Cylindrical image (Panorama)

April 23, 2024

1 Cylindrical coordiante system

Consider an orthonormal coordinate system $(A, \underbrace{[\vec{a}_1, \vec{a}_2, \vec{a}_3]}_{\alpha})$ and a cylinder defined by the set of points

$$\mathcal{C}_{(A,\alpha)} = \{e \mid e_{(A,\alpha)} = \begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix}^\top,\ \ e_1^2 + e_2^2 = 1\}$$

as is depicted in Figure 1.

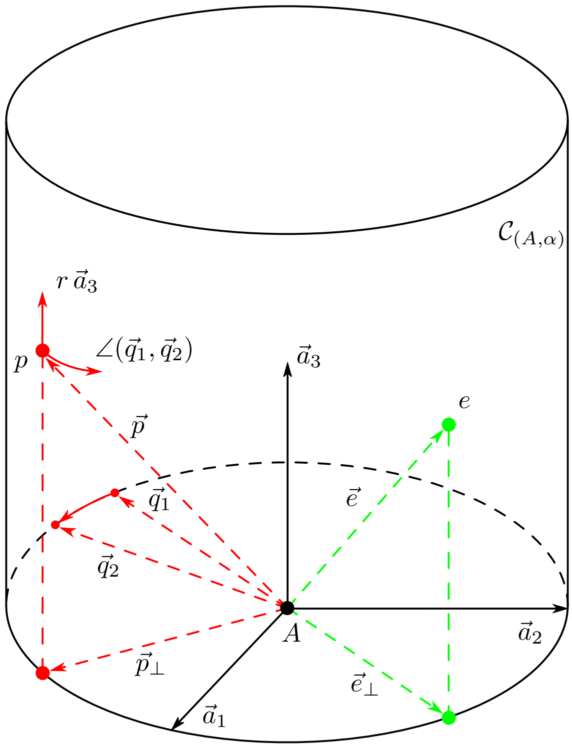

Figure 1: The cylinder and its coordinate system

Cylindrical coordinate system $\left(p, \underbrace{\begin{bmatrix} \angle(\vec{q}_1, \vec{q}_2) & r \end{bmatrix}}_{\psi}\right)$ of the cylinder $\mathcal{C}_{(A,\alpha)}$ consists of 3 elements:

1. The origin $p \in \mathcal{C}_{(A,\alpha)}$,

2. The angle resolution $\angle(\vec{q}_1, \vec{q}_2)$ defined by some $\vec{q}_1, \vec{q}_2 \perp \vec{a}_3$, $\vec{q}_1 \not\parallel \vec{q}_2$,
3. The vertical resolution $r \in \mathbb{R}\backslash\{0\}$ in α units.

In order to define the coordinates of a point $e \in \mathcal{C}_{(A,\alpha)}$ in a cylindrical coordinate system (p, ψ), we express $\vec{e}$ and $\vec{p}$ in α

$$\vec{e}_\alpha = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}, \quad \vec{p}_\alpha = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix},$$

form their projections onto the plane spanned by $\vec{a}_1$ and $\vec{a}_2$

$$\vec{e}_\perp = e_1\vec{a}_1 + e_2\vec{a}_2, \quad \vec{p}_\perp = p_1\vec{a}_1 + p_2\vec{a}_2$$

and define

$$e_{(p,\psi)} \stackrel{\text{def}}{=} \begin{bmatrix} \frac{\angle(\vec{p}_\perp, \vec{e}_\perp)}{\angle(\vec{q}_1, \vec{q}_2)} \\ \frac{e_3 - p_3}{r} \end{bmatrix} \tag{1}$$

If we denote

$$\mathcal{C} = \left\{ e_{(A,\alpha)} \mid e \in \mathcal{C}_{(A,\alpha)} \right\} \subset \mathbb{R}^3$$

then, depending on how we define the angle function $\angle(\cdot,\cdot)$, the function

$$\varphi \colon \mathcal{C} \to \mathbb{R}^2$$
$$e_{(A,\alpha)} \mapsto e_{(p,\psi)}$$

will have discontinuities at different lines on the cylinder. There are two common choices for the angle to make $\angle(\cdot,\cdot)$ to either belong to the interval $[0, 2\pi)$ or $(-\pi, \pi]$ (there is also a choice in the direction, which however doesn't influence the discontinuities). The vertical line across which the tearing happens in the case when $\angle(\cdot,\cdot) \in (-\pi, \pi]$ is shown in Figure 2 in blue.

