CS 83: Computer Vision **Winter 2024**

Quiz 4 — 02/07/2024

Prof. Pediredla ***Student: Amittai Siavava***

Credit Statement

I discussed solution ideas with:

1. Ivy (Aiwei) Zhang
2. Angelic McPherson

However, all typed work is my own, with reference to class notes especially on homographies and transformations. I also referred to some of my earlier notes on linear algebra (from **MATH 22**) on the interpretations of matrices and their null-spaces, column-spaces, and row-spaces.

Problem 1.

(i) Prove that there exists a homography $\mathbf{H}$ that satisfies

$$\mathbf{x}_1 \equiv \mathbf{H}\mathbf{x}_2 \tag{1.1}$$

between the 2D points (in homogeneous coordinates) $\mathbf{x}_1$ and $\mathbf{x}_2$ in the images of a *plane* Π captured by two 3×4 camera projection matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ respectively. The $\equiv$ symbol is equality up to scale. *Note: A degenerate case happens when the plane* Π *contains both cameras' centers, in which case there are infinite choices of* $\mathbf{H}$ *satisfying the above equation. You can ignore this special case in your answer.*

We can write the points $\mathbf{x}_1$ and $\mathbf{x}_2$ in terms of the camera projection matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ and the real-world 3D point $\mathbf{X}$:

$$\mathbf{x}_1 = \mathbf{P}_1\mathbf{X}$$

$$\mathbf{x}_2 = \mathbf{P}_2\mathbf{X}$$

Then, using the fact that the plane Π is imaged as a plane in both images, we can write $\mathbf{X} = \mathbf{H}\mathbf{X}_0$ for some $\mathbf{H}$ and $\mathbf{X}_0$, where $\mathbf{X}_0$ is a point in a reference frame of the plane Π. Substituting this into the above equations, we get

$$\mathbf{x}_1 = \mathbf{P}_1\mathbf{H}\mathbf{X}_0$$

$$\mathbf{x}_2 = \mathbf{P}_2\mathbf{H}\mathbf{X}_0$$

Expressing $\mathbf{P}_1$ and $\mathbf{P}_2$ in terms of their intrinsic and extrinsic parameters, we get

$$\mathbf{x}_1 = \mathbf{K}_1 \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} \mathbf{H}\mathbf{X}_0$$
$$\mathbf{x}_2 = \mathbf{K}_2 \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix} \mathbf{H}\mathbf{X}_0$$

Solving for $\mathbf{H}$ gives:

$$\mathbf{H} = \mathbf{K}_2 - \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix} \mathbf{K}_1$$

(ii) Prove that there exists a homography $\mathbf{H}$ that satisfies equaton 1.1 given two cameras separated by a pure rotation. That is, for camera 1, $\mathbf{x}_1 = \mathbf{K}_1 \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} \mathbf{X}$, and for camera 2, $\mathbf{x}_2 = \mathbf{K}_2 \begin{bmatrix} \mathbf{R} & \mathbf{0} \end{bmatrix} \mathbf{X}$. Note that $\mathbf{K}_1$ and $\mathbf{K}_2$ are the 3×3 intrinsic matrices of the two cameras and are different. $\mathbf{I}$ is the 3×3 identity matrix, $\mathbf{0}$ is the 3×1 zero vector, and $\mathbf{X}$ is a point in 3D space. $\mathbf{R}$ is the 3×3 rotation matrix of the camera.

First, to simplify the notation, note that:

$$\begin{bmatrix} \mathbf{R} & \mathbf{0} \end{bmatrix} \mathbf{X} = \mathbf{R} \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} \mathbf{X},$$

So, define $\mathbf{X}' = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} \mathbf{X}$. Then, we can write $\mathbf{x}_2 = \mathbf{K}_2\mathbf{R}\mathbf{X}'$ and $\mathbf{x}_1 = \mathbf{K}_1\mathbf{X}'$. Therefore;

$$\begin{aligned} \mathbf{x}_2 &= \mathbf{K}_2\mathbf{R}\mathbf{X}' \\ &= \mathbf{K}_2\mathbf{R}\underbrace{\mathbf{K}_1^{-1}\mathbf{K}_1}_{\mathbf{I}}\mathbf{X}' \\ &= \mathbf{K}_2\mathbf{R}\mathbf{K}_1^{-1}\underbrace{\mathbf{K}_1\mathbf{X}'}_{\mathbf{x}_1} \\ &= \mathbf{K}_2\mathbf{R}\mathbf{K}_1^{-1}\mathbf{x}_1 \end{aligned}$$

Thus, we can write $\mathbf{H} = \mathbf{K}_2\mathbf{R}\mathbf{K}_1^{-1}$ such that $\mathbf{x}_2 = \mathbf{H}\mathbf{x}_1$.

(iii) Suppose that a camera is rotating about its center $\mathbf{C}$, keeping the intrinsic parameters $\mathbf{K}$ constant. Let $\mathbf{H}$ be the homography that maps the view from one camera orientation to the view at a second orientation. Let θ be the angle of rotation between the two orientations. Show that $\mathbf{H}^2$ is the homography corresponding to a rotation of 2θ.

Let $\mathbf{H}_\theta$ be the homography corresponding to a rotation of θ, and let $\mathbf{H}_{2\theta}$ be the homography corresponding to a rotation of 2θ. As seen above, the homography $\mathbf{H}_\theta$ is given by $\mathbf{H}_\theta = \mathbf{K}\mathbf{R}(\theta)\mathbf{K}^{-1}$ and $\mathbf{H}_{2\theta} = \mathbf{K}\mathbf{R}(2\theta)\mathbf{K}^{-1}$.

