CSCI 677: Advanced Computer Vision HW 1 Solutions

Rijul Dhir, Jiaxin Chen

October 1, 2020

1 Q1 (5 points)

We offer three solutions

1.1 Solution A

The line l in the image is given by the parameters (a, b, c) i.e. the equation of the line in 2d plane is

$$ax + by + c = 0$$

The actual line L in the 3d would lie in a particular plane say π. We need to find the equation of π. Let π be given by four paramaters $(\alpha, \beta, \gamma, \delta)$ i.e. the equation of π is

$$\alpha x + \beta y + \gamma z + \delta$$

We further know that this plane has to go through the origin for any line lying on it form an image on l. Thus $\delta = 0$.

Further this plane intersects the plane at $z = 1$ at line l. The equation of the line where the plane π intersects the plane $z = 1$ is:

$$\alpha x + \beta y + \gamma = 0 \tag{1}$$

And this must be the same as the equation of the line l. Therefore we can choose $\alpha = a$, $\beta = b$ and $\gamma = c$. Therefore the equation of the plane π is given by

$$ax + by + cz = 0$$

1.2 Solution B

Denote the plane to derive as π and corresponding parameters to be (A, B, C, D), i.e.

$$Ax + By + Cz + D = 0 \tag{2}$$

Since π maps to a line l on the normalized plane, it has to go through origin point (0, 0, 0). We get

$$D = 0 \tag{3}$$

There exists two points $u = (0, -\frac{c}{b})$ and $v = (-\frac{c}{a}, 0)$ on line l. Since l lies on normalized place, the 3D coordinate of two points are $u = (0, -\frac{c}{b}, 1)$ and $v = (-\frac{c}{a}, 0, 1)$

Now we have three points on the plane π, origin $o = (0, 0, 0)$, u and v. The norm of π can be calculated by out product of vector $\overrightarrow{ou}$ and $\overrightarrow{ov}$

$$\begin{vmatrix} i & j & k \\ 0 & -\frac{c}{b} & 1 \\ -\frac{c}{a} & 0 & 1 \end{vmatrix} = -\frac{c}{b}i - \frac{c}{a}j - \frac{c^2}{ab}k = (-\frac{c}{b}, -\frac{c}{a}, -\frac{c^2}{ab}) \tag{4}$$

Hence, plane π is

$$-\frac{c}{b}x - \frac{c}{a}y - \frac{c^2}{ab}z = 0 \tag{5}$$

$$ax + by + cz = 0 \tag{6}$$

1.3 Solution C

For an arbitrary point $P = (X, Y, Z)$ (we use capital letter to indicate an instance of point) on plane π, its mapping on normalized plane is $p = (\frac{X}{Z}, \frac{Y}{Z}, 1)$. Point p's coordinate on normalized plane is $p = (\frac{X}{Z}, \frac{Y}{Z})$ and it goes through line l.

$$a\frac{X}{Z} + b\frac{Y}{Z} + c = 0 \tag{7}$$

Hence we know that for any point on π there is

$$aX + bY + cZ = 0 \tag{8}$$

Replacing the instance X, Y, Z with variable x, y, z, the equation of π is

$$ax + by + cz = 0 \tag{9}$$

2 Q2 (15 points)

2.1 Part A (7 points)

Focal length $f = 50$ mm. As the square pixel is 0.05 mm/pixel, $k = l = \frac{1}{0.05}$ pixel/mm.

$$\alpha = kf = 1000 \text{ pixel}, \tag{10}$$
$$\beta = lf = 1000 \text{ pixel}. \tag{11}$$

The flipping (retinal origin at upper left corner, the x-axis along the top-row and the y-axis points downward) and skewing are represented in the $\theta = -88°$. (negative sign encodes the flipping, while non-90° value encodes skewing). You can verify the skewing term $-\alpha\text{cot}\theta$ still holds, even if $\theta < 0$.