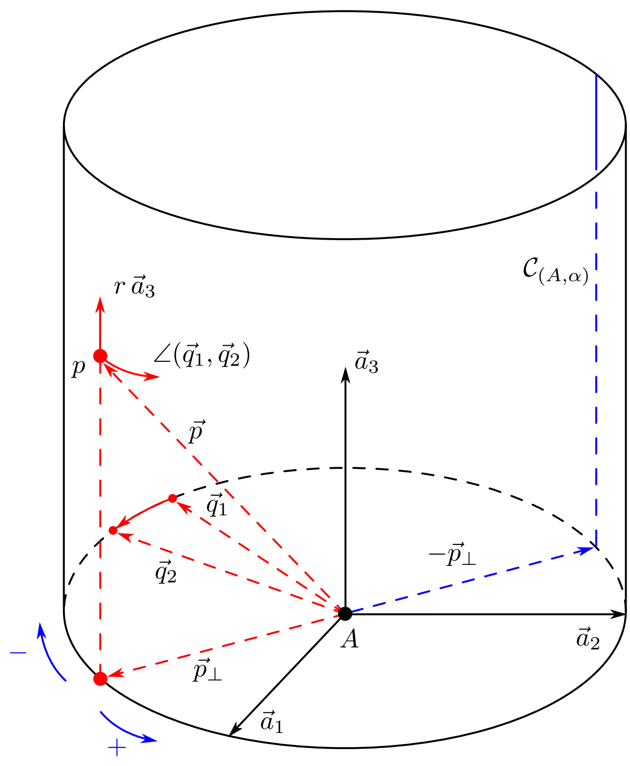

Figure 2: The angle function $\angle(\cdot,\cdot) \in (-\pi, \pi]$ causes the tearing of the cylinder along the blue vertical line.

In the case when $\angle(\cdot,\cdot) \in [0, 2\pi)$ the tearing would happen along the vertical line that passes through p.

We will see later that for constructing the cylindrical image, it is important to choose an appropriate definition of the angle function $\angle(\cdot,\cdot)$ in order to not tear the cylindrical image somewhere in the middle.

2 Projection to cylinder

If we have a general point x in space (not necessarily on $\mathcal{C}_{(A,\alpha)}$), we can project it along the ray that joins A and x denoted by $\vec{x}$ to $e \in \mathcal{C}_{(A,\alpha)}$. We are looking for λ such that

$$\vec{e} = \lambda \vec{x}$$

$$\vec{e}_\alpha = \lambda \vec{x}_\alpha$$

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Since $e \in \mathcal{C}_{(A,\alpha)}$, then we have $e_1^2 + e_2^2 = 1$, and hence

$$1 = e_1^2 + e_2^2 = \lambda^2 x_1^2 + \lambda^2 x_2^2 \iff \lambda^2 = \frac{1}{x_1^2 + x_2^2} \iff \lambda = \pm \frac{1}{\sqrt{x_1^2 + x_2^2}}$$

Having 2 values for λ corresponds to the fact that a ray defined by $\vec{x}$ intersects the cylinder $\mathcal{C}_{(A,\alpha)}$ at two different points represented by vectors

$$\vec{e}_1 = \frac{1}{\sqrt{x_1^2 + x_2^2}} \vec{x} \tag{2}$$

$$\vec{e}_2 = -\frac{1}{\sqrt{x_1^2 + x_2^2}} \vec{x} = -\vec{e}_1 \tag{3}$$

3 Constructing panorama

3.1 Cylindrical image surface

Having a projective camera with a cartesian camera coordinate system (C, γ), we first define the cylinder $\mathcal{C}_{(C,\gamma)}$ and its coordinate system. The cylinder is defined by the set of points

$$\mathcal{C}_{(C,\gamma)} = \{e \mid e_{(C,\gamma)} = \begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix}^\top,\ e_1^2 + e_3^2 = 1\}.$$

Notice that unlike in the previous sections, we define the cylinder here a bit differently: its axis goes along $\vec{c}_2$ (not $\vec{c}_3$). Hence Equations (1), (2) and (3) have to be changed appropriately.

The center p of the coordinate system of $\mathcal{C}_{(C,\gamma)}$ is defined to be the principal point of the camera. The horizontal resolution is defined to be the directed angle $\angle(\vec{c}_3, \vec{c}_3 + \vec{b}_1)$, since we would like to achieve approximately the same horizontal resolution in the cylindrical image as in the perspective image itself. As for the vertical resolution, we would like to get rid of the affine distortion in the perspective image caused by the non-orthonormality of $(\vec{b}_1, \vec{b}_2)$. In order to achieve this, we define the vertical resolution r to be the length of $\vec{b}_1$ in γ units, i.e. $r = \frac{\|\vec{b}_1\|}{f}$.

