Image Processing and Computer Vision (Module 2)

Last update: 16 April 2024

Academic Year 2023 – 2024

Alma Mater Studiorum · University of Bologna

Contents

1 Camera calibration

World reference frame (WRF) Coordinate system (X_W, Y_W, Z_W) of the real world relative to a reference point (e.g. a corner).

World reference frame (WRF)

Camera reference frame (CRF) Coordinate system (X_C, Y_C, Z_C) that characterizes a camera.

Camera reference frame (CRF)

Image reference frame (IRF) Coordinate system (U, V) of the image. They are obtained as a perspective projection of CRF coordinates as:

Image reference frame

$$u = \frac{f}{z} x_C \qquad v = \frac{f}{z} y_C$$

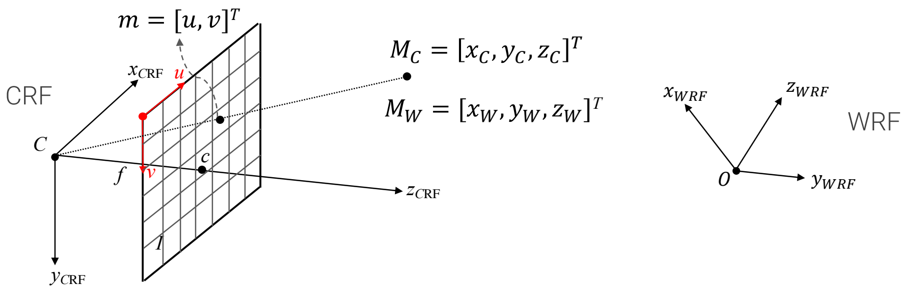

Figure 1.1: Example of WRF, CRF and IRF

1.1 Forward imaging model

1.1.1 Image pixelization (CRF to IRF)

The conversion from the camera reference frame to the image reference frame is done in two steps:

Image pixelization

Discetization Given the sizes (in mm) Δu and Δv of the pixels, it is sufficient to modify the perspective projection to map CRF coordinates into a discrete grid:

Discetization

$$u = \frac{1}{\Delta u} \frac{f}{z_C} x_C \qquad v = \frac{1}{\Delta v} \frac{f}{z_C} y_C$$

Origin translation To avoid negative pixels, the origin of the image has to be translated from the piercing point c to the top-left corner. This is done by adding an offset (u_0, v_0) to the projection (in the new system, $c = (u_0, v_0)$):

Origin translation

$$u = \frac{1}{\Delta u} \frac{f}{z_C} x_C + u_0 \qquad v = \frac{1}{\Delta v} \frac{f}{z_C} y_C + v_0$$

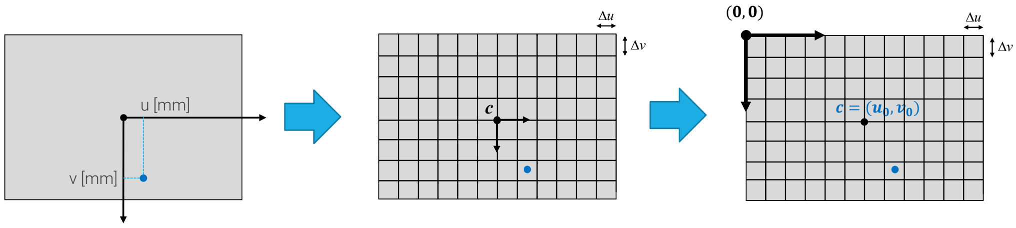

Figure 1.2: Pixelization process

Intrinsic parameters

Intrinsic parameters Parameters needed to convert from CRF to IRF.

By fixing $f_u = \frac{f}{\Delta u}$ and $f_v = \frac{f}{\Delta v}$, the projection can be rewritten as:

$$u = f_u \frac{x_C}{z_C} + u_0 \qquad v = f_v \frac{y_C}{z_C} + v_0$$

Therefore, there is a total of 4 parameters: f_u, f_v, u_0 and v_0.

1.1.2 Roto-translation (WRF to CRF)

Roto-translation

The conversion from the world reference system to the camera reference system is done through a roto-translation wrt the optical center.
Given:

- A WRF point $\mathbf{M}_W = (x_W, y_W, z_W)$,
- A rotation matrix $\boldsymbol{R}$,
- A translation vector $\mathbf{t}$,

the coordinates $\mathbf{M}_C$ in CRF corresponding to $\mathbf{M}_W$ are given by:

$$\mathbf{M}_C = \begin{pmatrix} x_C \\ y_C \\ z_C \end{pmatrix} = \boldsymbol{R}\mathbf{M}_W + \mathbf{t} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} x_W \\ y_W \\ z_W \end{pmatrix} + \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}$$

