Image Processing and Computer Vision (Module 2)

Last update: 20 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

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

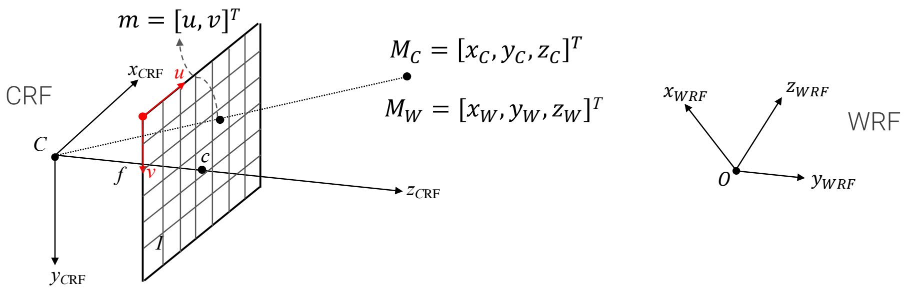


Figure 1.1: Example of WRF, CRF and IRF

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

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

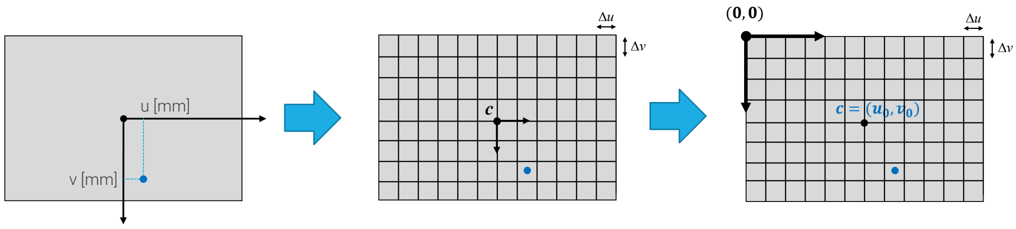


Figure 1.2: Pixelization process

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

Intrinsic parameters

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

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

Remark. A more general model includes a further parameter (skew) to account for non-orthogonality between the axes of the image sensor such as:

- Misplacement of the sensor so that it is not perpendicular to the optical axis.
- Manufacturing issues.

Nevertheless, in practice skew is always 0.

1.1.2 Roto-translation (WRF to CRF)

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

Roto-translation

- 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:

$$\mathbf{M}_C = \begin{bmatrix} x_C \\ y_C \\ z_C \end{bmatrix} = \boldsymbol{R}\mathbf{M}_W + \mathbf{t} = \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} \\ r_{2,1} & r_{2,2} & r_{2,3} \\ r_{3,1} & r_{3,2} & r_{3,3} \end{bmatrix} \begin{bmatrix} x_W \\ y_W \\ z_W \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}$$

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.

$$u = f_u \frac{r_{1,1}x_W + r_{1,2}y_W + r_{1,3}z_W + t_1}{r_{3,1}x_W + r_{3,2}y_W + r_{3,3}z_W + t_3} + u_0 \qquad v = f_v \frac{r_{2,1}x_W + r_{2,2}y_W + r_{2,3}z_W + t_2}{r_{3,1}x_W + r_{3,2}y_W + r_{3,3}z_W + t_3} + v_0$$

1.2 Projective space

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

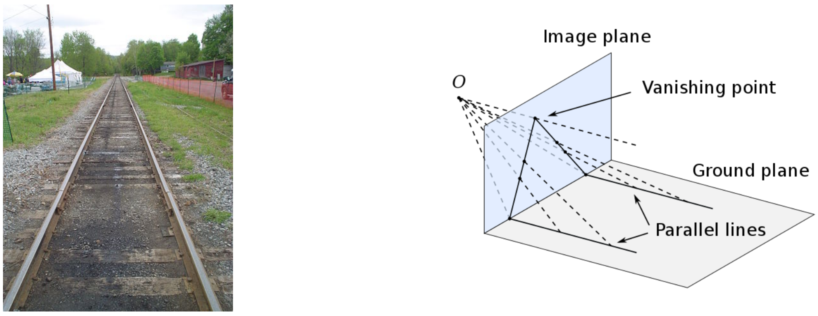


Figure 1.3: Example of point at infinity

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 $\mathbb{R}^n$.

