School of Computing
National University of Singapore
CS4277/CS5477: 3D Computer Vision
Semester 2, AY 2022/23

Exercise 2

Question 1

Figure 1.1 shows three views of a camera, where $\mathbf{x}_1 \leftrightarrow \mathbf{l}_1' \leftrightarrow \mathbf{x}_1''$ and $\mathbf{x}_2 \leftrightarrow \mathbf{l}_2' \leftrightarrow \mathbf{x}_2''$ are two point-line-point image correspondence over the three views. The camera normalized homogeneous coordinates of the points and lines are:

$\mathbf{x}_1 = [\,2.1794\,,\ -1.4935,\ \ 1.0000]^\top$; $\mathbf{x}_2 = [\,1.6132,\ \ 1.9832,\ \ 1.0000]^\top$;

$\mathbf{l}_1' = [\,-0.0236,\ -0.7193,\ -1.0475]^\top$; $\mathbf{l}_2' = [\,-0.0433,\ -2.0582,\ \ 4.2798]^\top$;

$\mathbf{x}_1'' = [2.1281,\ -1.4583,\ 1.0000]^\top$; $\mathbf{x}_2'' = [1.5667,\ \ 1.9260,\ \ 1.0000]^\top$.

Given that the intrinsics normalized camera projection matrices of the three views are given by $\mathrm{P} = [\mathrm{R} \mid \mathbf{t}]$, $\mathrm{P}' = [\mathrm{R}' \mid \mathbf{t}']$, and $\mathrm{P}'' = [\mathrm{R}'' \mid \mathbf{t}'']$, where $\mathrm{R} = \mathrm{I}_{3\times3}$ and $\mathbf{t} = 0_{3\times1}$, $\mathrm{R}' = \mathrm{I}_{3\times3}$ and $\mathbf{t}' = [t_x', 0, t_z']^\top$, and $\mathrm{R}'' = \mathrm{I}_{3\times3}$ and $\mathbf{t}'' = [0, 0, t_z'']^\top$. Find the three unknown elements t_x', t_z', t_z'' in the translation vectors. **Show all your workings clearly.**

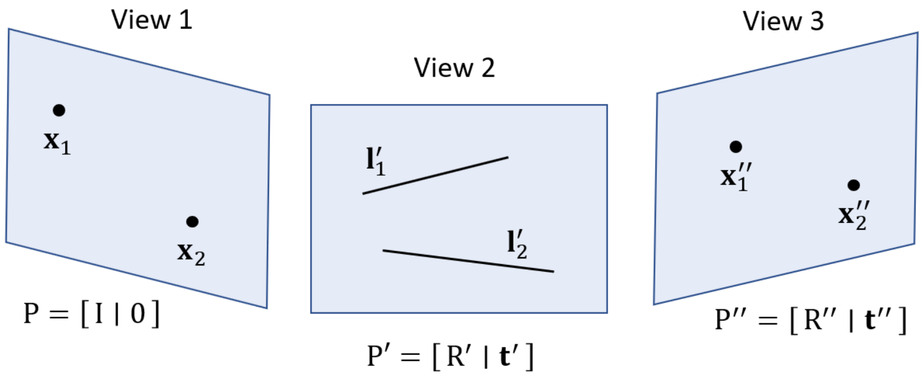


Figure 1.1

(25 marks)

Solution

Get trifocal tensor $[\mathrm{T}_1, \mathrm{T}_2, \mathrm{T}_3]$ using

$$\mathrm{T}_i = \mathbf{a}_i \mathbf{b}_4^\top - \mathbf{a}_4 \mathbf{b}_i^\top ,$$

where $\mathrm{P}' = [\mathrm{a}_1, \mathrm{a}_2, \mathrm{a}_3, \mathrm{a}_4]$ an $\mathrm{P}'' = [\mathrm{b}_1, \mathrm{b}_2, \mathrm{b}_3, \mathrm{b}_4]$.

We have:

$$\mathrm{T}_1 = \begin{bmatrix} -t'_x & 0 & t''_z \\ 0 & 0 & 0 \\ -t'_z & 0 & 0 \end{bmatrix}, \mathrm{T}_2 = \begin{bmatrix} 0 & -t'_x & 0 \\ 0 & 0 & t''_z \\ 0 & -t'_x & 0 \end{bmatrix}, \mathrm{T}_3 = \begin{bmatrix} 0 & 0 & -t'_x \\ 0 & 0 & 0 \\ 0 & 0 & t''_z - t'_z \end{bmatrix}.$$

Given a line $\mathbf{l}' = [l'_x, l'_y, l'_w]^\top$ in the second view, the homography from first to third view of the point-line-point correspondence is given by:

$$\mathrm{H}_{13}(\mathbf{l}') = [\mathrm{T}_1^\top, \mathrm{T}_2^\top, \mathrm{T}_3^\top]\mathrm{l}' = \begin{bmatrix} -l'_x * t'_x - l'_w * t'_z & 0 & 0 \\ 0 & -l'_x * t'_x - l'_w * t'_z & 0 \\ l'_x * t''_z & l'_y * t''_z & -l'_x * t'_x - l'_w * (t'_z - t''_z) \end{bmatrix}.$$

We have the following relations between the three views: $\mathbf{x}'' = \mathrm{H}_{13}(\mathbf{l}')\mathbf{x}$. Taking the cross product, we get:

$$\mathbf{x}'' \times (\mathrm{H}_{13}(\mathbf{l}')\mathbf{x}) = \mathbf{0}$$

$$\begin{bmatrix} 0 & -1 & x''_y \\ 1 & 0 & -x''_x \\ -x''_y & x''_x & 0 \end{bmatrix} \begin{bmatrix} -l'_x * t'_x - l'_w * t'_z & 0 & 0 \\ 0 & -l'_x * t'_x - l'_w * t'_z & 0 \\ l'_x * t''_z & l'_y * t''_z & -l'_x * t'_x - l'_w * (t'_z - t''_z) \end{bmatrix} \begin{bmatrix} x_x \\ x_y \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & -1 & x''_y \\ 1 & 0 & -x''_x \\ -x''_y & x''_x & 0 \end{bmatrix} \begin{bmatrix} x_x(-l'_x * t'_x - l'_w * t'_z) \\ x_y(-l'_x * t'_x - l'_w * t'_z) \\ x_x(l'_x * t''_z) + x_y(l'_y * t''_z) + (-l'_x * t'_x - l'_w * (t'_z - t''_z)) \end{bmatrix}$$

$$= \begin{bmatrix} (l'_x * x_y - l'_x * x''_y) * t'_x + (l'_w * x_y - l'_w * x''_y) * t'_z + x''_y * (l'_w + l'_x * x_x + l'_y * x_y) * t''_z \\ (l'_x * x''_x - l'_x * x_x) * t'_x + (l'_w * x''_x - l'_w * x_x) * t'_z + (-x''_x * (l'_w + l'_x * x_x + l'_y * x_y)) * t''_z \\ (l'_x * x_x * x''_y - l'_x * x_y * x''_x) * t'_x + (l_w * x_x * x''_y - l'_w * x_y * x''_x) * t'_z \end{bmatrix}$$

$$= \mathbf{0}$$

Only first two equations are independent. Thus, we get:

$$\begin{bmatrix} (l'_x * x_y - l'_x * x''_y) & (l'_w * x_y - l'_w * x''_y) & x''_y * (l'_w + l'_x * x_x + l'_y * x_y) \\ (l'_x * x''_x - l'_x * x_x) & (l'_w * x''_x - l'_w * x_x) & (-x''_x * (l'_w + l'_x * x_x + l'_y * x_y)) \end{bmatrix} \begin{bmatrix} t'_x \\ t'_z \\ t''_z \end{bmatrix}$$

$$= \mathbf{0}.$$

We need three equations to solve for the three unknowns. Substituting the first correspondence values into the two equations, we get:

$$(-0.0236 * -1.4935 \; - (-0.0236 * -1.4583)) * t'_x \\ + \left(-1.0475 * -1.4935 - (- 1.0475 * -1.4583)\right) * t'_z \; - 1.4583 \\ * \left(-1.0475 \; + (-0.0236 * \; 2.1794) + (-0.7193 * -1.4935)\right) * t''_z = 0$$

$0.0008 * t'_x \; + \; 0.0369 * t'_z \; + 0.0360 * t''_z = 0$ -- (1)

$$(-0.0236 * 2.1281 \; - (-0.0236 * 2.1794)) * t'_x \\ + \left(-1.0475 * 2.1281 \; - (-1.0475 * \; 2.1794)\right) * t'_z \\ + \left(-2.1281 \\ * \left(-1.0475 \; + (-0.0236 * 2.1794) + (-0.7193 * -1.4935)\right)\right) * t''_z$$

$0.0012 * t'_x \; + \; 0.0537 * t'_z \; + 0.0525 * t''_z = 0$ -- (2)

Substituting the second correspondence values into the first equation, we get:

$$(-0.0433 * \; 1.9832 \; - (-0.0433 * 1.9260)) * t'_x \\ + \; (4.2798 * 1.9832 \; - \; 4.2798 * 1.9260) * t'_z \; + \; 1.9260 \\ * (4.2798 \; + (-0.0433 * 1.6132) \; + (-2.0582 * 1.9832)) * t''_z$$

$-0.0025 * t'_x \; + \; 0.2448 * t'_z \; + 0.2468 * t''_z = 0$ -- (3)

Solving for the unknowns in equations (1), (2) and (3) give us: $\boldsymbol{t'_x = 0.6}$, $\boldsymbol{t'_z = -0.6}$, $\boldsymbol{t''_z = 0.6}$.