Before looking at the horizontal resolution we define the angle function $\angle(\cdot,\cdot)$. Since we defined p to be a principal point and the angle resolution to be $\angle(\vec{c}_3, \vec{c}_3 + \vec{b}_1)$, it will be sufficient for us to define $\angle(\vec{c}_3, \vec{v})$ for $\vec{v} \in \langle \vec{c}_1, \vec{c}_3 \rangle$, since this is all we need to evaluate Equation (1). If

$$\vec{v}_\gamma = \begin{bmatrix} v_1 \\ 0 \\ v_3 \end{bmatrix}$$

then we define

$$\angle(\vec{c}_3, \vec{v}) \stackrel{\text{def}}{=} \operatorname{atan2}(v_1, v_3) \in (-\pi, \pi]$$

The geometry of such a definition is visualized in Figure 3 in magenta color. The tearing in the cylindrical coordinates happens along the line ℓ_1, and as a consequence when we project the image to the cylinder it will not be teared when visualized in the cylindrical coordinates, because ℓ_1 is behind the camera. If we used another common definition of the angle $\angle(\vec{c}_3, \vec{v}) \in [0, 2\pi)$, we would tear the cylindrical image along the line ℓ_2, because ℓ_2 is in front of the camera.

By looking at the horizontal resolution of the cylindrical coordinate system, we see that

$$\vec{c}_3 \perp \vec{b}_1 \Rightarrow \angle(\vec{c}_3, \vec{c}_3 + \vec{b}_1) = \arctan\left(\frac{\|\vec{b}_1\|}{\|\vec{c}_3\|}\right)$$

which, since $\|\vec{c}_3\| \gg \|\vec{b}_1\|$, can be written approximately as

$$\angle(\vec{c}_3, \vec{c}_3 + \vec{b}_1) \approx \frac{\|\vec{b}_1\|}{\|\vec{c}_3\|} = \frac{\|\vec{b}_1\|}{f} = \frac{1}{k_{11}}$$

