Group 27

Computer Vision - HW1 Camera Calibration

A. Introduction

In this assignment, we practice how to implement camera calibration. We take photo from different angles of 2D chessboard. Derive intrinsic matrix and extrinsic matrices from these pictures.

B. Implementation Procedure

Step 0. Choose good point pairs by RANSAC

Randomly select some points as inner group first, and then calculate their fitting line. Check if the points that are chosen fit the line. If yes, put them into inner group. Finally, calculate the number of points in the inner groups.
And we choose the most number after all iterations finish.

Step 1. Find Homography matrix H

In each plane,we have image points $p_i = \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix}$ and object points $P_i = \begin{bmatrix} U_i \\ V_i \\ 0 \\ 1 \end{bmatrix}$

because the 3rd row of Pi is 0 , we kick it out.

So according to the definition of H, we have $p_i = HP_i$ where H = $\begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}$

Rewrite the relation to an equation form, we will get $\begin{cases} u_i(h_3P_i) - h_1P_i = 0 \\ v_i(h_3P_i) - h_2P_i = 0 \end{cases}$ for 1 pair of points, and we have n pairs of points each plane. So we will have n equations and rewrite them to the matrix form.

$$\begin{bmatrix} P_1^T & 0 & -u_1P_1^T \\ 0 & P_1^T & -v_1P_1^T \\ \vdots & \vdots & \vdots \\ P_n^T & 0 & -u_nP_n^T \\ 0 & P_n^T & -v_nP_n^T \end{bmatrix} \begin{bmatrix} h_1^T \\ h_2^T \\ h_3^T \end{bmatrix} = \boldsymbol{X}h = 0 \xRightarrow[\text{may not be 0}]{\text{Because of noise,Xh}} \text{minimize}\|\boldsymbol{X}h\|^2$$

subject to $\|h\|^2 = 1$

And we can solve this by SVD

$\boldsymbol{X}h_{2n*9} = \boldsymbol{U}_{2n*9}\boldsymbol{D}_{9*9}\boldsymbol{V}_{9*9}^T$, and we set h equal to the last column of V(the last row of V^T), and resize h(1*9) to H(3*3)

Step 2. Find Intrinsic Matrix K

H can be decomposed into intrinsic matrix K and extrinsic matrix.
That is,

$$H = (\mathrm{h}1, \mathrm{h}2, \mathrm{h}3) = \begin{bmatrix} f/s_x & 0 & \mathrm{o_x} \\ 0 & \mathrm{f}/s_y & \mathrm{o_y} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & t_1 \\ r_{21} & r_{22} & t_2 \\ r_{31} & r_{32} & t_3 \end{bmatrix}$$

Therefore, for each H, we can derive into two equations by following.

$$H = (\mathrm{h}1, \mathrm{h}2, \mathrm{h}3) = \mathrm{K}(\mathrm{r}1, \mathrm{r}2, \mathrm{t})$$

$$r_1 = K^{-1}h_1 \text{ and } r_2 = K^{-1}h_2$$

Because (r_1, r_2, r_3) form an orthonormal basis,
We can know that

$$\begin{cases} r_1{}^T r_2 = 0 \\ \|r_1\| = \|r_2\| = 1 \end{cases}$$

$$=> \begin{cases} h_1{}^T K^{-T} K^{-1} h_2 = 0 \\ h_1{}^T K^{-T} K^{-1} h_1 = h_2{}^T K^{-T} K^{-1} h_2 \end{cases}$$

To solve K, we define $\mathrm{B} := K^{-T}K^{-1} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$, where B is symmetric and positive definite.

Define $\mathrm{b} = (b_{11}, b_{12}, b_{13}, b_{22}, b_{23}, b_{33})$,
and leads to the system of $V_i\mathrm{b} = 0$ for each H_i
expand and derive the two equation which got from H_i,
and we can get V_i as following.

$$V_i b = \begin{bmatrix} v_{12}{}^T \\ (v_{11} v_{22})^T \end{bmatrix} b = 0$$

, where $v_{ij}{}^T = [h_{i1}h_{j1} \;\; h_{i1}h_{j2} + h_{i2}h_{j1} \;\; h_{i2}h_{j2} \;\; h_{i3}h_{j1} + h_{i1}h_{j3} \;\; h_{i3}h_{j2} + h_{i2}h_{j3} \;\; h_{i3}h_{i3} \;\;]$

And $V = \begin{bmatrix} V_1 \\ \vdots \\ V_n \end{bmatrix}$, where n is the number of plane.