Remark: Since scale is unknown, the acceptable answers are $t'_x : t'_z : t''_z = 1 : -1 : 1$.

Question 2

A research engineer mounts a camera on a ground robot. She drove the robot on a flat planar ground while the camera captures two images I and I' with overlapping field-of-view at two different times. Let the camera reference frames where the images are taken be $(\mathrm{x}, \mathrm{y}, \mathrm{z})$ and $(\mathrm{x}', \mathrm{y}', \mathrm{z}')$. As shown in Fig. 2.1, the relative motion between these two frames is denoted by $(t_x, t_z, \theta_y = 0)$, where (t_x, t_z) is the translation vector on the xz-plane and θ_y is the rotation angle around the y-axis.

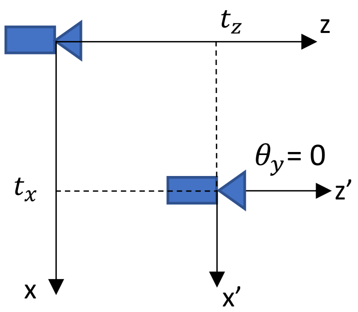


Fig. 2.1

(a) Write the expression for the essential matrix E between I and I' in terms of (t_x, t_z, θ_y), where $\theta_y = 0$.

(b) The camera focal length and principle point are given by (f_x, f_y) and $(C_x = 0, C_y = 0)$ respectively. A 3D point P in the scene appears as $p = (p_x, p_y)$ in image I, write the expression for the corresponding epipolar line L' in image I' in terms of (f_x, f_y), (C_x, C_y), (p_x, p_y) and $(t_x, t_z, \theta_y = 0)$. Show your working clearly.

(c) Given a pair of point correspondence $p \leftrightarrow p'$ i.e. $(p_x, p_y) \leftrightarrow (p'_x, p'_y)$ from I and I',

(i) show that the relationship between $p \leftrightarrow p'$ is given by

$$p_y(f_x t_x - p'_x t_z) - p'_y(f_x t_x - p_x t_z) = 0.$$

(ii) What is the minimum number of image correspondences needed to solve for the relative motion $(t_x, t_z, \theta_y = 0)$?

(d) Suppose four 3D points lying on a plane in the scene are concurrently seen by I and I'. If the plane is parallel to both the image planes and is located at d distance away from the reference frame of the first camera $(\mathrm{x}, \mathrm{y}, \mathrm{z})$, find the homography in terms of $(t_x, t_z, \theta_y = 0)$ and d that relates I, I' and the 3D plane.

Solution

(a) Essential matrix:

$$E = \lfloor t \rfloor_\times R = \begin{bmatrix} 0 & -t_z & 0 \\ t_z & 0 & -t_x \\ 0 & t_x & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -t_z & 0 \\ t_z & 0 & -t_x \\ 0 & t_x & 0 \end{bmatrix}$$

(b) Camera intrinsic matrix:

$$K = \begin{bmatrix} f_x & 0 & 0 \\ 0 & f_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \quad K^{-1} = \begin{bmatrix} 1/f_x & 0 & 0 \\ 0 & 1/f_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Fundamental matrix:

$$F = K^{-\top} E K^{-1} = \begin{bmatrix} 1/f_x & 0 & 0 \\ 0 & 1/f_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -t_z & 0 \\ t_z & 0 & -t_x \\ 0 & t_x & 0 \end{bmatrix} \begin{bmatrix} 1/f_x & 0 & 0 \\ 0 & 1/f_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & -t_z/(f_x f_y) & 0 \\ t_z/(f_x f_y) & 0 & -t_x/f_y \\ 0 & t_x/f_y & 0 \end{bmatrix}$$

Epipolar Line:

$$L' = Fp = \begin{bmatrix} 0 & -t_z/(f_x f_y) & 0 \\ t_z/(f_x f_y) & 0 & -t_x/f_y \\ 0 & t_x/f_y & 0 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ 1 \end{bmatrix} = \begin{bmatrix} -\dfrac{t_z p_y}{f_x f_y} \\ \dfrac{t_z p_x}{f_x f_y} - \dfrac{t_x}{f_y} \\ \dfrac{t_x p_y}{f_y} \end{bmatrix}$$

(ci) Epipolar geometry:

$$p'^\top F p = 0$$

$$\begin{bmatrix} p'_x & p'_y & 1 \end{bmatrix} \begin{bmatrix} 0 & -t_z/(f_x f_y) & 0 \\ t_z/(f_x f_y) & 0 & -t_x/f_y \\ 0 & t_x/f_y & 0 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ 1 \end{bmatrix} = 0$$

$$\begin{bmatrix} \dfrac{p'_y t_z}{f_x f_y} & -\dfrac{t_z p'_x}{f_x f_y} + \dfrac{t_x}{f_y} & -\dfrac{t_x p'_y}{f_y} \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ 1 \end{bmatrix} = 0$$

$$\frac{p_x p'_y t_z}{f_x f_y} - \frac{t_z p_y p'_x}{f_x f_y} + \frac{p_y t_x}{f_y} - \frac{t_x p'_y}{f_y} = 0$$

Multiple by $f_x f_y$,

$$p_x p_y' t_z - t_z p_y p_x' + p_y t_x f_x - t_x p_y' f_x = 0$$

$$p_y(f_x t_x - p_x' t_z) - p_y'(f_x t_x - p_x t_z) = 0 \qquad \text{(Shown)}$$

(cii) 1 images correspondence (2 unknowns, but only up to scale)

(d) Unit normal vector of 3D plane to first camera center:

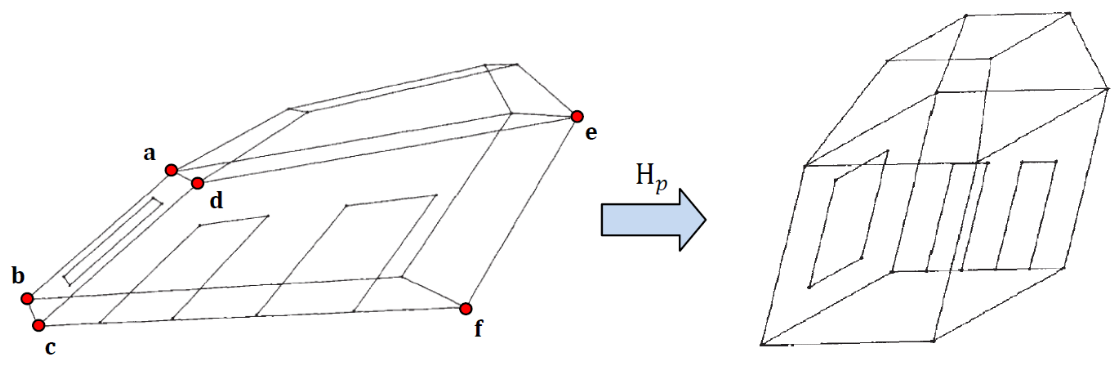
$$N^\top = \begin{bmatrix} 0 & 0 & -1 \end{bmatrix}$$

Homography:

$$H = R + \frac{tN^\top}{d} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \frac{1}{d}\begin{bmatrix} t_x \\ 0 \\ t_z \end{bmatrix}\begin{bmatrix} 0 & 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \frac{1}{d}\begin{bmatrix} 0 & 0 & -t_x \\ 0 & 0 & 0 \\ 0 & 0 & -t_z \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\frac{t_x}{d} \\ 0 & 1 & 0 \\ 0 & 0 & 1 - \frac{t_z}{d} \end{bmatrix}$$

Question 3



(a) Projective reconstruction, $\mathbf{X}_p$

(b) Affine reconstruction, $\mathbf{X}_a$

Figure 3.1

Fig. 3.1(a) shows the 3D reconstruction $\mathbf{X}_p$ from a Fundamental matrix F with projective ambiguity. Given the homogenous coordinates of the points $\mathbf{a} = [13,16,27,-186]^\top$, $\mathbf{b} = [4,10,6,-75]^\top$, $\mathbf{c} = [0,0,0,1]^\top$, $\mathbf{d} = [0.8182, \ 0.5455, \ 1.9091, \ -10.0000]^\top$, $\mathbf{e} = [34,11,31,-215]^\top$, and $\mathbf{f} = [25,5,10,-104]^\top$ in the reconstructed structure, and we further know that the following pairs of lines $\mathbf{L}_{ad} \parallel \mathbf{L}_{bc}$, $\mathbf{L}_{de} \parallel \mathbf{L}_{cf}$ and $\mathbf{L}_{ab} \parallel \mathbf{L}_{dc}$ are parallel in the affine reconstruction $\mathbf{X}_a$ shown in Fig. 3.1(b).

