Unit II: Geometry

Lecture: Xu Zhao `zhaoxu@sjtu.edu.cn`

Student: Qingyue Tong(123032910106) `tongqingyue@sjtu.edu.cn`

Problem 1. Projective transformation

A projective transformation can be decomposed into a chain of transformations, where each matrix in the chain represents a transformation higher in the hierarchy than the previous one. As for 2D-2D transformation,

$$H = H_S H_A H_P = \begin{bmatrix} sR & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} K & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} I & \mathbf{0} \\ \mathbf{v}^T & v \end{bmatrix}, \tag{1}$$

where R is a 2D rotation matrix, $\mathbf{t}$ and $\mathbf{v}$ are 2D vectors, and K is an upper-triangular matrix normalized as $\det(K) = 1$. H_S represents a similarity transformation (4 DoF), H_A represents an affine transformation (2 DoF), and H_P represents a projective transformation (2 DoF).

Given a known projective transformation H as

$$H = \begin{bmatrix} 2.75 & 2.7 & 1.0 \\ 2.15 & 1.8 & 0.6 \\ 0.3 & 0.5 & 0.2 \end{bmatrix} \tag{2}$$

Please find its decomposition H_S, H_A, and H_P. Discuss when the decomposition is unique. (15 points)

Grading standards. Totally 15 points, 5 for each matrix.

Answer here.

$$H = H_S H_A H_P = \begin{bmatrix} sR & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} K & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} I & \mathbf{0} \\ \mathbf{v}^T & v \end{bmatrix} = \begin{bmatrix} A & v\mathbf{t} \\ \mathbf{v}^T & v \end{bmatrix},$$

where $A = sRK + \mathbf{t}\mathbf{v}^T$.

Observing the above equation, we have $H_P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$, and

$$sRK = A - \mathbf{t}\mathbf{v}^T = \begin{bmatrix} 2.75 & 2.7 \\ 2.15 & 1.8 \end{bmatrix} - \frac{1}{0.2} \begin{bmatrix} 1.0 \\ 0.6 \end{bmatrix} \begin{bmatrix} 0.3 & 0.5 \end{bmatrix} = \begin{bmatrix} 1.25 & 0.2 \\ 1.25 & 0.3 \end{bmatrix}$$

Since $\det(R) = \det(K) = 1$, we have

$$|sRK| = |s|^2 |R||K| = |s|^2 = \begin{vmatrix} 1.25 & 0.2 \\ 1.25 & 0.3 \end{vmatrix} = 0.125$$

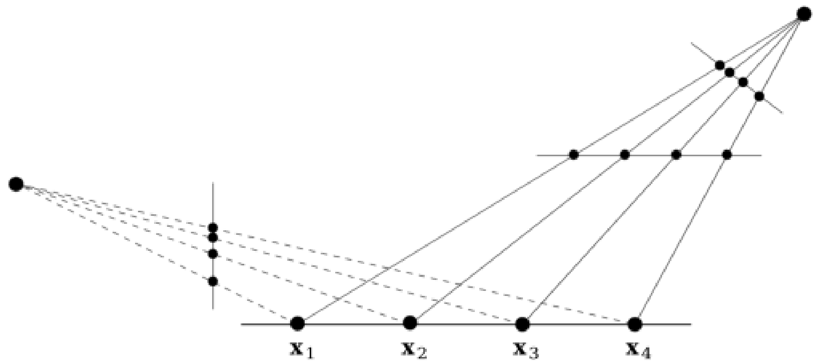

Figure 1: Projective transformations between lines with equivalent cross ratios.

If s is positive, then

$$RK = \frac{1}{\sqrt{0.125}} \begin{bmatrix} 1.25 & 0.2 \\ 1.25 & 0.3 \end{bmatrix} = \sqrt{2} \begin{bmatrix} 2.5 & 0.4 \\ 2.5 & 0.6 \end{bmatrix}$$

Let a is positive, solve the following equation

$$\begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} \begin{bmatrix} a & b \\ 0 & 1/a \end{bmatrix} = \sqrt{2} \begin{bmatrix} 2.5 & 0.4 \\ 2.5 & 0.6 \end{bmatrix}$$

There is a unique solution

$$R = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, K = \begin{bmatrix} 5 & 1 \\ 0 & 0.2 \end{bmatrix}$$

To sum up,

$$H = H_S H_A H_P = \begin{bmatrix} 0.25 & -0.25 & 5 \\ 0.25 & 0.25 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & 1 & 0 \\ 0 & 0.2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$$

If the diagonal elements of matrix K are positive, the decomposition is unique.