Then we use SVD to get $\mathrm{b} = \underset{b}{\mathrm{argmin}}\, Vb$,

$$V_{2n*6} = U_{2n*2n} D_{2n*6} V_{6*6}^T$$

, and we set b equals to the last column V so that we have B.
Finally, K can be calculated from B using Cholesky factorization.
$\mathrm{B} = LL^H$ and $\mathrm{B} = K^{-T}K^{-1}$
So we get Intrinsic matrix $\mathrm{K} = L^{-T}$

Step 3. Find Extrinsic Matrix

While setting all W=0 in 2D calibration, we have $\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = cH \begin{bmatrix} U \\ V \\ 1 \end{bmatrix}$, where c is a non-zero constant. According to step 1 and 2, homography H and intrinsic matrix K are known. Then, we can easily derive extrinsic matrix by following.

$$cH = K[r1 \;\; r2 \;\; t]$$

$$cK^{-1}H = cK^{-1}[h1 \;\; h2 \;\; h3] = [r1 \;\; r2 \;\; t]$$

$$\mathrm{c} = \frac{1}{\|K^{-1}h1\|}$$

$$\mathrm{r1} = \mathrm{c}K^{-1}h1$$

$$\mathrm{r2} = \mathrm{c}K^{-1}h2$$

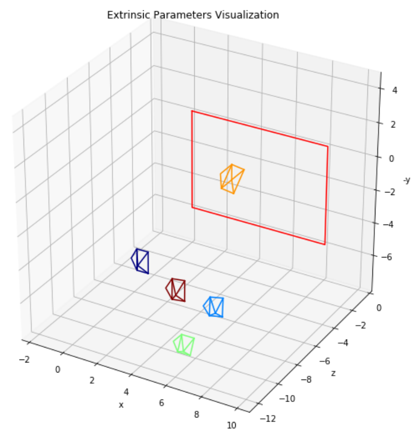
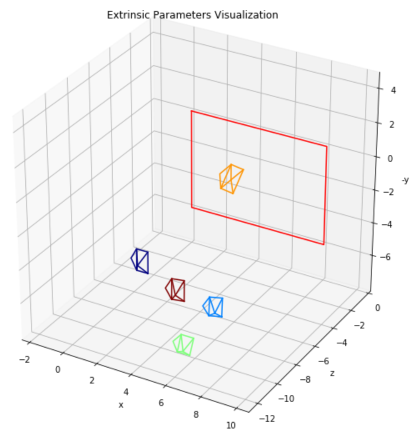
$$\mathrm{t} = cK^{-1}h3$$

We can then get $\mathrm{r3} = \mathrm{r1} \times \mathrm{r2}$, since it is a rotation matrix $\mathrm{R} = [r1 \;\; r2 \;\; r3]$. Finally, we have all values needed for calibration. It can now be applied in 3D real world space.

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathrm{K} \begin{bmatrix} & \mathrm{R} & & t \\ 0 & 0 & 0 & 1 \end{bmatrix}_{4\times4} \begin{bmatrix} U \\ V \\ W \\ 1 \end{bmatrix}$$

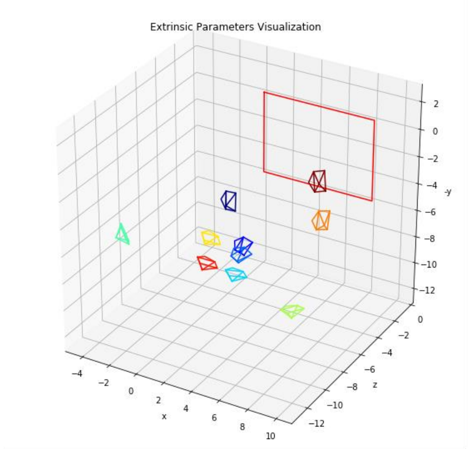
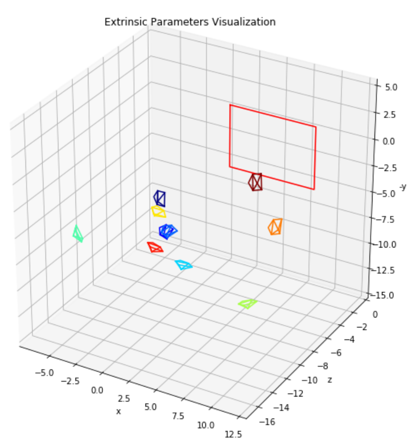
C. Experimental Result