Assuming no noise in the given points, compute the projective transformation H_p that recovers the affine structure $\mathbf{X}_a$ from the 3D reconstruction $\mathbf{X}_p$. Show all your workings clearly.

(**Hint:** use Plücker lines).

Solution

Get the Plücker lines:

$\mathbf{L}_{ad} = [\mathbf{q}_{ab}, \mathbf{q}'_{ab}]^\top$, where $\mathbf{q}_{ab} = \frac{\mathbf{b}-\mathbf{a}}{||\mathbf{b}-\mathbf{a}||}$ and $\mathbf{q}'_{ab} = \mathbf{a} \times \mathbf{q}_{ab}$.

$\Rightarrow \mathbf{L}_{ad} = [-0.2100 \quad 0.5542 \quad -0.8055 \quad 0.1497 \quad -0.0258 \quad -0.0568]^\top$.

$\mathbf{L}_{bc} = [\mathbf{q}_{bc}, \mathbf{q}'_{bc}]^\top$, where $\mathbf{q}_{bc} = \frac{\mathbf{c}-\mathbf{b}}{||\mathbf{c}-\mathbf{b}||}$ and $\mathbf{q}'_{bc} = \mathbf{b} \times \mathbf{q}_{bc}$.

$\Rightarrow \mathbf{L}_{bc} = [0.3244 \quad 0.8111 \quad 0.4867 \qquad 0 \quad -0.0000 \quad 0.0000]^\top$.

$\mathbf{L}_{de} = [\mathbf{q}_{de}, \mathbf{q}'_{de}]^\top$, where $\mathbf{q}_{de} = \frac{\mathbf{e}-\mathbf{d}}{||\mathbf{e}-\mathbf{d}||}$ and $\mathbf{q}'_{de} = \mathbf{d} \times \mathbf{q}_{de}$.

$\Rightarrow \mathbf{L}_{de} = [-0.8523 \quad 0.0378 \quad 0.5217 \quad -0.0212 \quad 0.2054 \quad -0.0496]^\top$.

$\mathbf{L}_{cf} = \left[\mathbf{q}_{cf}, \mathbf{q}'_{cf}\right]^\top$, where $\mathbf{q}_{cf} = \frac{\mathbf{f}-\mathbf{c}}{||\mathbf{f}-\mathbf{c}||}$ and $\mathbf{q}'_{cf} = \mathbf{c} \times \mathbf{q}_{cf}$.

$\Rightarrow \mathbf{L}_{cf} = [-0.9129 \quad -0.1826 \quad -0.3651 \quad 0 \quad 0 \quad 0]^\top$.

$\mathbf{L}_{ab} = [\mathbf{q}_{ab}, \mathbf{q}'_{ab}]^\top$, where $\mathbf{q}_{ab} = \frac{\mathbf{b}-\mathbf{a}}{||\mathbf{b}-\mathbf{a}||}$ and $\mathbf{q}'_{ab} = \mathbf{b} \times \mathbf{q}_{ab}$.

$\Rightarrow \mathbf{L}_{ab} = [0.2014 \quad -0.5755 \quad 0.7926 \quad -0.1517 \quad 0.0262 \quad 0.0575]^\top$.

$\mathbf{L}_{dc} = [\mathbf{q}_{dc}, \mathbf{q}'_{dc}]^\top$, where $\mathbf{q}_{dc} = \frac{\mathbf{c}-\mathbf{d}}{||\mathbf{c}-\mathbf{d}||}$ and $\mathbf{q}'_{dc} = \mathbf{c} \times \mathbf{q}_{dc}$.

$\Rightarrow \mathbf{L}_{dc} = [0.3810 \quad 0.2540 \quad 0.8890 \quad -0.0000 \quad 0.0000 \quad 0]^\top$.

For a pair of parallel lines $\mathbf{L}_1 \parallel \mathbf{L}_2$, the vanishing point is given by their intersection. Let us denote the points on each line as:

$\mathbf{p}_1 = (\mathbf{q}_1 \times \mathbf{q}'_1) + \alpha_1 \mathbf{q}_1$ and $\mathbf{p}_2 = (\mathbf{q}_2 \times \mathbf{q}'_2) + \alpha_2 \mathbf{q}_2$.

The vanishing point is an intersection, i.e. a common point of both lines, it is thus given by solving α_1 and α_2 from:

$(\mathbf{q}_1 \times \mathbf{q}'_1) + \alpha_1 \mathbf{q}_1 = (\mathbf{q}_2 \times \mathbf{q}'_2) + \alpha_2 \mathbf{q}_2$,

which can be written into an over-determinate linear equations: $\mathrm{A}[\alpha_1 \quad \alpha_2]^\top = \mathbf{b}$, where A is a 6×2 matrix and $\mathbf{b}$ is a 6×1 vector made up of the entries in $\mathbf{L}_1$ and $\mathbf{L}_2$. Since the given points are assumed to be without noise, we can just take the first two equations in the over-determinate linear equations to solve for α_1 and α_2.

Thus, we have:

$(\mathbf{q}_{ad} \times \mathbf{q}'_{ad}) + \alpha_{ad}\mathbf{q}_{ad} = (\mathbf{q}_{bc} \times \mathbf{q}'_{bc}) + \alpha_{bc}\mathbf{q}_{bc} \Rightarrow [\alpha_{ad} \; \alpha_{bc}] = [0.0017 \quad -0.1622]$.

The first vanishing point is then:

$\mathbf{v}_1 = (\mathbf{q}_{ad} \times \mathbf{q}'_{ad}) + \alpha_{ad}\mathbf{q}_{ad} = (\mathbf{q}_{bc} \times \mathbf{q}'_{bc}) + \alpha_{bc}\mathbf{q}_{bc}$

$= [-0.0526 \quad -0.1316 \quad -0.0789]^\top$.

$(\mathbf{q}_{de} \times \mathbf{q}'_{de}) + \alpha_{de}\mathbf{q}_{de} = \left(\mathbf{q}_{cf} \times \mathbf{q}'_{cf}\right) + \alpha_{cf}\mathbf{q}_{cf} \Rightarrow \left[\alpha_{de} \; \alpha_{cf}\right] = [0.1514 \quad 0.2608]$.

The second vanishing point is then:

$$\mathbf{v}_2 = (\mathbf{q}_{de} \times \mathbf{q}'_{de}) + \alpha_{de}\mathbf{q}_{de} = \left(\mathbf{q}_{cf} \times \mathbf{q}'_{cf}\right) + \alpha_{cf}\mathbf{q}_{cf}$$

$$= [-0.2381 \quad -0.0476 \quad -0.0952]^\top.$$

$$(\mathbf{q}_{ab} \times \mathbf{q}'_{ab}) + \alpha_{ab}\mathbf{q}_{ab} = (\mathbf{q}_{dc} \times \mathbf{q}'_{dc}) + \alpha_{dc}\mathbf{q}_{dc} \Rightarrow [\alpha_{ab}\ \alpha_{dc}] = [-0.1352 \quad -0.2128].$$

The third vanishing point is then:

$$\mathbf{v}_3 = (\mathbf{q}_{ab} \times \mathbf{q}'_{ab}) + \alpha_{ab}\mathbf{q}_{ab} = (\mathbf{q}_{dc} \times \mathbf{q}'_{dc}) + \alpha_{dc}\mathbf{q}_{dc}$$

$$= [-0.0811 \quad -0.0541 \quad -0.1892]^\top.$$

We can now get the plane formed by the three vanishing points:

$$\mathbf{L}_1 = \mathbf{v}_1 - \mathbf{v}_3 = [0.0284 \quad -0.0775 \quad 0.1102]^\top$$

$$\mathbf{L}_2 = \mathbf{v}_2 - \mathbf{v}_3 = [-0.1570 \quad 0.0064 \quad 0.0940]^\top$$

$$\tilde{\boldsymbol{\pi}} = \frac{\mathbf{L_1} \times \mathbf{L_2}}{||\mathbf{L_1} \times \mathbf{L_2}||} = [-0.0080 \quad -0.0200 \quad -0.0120]^\top$$

Distance of plane to origin: $d = -\mathbf{v}_3^\top(\mathbf{v}_1 \times \mathbf{v}_2) = -0.0040.$

Thus, we get the plane as:

$$\boldsymbol{\pi} = \frac{[\tilde{\boldsymbol{\pi}}\ \ d]}{||[\tilde{\boldsymbol{\pi}}\ \ d]||} = [2.0000 \quad 5.0000 \quad 3.0000 \quad 1.0000]^\top.$$

The homography that transforms the projective reconstruction to the affine reconstruction is given by:

$$\mathrm{H}_{p\to A} = \begin{bmatrix} \mathrm{I}_{3\times 4} \\ \boldsymbol{\pi}^\top \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2 & 5 & 3 & 1 \end{bmatrix}.$$

$$\mathbf{X}_a = \mathrm{H}_{p\to A}\mathbf{X}_p \text{¿} \mathbf{X}_a = \frac{\mathbf{x}_a}{\mathbf{x}_{a(4)}}.$$

We get:

$$\mathbf{a}_A = [13.0000 \quad 16.0000 \quad 27.0000 \quad 1.0000]^\top,$$

$\mathbf{b}_A = [4.0000 \quad 10.0000 \quad 6.0000 \quad 1.0000]^\top$,

$\mathbf{c}_A = [\ 0 \quad 0 \quad 0 \quad 1.0000]^\top$,

$\mathbf{d}_A = [9.0000 \quad 6.0000 \quad 21.0000 \quad 1.0000]^\top$,

$\mathbf{e}_A = [34.0000 \quad 11.0000 \quad 31.0000 \quad 1.0000]^\top$,

$\mathbf{f}_A = [\ 25.0000 \quad 5.0000 \quad 10.0000 \quad 1.0000]^\top$.

