Camera Notation Tutorial

Roberto Pellerito

September 14, 2026

This document is meant to summarize the explanations given during the Exercise Session. For extensive explanations, please watch the video recording of the tutorial.

1 Notation

In a mobile robotics setting, we typically encounter a (fixed, stationary) world frame $\mathcal{W}$ and a (potentially moving) camera frame $\mathcal{C}$. Let us consider a point ${}_{\mathcal{W}}\mathbf{p}$, where the pre-subscript $\mathcal{W}$ indicates that this coordinate is given is world frame. If we want to express the point in the camera coordinate frame, we need to apply a transformation, see fig. 1. This transformation can be described by a rotation matrix $\mathbf{R}_{\mathcal{CW}}$ and a translation ${}_{\mathcal{C}}\mathbf{t}_{\mathcal{CW}}$ and we say "we transform the point from the world to the camera coordinate frame" or, equivalently, "we express the point in the camera frame". Mathematically,

$$ {}_{\mathcal{C}}\mathbf{p} = \mathbf{R}_{\mathcal{CW}}\, {}_{\mathcal{W}}\mathbf{p} + {}_{\mathcal{C}}\mathbf{t}_{\mathcal{CW}} $$

All the letters written in red correspond to the frame we want to express the point in. The letters in blue are the frame the point is currently expressed in. All red and blue letters must match for the transform to be valid!

We can also unify rotation and transformation into one operation by using *homogeneous coordinates*. Let $\mathbf{p} = \begin{bmatrix} r & s & t \end{bmatrix}^\top$, then we can convert this point to homogeneous coordinates by adding a 1 to the vector, as follows $\tilde{\mathbf{p}} = \begin{bmatrix} r & s & t & 1 \end{bmatrix}^\top$. We can now rewrite the transformation above as

$$ {}_{\mathcal{C}}\mathbf{p} = \begin{bmatrix} \mathbf{R}_{\mathcal{CW}} | \, {}_{\mathcal{C}}\mathbf{t}_{\mathcal{CW}} \end{bmatrix} {}_{\mathcal{W}}\tilde{\mathbf{p}} $$

If we want to ensure the result of the transform is also in homogeneous coordinates, we append a row to the transformation $\mathbf{T}_{\mathcal{CW}}$ matrix. We get

$$ \mathbf{T}_{\mathcal{CW}} = \begin{bmatrix} \mathbf{R}_{\mathcal{CW}} \in \mathbb{R}^{3\times 3} & {}_{\mathcal{C}}\mathbf{t}_{\mathcal{CW}} \in \mathbb{R}^{3\times 1} \\ \mathbf{0} \in \mathbb{R}^{1\times 3} & 1 \end{bmatrix}, \qquad {}_{\mathcal{C}}\tilde{\mathbf{p}} = \mathbf{T}_{\mathcal{CW}}\, {}_{\mathcal{W}}\tilde{\mathbf{p}} $$

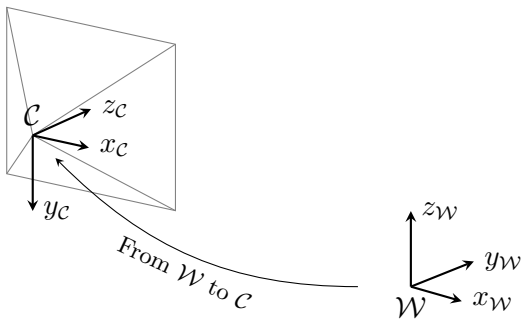


Figure 1: The camera frame $\mathcal{C}$ and world frame $\mathcal{W}$. The arrow "From $\mathcal{W}$ to $\mathcal{C}$" denotes the coordinate transformation ${}_{\mathcal{C}}\tilde{\mathbf{p}} = \mathbf{T}_{\mathcal{CW}}\, {}_{\mathcal{W}}\tilde{\mathbf{p}}$.