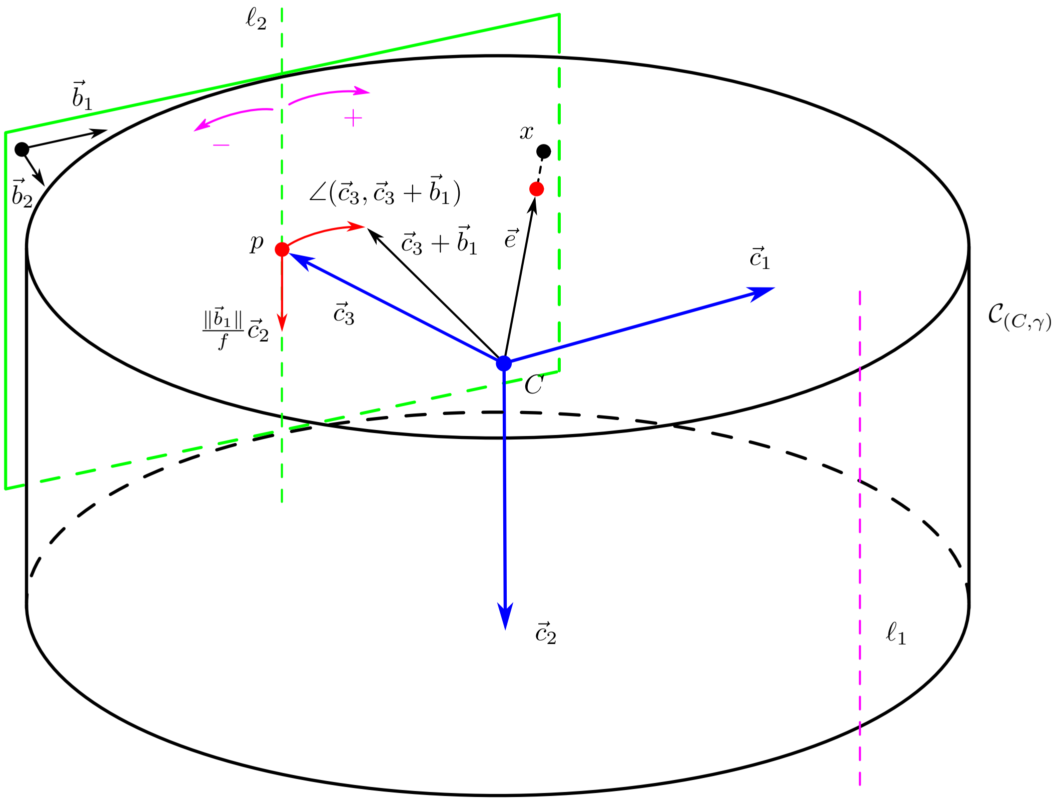

Figure 3: The cylindrical image surface $\mathcal{C}_{(C,\gamma)}$ and its coordinate system

The vertical resolution is

$$r = \frac{\|\vec{b}_1\|}{f} = \frac{1}{k_{11}}$$

and the coordinates $p_{(C,\gamma)}$ of the center of the cylindrical coordinate system in the camera cartesian coordinate system are

$$p_{(C,\gamma)} = \vec{c}_{3_\gamma} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

By looking at a general point x in the image, we project it along $\vec{x}$ to the cylinder and get a point $e \in \mathcal{C}_{(C,\gamma)}$ represented by vector $\vec{e}$. To get its coordinates in γ, we compute

$$\vec{x}_\gamma = \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix} \Rightarrow \vec{e}_\gamma = \frac{1}{\sqrt{x_1^2+1}} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}, \quad \vec{e}_\perp \sim x_1\vec{c}_1 + \vec{c}_3$$

where the scale λ according to Section 2 was chosen to be positive in order to obtain the projected point in front of the camera. As for the angle between $\vec{c}_3$ and $\vec{e}_\perp$ we have

$$\angle(\vec{c}_3, \vec{e}_\perp) = \text{atan2}\,(x_1, 1)$$

The coordinates of e in the cylindrical coordinate system defined above can be now obtained as

$$e_{(p,\psi)} = \begin{bmatrix} \frac{\angle(\vec{c}_3,\vec{e}_\perp)}{\angle(\vec{c}_3,\vec{c}_3+\vec{b}_1)} \\ \frac{\vec{e}_{\gamma,2}-\vec{c}_{3\gamma,2}}{r} \end{bmatrix} = k_{11} \cdot \begin{bmatrix} \text{atan2}\,(x_1,1) \\ \frac{x_2}{\sqrt{x_1^2+1}} \end{bmatrix}$$

3.2 Gluing images on the cylinder

Suppose we took n images with n cameras that all have the same projection center C. Let's order the cameras in a way that the cylinder will be defined in the coordinate system of the first camera. Let us take the j-th camera and show how to express the projections of its image points in the coordinate system of $\mathcal{C}_{(C,\gamma_1)}$.