Question 4

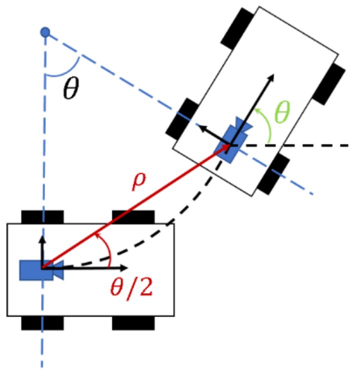


Figure 4.1

Fig. 4.1 shows two consecutive frames of a camera mounted on a moving car at time k and $k+1$. Assuming the car follows the Ackermann motion model, hence the relative pose (rotation and translation) of the camera can be expressed as:

$$\mathrm{R} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{t} = \rho \begin{bmatrix} \cos(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) \\ 0 \end{bmatrix},$$

Where ρ is the vehicle displacement and θ is the yaw angle as shown in Fig 4.1. Given a point correspondence $[x, y, 1]^\top \leftrightarrow [x', y', 1]^\top$ in the two views**,** find the relative pose $(\mathrm{R}, \mathbf{t})$ in terms of the point correspondence. Use $\mathrm{K} = \mathrm{I}_{3\times3}$ as the camera intrinsic matrix. Show all your workings clearly.

Note the following trigonometric identities:

$$\cos(A - B) = \cos A \cos B + \sin A \sin B,$$
$$\sin(A - B) = \sin A \cos B - \cos A \sin B,$$
$$\sin A = \tan A \cos A.$$

Solution

The essential matrix is given by:

$$\mathrm{E} = [\mathbf{t}]_\times \mathrm{R} = \begin{bmatrix} 0 & 0 & \rho \sin\left(\frac{\theta}{2}\right) \\ 0 & 0 & -\rho\cos\left(\frac{\theta}{2}\right) \\ \rho\cos\left(\frac{\theta}{2}\right)\sin(\theta) - \rho\sin\left(\frac{\theta}{2}\right)\cos\theta & \rho\sin\left(\frac{\theta}{2}\right)\sin\theta + \rho\cos\left(\frac{\theta}{2}\right)\cos\theta & 0 \end{bmatrix}.$$

The essential matrix can be further simplified using the double angle formula into:

$$\mathrm{E} = \begin{bmatrix} 0 & 0 & \rho \sin\left(\frac{\theta}{2}\right) \\ 0 & 0 & -\rho\cos\left(\frac{\theta}{2}\right) \\ \rho\sin(\frac{\theta}{2}) & \rho\cos(\frac{\theta}{2}) & 0 \end{bmatrix}.$$

Putting E into the epipolar geometry relation, we get:

$$\mathbf{x'}^\top \mathrm{E}\mathbf{x} = 0.$$

We can observe that ρ can be factorized out of E, and thus any value for ρ satisfies the epipolar geometry.

Putting $[x, y, 1]^\top \leftrightarrow [x', y', 1]^\top$ into the epipolar geometry, we get:

$$\rho\left(y'\cos\left(\frac{\theta}{2}\right) - y\cos\left(\frac{\theta}{2}\right) + x\sin\left(\frac{\theta}{2}\right) + x'\sin\left(\frac{\theta}{2}\right)\right) = 0,$$

ρ takes any value. Solving for $\frac{\theta}{2}$, we get:

$$(y' - y)\cos\left(\frac{\theta}{2}\right) + (x - x')\sin\left(\frac{\theta}{2}\right) = 0$$

$$\tan\frac{\theta}{2} = -\frac{y'-y}{x-x'} \Rightarrow \theta = 2 * \mathrm{atan2}\left(-\frac{y'-y}{x-x'}\right).$$

Question 5

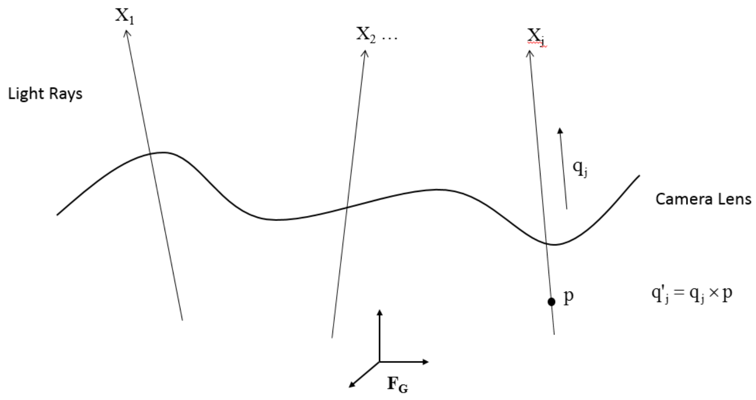


Figure 5.1.: A generalized camera.

The main difference between a generalized camera shown in Figure 5.1 and a commonly used pinhole camera is that light rays projected from the 3D points $X_1, X_2, \ldots, X_j$ do not meet at a single center of projection. We can express a light ray as a 6-vector Plücker line $L_j = \begin{bmatrix} q_j^\top & q'^\top_j \end{bmatrix}^\top$, where q_j is the unit direction vector of the light ray and $q'_j = q_j \times p$. Here, p can be any point on the line. q_j and p are expressed in the camera reference frame F_G. As a result, we can express all light rays with respect to a common frame F_G.

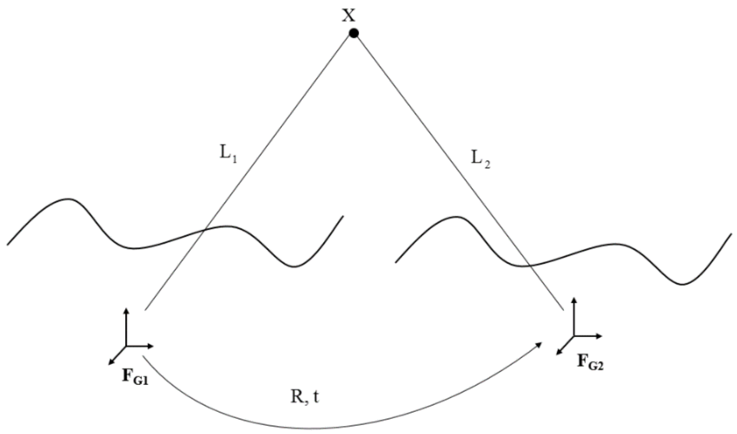


Figure 5.2.: Generalized epipolar constraint.

Figure 5.2 shows a generalized camera that has moved from F_{G1} to F_{G2} over a transformation described by a rotation matrix R and translation vector t. Two light rays $L_1$and $L_2$from F_{G1} and

F_{G2} respectively sees a common 3D point, i.e. the two rays intersect. A pair of Plücker lines expressed in the same reference frame intersects if and only if

$$q_2 . q'_1 + q'_2 . q_1 = 0. \tag{5.1}$$

We know that L_1 can be transformed into the same reference frame as L_2, i.e. F_{G2} via

$$\begin{bmatrix} R & 0 \\ \lfloor t \rfloor_\times R & R \end{bmatrix} L_1, \tag{5.2}$$

where $\lfloor t \rfloor_\times$is the skew symmetric matrix of t.

$$\lfloor t \rfloor_\times = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \tag{5.3}$$

(a) Show that the generalized epipolar constraint is given by

$$L_2^\top \begin{bmatrix} E & R \\ R & 0 \end{bmatrix} L_1 = 0, \tag{5.4}$$

where $E = \lfloor t \rfloor_\times R$ is the essential matrix for a pinhole camera.