Let us expand the rotation matrices and find an expression for $\mathbf{R}(2\theta)$ in terms of $\mathbf{R}(\theta)$.:

$$\mathbf{R}(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$\mathbf{R}(2\theta) = \begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix}$$

Using the double-angle formulae for sine and cosine,

$$\sin 2\theta = 2\sin\theta\cos\theta$$

$$\cos 2\theta = \cos^2\theta - \sin^2\theta.$$

Substituting these into the expression for $\mathbf{R}(2\theta)$ gives

$$\begin{aligned} \mathbf{R}(2\theta) &= \begin{bmatrix} \cos^2\theta - \sin^2\theta & -2\sin\theta\cos\theta \\ 2\sin\theta\cos\theta & \cos^2\theta - \sin^2\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \\ &= \mathbf{R}(\theta)^2 \end{aligned}$$

Substituting this into the original expression for $\mathbf{H}_{2\theta}$ gives

$$\begin{aligned} \mathbf{H}_{2\theta} &= \mathbf{K}\mathbf{R}(2\theta)\mathbf{K}^{-1} \\ &= \mathbf{K}\mathbf{R}(\theta)^2\mathbf{K}^{-1} \\ &= \mathbf{K}\mathbf{R}(\theta)\mathbf{R}(\theta)\mathbf{K}^{-1} \\ &= \mathbf{K}\mathbf{R}(\theta)\underbrace{bK^{-1}\mathbf{K}}_{\mathbf{I}}\mathbf{R}(\theta)\mathbf{K}^{-1} \\ &= \mathbf{K}\mathbf{R}(\theta)\mathbf{K}^{-1}\mathbf{K}\mathbf{R}(\theta)\mathbf{K}^{-1} \\ &= \left(\mathbf{K}\mathbf{R}(\theta)\mathbf{K}^{-1}\right)^2 \\ &= \mathbf{H}_\theta^2 \end{aligned}$$

Problem 2.

In class, we say that a camera matrix satisfies the equation $\mathbf{x}_i = \mathbf{P}\mathbf{X}_i$, and that six 3D-2D matches $\mathbf{x} \leftrightarrow \mathbf{X}$ are sufficient to recover $\mathbf{P}$ using a linear (non-iterative) algorithm.

Find a linear algorithm for computing the camera matrix $\mathbf{P}$ in the special case when the camera location (but not orientation) is known. Ignoring degenerate configurations, how many 2D-3D matches are required for there to be a unique solution? Justify your answer.

In the special case when the camera location (but not orientation) is known, we can write the camera matrix $\mathbf{P}$ as:

$$\mathbf{P} = \mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}$$

where $\mathbf{K}$ is the 3×3 intrinsic matrix of the camera, $\mathbf{R}$ is the 3×3 rotation matrix, and $\mathbf{t}$ is the 3×1 translation vector, which is known.

Our goal is to find $\mathbf{K}$ and $\mathbf{R}$ from the 3D-2D matches $\mathbf{x} \leftrightarrow \mathbf{X}$. Suppose we have N 3D-2D matches $\mathbf{x} \leftrightarrow \mathbf{X}$. For any one correspondence $\mathbf{x}_i \leftrightarrow \mathbf{X}_i$, we can write:

$$\mathbf{x}_i = \mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\mathbf{X}_i$$

$$\begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} r_1 & r_2 & r_3 & t_1 \\ r_4 & r_5 & r_6 & t_2 \\ r_7 & r_8 & r_9 & 1 \end{bmatrix}\begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix}$$

Thus, we get the following equations:

$$x_i = f_x r_1 X_i + f_x r_2 Y_i + f_x r_3 Z_i + (f_x t_1 + c_x) Z_i$$

$$y_i = f_y r_4 X_i + f_y r_5 Y_i + f_y r_6 Z_i + (f_y t_2 + c_y) Z_i$$

Ignoring degenerate configurations where $\mathbf{K}$ changes, and ignoring the scale factor, we would have $12 - 1 = 11$ unknowns in the matrix $\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}$. However, since $\mathbf{t} \equiv \begin{bmatrix} t_1 \\ t_2 \\ 1 \end{bmatrix}$ is known, we have $11 - 2 = 9$ remaining unknowns, which we can solve using at least 9 equations. Since each $\mathbf{x}_i \leftrightarrow \mathbf{X}_i$ gives us 2 equations, we need at least 5 3D-2D matches to have a unique solution.

Algorithm

1. For each 3D-2D match $\mathbf{x}_i \leftrightarrow \mathbf{X}_i$, write the equations

$$x_i = f_x r_1 X_i + f_x r_2 Y_i + f_x r_3 Z_i + (f_x t_1 + c_x) Z_i$$

$$y_i = f_y r_4 X_i + f_y r_5 Y_i + f_y r_6 Z_i + (f_y t_2 + c_y) Z_i$$

2. Stack the equations for all $N = 5$ 3D-2D matches to get a system of linear equations

$$\mathbf{A}\mathbf{R} = \mathbf{b}$$

3. Solve for $\mathbf{R}$ using the linear least squares method.
4. Once $\mathbf{R}$ is known, solve for $\mathbf{P}$ using the equation

$$\mathbf{P} = \mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}$$