CIS 580: Machine Perception, Spring 2021
Homework 4 Solutions
Due: Mar 4, 2021 at 11:59pm

Instructions

- This is an individual homework.
- You must submit your solutions on Gradescope, the entry code is MB8ZJP. We recommend that you use LaTeX, but we will accept scanned solutions as well. Please box your answers if you submit scanned versions.
- Start early! If you get stuck, please post your questions on Piazza or come to office hours!

1 Problem 1, 30pts

There are four $2D$ points $(0, 1)$, $(1, 0)$, $(1, 1)$ and $(3, 4)$. After same rotation and translation they become $(1.15, 1.28)$, $(0.28, 0.17)$, $(1.27, 0.29)$, $(4.5, -1.32)$.

(i) Could you use the solution for orthogonal Procrustes problem to obtain the rotation matrix?

(ii) Find a solution that solves directly for the rotation angle θ and a translation t_x, t_y by minimizing

$$\sum_{i=1}^{N}\{(x_i' - x_i \cos\theta + y_i \sin\theta - t_x)^2 + (y_i' - x_i \sin\theta - y_i \cos\theta - t_y)^2\}.$$

Answer:

$$\bar{A} = \sum_{i=1}^{4} A_i = \begin{bmatrix} 1.8 \\ 0.105 \end{bmatrix}$$

$$\bar{B} = \sum_{i=1}^{4} B_i = \begin{bmatrix} 1.25 \\ 1.5 \end{bmatrix}$$

Using the method from the slides we get

$$R = \begin{bmatrix} 0.12434798 & 0.99223867 \\ -0.99223867 & 0.12434798 \end{bmatrix}$$

and then

$$T = \bar{A} - R\bar{B} = \begin{bmatrix} 0.15620702 & 1.15877638 \end{bmatrix}^T$$

ii) We first minimize with respect to t_x, t_y, by taking their partial derivatives to zeros. So for t_x we get:

$$-2\sum_{i=1}^{N}\left(x_i' - x_i\cos(\theta) + y_i\sin(\theta) - t_x\right) = 0 \implies$$

$$t_x = \frac{\sum_{i=1}^{N} x_i' - x_i\cos(\theta) + y\sin(\theta)}{n}$$

Similar for t_y we get:

$$t_y = \frac{\sum_{i=1}^{N} y_i' - x_i\sin(\theta) - y_i\cos(\theta)}{n}$$

We substitute t_x, t_y back to the original objective. We also define the variables:

$$\begin{aligned}
\tilde{x}_i &= x_i - \bar{x} \\
\tilde{y}_i &= y_i - \bar{y} \\
\tilde{x}_i' &= x_i' - \bar{x}' \\
\tilde{y}_i' &= y_i' - \bar{y}'
\end{aligned}$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{N} x_i$ and similarly for $\bar{y}, \bar{x}', \bar{y}'$.
After we replace t_x and t_y and we write the minimization objective in terms of the new defined variables we can write the objective as

$$\sum_{i=1}^{N}\left\{(\tilde{x}_i' - \tilde{x}_i\cos(\theta) + \tilde{y}_i\cos(\theta))^2 + (\tilde{y}_i' - \tilde{x}_i\sin(\theta) - \tilde{y}_i\cos(\theta))^2\right\}$$

Setting the derivative w.r.t. θ to be equal to 0 gives us:

$$2\sum_{i=1}^{N}\left\{(\tilde{x}_i' - \tilde{x}_i\cos(\theta) + \tilde{y}_i\sin(\theta))(\tilde{x}_i\sin(\theta) + \tilde{y}_i\cos(\theta)) + (\tilde{y}_i' - \tilde{x}_i\sin(\theta) - \tilde{y}_i\cos(\theta))(-\tilde{x}_i\cos(\theta) + \tilde{y}_i\sin(\theta))\right\} = 0$$

$$\implies \sum_{i=1}^{N}\left\{\tilde{x}_i'\tilde{x}_i\sin(\theta) + \tilde{x}_i'\tilde{y}_i\cos(\theta) - \tilde{y}_i'\tilde{x}_i\cos(\theta) + \tilde{y}_i'\tilde{y}_i\sin(\theta)\right\} = 0$$