(b) Given a pair of Plücker line correspondence $L_1 \leftrightarrow L_2$, the generalized epipolar constraint can be written into a homogenous system of linear equations $Ap = 0$, where p is a vector that contains all the unknown elements from E and R. How many Plücker line correspondences are needed to solve ***linearly*** for the motion transformation R and t? Explain the steps to solve for R and t.

(c) Figure 3.3 shows a special motion case where the generalized camera undergoes pure planar motion. Show that the generalized epipolar constraint is given by

$$\frac{L_2}{1+q^2} \begin{bmatrix} 0 & 0 & t_y(1+q^2) & 1-q^2 & -2q & 0 \\ 0 & 0 & -t_x(1+q^2) & 2q & 1-q^2 & 0 \\ 2t_x q - t_y(1-q^2) & 2t_y q + t_x(1-q^2) & 0 & 0 & 0 & 1+q^2 \\ 1-q^2 & -2q & 0 & 0 & 0 & 0 \\ 2q & 1-q^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1+q^2 & 0 & 0 & 0 \end{bmatrix} L_1 = 0 \tag{3.5}$$

Hint: Let $q = \tan\left(\frac{\theta}{2}\right)$, hence $\cos(\theta) = \frac{1-q^2}{1+q^2}$ and $\sin(\theta) = \frac{2q}{1+q^2}$.

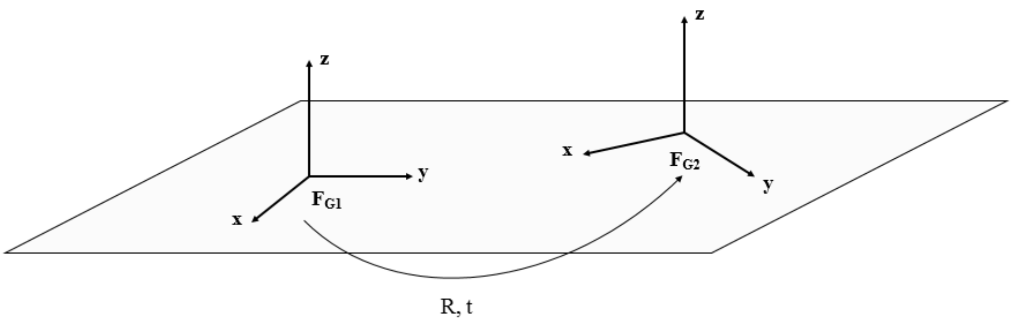


Figure 3.3.: Generalized camera undergoes pure planar motion.

Solution

(a) Refer to the slides of Lecture 12.

(b) 18 unknowns, 9 in each of E and R. Therefore, 17 Plücker line correspondences are needed to solve for R and t linearly.

(c) Planar motion means $t_z = 0$, roll and pitch angles are 0, i.e.

$$\lfloor \mathbf{t} \rfloor_\times = \begin{bmatrix} 0 & 0 & t_y \\ 0 & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}, \mathrm{R} = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We can then compute:

$$\mathrm{E} = \lfloor \mathbf{t} \rfloor_\times \mathrm{R} = \begin{bmatrix} 0 & 0 & t_y \\ 0 & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and putting E and R into the generalized essential matrix, and substituting $\cos(\theta) = \frac{1-q^2}{1+q^2}$ and $\sin(\theta) = \frac{2q}{1+q^2}$, we can get:

$$\begin{bmatrix} E & R \\ R & 0 \end{bmatrix} =$$

$$\frac{1}{1+q^2}\begin{bmatrix} 0 & 0 & t_y(1+q^2) & 1-q^2 & -2q & 0 \\ 0 & 0 & -t_x(1+q^2) & 2q & 1-q^2 & 0 \\ 2t_xq - t_y(1-q^2) & 2t_yq + t_x(1-q^2) & 0 & 0 & 0 & 1+q^2 \\ 1-q^2 & -2q & 0 & 0 & 0 & 0 \\ 2q & 1-q^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1+q^2 & 0 & 0 & 0 \end{bmatrix}$$

Question 6

M lines given in a fixed world frame F_w are represented as the null-spaces of the 2×4 matrices in $\mathbb{P}^3$:

$$\mathrm{W}_i = \begin{bmatrix} \mathbf{A}_i^\top \\ \mathbf{B}_i^\top \end{bmatrix} \text{ for } i = 1, \ldots, M.$$

The projections of these lines in an image taken by a camera with a projection matrix:

$$\mathrm{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix}$$

are: $\mathbf{l}_i = [l_{i1} \quad l_{i2} \quad l_{i3}]^\top$ for $i = 1, \ldots, M$.

The 2D-3D line correspondences with the camera projection matrix can be expressed as the homogeneous linear equation $\mathrm{D}\mathbf{p} = 0$, where $\mathbf{p}$ is the 12×1 vector consists of the 12 elements in P, and D is an $2M \times 12$ matrix in terms of W_i, $\mathbf{l}_i$ and P for $i = 1, \ldots, M$.

a) What is the minimum number of 2D-3D line correspondences needed to solve for the *full* 3×4 camera projection matrix P?

b) Derive the expression for the $2M \times 12$ matrix D in terms of W_i, $\mathbf{l}_i$ and P for $i = 1, \ldots, M$.

c) Given $f_x = f_y = f$ is the focal length of the camera, the skew is zero, $c_x = c_y = 0$ is the principal point, and two 2D-3D line correspondences:

$$\mathrm{W}_1 = \begin{bmatrix} 6.0 & 10.0 & 16.0 & 2.0 \\ 15.0 & 24.0 & 6.0 & 3.0 \end{bmatrix} \leftrightarrow \mathbf{l}_1 = \begin{bmatrix} -148.5 \\ 76.0 \\ -540.0 \end{bmatrix} \text{ and}$$

$$\mathrm{W}_2 = \begin{bmatrix} 16.0 & 12.0 & 24.0 & 4.0 \\ 2.0 & 4.0 & 14.0 & 2.0 \end{bmatrix} \leftrightarrow \mathbf{l}_2 = \begin{bmatrix} 15.0 \\ -31.1 \\ 800.0 \end{bmatrix},$$

find the camera intrinsic matrix K and the pose of the camera $\mathrm{R} \in SO(3)$ and $\mathbf{t} \in \mathbb{R}^3$ with respect to F_w. We further know that K has zero skew $s = 0$ and zero principal point $c_x = c_y = 0$, and no rotation $\mathrm{R} = \mathrm{I}_{3\times 3}$ and a translation of t_y and t_z along the y and z axes of F_W.

Solution

a)

Each 2D-3D line correspondence gives 2 constraints. Therefore, six 2D-3D line correspondences are needed to solve for the 12 unknowns (up to scale) in the camera projection matrix.

b)

A 2D line $\mathbf{l}_i = [l_{i1} \quad l_{i2} \quad l_{i3}]^\top$ is back-projected to a plane: $\boldsymbol{\pi}_i^\top = \boldsymbol{l}_i^\top \mathrm{P}$. The plane $\boldsymbol{\pi}_i$ contains the 3D line W_i, i.e. $\mathrm{W}_i \boldsymbol{\pi}_i = 0$.

Thus, we get:

$$\begin{bmatrix} \mathbf{A}_i^\top \\ \mathbf{B}_i^\top \end{bmatrix} \boldsymbol{\pi}_i = 0 \Rightarrow \mathbf{A}_i^\top \boldsymbol{\pi}_i = 0 \text{ and } \mathbf{B}_i^\top \boldsymbol{\pi}_i = 0.$$

$$\boldsymbol{\pi}_i^\top = \boldsymbol{l}_i^\top \mathrm{P} = [l_{i1} \quad l_{i2} \quad l_{i3}] \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix}$$

$$= [l_{i1}p_{11} + l_{i2}p_{21} + l_{i3}p_{31}, \quad l_{i1}p_{12} + l_{i2}p_{22} + l_{i3}p_{32},$$
$$l_{i1}p_{13} + l_{i2}p_{23} + l_{i3}p_{33}, \quad l_{i1}p_{14} + l_{i2}p_{24} + l_{i3}p_{34}].$$

$$\mathbf{A}_i^\top \boldsymbol{\pi}_i = 0$$

$$\Rightarrow [a_{i1} \quad a_{i2} \quad a_{i3} \quad a_{i4}][l_{i1}p_{11} + l_{i2}p_{21} + l_{i3}p_{31}, \quad l_{i1}p_{12} + l_{i2}p_{22} + l_{i3}p_{32},$$
$$l_{i1}p_{13} + l_{i2}p_{23} + l_{i3}p_{33}, \quad l_{i1}p_{14} + l_{i2}p_{24} + l_{i3}p_{34}]^\top = 0$$

$$a_{i1}(l_{i1}p_{11} + l_{i2}p_{21} + l_{i3}p_{31}) + a_{i2}(l_{i1}p_{12} + l_{i2}p_{22} + l_{i3}p_{32})$$
$$+ a_{i3}(l_{i1}p_{13} + l_{i2}p_{23} + l_{i3}p_{33}) + a_{i4}(l_{i1}p_{14} + l_{i2}p_{24} + l_{i3}p_{34}) = 0$$

$$\mathbf{B}_i^\top \boldsymbol{\pi}_i = 0$$

$$\Rightarrow [b_{i1} \quad b_{i2} \quad b_{i3} \quad b_{i4}][l_{i1}p_{11} + l_{i2}p_{21} + l_{i3}p_{31}, \quad l_{i1}p_{12} + l_{i2}p_{22} + l_{i3}p_{32},$$
$$l_{i1}p_{13} + l_{i2}p_{23} + l_{i3}p_{33}, \quad l_{i1}p_{14} + l_{i2}p_{24} + l_{i3}p_{34}]^\top = 0$$

$$b_{i1}(l_{i1}p_{11} + l_{i2}p_{21} + l_{i3}p_{31}) + b_{i2}(l_{i1}p_{12} + l_{i2}p_{22} + l_{i3}p_{32})$$
$$+ b_{i3}(l_{i1}p_{13} + l_{i2}p_{23} + l_{i3}p_{33}) + b_{i4}(l_{i1}p_{14} + l_{i2}p_{24} + l_{i3}p_{34}) = 0$$