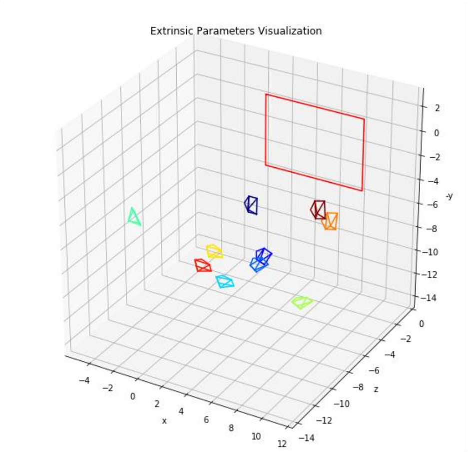
- Our own dataset

	cv2	Our Method
extrinsic visualization	 	
intrinsic value (K)	[[3.30823942e+03 0.00000000e+00 1.51772446e+03] [0.00000000e+00 3.31643064e+03 1.97399635e+03] [0.00000000e+00 0.00000000e+00 1.00000000e+00]]	[[3.38260824e+03 1.86701618e+00 1.51365451e+03] [0.00000000e+00 3.39268217e+03 1.98333014e+03] [0.00000000e+00 0.00000000e+00 1.00000000e+00]]
H of first input image	[[-2.23484537e+01 -3.32207272e+03 2.64720009e+04] [3.24881747e+03 8.87107484e+00 1.18677991e+04] [-3.32615308e-02 -9.28166827e-03 1.06884713e+01]]	[[-1.24175777e+00 -3.15452024e+02 2.49276585e+03] [3.10207845e+02 1.58132212e+00 1.09498313e+03] [-2.67609364e-03 -4.61942858e-04 1.00000000e+00]]

Group 27

- Dataset provided by TA

	cv2	Our Method without RANSAC
extrinsic visualization		
intrinsic value (K)	[[2.70192979e+03 0.00000000e+00 1.53820705e+03] [0.00000000e+00 2.73809172e+03 1.96013469e+03] [0.00000000e+00 0.00000000e+00 1.00000000e+00]]	[[3.40033699e+03 -3.52543622e+01 1.47568443e+03] [0.00000000e+00 3.34935944e+03 1.40822868e+03] [0.00000000e+00 0.00000000e+00 1.00000000e+00]]
H of first input image	[[-1.68199319e+02 -2.65447800e+03 2.16459347e+04] [2.47102197e+03 9.38760734e+01 1.22202073e+04] [-1.25201677e-01 2.99170043e-02 9.36939927e+00]]	[[-4.66914544e+00 -2.91652458e+02 2.32764899e+03] [2.84825821e+02 7.71281888e+00 1.29354235e+03] [-4.95331316e-03 2.58358143e-03 1.00000000e+00]]

	Our Method with RANSAC
extrinsic visualization	
intrinsic value (K)	[[2.76155672e+03 -4.72892162e+01 1.48394933e+03] [0.00000000e+00 2.80235901e+03 2.73170627e+03] [0.00000000e+00 0.00000000e+00 1.00000000e+00]]
H of first input image	[[-3.18534430e+00 -2.92553544e+02 2.32814727e+03] [2.87345538e+02 6.41483894e+00 1.29354630e+03] [-3.55297131e-03 1.72380444e-03 1.00000000e+00]]

D. Discussion

1. By observing the difference between our result and cv2, we find out there are some differences between our K matrix. cv2 set skew factor to be zero and stay zero. However, we do not set this limitation for our method. Thus, there exists a small value in K[1][2] in our result. Pixels of cameras nowadays have perfectly rectangular, so the skew factor should be zero or really close to zero. This matches our result.
2. We took some pictures of 2D calibration ourselves. There are some groups of pictures that do not work well. We found non positive definite B matrix, so we can not do SVD to them. We think this also results from error of detecting corners' coordinate positions. We first tried to random pick some pictures from our own datasets. It works. Though random pick dataset may lead to less data points, the result become better.
3. To solve the problem mentioned in 2. We use RANSAC to choose good corner detection points before finding H. This make it stable. At the end, we pick only one group of pictures to be the result shown above.
4. We do not take distortion into consideration.

Things we have done to improve our result：

1. We search on the Internet to see how others implement camera calibration. There are many ways to implement calibration. We think that eliminating points with bad detection is an important step.
2. Because some image points may be out of border, we should eliminate these points. And we implement RANSAC algorithm in step 0 to select better points. The concept is finding the most points to match the fitting line in each iteration.

E. Conclusion

1. According to experimental results, we discover that our intrinsic matrices are good, but homographies are with larger error. The reason is that there is a great error in finding homography due to the error on detecting chessboard corners. However, according to multiple homographies from different chess boards, we minimize the error, and get a good intrinsic matrix. The error of detecting then shows on extrinsic matrix.

2. We choose the last column of V to be the value of H. This is due to the reason that this column match to the smallest eigenvalues, which means that the last column would have the largest value in V.
3. We can easily see that with RANSAC to choose good corner pairs in input data has much better result than using all detection without choosing. This method improve the accuracy in extrinsic matrices.
4. While using RANSAC, it is really important to set different threshold to meet the requirement. It is important to make sure that this process does not eliminate most of the points, and thus lead to a larger error. Also, it is necessary not to set the threshold too large that it just include all points, which means that RANSAC does not make any effect to our input data points. We choose 30 as the threshold for the result shown above.

F. Work Assignment