Problem 2. Projective invariant

The most fundamental projective invariant is the cross ratio of four collinear points. As show in Figure 1, given 4 points $\mathbf{x}_i$ the cross ratio is defined as

$$\mathrm{Cross}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = \frac{|\mathbf{x}_1\mathbf{x}_2||\mathbf{x}_3\mathbf{x}_4|}{|\mathbf{x}_1\mathbf{x}_3||\mathbf{x}_2\mathbf{x}_4|} = \frac{|\mathbf{x}_1\mathbf{x}_2|}{|\mathbf{x}_1\mathbf{x}_3|} : \frac{|\mathbf{x}_2\mathbf{x}_4|}{|\mathbf{x}_3\mathbf{x}_4|}, \tag{3}$$

where $\mathbf{x}$ is represented by homogeneous coordinates $(x_1, x_2)^T$, and $|\mathbf{x}_i\mathbf{x}_j|$ represents the signed distance from $\mathbf{x}_i$ to $\mathbf{x}_j$ if $x_2 = 1$. Note that the definition of cross ratio is not unique. Please prove that the value of the cross ratio is invariant under any projective transformation, $H_{2\times 2}$, of the line. (10 points)

Hint. if $\mathbf{x}' = H\mathbf{x}$, then

$$\mathrm{Cross}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = \mathrm{Cross}(\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}'_3, \mathbf{x}'_4) \tag{4}$$

Grading standards. Totally 10 points.

Answer here.

Proof: If $\mathbf{x}' = H\mathbf{x}$, according to the properties of the determinant

$$|\mathbf{x}_1'\mathbf{x}_2'| = |(H\mathbf{x}_1)(H\mathbf{x}_2)| = |H(\mathbf{x}_1\mathbf{x}_2)| = |H||\mathbf{x}_1\mathbf{x}_2|$$

Then

$$\text{Cross}(\mathbf{x}_1', \mathbf{x}_2', \mathbf{x}_3', \mathbf{x}_4') = \frac{|\mathbf{x}_1'\mathbf{x}_2'||\mathbf{x}_3'\mathbf{x}_4'|}{|\mathbf{x}_1'\mathbf{x}_3'||\mathbf{x}_2'\mathbf{x}_4'|} = \frac{|H|^2|\mathbf{x}_1\mathbf{x}_2||\mathbf{x}_3\mathbf{x}_4|}{|H|^2|\mathbf{x}_1\mathbf{x}_3||\mathbf{x}_2\mathbf{x}_4|} = \frac{|\mathbf{x}_1\mathbf{x}_2||\mathbf{x}_3\mathbf{x}_4|}{|\mathbf{x}_1\mathbf{x}_3||\mathbf{x}_2\mathbf{x}_4|} = \text{Cross}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4),$$

i.e. the value of the cross ratio is invariant under any projective transformation of the line.

Problem 3. Camera matrix

In the process of camera calibration, suppose that we have obtained a 3×4 camera matrix M, which transforms homogeneous 3D world coordinates to homogeneous 2D image coordinates. Then we want to decompose M into an intrinsic matrix, K, and an extrinsic matrix $[R| - RC]$:

$$M = K[R| - RC], \tag{5}$$

where C is a column-vector representing the camera's position in world coordinates.

1) Please give the decomposition of the following camera matrix M. (15 points)
2) Discuss how to make sure that the matrices, K and R, are valid. (5 points)

$$M = \begin{bmatrix} -50 & 0 & 100 & -2500 \\ -60 & 100 & 0 & -1400 \\ -1 & 0 & 0 & 10 \end{bmatrix} \tag{6}$$

Grading standards. Totally 20 points, 15 for the decomposition results (K, R, C) and 5 for the discussion.