So solving for θ we get

$$\tan(\theta) = \frac{\sum_{i=1}^{N}\{\tilde{y}_i'\tilde{x}_i - \tilde{x}_i'\tilde{y}_i\}}{\sum_{i=1}^{N}\{\tilde{x}_i'\tilde{x}_i + \tilde{y}_i'\tilde{y}_i\}} \implies \theta = \arctan\left(\frac{\sum_{i=1}^{N}\{\tilde{y}_i'\tilde{x}_i - \tilde{x}_i'\tilde{y}_i\}}{\sum_{i=1}^{N}\{\tilde{x}_i'\tilde{x}_i + \tilde{y}_i'\tilde{y}_i\}}\right)$$

2 Problem 2, 40pts

You are holding your phone vertically with the optical axis Z_c parallel to the ground. Assume that your camera coordinates are calibrated $K = I$. You see two points on the ground whose world coordinates we assume to be (0,0,0) and (a,0,0). For example, they can be end points of a rod of length a. Their projections are (x_1, y_1) and (x_2, y_2), respectively.

1. Assume that the relative angle between the x-axis of the world and the x-axis of your phone camera is θ, also called the yaw angle. Assume that the translation vector from the projection center to the world origin (0,0,0) is $T = (T_x, T_y, T_z)$. Write the projection equations for (x_1, y_1) and (x_2, y_2) (4 equations total, left hand side must be (x_1, y_1) and (x_2, y_2), right hand side should include only a, θ, T).
2. Solve these 4 equations for the yaw angle θ and (T_x, T_y, T_z).
3. What are the conditions on the camera position in order to obtain a unique or finite number of solutions.

Answer:

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K[RT] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

where $T = [T_x, T_y, T_z]^T$ is the translation from world coordinate to camera coordinate. Since $K = I$ and Z_c is parallel to the ground, there only exists rotation around the y axis,

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & \sin\theta & T_x \\ 0 & 1 & 0 & T_y \\ -\sin\theta & 0 & \cos\theta & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Since the projections of (0,0,0) and $(a, 0, 0)$ are (x_1, y_1) and (x_2, y_2) ,

$$\lambda \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}$$

$$\lambda \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} a\cos\theta + T_x \\ T_y \\ -a\sin\theta + T_z \end{bmatrix}$$

Therefore, $\lambda = -a\sin\theta + T_z$, the projection equations are

$$x_1 = \frac{T_x}{T_z}$$

$$y_1 = \frac{T_y}{T_z}$$

$$x_2 = \frac{a\cos\theta + T_x}{-a\sin\theta + T_z}$$

$$y_2 = \frac{T_y}{-a\sin\theta + T_z}$$

Part b),

$$\frac{x_2}{y_2} = \frac{a\cos\theta + T_x}{T_y} \implies \frac{x_2}{y_2} - \frac{x_1}{y_1} = \frac{a\cos\theta}{T_y}$$

$$T_y = y_1 T_z = y_2\left(-a\sin\theta + T_z\right) \implies T_z = \frac{-a y_2 \sin\theta}{y_1 - y_2}$$

Solving the above equations gives

$$\tan\theta = \frac{y_2 - y_1}{y_1 x_2 - y_2 x_1}$$

Note that $\sin(\arctan(x)) = \frac{x}{\sqrt{x^2+1}}$,

$$\sin\theta = \sqrt{\frac{\left(\frac{y_2-y_1}{y_1x_2-y_2x_1}\right)^2}{\left(\frac{y_2-y_1}{y_1x_2-y_2x_1}\right)^2+1}}$$

Substitute into $T_z = \frac{-a y_2 \sin\theta}{y_1 - y_2}$, we have

$$\begin{aligned}
T_z &= \frac{|a y_2|}{\sqrt{(y_2-y_1)^2 + (y_1x_2 - x_1y_2)^2}} \\
T_x = x_1 T_z &= \frac{|a y_2|\, x_1}{\sqrt{(y_2-y_1)^2 + (y_1x_2 - x_1y_2)^2}} \\
T_y = y_1 T_z &= \frac{|a y_2|\, y_1}{\sqrt{(y_2-y_1)^2 + (y_1x_2 - x_1y_2)^2}}
\end{aligned}$$

Part c), when $x_1 = x_2, y_1 = y_2$, then

$$\frac{T_y}{T_z} = \frac{T_y}{-a\sin\theta + T_z}$$

Thus, $\sin\theta = 0$. Moreover, since

$$\frac{T_x}{T_z} = \frac{a\cos\theta + T_z}{-a\sin\theta + T_z}$$

we have $\cos\theta = 0$, which contradicts with the fact that $\sin\theta = 0$. Therefore, we must have $T_y = 0$ and $\frac{T_x}{T_z} = -\cot(\theta)$. There exist infinite solutions because we only know T_x, T_z to a scale.
In general we need that $T_y \neq 0$ because otherwise the projection center is on the same plane as the two points and we get infinite solutions for our position.