$D =$

$$\begin{bmatrix} \\ a_{i1}l_{i1} & a_{i1}l_{i1} & a_{i1}l_{i1} & a_{i1}l_{i1} & a_{i2}l_{i2} & a_{i2}l_{i3} & a_{i3}l_{i1} & a_{i3}l_{i2} & a_{i3}l_{i3} & a_{i4}l_{i1} & a_{i4}l_{i2} & a_{i4}l_{i2} \\ b_{i1}l_{i1} & b_{i1}l_{i1} & b_{i1}l_{i1} & b_{i1}l_{i1} & b_{i2}l_{i2} & b_{i2}l_{i3} & b_{i3}l_{i1} & b_{i3}l_{i2} & b_{i3}l_{i3} & b_{i4}l_{i1} & b_{i4}l_{i2} & b_{i4}l_{i2} \\ \\ \end{bmatrix}$$

c)

Get camera projection matrix in terms of f_x, f_y, α, t_z :

$$\mathrm{K} = \begin{bmatrix} f_x & 0 & 0 \\ 0 & f_y & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathrm{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{t} = \begin{bmatrix} 0 \\ t_y \\ t_z \end{bmatrix}.$$

$$\Rightarrow \mathrm{P} = \mathrm{k}[\mathrm{R} \quad \mathrm{t}] = \begin{bmatrix} f_x & 0 & 0 \\ 0 & f_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \end{bmatrix} = \begin{bmatrix} f_x & 0 & 0 & 0 \\ 0 & f_y & 0 & f_y t_y \\ 0 & 0 & 1 & t_z \end{bmatrix}$$

$$\mathrm{W}_1 = \begin{bmatrix} 6.0 & 10.0 & 16.0 & 2.0 \\ 15.0 & 24.0 & 6.0 & 3.0 \end{bmatrix} \leftrightarrow \mathbf{l}_1 = \begin{bmatrix} -148.5 \\ 76.0 \\ -540.0 \end{bmatrix}$$

Get back-projected planes by the 2D line:

First line:

$$\boldsymbol{\pi}_1^\top = \boldsymbol{l}_1^\top \mathrm{P} = \begin{bmatrix} -148.5 \\ 76.0 \\ -540.0 \end{bmatrix}^\top \begin{bmatrix} f_x & 0 & 0 & 0 \\ 0 & f_y & 0 & f_y t_y \\ 0 & 0 & 1 & t_z \end{bmatrix}$$

$$= [-(297 * f_x)/2, \quad 76 * f_y, \quad -540, \quad 76 * f_y * t_y - 540 * t_z]$$

$$\mathbf{A}_1^\top \boldsymbol{\pi}_1 = 0$$

$$\Rightarrow [6.0 \quad 10.0 \quad 16.0 \quad 2.0]\left[-\frac{297*f_x}{2},\ 76*f_y,\ -540, 76*f_y*t_y,\ -540*t_z\right]^\top = 0$$

$$\Rightarrow 152*f_y*t_y\ -\ 891*f_x\ +\ 760*f_y\ -\ 1080*t_z\ -\ 8640 = 0$$ **--(1)**

$B_1^\top \boldsymbol{\pi}_1 = 0$

$$\Rightarrow [15.0 \quad 24.0 \quad 6.0 \quad 3.0]\left[-\frac{297*f_x}{2},\ 76*f_y,\ -540, 76*f_y*t_y,\ -540*t_z\right]^\top = 0$$

$$\Rightarrow 228*f_y*t_y\ -\frac{4455*f_x}{2}\ +\ 1824*f_y\ -\ 1620*t_z\ -\ 3240 = 0$$ **--(2)**

Second line:

$$\boldsymbol{\pi}_2^\top = \boldsymbol{l}_2^\top \mathrm{P} = [15.0 \quad -31.1 \quad 800.0]\begin{bmatrix} f_x & 0 & 0 & 0 \\ 0 & f_y & 0 & f_y t_y \\ 0 & 0 & 1 & t_z \end{bmatrix}$$

$$= [15*f_x, \qquad -\frac{280*f_y}{9}, \qquad 800, \qquad 800*t_z - \frac{280*f_y*t_y}{9}]$$

$\mathbf{A}_2^\top \boldsymbol{\pi}_2 = 0$

$$\Rightarrow [16.0 \quad 12.0 \quad 24.0 \quad 4.0][15*f_x,\ -\frac{280*f_y}{9},\ 800,\ 800*t_z\ -\frac{280*f_y*t_y}{9}]^\top = 0$$

$$\Rightarrow 240*f_x\ -\frac{1120*f_y*t_y}{9}\ -\frac{1120*f_y}{3}\ +\ 3200*t_z\ +\ 19200 = 0$$ **-- (3)**

$\mathbf{B}_2^\top \boldsymbol{\pi}_2 = 0$

$$\Rightarrow [2.0 \quad 4.0 \quad 14.0 \quad 2.0]\left[15*f_x,\ -\frac{280*f_y}{9},\ 800,\ 800*t_z\ -\frac{280*f_y*t_y}{9}\right]^\top = 0$$

$$\Rightarrow 30*f_x\ -\frac{560*f_y*t_y}{9}\ -\frac{1120*f_y}{9}\ +\ 1600*t_z\ +\ 11200 = 0$$ **-- (4)**

Solving for the fours unknows using the four equations, we get

$$\begin{bmatrix} 152 & -891 & 760 & -1080 \\ 228 & -\frac{4455}{2} & 1824 & -1620 \\ -\frac{1120}{9} & 240 & -\frac{1120}{3} & 3200 \\ -\frac{560}{9} & 30 & -\frac{1120}{9} & 1600 \end{bmatrix} \begin{bmatrix} f_y * t_y \\ f_x \\ f_y \\ t_z \end{bmatrix} = \begin{bmatrix} 8640 \\ 3240 \\ -19200 \\ -11200 \end{bmatrix}$$

We get:

$$\begin{bmatrix} f_y * t_y \\ f_x \\ f_y \\ t_z \end{bmatrix} = \begin{bmatrix} 90.0 \\ 80.0 \\ 90.0 \\ 2.0 \end{bmatrix}$$

$$\boldsymbol{t_y = \frac{90.0}{90.0} = 1.0,\quad f_x = 80.0,\quad f_y = 90.0,\quad t_z = 2.0.}$$

Question 7

a) Given the following camera normalized three-view point-line-line and point-point-line correspondences:

$$\mathbf{x}_1 = \begin{bmatrix} 1.4865 \\ -1.8908 \\ 1.0000 \end{bmatrix} \leftrightarrow \mathbf{l}'_1 = \begin{bmatrix} -0.1771 \\ -0.4139 \\ -0.3600 \end{bmatrix} \leftrightarrow \mathbf{l}''_1 = \begin{bmatrix} -0.0607 \\ -0.6057 \\ -1.2442 \end{bmatrix} \text{ and}$$

$$\mathbf{x}_2 = \begin{bmatrix} -1.1017 \\ -1.6739 \\ 1.0000 \end{bmatrix} \leftrightarrow \mathbf{x}'_2 = \begin{bmatrix} -0.8979 \\ -1.5116 \\ 1.0000 \end{bmatrix} \leftrightarrow \mathbf{l}''_2 = \begin{bmatrix} -1.0853 \\ -0.4701 \\ -2.1646 \end{bmatrix}.$$

Find the unknowns a, b, c and d in the trifocal tensor that satisfy the point-line-line and point-point-line correspondences:

$$\mathrm{T}_1 = \begin{bmatrix} -5 & a & -5 \\ 0 & 0 & 0 \\ -5 & 0 & 0 \end{bmatrix},\quad \mathrm{T}_2 = \begin{bmatrix} 0 & b & 0 \\ 0 & 5 & -5 \\ 0 & -5 & 0 \end{bmatrix},\quad \mathrm{T}_2 = \begin{bmatrix} 0 & 0 & c \\ 0 & 0 & 0 \\ 0 & 5 & d \end{bmatrix}.$$

b) Find the epipolar lines $\mathbf{l}'_e$ and $\mathbf{l}''_e$ of $\mathbf{x}_1$ in the second and third image views.

c) Given the following point-line correspondence in the first and second views:

$$\mathbf{x} = \begin{bmatrix} 0.8822 \\ -2.0007 \\ 1.0000 \end{bmatrix} \leftrightarrow \mathbf{l}' = \begin{bmatrix} -1.4238 \\ 0.3755 \\ 1.9520 \end{bmatrix},$$

find the point correspondence $\mathbf{x}''$ in the third view.

d) Given the following line-line correspondences in the first and third views:

$$\mathbf{l} = \begin{bmatrix} -2.8434 \\ 1.4639 \\ 3.3164 \end{bmatrix} \leftrightarrow \mathbf{l}'' = \begin{bmatrix} -2.9907 \\ 1.4462 \\ 3.5648 \end{bmatrix},$$

find the line correspondence $\mathbf{l}'$ in the second view.