Figure 2 gives a visual practical example: the blue and red arrows locate the same point P from two different origins. The purple arrow is the same displacement $\mathbf{v} = \overrightarrow{PQ}$, expressed in world and camera coordinates.

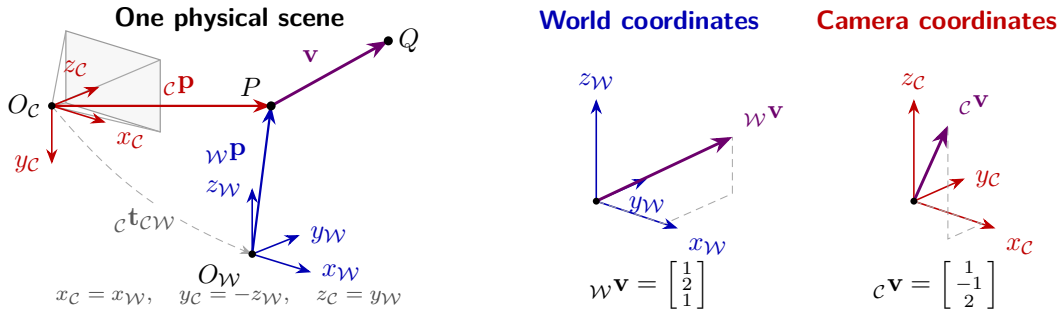


Figure 2: Left: two position arrows to the same point P and the displacement $\mathbf{v} = \overrightarrow{PQ}$. Right: the same displacement expressed in each frame. Since translation cancels in a point difference, $_C\mathbf{v} = \mathbf{R}_{CW}\,{}_W\mathbf{v}$.

2 Example in 3D

In this example the goal is to find the transformation from coordinate system A to C. That is, by applying our transformation we want to be able to express a point from coordinate system A in coordinate system C. See fig. 3 for illustration and specification.

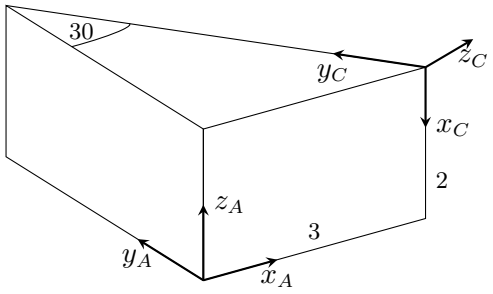


Figure 3: Setting considered for the exercise: Find the transformation from frame A to C.

2.1 Option 1: Concatenating transforms

The mathematically most rigorous option to solve this task is by concatenating transformations. In this example we will use *passive rotations*, that is we rotate the coordinate frames (so all our matrices are the transposed w.r.t the wikipedia[1]). In this case our job is to transform frame A into C with small steps. In this iterative approach we have to append the transformation matrices *to the left*. Verbally, the order of operations is the following:

1. Translation: Shift coordinate system A by 3 units into positive x_A and by 2 units in positive z_A. Note that if we start with the translation, this is tricky in one aspect: we always need to express the vector pointing from C to A for the translation, if we want to transform from frame A to frame C. So the translation is not $\begin{bmatrix} 3 & 0 & 2 \end{bmatrix}^\top$ but actually $\begin{bmatrix} -3 & 0 & -2 \end{bmatrix}^\top$!
2. Rotation 1: Rotate A by 90 degrees around its y_A axis. This will ensure that the x_A and the x_C axis align. We refer to this intermediate frame as A'.

[1] `https://en.wikipedia.org/wiki/Rotation_matrix`

3. Rotation 2: Rotate A' by negative 30 degrees around its $x_{A'}$ axis. Note, this rotation is around the new x-axis $x_{A'}$! And because this axis points downwards, it's negative 30 degrees.