The optical center is indeed in the center of the square sensor, however its value (x_0, y_0) are the coordinates w.r.t. the "skewed" axes [1]. An illustration is shown in Fig.1.

$$x_0 = 450 - 450\text{cot}(|\theta|) = 434.286 \text{ pixel} \tag{12}$$
$$y_0 = \frac{450}{\sin(|\theta|)} = 450.274 \text{ pixel} \tag{13}$$

Therefore,

$$\mathbf{K} = \begin{bmatrix} \alpha & -\alpha\text{cot}\theta & x_0 \\ 0 & \frac{\beta}{\sin\theta} & y_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1000 & 34.921 & 434.286 \\ 0 & -1000.609 & 450.274 \\ 0 & 0 & 1 \end{bmatrix}$$

We aim to solve for extrinsic matrix, which consists of rotation $\mathbf{R}$ and translation $\mathbf{T}$ that convert any point from world coordinate system $\{\mathcal{W}\}$ to camera coordinate system $\{\mathcal{C}\}$, i.e. $\mathbf{P}^{\mathcal{C}} = \mathbf{R}\mathbf{P}^{\mathcal{W}} + \mathbf{T}$

The rotation matrix, for rotating along x-axis by $\phi = -10°$, is:

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.985 & -0.173 \\ 0 & 0.173 & 0.985 \end{bmatrix}$$

We know the origin of camera coordinate in $\{\mathcal{W}\}$ is $\mathbf{O}^{\mathcal{W}} = (2, 3, -5)^T$. Further, the origin of camera coordinate in $\{\mathcal{C}\}$ is $\mathbf{O}^{\mathcal{C}} = (0, 0, 0)^T$. Given $\{\mathcal{C}\}$, i.e. $\mathbf{O}^{\mathcal{C}} = \mathbf{R}\mathbf{O}^{\mathcal{W}} + \mathbf{T}$, we can find (in unit of meter)

$$\mathbf{T} = -\mathbf{R}\mathbf{O}^{\mathcal{W}} \tag{14}$$
$$= -\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.985 & -0.173 \\ 0 & 0.173 & 0.985 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ -5 \end{bmatrix} = \begin{bmatrix} -2 \\ -3.822 \\ 4.403 \end{bmatrix} \tag{15}$$

[1] **note**: this mistake is marked as "Skewed offset issue" in your graded homework.

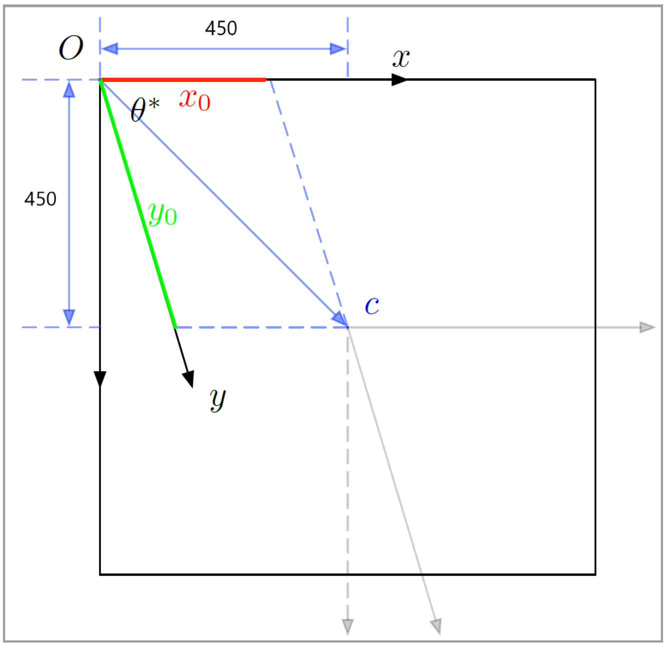

Figure 1: The offset coordinates $c = (x_0, y_0)$ corresponds to the possibly skewed retina coordinate system. $\theta^* = |\theta| = 88°$

Therefore, projection matrix (in meter) is

$$\mathbf{M} = \mathbf{K}\,[\mathbf{R},\ \mathbf{T}] \tag{16}$$

$$= \begin{bmatrix} 1000 & 109.803 & 421.624 & -221.289 \\ 0 & -907.218 & 617.187 & 5807.591 \\ 0 & 0.173 & 0.985 & 4.403 \end{bmatrix} \tag{17}$$

If you are using millimeter, then projection matrix is

$$\mathbf{M} = \mathbf{K}\,[\mathbf{R},\ \mathbf{T}] \tag{18}$$

$$= \begin{bmatrix} 1000 & 109.803 & 421.624 & -221289.6 \\ 0 & -907.218 & 617.187 & 5807591.3 \\ 0 & 0.173 & 0.985 & 4403 \end{bmatrix} \tag{19}$$

But you will probably find out the unit of world coordinate does not affect image coordinate ("scale ambiguity").