Show all your workings clearly.

<u>Solution</u>

a)

Point-line-line correspondence give:

$$\mathbf{l}'^{\top} \left(\sum_i x^i \mathrm{T}_i \right) \mathbf{l}'' = 0$$

$$\Rightarrow \begin{bmatrix} -0.1771 \\ -0.4139 \\ -0.3600 \end{bmatrix}^{\top} \left(1.4865 \begin{bmatrix} -5 & a & -5 \\ 0 & 0 & 0 \\ -5 & 0 & 0 \end{bmatrix} - 1.8908 \begin{bmatrix} 0 & b & 0 \\ 0 & 5 & -5 \\ 0 & -5 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & c \\ 0 & 0 & 0 \\ 0 & 5 & d \end{bmatrix} \right) \begin{bmatrix} -0.0607 \\ -0.6057 \\ -1.2442 \end{bmatrix} = 0$$

$$\Rightarrow 0.1595a - 0.2028b + 0.2203c + 0.4479d + 3.7701 = 0 \qquad \textbf{-- (1)}$$

Point-point-line correspondences give:

$$[\mathbf{x}']_{\times} \left(\sum_i x^i \mathrm{T}_i \right) \mathbf{l}'' = 0$$

$$\begin{bmatrix} 0 & -1.0000 & -1.5116 \\ 1.0000 & 0 & 0.8979 \\ 1.5116 & -0.8979 & 0 \end{bmatrix} \left(-1.1017 \begin{bmatrix} -5 & a & -5 \\ 0 & 0 & 0 \\ -5 & 0 & 0 \end{bmatrix} - 1.6739 \begin{bmatrix} 0 & b & 0 \\ 0 & 5 & -5 \\ 0 & -5 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & c \\ 0 & 0 & 0 \\ 0 & 5 & d \end{bmatrix} \right) \begin{bmatrix} -1.0853 \\ -0.4701 \\ -2.1646 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow 3.2720d + 32.7194 = 0 \Rightarrow \boldsymbol{d} = \mathbf{-10.0} \qquad \textbf{--(2)}$$

$$\Rightarrow 0.5179a + 0.7869b - 2.1646c - 1.9436d - 28.9134 = 0 \qquad \textbf{--(3)}$$

$$\Rightarrow 0.7829a + 1.1895b - 3.2720c - 14.3267 = 0 \qquad \textbf{--(4)}$$

Substitute (2) into (1), (3) and (4), we get:

$\begin{bmatrix} 0.1595 & -0.2028 & 0.2203 \\ 0.5179 & 0.7869 & -2.1646 \\ 0.7829 & 1.1895 & -3.2720 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0.7100 \\ 9.4780 \\ 14.3270 \end{bmatrix}$**,** solving for the unknowns, we get:

$\boldsymbol{a = 5; b = -5; c = -5}$

b)

Epipolar line in the second view is given by:

$$\mathbf{l}_e'^{\top} \sum_i x^i \mathrm{T}_i = \mathbf{0}^{\top}$$

$$\begin{bmatrix} l'_{e1} \\ l'_{e2} \\ l'_{e3} \end{bmatrix}^{\top} \left(1.4865 \begin{bmatrix} -5 & a & -5 \\ 0 & 0 & 0 \\ -5 & 0 & 0 \end{bmatrix} - 1.8908 \begin{bmatrix} 0 & b & 0 \\ 0 & 5 & -5 \\ 0 & -5 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & c \\ 0 & 0 & 0 \\ 0 & 5 & d \end{bmatrix} \right) = \mathbf{0}^{\top}$$

$$\begin{bmatrix} l'_{e1} \\ l'_{e2} \\ l'_{e3} \end{bmatrix}^{\top} \begin{bmatrix} -7.4325 & 16.8865 & -12.4325 \\ 0 & -9.4540 & 9.4540 \\ -7.4325 & 14.4540 & -10.000 \end{bmatrix} = \mathbf{0}^{\top}$$

3 equations, 3 unknowns, we get:

$$\Rightarrow \mathbf{l}'_e = \begin{bmatrix} l'_{e1} \\ l'_{e2} \\ l'_{e3} \end{bmatrix} = \begin{bmatrix} \mathbf{-0.6957} \\ \mathbf{-0.1790} \\ \mathbf{0.6957} \end{bmatrix}.$$

Epipolar line in the third view is given by:

$$\sum_i x^i \mathrm{T}_i \, \mathbf{l}''_e = 0$$

$$\begin{bmatrix} -7.4325 & 16.8865 & -12.4325 \\ 0 & -9.4540 & 9.4540 \\ -7.4325 & 14.4540 & -10.000 \end{bmatrix} \begin{bmatrix} l''_{e1} \\ l''_{e2} \\ l''_{e3} \end{bmatrix} = \mathbf{0}$$

3 equations, 3 unknowns, we get:

$$\Rightarrow \mathbf{l}''_e = \begin{bmatrix} l''_{e1} \\ l''_{e2} \\ l''_{e3} \end{bmatrix} = \begin{bmatrix} \mathbf{-0.3902} \\ \mathbf{-0.6511} \\ \mathbf{-0.6511} \end{bmatrix}.$$

c)

$$H_{13} = [T_1^{\top}, T_2^{\top}, T_3^{\top}]\mathbf{l}'$$

$$= [\begin{bmatrix} -5 & a & -5 \\ 0 & 0 & 0 \\ -5 & 0 & 0 \end{bmatrix}^{\top} \begin{bmatrix} -1.4238 \\ 0.3755 \\ 1.9520 \end{bmatrix}, \begin{bmatrix} 0 & b & 0 \\ 0 & 5 & -5 \\ 0 & -5 & 0 \end{bmatrix}^{\top} \begin{bmatrix} -1.4238 \\ 0.3755 \\ 1.9520 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & c \\ 0 & 0 & 0 \\ 0 & 5 & d \end{bmatrix}^{\top} \begin{bmatrix} -1.4238 \\ 0.3755 \\ 1.9520 \end{bmatrix}]$$

$$H_{13} = \begin{bmatrix} -2.6410 & 0 & 0 \\ -7.1190 & -0.7635 & 9.7600 \\ -7.1190 & -1.8775 & -12.4010 \end{bmatrix}$$

$$\mathbf{x}'' = \begin{bmatrix} -2.6410 & 0 & 0 \\ -7.1190 & -0.7635 & 9.7600 \\ -7.1190 & -1.8775 & -12.4010 \end{bmatrix} \begin{bmatrix} 0.8822 \\ -2.0007 \\ 1.0000 \end{bmatrix} = \begin{bmatrix} -2.3299 \\ 5.0072 \\ -2.6343 \end{bmatrix}$$

$$\mathbf{x}'' = \frac{\mathbf{x}''}{\mathbf{x}''(\mathbf{3})} = \begin{bmatrix} \mathbf{0.9854} \\ \mathbf{-2.1178} \\ \mathbf{1.000} \end{bmatrix}$$

d)

$$H_{12} = [T_1, T_2, T_3]\mathbf{l}''$$

$$= [\begin{bmatrix} -5 & a & -5 \\ 0 & 0 & 0 \\ -5 & 0 & 0 \end{bmatrix} \begin{bmatrix} -2.9907 \\ 1.4462 \\ 3.5648 \end{bmatrix}, \quad \begin{bmatrix} 0 & b & 0 \\ 0 & 5 & -5 \\ 0 & -5 & 0 \end{bmatrix} \begin{bmatrix} -2.9907 \\ 1.4462 \\ 3.5648 \end{bmatrix}, \begin{bmatrix} 0 & 0 & c \\ 0 & 0 & 0 \\ 0 & 5 & d \end{bmatrix} \begin{bmatrix} -2.9907 \\ 1.4462 \\ 3.5648 \end{bmatrix}]$$

$$H_{12} = \begin{bmatrix} 4.3605 & -7.2310 & -17.8240 \\ 0 & -10.5930 & 0 \\ 14.9535 & -7.2310 & -28.4170 \end{bmatrix}$$

$$\mathbf{l} = H_{12}^{-\top} \begin{bmatrix} -2.8434 \\ 1.4639 \\ 3.3164 \end{bmatrix} = \begin{bmatrix} 0.2030 \\ -0.1077 \\ -0.2342 \end{bmatrix}$$

$$\mathbf{l} = \frac{\mathbf{l}}{\mathbf{l}(3)} = \begin{bmatrix} \mathbf{-0.8667} \\ \mathbf{0.4497} \\ \mathbf{1.0000} \end{bmatrix}.$$