Remark. The coordinates $\mathbf{C}_W$ of the optical center $\mathbf{C}$ are obtained as:

$$\bar{\mathbf{0}} = \boldsymbol{R}\mathbf{C}_W + \mathbf{t} \iff (\bar{\mathbf{0}} - \mathbf{t}) = \boldsymbol{R}\mathbf{C}_W \iff \mathbf{C}_W = \boldsymbol{R}^T(\bar{\mathbf{0}} - \mathbf{t}) \iff \mathbf{C}_W = -\boldsymbol{R}^T\mathbf{t}$$

Extrinsic parameters

Extrinsic parameters

- The rotation matrix $\boldsymbol{R}$ has 9 elements of which 3 are independent (i.e. the rotation angles around the axes).
- The translation matrix $\mathbf{t}$ has 3 elements.

Therefore, there is a total of 6 parameters.

Remark. It is not possible to combine the intrinsic camera model and the extrinsic roto-translation to create a linear model for the forward imaging model.

$$u = f_u \frac{r_{11}x_W + r_{12}y_W + r_{13}z_W + t_1}{r_{31}x_W + r_{32}y_W + r_{33}z_W + t_3} + u_0 \qquad v = f_v \frac{r_{21}x_W + r_{22}y_W + r_{23}z_W + t_2}{r_{31}x_W + r_{32}y_W + r_{33}z_W + t_3} + v_0$$

1.2 Projective space

Remark. In the 2D Euclidean plane $\mathbb{R}^2$, parallel lines never intersect and points at infinity cannot be represented.

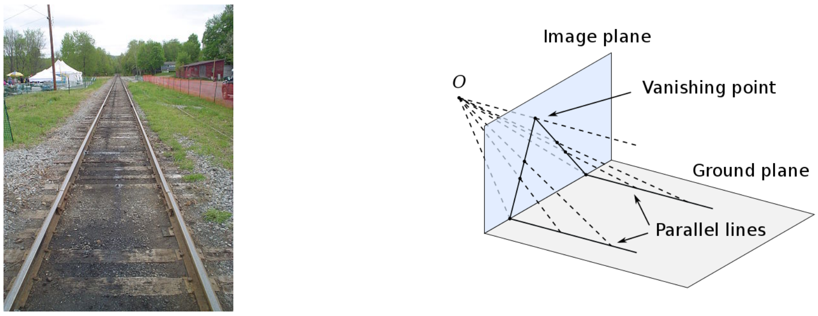

Figure 1.3: Example of point at infinity

Remark. Point at infinity is a point in space while the vanishing point is in the image plane.

Homogeneous coordinates

Homogeneous coordinates Without loss of generality, consider the 2D Euclidean space $\mathbb{R}^2$.

Given a coordinate (u, v) in Euclidean space, its homogeneous coordinates have an additional dimension such that:

$$(u, v) \equiv (ku, kv, k) \, \forall k \neq 0$$

In other words, a 2D Euclidean point is represented by an equivalence class of 3D points.

Projective space

Projective space Space $\mathbb{P}^n$ associated with the homogeneous coordinates of an Euclidean space $\mathbb{R}^n$.

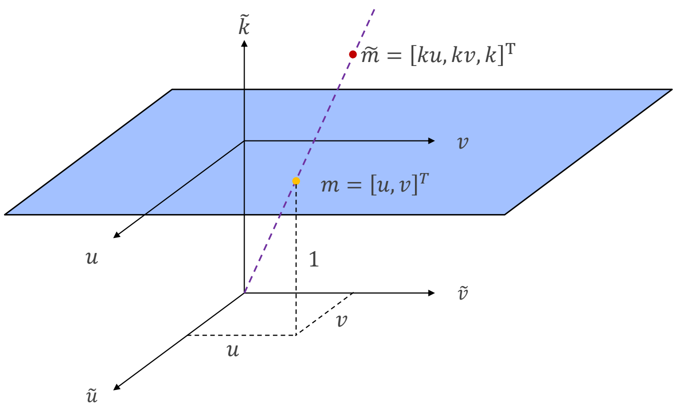

Figure 1.4: Example of projective space $\mathbb{P}^2$

Remark. $\bar{\mathbf{0}}$ is not a valid point in $\mathbb{P}^n$.

Remark. A projective space allows to homogeneously handle both ordinary (image) and ideal (scene) points without introducing additional complexity.