The operations describe above need to be composed from right to left, e.g. the first operation is the *rightmost*! Mathematically (using homegeneous coordinates):

$$\begin{aligned}\mathbf{T}_{CA} &= \underbrace{\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & \cos(-\pi/6) & \sin(-\pi/6) & 0\\ 0 & -\sin(-\pi/6) & \cos(-\pi/6) & 0\\ 0 & 0 & 0 & 1\end{pmatrix}}_{\text{Rotation 2}} \times \underbrace{\begin{pmatrix}\cos(\pi/2) & 0 & -\sin(\pi/2) & 0\\ 0 & 1 & 0 & 0\\ \sin(\pi/2) & 0 & \cos(\pi/2) & 0\\ 0 & 0 & 0 & 1\end{pmatrix}}_{\text{Rotation 1}} \times \underbrace{\begin{pmatrix}1 & 0 & 0 & -3\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & -2\\ 0 & 0 & 0 & 1\end{pmatrix}}_{\text{Translation}}\\ &= \begin{pmatrix}0 & 0 & -1 & 2\\ \sin(-\pi/6) & \cos(-\pi/6) & 0 & -3\sin(-\pi/6)\\ \cos(-\pi/6) & -\sin(-\pi/6) & 0 & -3\cos(-\pi/6)\\ 0 & 0 & 0 & 1\end{pmatrix}\end{aligned}$$

The signs in the above example are only written for illustrative purposes. Of course, in an exam, you would be expected to simplify for example $-\sin(-\pi/6) = \sin(\pi/6) = 0.5$.

2.2 Option 2: Shortcut

Instead of concatenating all the transforms an performing many matrix multiplications, an easier way is to do the following:

1. To find the translation, express the origin of coordinate system A in the coordinates of C.
2. To find the rotation, write the basis vectors of A in the coordinates of C and put them as *columns* in the rotation matrix.

For the translation, we readily obtain (note the angle is now $\pi/6$ as we can't differentiate whether there's individual minus signs anywhere!)

$${}_C\mathbf{t}_{CA} = \begin{bmatrix}2 & 3\sin(\pi/6) & -3\cos(\pi/6)\end{bmatrix}^\top$$

For the rotation, we obtain

$$\begin{aligned}\mathbf{R}_{CA} &= \begin{bmatrix}{}_C\mathbf{x}_A & {}_C\mathbf{y}_A & {}_C\mathbf{z}_A\end{bmatrix}\\ &= \begin{bmatrix}0 & 0 & -1\\ -\sin(\pi/6) & \cos(\pi/6) & 0\\ \cos(\pi/6) & \sin(\pi/6) & 0\end{bmatrix}\end{aligned}$$

All that remains is to compose the final transformation as shown above

$$\mathbf{T}_{CA} = \begin{bmatrix}\mathbf{R}_{CA} \in \mathbb{R}^{3\times3} & {}_C\mathbf{t}_{CA} \in \mathbb{R}^{3\times1}\\ \mathbf{0} \in \mathbb{R}^{1\times3} & 1\end{bmatrix}$$

3 Example: From one camera view to another

Imagine a robot observing the same stationary landmark P at two different poses. We call the camera frames $\mathcal{C}_1$ and $\mathcal{C}_2$. In each camera, x points right, y down and z forward. Our goal is to compute ${}_{\mathcal{C}_2}\mathbf{p}$ from ${}_{\mathcal{C}_1}\mathbf{p}$: the landmark stays still, while its coordinates change.

3.1 Move the camera, then change coordinates

The second camera is 4 m to the right of the first and turns right by $\theta = \arctan(3/4) \approx 36.87°$. This angle gives $\cos\theta = 4/5$ and $\sin\theta = 3/5$, making the arithmetic simple. The position of its origin and the landmark, both expressed in $\mathcal{C}_1$, are

$$\boxed{\mathbf{c} = {}_{\mathcal{C}_1}\mathbf{t}_{\mathcal{C}_1\mathcal{C}_2} = \begin{bmatrix}4 & 0 & 0\end{bmatrix}^\top}, \qquad {}_{\mathcal{C}_1}\mathbf{p} = \begin{bmatrix}4 & 1 & 4\end{bmatrix}^\top \quad \text{(metres)}.$$

Notation note. $\mathbf{c}$ is shorthand for ${}_{\mathcal{C}_1}\mathbf{t}_{\mathcal{C}_1\mathcal{C}_2}$. It points **from camera 1's origin to camera 2's origin** and is expressed along **camera 1's axes**.