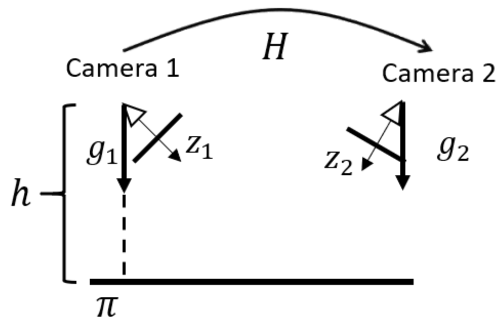
3 Problem 3 (30pts)

Assume that we have two cameras looking at the same ground plane π. Camera 1 is at height $h = 2$ above the ground plane. We also know the gravity vector $g_1 = \left[0, \sqrt{3}/2, 1/2\right]^T$, in the camera's 1 coordinate system and the gravity vector $g_2 = \left[0, 1/2, \sqrt{3}/2\right]^T$ in the camera's 2 coordinate system. The gravity vector is vertical to the ground plane π. Both cameras have camera matrix $K = I$.

Your are given homography H that maps points from camera's 1 image plane to camera's 2 image plane.

$$H = \frac{1}{8}\begin{bmatrix} -4 & 0 & 8 \\ 6 & \sqrt{3} - 12 & 5 - 4\sqrt{3} \\ -2\sqrt{3} & 7 + 4\sqrt{3} & \sqrt{3} + 4 \end{bmatrix}$$

Decompose H so that you find the rotation R and the translation T from camera 1 to camera 2.



Answer:
We want to find rotations R_1 and R_2 so that we transform the camera coordinates, so that y_1 and y_2 are aligned with the direction of gravity. So we want

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = R_1 \begin{bmatrix} x_1' \\ y_1' \\ z_1' \end{bmatrix}$$

where

$$R_1 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = g_1 = [0, \sqrt{3}/2, 1/2]^T$$

We can get such rotation when we rotate 30 degrees along x-axis so

$$R_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{3}/2 & -1/2 \\ 0 & 1/2 & \sqrt{3}/2 \end{bmatrix}$$

Similar for the R_2 we can find that it is the rotation of 60 degrees along the x-axis so

$$R_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & -\sqrt{3}/2 \\ 0 & \sqrt{3}/2 & 1/2 \end{bmatrix}$$

Then for the homography we have that

$$p_2 \sim H p_1 \implies R_2 p_2' \sim H R_1 p_1' \implies p_2' \sim R_2^T H R_1 p_1'$$

We know that $H = (R + tN^T/d)$. So

$$H' = R_2^T H R_1 = (R_2^T R R_1 + (R_2^T t)(R_1^T N)^T/d) = (R' + t' N'^T/d)$$

For the transformed coordinate system we know that the y-axes are vertical to the plane so $N' = [0, 1, 0]^T$ and also the camera frames (in the transformed coordinates) differ in orientation only by a rotation along the y-axis. So

$$R' = R_{y,\theta} = \begin{bmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{bmatrix}$$

and

$$t'N'^T/d = \begin{bmatrix} 0 & t'_x/d & 0 \\ 0 & t'_y/d & 0 \\ 0 & t'_z/d & 0 \end{bmatrix}$$

So given that $d = h = 2$

$$H' = R_2^T H R_1 = \begin{bmatrix} -1/2 & 1/2 & \sqrt{3}/2 \\ 0 & 1 & 0 \\ -\sqrt{3}/2 & 2 & -1/2 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & t'_x/2 & \sin(\theta) \\ 0 & 1 + t'_y/2 & 0 \\ -\sin(\theta) & t'_z/2 & \cos(\theta) \end{bmatrix}$$

it is easy to see that $\theta = 2\pi/3$ and $t' = [1, 0, 4]^T$. Then we have that $R' = R_{y,2\pi/3} = R_2^T R R_1 \implies R = R_2 R_{y,2\pi/3} R_1^T$ and $t = R_2 t'$. If we do the multiplications we get

$$R = \frac{1}{8} \begin{bmatrix} -4 & -2\sqrt{3} & 6 \\ 6 & \sqrt{3} & 5 \\ -2\sqrt{3} & 7 & \sqrt{3} \end{bmatrix}$$

$$t = \begin{bmatrix} 1 \\ -2\sqrt{3} \\ 2 \end{bmatrix}$$