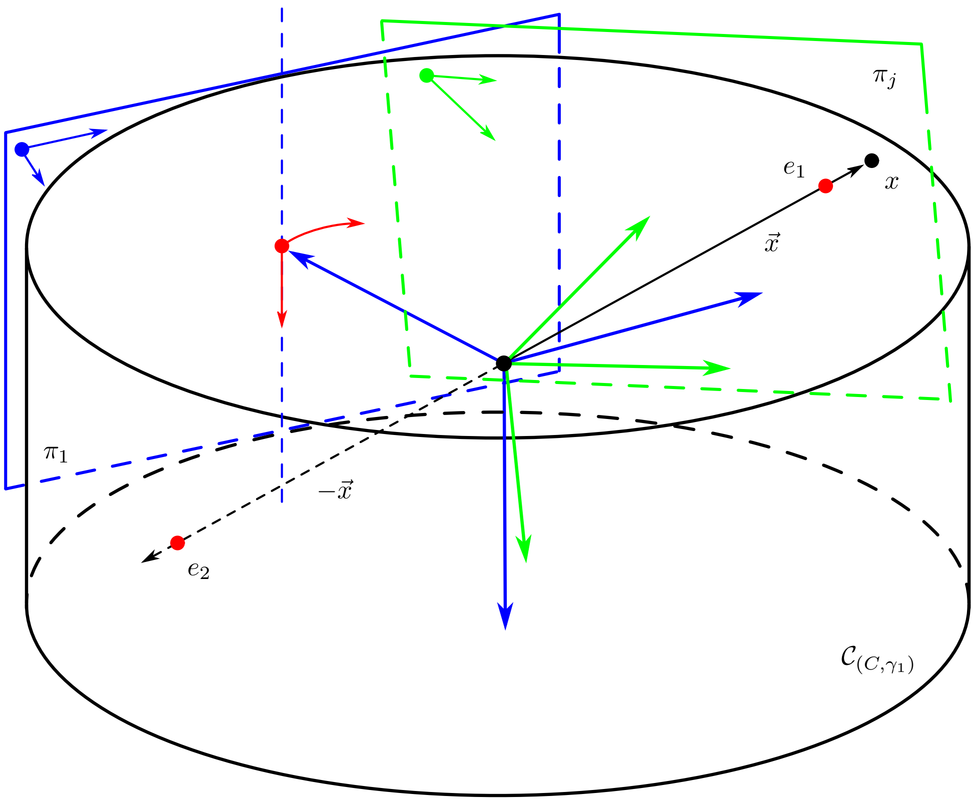

Figure 4: Projecting the image points in π_j onto the cylinder $\mathcal{C}_{(C,\gamma_1)}$

Since $C_1 = C_j = C$, then according to [1, Section 8.1], there is a homography $\mathtt{H}_j$ that transforms the coordinates from β_j to β_1 (see [1, Equation (8.4)]). Further, we have

$$\mathtt{H}_j = \mathtt{T}_{\beta_j\to\beta_1} = \mathtt{T}_{\gamma_1\to\beta_1}\mathtt{T}_{\gamma_j\to\gamma_1}\mathtt{T}_{\beta_j\to\gamma_j} = \mathtt{K}_1\, s\mathtt{R}\,\mathtt{K}_j^{-1}$$

where $\mathtt{K}_1, \mathtt{K}_j$ are the camera calibration matrices of the 1-st and j-th cameras, and $s\mathtt{R}$ (scaled rotation) is the transition matrix from γ_j to γ_1. Notice that $s > 0$, since γ_1 and γ_j are both right-handed (and thus a transition matrix between them must have positive determinant). Since we also have $\det \mathtt{K}_1 > 0$ and $\det \mathtt{K}_j > 0$ (by the choice made in this course, namely, $k_{11} > 0$ and $k_{22} > 0$), then we have a semi-algebraic constraint on $\mathtt{H}_j$:

$$\det \mathtt{H}_j = s^3 \det \mathtt{K}_1 \det \mathtt{R} \frac{1}{\det \mathtt{K}_j} = s^3 \frac{\det \mathtt{K}_1}{\det \mathtt{K}_j} > 0$$

If we have recovered only a multiple $\mathtt{G}_j = \tau \mathtt{H}_j$, then we can obtain a multiple of the transition matrix from γ_j to γ_1:

$$\tau \mathtt{T}_{\gamma_j \to \gamma_1} = \tau s \mathtt{R} = \tau \mathtt{K}_1^{-1} \mathtt{H}_j \mathtt{K}_j = \mathtt{K}_1^{-1} \mathtt{G}_j \mathtt{K}_j$$

In order to project a general point $x \in \pi_j$ to the cylinder $\mathcal{C}_{(C,\gamma_1)}$, we express a vector $\vec{x}$ that represents x in (C, γ_j) in γ_1 as

$$(\tau \vec{x})_{\gamma_1} = \tau \vec{x}_{\gamma_1} = \tau \mathtt{T}_{\gamma_j \to \gamma_1} \vec{x}_{\gamma_j} = \mathtt{K}_1^{-1} \mathtt{G}_j \mathtt{K}_j \vec{x}_{\gamma_j}$$

Our aim is to obtain the projection of x onto the cylinder $\mathcal{C}_{(C,\gamma_1)}$ that will be in front of the j-th camera. As is shown in Figure 4, we are interested in e_1, and not in e_2. For this, we need to apply the projection of $\tau \vec{x}$ for $\tau > 0$ according to Equation (2). All that is left is to obtain a positive multiple of $\mathtt{H}_j$ from $\mathtt{G}_j$. This can be done by considering

$$\det \mathtt{G}_j \cdot \mathtt{G}_j = \tau^3 \cdot \det \mathtt{H}_j \cdot \tau \mathtt{H}_j = \underbrace{\tau^4 \cdot \det \mathtt{H}_j}_{>0} \cdot \mathtt{H}_j$$

References

[1] Tomas Pajdla, *Elements of geometry for computer vision*, `https://cw.fel.cvut.cz/wiki/_media/courses/gvg/pajdla-gvg-lecture-2021.pdf`.