remember you are practicing a camera calibration, so you should pretend you know nothing about the camera parameters (i.e., you cannot use the ground truth data in your calibration process). However, in the direct calibration method, you could use the knowledge of the image center (in the homogeneous system to find extrinsic parameters) and the aspect ratio (in the Orthocenter theorem method to find image center). (15 points)

Ans: The detailed algorithm please refer to attached script. The recovered parameters output is shown below:

```
estimating R:  [[ 7.66044443e-01 -6.42787610e-01 -2.50530419e-16]
 [-3.21393805e-01 -3.83022222e-01  8.66025404e-01]
 [-5.56670399e-01 -6.63413948e-01 -5.00000000e-01]]
original R:  [[ 0.76604444 -0.64278761  0.        ]
 [-0.3213938  -0.38302222  0.8660254 ]
 [-0.5566704  -0.66341395 -0.5       ]]
the sum of error between estimated R and oringal R:  3.137110282783279e-15
estimating T: [6.114900252818246e-16, 7.806255641895628e-16, 4.999999999999971] original T:  [[0]
 [0]
 [5]]
estimating fx:  930.9090909090802 original fx:  930.9090909090909
estimating fy:  1241.2121212121078 original fy:  1241.2121212121212
estimating alpha, beta, gammar:  -59.99999999999993 1.4354335634471214e-14 40.000000000000014
original alpha, beta, gammar:  -119.99999999999999 0.0 40.0
```

The discrepancy in alpha rotation angle comes from the same sin value of -60 and -120 degree and the codomain of arcsin.

C(ii). Study whether the unknown aspect ratio matters in estimating the image center (5 points), and how the initial estimation of image center affects the estimating of the remaining parameters (5 points), by experimental results. Give a solution to solve the problems if any (5 points).

Ans: The aspect ratio change will affect the position of the image center, and the image center position will affect the remaining parameters results. In the experiment, I gave a wrong image center, saying, (0, 0), the output as shown below. We can find a great impact on all parameter’s estimations.

```
estimating R:  [[ 0.89261239 -0.4378804   0.1072561 ]
 [-0.22631655 -0.2294695   0.94663856]
 [-0.38990247 -0.86925514 -0.3039269 ]]
original R:  [[ 0.76604444 -0.64278761  0.        ]
 [-0.3213938  -0.38302222  0.8660254 ]
 [-0.5566704  -0.66341395 -0.5       ]]
the sum of error between estimated R and oringal R:  -0.9249742357520413
estimating T: [-1.3257824151052173, -1.0099916733995769, 4.11997439251899] original T:  [[0]
 [0]
 [5]]
estimating fx:  741.2174490084019 original fx:  930.9090909090909
estimating fy:  972.9714467415338 original fy:  1241.2121212121212
```

C(iii). Accuracy Issues. Add in some random noises to the simulated data and run the calibration algorithms again. See how the “design tolerance” of the calibration target and the localization errors of 2D image points affect the calibration accuracy. For example, you can add 0.1 mm (or more) random error to 3D points and 0.5 pixel (or more) random error to 2D points. Also analyze how sensitive of the Orthocenter method is to the extrinsic parameters in imaging the three sets of the orthogonal parallel lines. (* extra points:10)

Ans: With 0.1 to 0.2 mm random error adding to 3D points and 1 to 1.5 pixel random error adding to 2D points, we got following error sensitivity:

```
Error sensitive:
R:  -37.40665165443795
alpha:  -0.0021348631251016294
Tx:  [-0.00195628]
Ty:  [0.00234256]
Tz:  [-0.12514957]
fx:  -122217.58380245918
fy:  -162960.30969996107
```

Obviously, fx and fy are most sensitive.

In order to analyze the impact of image center error in Orthocenter method, 0.5 pixel error are added ox and oy separately.

```
Image center error sensitive (ox):
R:  0.001980612036512412
Tx:  [-0.00537109]
Ty:  [1.80411485e-15]
Tz:  [0.00184655]
```

```
Image center error sensitive (oy):
R:  0.0015162657870579355
Tx:  [2.91433937e-15]
Ty:  [-0.00402832]
Tz:  [-0.00168591]
```

It shows that the impact on both the rotation matrix and Tz are limited no matter which direction appear error, while the Tx and Ty only be affected by corresponding direction where appear error.