<u>Question 8</u>

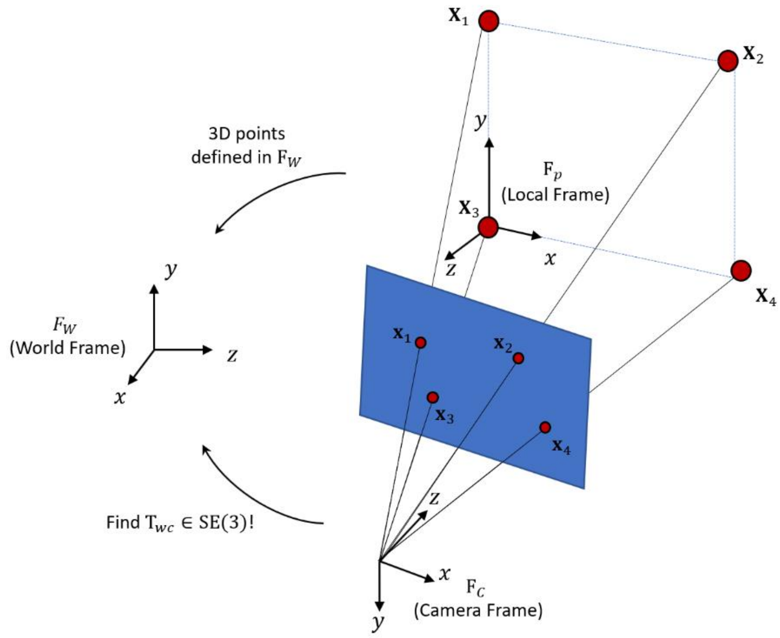


Figure 1.1.

Figure 1.1 shows four corner points of a rectangle that lies on a plane. The homogenous coordinates of these four points in a fixed world frame F_w are given by:

$\mathbf{X}_1 = [1.00,\ \ 10.00,\ \ 0.00,\ \ 1.00]^\top$, $\mathbf{X}_2 = [9.78,\ \ 10.00,\ \ -4.79,\ \ 1.00]^\top$,

$\mathbf{X}_3 = [1.00,\ \ 0.00,\ \ 0.00,\ \ 1.00]^\top$, $\mathbf{X}_4 = [9.78,\ \ 0.00,\ \ -4.79,\ \ 1.00]^\top$.

Let F_p be the local reference coordinate frame rigidly attached to the plane, where the origin of F_p lies on $\mathbf{X}_3$, the x-axis points towards $\mathbf{X}_4$ and the y-axis points towards $\mathbf{X}_1$. Further note that the transformation $T_{wp} \in SE(3)$ from F_p to F_w consists of only a rotation around the y-axis and a translation along the x-axis.

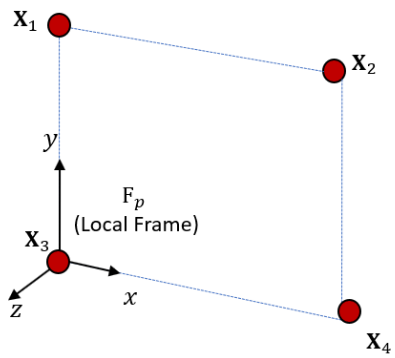
The image coordinates of the four points expressed in the local reference coordinate frame F_p is given by:

$\mathbf{x}_1 = [-188.70 \quad 1490.10 \quad 1.00]^\top$, $\mathbf{x}_2 = [791.4 \quad 1788.10 \quad 1.00]^\top$,

$\mathbf{x}_3 = [10.00 \quad 20.00 \quad 1.00]^\top$, $\mathbf{x}_4 = [990.10 \quad 318.00 \quad 1.00]^\top$.

Given that the camera focal length is $f_x = 100$ and $f_y = 150$, and principal point is $c_x = 10$, $c_x = 20$, find the transformation $T_{wc} \in SE(3)$ that brings the camera frame F_c to the world frame F_w. Note that T_{wc} consists of only a rotation around the z-axis and a translation along the z-axis.

Solution



Get 3D points in the local frame F_p. Let $\mathbf{x}_3^p = [0 \quad 0 \quad 0 \quad 1]^\top$. We then compute the distances of:

$d(\mathbf{X}_1, \mathbf{X}_3) = \sqrt{(1-1)^2 + (10-1)^2 + (0-0)^2} = 10.0$;

$d(\mathbf{X}_3, \mathbf{X}_4) = \sqrt{(1-9.78)^2 + (0-0)^2 + (0+4.79)^2} = 10.0.$

Thus, the coordinates of the points in F_p are given by:

$\mathbf{X}_1 = [0.00, \ 10.00, \ 0.00, \ 1.00]^\top$, $\mathbf{X}_2 = [10.00, \ 10.00, \ 0.00, \ 1.00]^\top$,

$\mathbf{X}_3 = [0.00, \ 0.00, \ 0.00, \ 1.00]^\top$, $\mathbf{X}_4 = [10.0, \ 0.00, \ 0.00, \ 1.00]^\top$.

From the camera projection equation, we get:

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} f_x & 0 & 0 \\ 0 & f_y & 0 \\ 0 & 0 & 1 \end{bmatrix} [r_1 \quad r_2 \quad r_3 \quad \mathbf{t}] \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} [r_1 \quad r_2 \quad \mathbf{t}] \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}.$$

Since T_{cp} consists of only a rotation around the z-axis and a translation along the z-axis, we have

$$r_1 = \begin{bmatrix} \cos\gamma \\ \sin\gamma \\ 0 \end{bmatrix}, r_2 = \begin{bmatrix} -\sin\gamma \\ \cos\gamma \\ 0 \end{bmatrix}, \mathbf{t} = \begin{bmatrix} 0 \\ 0 \\ t_z \end{bmatrix}.$$

Therefore, we get:

$c_x * t_z \ + X * \cos\gamma * f_x \ - \ Y * f_x * \sin\gamma = x$ -- (1)

$c_y * t_z \ + Y * \cos\gamma * f_y \ + \ X * f_y * \sin\gamma = y$ -- (2)

$t_z = w$ -- (3)

Putting $\mathbf{X}_3$ into Equation (1), we get:

$10 * t_z \ + 0 \ - 0 = 10 \Rightarrow \boldsymbol{t_z} = \mathbf{1}$

Putting $\mathbf{X}_1$ into Equation (1), we get:

$(10) * (1) \ + \ 0 \ - (10) * (100) * \sin\gamma = -188.70$

$$\Rightarrow \ \sin\gamma = \frac{-188.70 - 10}{(10) * (100)} \Rightarrow \boldsymbol{\gamma = 0.2\mathbf{rad}}$$

$\mathrm{T}_{wp} = \begin{bmatrix} \cos\beta & 0 & \sin\beta & t_x \\ 0 & 1 & 0 & 0 \\ -\sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ and the transformation of a point from the local coordinate frame of the plane F_p to the world frame F_W is given by:

$\mathbf{X}_w = \mathrm{T}_{wp}\mathbf{X}_p \Rightarrow \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\beta & 0 & \sin\beta & t_x \\ 0 & 1 & 0 & 0 \\ -\sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_p \\ Y_p \\ Z_p \\ 1 \end{bmatrix}$. Thus, we get the following constraints:

$X_w = X_p \cos\beta + Z_p \sin\beta + t_x$ -- (1)

$Z_w = -X_p \sin\beta + Z_p \cos\beta$ -- (2)

Using $\mathbf{X}_3$, we get:

$1 = (0)\cos\beta + (0)\sin\beta + t_x \Rightarrow t_x = 1.00$

Using $\mathbf{X}_4$, we get:

$9.78 = (10)\cos\beta + (0)\sin\beta + 1 \Rightarrow \beta = 0.5\text{rad}.$

$\mathrm{T}_{wc} = \mathrm{T}_{wp}\mathrm{T}_{pc} = \begin{bmatrix} \cos 0.5 & 0 & \sin 0.5 & 1 \\ 0 & 1 & 0 & 0 \\ -\sin 0.5 & 0 & \cos 0.5 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos 0.2 & -\sin 0.2 & 0 & 0 \\ \sin 0.2 & \cos 0.2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1}.$