Common Mistakes:

1. **Wrong calculation of alpha or beta**: 1 point deducted
2. **"Skewed offset issue"** or **"SO issue"**: 1 point deducted. Some cases take 0.5 points since the student is aware of this issue, however didn't get it completely correct.
3. **use** $\theta = 88°$: 1 point deducted.
4. **intrinsic in unit of mm or meter**: 1 point deducted.
5. **rotation direction incorrect**: 0.5 point deducted.
6. **translation incorrect**: 1 point deducted.
7. The mistakes due to wrong results from previous parts are not repeatedly punished.

2.2 Part B,C,D (8 points)

2.2.1 Part (ii) (3 points)

The direction of all vertical lines in $\{\mathcal{W}\}$ is $(0, s_y, 0)^T$, with arbitrary $s_y \neq 0$. Following page 12 of lecture slide 5, all vertical lines intersects Π_∞ at $(0, s_y, 0, 0)^T$.

Therefore the corresponding vanishing point is

$$\widetilde{\mathbf{V}} = \mathbf{M} \begin{bmatrix} 0 \\ s_y \\ 0 \\ 0 \end{bmatrix} = s_y \begin{bmatrix} 109.803 \\ -907.218 \\ 0.173 \end{bmatrix} \tag{20}$$

Then we compute the image coordinate:

$$\widetilde{\mathbf{v}} = \widetilde{\mathbf{V}}/Z = \begin{bmatrix} 634.69 \\ -5244.03 \\ 1 \end{bmatrix} \tag{21}$$

2.2.2 Part (iii) (3 points)

Similar to part (ii), the direction of all lines in the horizontal plane in $\{\mathcal{W}\}$ is $(s_x, 0, s_z)^T$, with arbitrary s_x and s_y which cannot be zero at the same time. Following page 12 of lecture slide 5, all vertical lines intersects Π_∞ at $(s_x, 0, s_z, 0)^T$.

Therefore the corresponding vanishing point is

$$\widetilde{\mathbf{V}} = \mathbf{M} \begin{bmatrix} s_x \\ 0 \\ s_z \\ 0 \end{bmatrix} = s_x \begin{bmatrix} 1000 \\ 0 \\ 0 \end{bmatrix} + s_z \begin{bmatrix} 421.624 \\ 617.187 \\ 0.985 \end{bmatrix} \tag{22}$$

Then we compute the image coordinate:

$$\widetilde{\mathbf{v}} = \widetilde{\mathbf{V}}/Z = \begin{bmatrix} 1015.228\frac{s_x}{s_z} + 428.04 \\ 626.58 \\ 1 \end{bmatrix} \tag{23}$$

2.2.3 Part (iv) (2 points)

From part (iii), we can see that the vanishing points of lines, which are parallel to the horizontal plane, has a form (in non-homogeneous coordinate):

$$\mathbf{v} = \begin{bmatrix} 1015.228\frac{s_x}{s_z} + 428.04 \\ 626.58 \end{bmatrix} = \begin{bmatrix} 1015.228s + 428.04 \\ 626.58 \end{bmatrix} \tag{24}$$

Where $s = \frac{s_x}{s_z}$ can take arbitrary value. we can see all vanishing points lie in a "horizontal" line in the image coordinate system, where $y = 626.58$.

Notes:

1. Some students approached to vanishing point/line by finding the limit of the projected image coordinate as moving object to infinitely far. It's also a good approach as you can see how v.p./v.l. is formed. As you might define the line as $\mathbf{X} = \mathbf{X}_0 + k\mathbf{V}$ in world coordinate, you should aim to find its image coordinate using both intrinsic $\mathbf{K}$ and extrinsic $[\mathbf{R}, \mathbf{T}]$, and find the limit as $k \to \infty$. Try to use matrix operation (avoiding detail calculation with particular entries) as it can save you a lot of work. You can show this approach is equivalent to the projective geometry formulation. Also, you will find the v.p. is irrelevant to $\mathbf{X}_0$.