Point at infinity Given the parametric equation of a 2D line defined as:

Point at infinity

$$\mathbf{m} = \mathbf{m}_0 + \lambda \mathbf{d} = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} + \lambda \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} u_0 + \lambda a \\ v_0 + \lambda b \end{pmatrix}$$

It is possible to define a generic point in the projective space along the line m as:

$$\tilde{\mathbf{m}} \equiv \begin{pmatrix} \mathbf{m} \\ 1 \end{pmatrix} \equiv \begin{pmatrix} u_0 + \lambda a \\ v_0 + \lambda b \\ 1 \end{pmatrix} \equiv \begin{pmatrix} \frac{u_0}{\lambda} + a \\ \frac{v_0}{\lambda} + b \\ \frac{1}{\lambda} \end{pmatrix}$$

The projective coordinates $\tilde{\mathbf{m}}_\infty$ of the point at infinity of a line m is given by:

$$\tilde{\mathbf{m}}_\infty = \lim_{\lambda \to \infty} \tilde{\mathbf{m}} \equiv \begin{pmatrix} a \\ b \\ 0 \end{pmatrix}$$

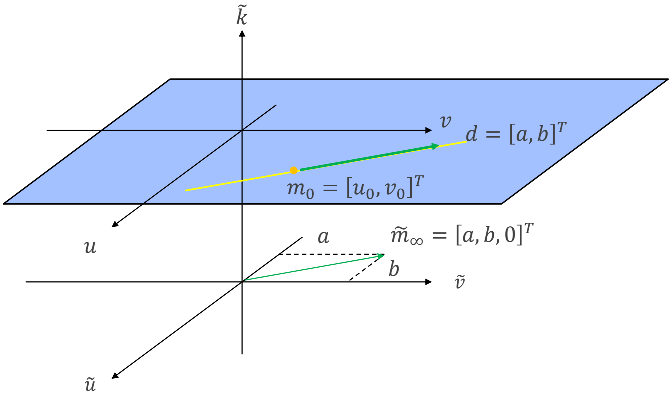

Figure 1.5: Example of infinity point in $\mathbb{P}^2$

In 3D, the definition is trivially extended as:

$$\tilde{\mathbf{M}}_\infty = \lim_{\lambda \to \infty} \begin{pmatrix} \frac{x_0}{\lambda} + a \\ \frac{y_0}{\lambda} + b \\ \frac{z_0}{\lambda} + c \\ \frac{1}{\lambda} \end{pmatrix} \equiv \begin{pmatrix} a \\ b \\ c \\ 0 \end{pmatrix}$$

Perspective projection Given a point $\mathbf{M}_C = (x_C, y_C, z_C)$ in the CRF and its corresponding point $\mathbf{m} = (u, v)$ in the image, the non-linear perspective projection in Euclidean space can be done linearly in the projective space as:

Perspective projection in projective space

$$\begin{aligned}
\tilde{\mathbf{m}} \equiv \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} &\equiv \begin{pmatrix} f_u \frac{x_C}{z_C} + u_0 \\ f_v \frac{y_C}{z_C} + v_0 \\ 1 \end{pmatrix} \equiv z_C \begin{pmatrix} f_u \frac{x_C}{z_C} + u_0 \\ f_v \frac{y_C}{z_C} + v_0 \\ 1 \end{pmatrix} \\
&\equiv \begin{pmatrix} f_u x_C + z_C u_0 \\ f_v y_C + z_C v_0 \\ z_C \end{pmatrix} \equiv \begin{pmatrix} f_u & 0 & u_0 & 0 \\ 0 & f_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_C \\ y_C \\ z_C \\ 1 \end{pmatrix} \equiv \boldsymbol{P}_{\text{int}} \tilde{\mathbf{M}}_C
\end{aligned}$$

Remark. The equation can be written to take account of the arbitrary scale factor k as:

$$k\tilde{\mathbf{m}} = \boldsymbol{P}_{\text{int}}\tilde{\mathbf{M}}_C$$

or, if k is omitted, as:

$$\tilde{\mathbf{m}} \approx \boldsymbol{P}_{\text{int}}\tilde{\mathbf{M}}_C$$

Remark. In projective space, we can also project in Euclidean space the point at infinity of parallel 3D lines in CRF with direction (a, b, c):

$$\tilde{\mathbf{m}}_\infty \equiv \boldsymbol{P}_{\text{int}} \begin{pmatrix} a \\ b \\ c \\ 0 \end{pmatrix} \equiv \begin{pmatrix} f_u & 0 & u_0 & 0 \\ 0 & f_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ 0 \end{pmatrix} \equiv \begin{pmatrix} f_u a + c u_0 \\ f_v b + c v_0 \\ c \end{pmatrix} \equiv c \begin{pmatrix} f_u \frac{a}{c} + u_0 \\ f_v \frac{b}{c} + v_0 \\ 1 \end{pmatrix}$$

Therefore, the Euclidean coordinates are:

$$\mathbf{m}_\infty = \begin{pmatrix} f_u \frac{a}{c} + u_0 \\ f_v \frac{b}{c} + v_0 \end{pmatrix}$$

Note that this is not possible when $c = 0$.