Answer here.

1)

$$M = [KR| - KRC] = \left[\begin{array}{ccc|c} -50 & 0 & 100 & -2500 \\ -60 & 100 & 0 & -1400 \\ -1 & 0 & 0 & 10 \end{array}\right]$$

Do RQ decomposition for matrix KR

$$KR = \begin{bmatrix} 100 & 0 & 50 \\ 0 & 100 & 60 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

And solve the following equation

$$\begin{bmatrix} -50 & 0 & 100 \\ -60 & 100 & 0 \\ -1 & 0 & 0 \end{bmatrix} C = \begin{bmatrix} 2500 \\ 1400 \\ 10 \end{bmatrix} \Rightarrow C = \begin{bmatrix} -10 \\ 8 \\ 20 \end{bmatrix}$$

2) Let the diagonal elements of matrix K are positive, then the RQ decomposition is unique. Furthermore, if the last diagonal element of K is 1 and R is a unit orthogonal matrix, we can say K and R is valid.

Problem 4. Estimation of 2D homography (Code & Report)

In this problem, we discuss the projective transformation between two planes. Given three sets of 2D correspondences, which are with different level of noise. Please use the DLT (Direct Linear Transformation) algorithm to estimate the 2D homography for each set of correspondences. Then calculate the reprojection error of each dataset and discuss the results. Can you come up with a method to reduce the reprojection error? Please implement your method and explain why it works. (25 points)

Hint.

- The data files can be found on canvas. We have provided a template of Jupyter Notebook on canvas to help you finish this problem. Please submit your reproducible code and summarize your results in the report.
- We recommend you to try implementing the Gold Standard algorithm, as introduced in [1]. You can also use *cv2.findHomography* function, but please explain the chosen algorithm and the parameters in your report.

Grading standards. Totally 25 points, 10 for the basic DLT algorithm, 5 for the comparison of reprojection error, 10 for the second method and the corresponding explanation. Grades will take into account both the reproducibility of the code and the quality of the report.

Answer here and code in Jupyter Notebook.

Both code and report are shown in Jupyter Notebook.

Problem 5. Camera calibration (Code & Report)

Camera calibration is the process of determining the intrinsic and extrinsic parameters of a camera. One of the most commonly used methods is proposed by Zhang [2] in 1999. In this problem, let's simply simulate the process of this algorithm. We provide a copy of python code for the simulation of a perspective camera model and a 3D chessboard, which can be used to generate the corner points of chessboard images from different viewpoints. In this way, we can skip the process of corner detection and focus on the calibration processes.

1) *Intrinsic parameters calibration.* Please generate 2D-3D correspondences using the provided code. Then use them to estimate the intrinsic parameters of the pre-defined camera. Compare the results with the ground-truth values, and discuss how should we place the chessboard (or the camera) during the calibration to get better results. (15 points)
2) *Extrinsic parameters calibration.* Assume that we have captured two images with the **same** chessboard and camera from different viewpoints. However, the detected corner points are somehow noised. We use the 3D coordinates of chessboard as the world coordinates. Please design a method to estimate the extrinsic camera parameters of the two images. For qualitative and quantitative evaluation, please visualize your results by projecting the 3D chessboard corner points onto the two images with the estimated camera parameters, and calculate the reprojection error. (15 points)

Hint.

- The code and data files can be found on canvas. We recommend you to write code in the provided Jupyter Notebook file. Please submit your reproducible code and summarize your results in the report.
- You can use *cv2.calibrateCamera* and *cv2.solvePnP* functions for camera calibration. We also encourage students to write your own code, which may earn more points.

Grading standards. Totally 30 points, 15 for intrinsic calibration and 15 for extrinsic calibration. Grades will take into account both the reproducibility of the code and the quality of the report.

Answer here and code in Jupyter Notebook.

Both code and report are shown in Jupyter Notebook.

Reference

[1] Hartley R, Zisserman A. Multiple view geometry in computer vision[M]. Cambridge university press, 2003.

[2] Zhang Z. Flexible camera calibration by viewing a plane from unknown orientations[C]. ICCV, 1999.