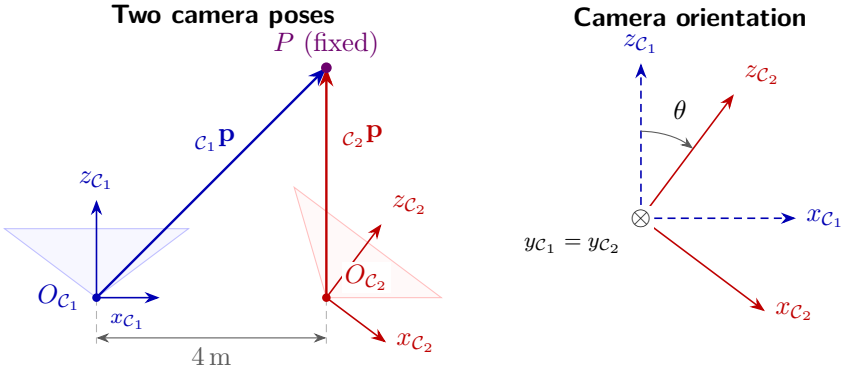


Figure 4: Top view: both arrows end at the same landmark. The landmark is also 1 m below the cameras, which is hidden in this view. On the right, the axes are overlaid at a common origin to show the rotation; $\otimes$ means that positive y points into the page.

Step 1: Shift the origin. The vector from $O_{\mathcal{C}_2}$ to P, still expressed along the axes of $\mathcal{C}_1$, is

$${}_{\mathcal{C}_1}\mathbf{p} - \mathbf{c} = \begin{bmatrix}4 & 1 & 4\end{bmatrix}^\top - \begin{bmatrix}4 & 0 & 0\end{bmatrix}^\top = \begin{bmatrix}0 & 1 & 4\end{bmatrix}^\top.$$

Moving the camera to the right makes the landmark's rightward displacement smaller.

Step 2: Express this vector along the new axes. Taking its dot product with each camera-2 unit axis gives the new components:

$$\mathbf{R}_{\mathcal{C}_2\mathcal{C}_1} = \begin{bmatrix}4/5 & 0 & -3/5\\ 0 & 1 & 0\\ 3/5 & 0 & 4/5\end{bmatrix}, \qquad {}_{\mathcal{C}_2}\mathbf{p} = \mathbf{R}_{\mathcal{C}_2\mathcal{C}_1}\begin{bmatrix}0\\1\\4\end{bmatrix} = \boxed{\begin{bmatrix}-12/5\\1\\16/5\end{bmatrix}}.$$

The landmark is now 2.4 m to the left and 1 m below the camera, with a forward depth of 3.2 m.

Step 3: Combine the two operations. Expanding ${}_{\mathcal{C}_2}\mathbf{p} = \mathbf{R}_{\mathcal{C}_2\mathcal{C}_1}({}_{\mathcal{C}_1}\mathbf{p} - \mathbf{c})$ gives

$${}_{\mathcal{C}_2}\mathbf{t}_{\mathcal{C}_2\mathcal{C}_1} = -\mathbf{R}_{\mathcal{C}_2\mathcal{C}_1}\mathbf{c} = \begin{bmatrix}-16/5\\0\\-12/5\end{bmatrix}, \qquad \boxed{\mathbf{T}_{\mathcal{C}_2\mathcal{C}_1} = \begin{bmatrix}4/5 & 0 & -3/5 & -16/5\\ 0 & 1 & 0 & 0\\ 3/5 & 0 & 4/5 & -12/5\\ 0 & 0 & 0 & 1\end{bmatrix}}.$$

The translation column is the position of $O_{\mathcal{C}_1}$ expressed in $\mathcal{C}_2$. This is why we both negate and rotate the camera displacement $\mathbf{c}$.