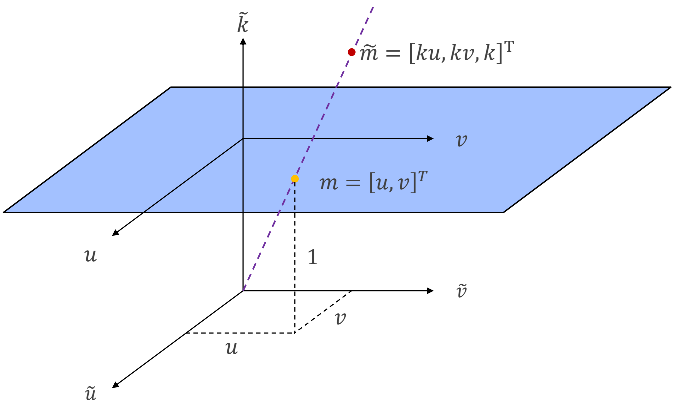


Figure 1.4: Example of projective space $\mathbb{P}^2$

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

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

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

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

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

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

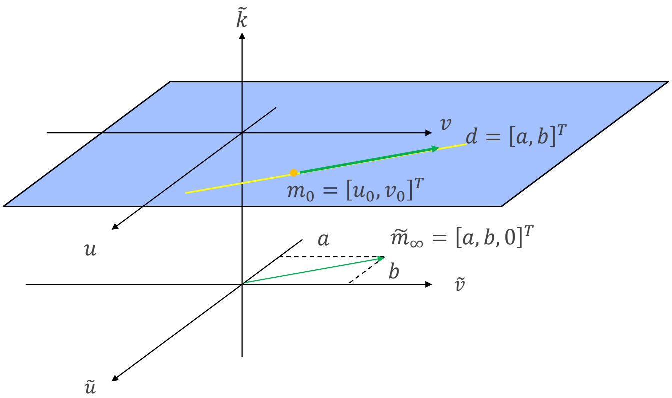


Figure 1.5: Example of infinity point in $\mathbb{P}^2$

In 3D, the definition is trivially extended as:

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

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

$$\tilde{\mathbf{m}} \equiv \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \equiv \begin{bmatrix} f_u \frac{x_C}{z_C} + u_0 \\ f_v \frac{y_C}{z_C} + v_0 \\ 1 \end{bmatrix} \equiv z_C \begin{bmatrix} f_u \frac{x_C}{z_C} + u_0 \\ f_v \frac{y_C}{z_C} + v_0 \\ 1 \end{bmatrix}$$

$$\equiv \begin{bmatrix} f_u x_C + z_C u_0 \\ f_v y_C + z_C v_0 \\ z_C \end{bmatrix} \equiv \begin{bmatrix} f_u & 0 & u_0 & 0 \\ 0 & f_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_C \\ y_C \\ z_C \\ 1 \end{bmatrix} \equiv \boldsymbol{P}_{\text{int}} \tilde{\mathbf{M}}_C$$

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):

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

Therefore, the Euclidean coordinates are:

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

Note that this is not possible when $c = 0$ (i.e. the line is parallel to the image plane).

Intrinsic parameter matrix The intrinsic transformation can be expressed through a matrix:

Intrinsic parameter matrix

$$\boldsymbol{A} = \begin{bmatrix} f_u & 0 & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

$\boldsymbol{A}$ is always upper right triangular and models the characteristics of the imaging device.

Remark. If skew is considered, it would be at position $(1, 2)$.

Extrinsic parameter matrix The extrinsic transformation can be expressed through a matrix:

Extrinsic parameter matrix

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{R} & \mathbf{t} \\ \bar{\mathbf{0}} & 1 \end{bmatrix} = \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} & t_1 \\ r_{2,1} & r_{2,2} & r_{2,3} & t_2 \\ r_{3,1} & r_{3,2} & r_{3,3} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Perspective projection matrix (PPM) As the following hold:

Perspective projection matrix

$$\boldsymbol{P}_{\text{int}} = [\boldsymbol{A}|\bar{\mathbf{0}}] \qquad \tilde{\mathbf{M}}_C \equiv \boldsymbol{G}\tilde{\mathbf{M}}_W$$

The perspective projection can be represented in matrix form as:

$$\tilde{\mathbf{m}} \equiv \boldsymbol{P}_{\text{int}}\tilde{\mathbf{M}}_C \equiv \boldsymbol{P}_{\text{int}}\boldsymbol{G}\tilde{\mathbf{M}}_W \equiv \boldsymbol{P}\tilde{\mathbf{M}}_W$$

where $\boldsymbol{P} = \boldsymbol{P}_{\text{int}}\boldsymbol{G}$ is the perspective projection matrix. It is full-rank and has shape 3×4.

Remark. Every full-rank 3×4 matrix is a PPM.

Canonical perspective projection PPM of form:

Canonical perspective projection

$$\boldsymbol{P} \equiv [\boldsymbol{I}|\bar{\mathbf{0}}]$$

It is useful to represent the core operations carried out by a perspective projection as any general PPM can be factorized as:

$$\boldsymbol{P} \equiv \boldsymbol{A}[\boldsymbol{I}|\bar{\mathbf{0}}]\boldsymbol{G}$$

where:

- $\boldsymbol{G}$ converts from WRT to CRF.
- $[\boldsymbol{I}|\bar{\mathbf{0}}]$ performs the canonical perspective projection (i.e. divide by the third coordinate).
- $\boldsymbol{A}$ applies camera specific transformations.

A further factorization is:

$$\boldsymbol{P} \equiv \boldsymbol{A}[\boldsymbol{I}|\bar{\mathbf{0}}]\boldsymbol{G} \equiv \boldsymbol{A}[\boldsymbol{I}|\bar{\mathbf{0}}]\begin{bmatrix} \boldsymbol{R} & \mathbf{t} \\ \bar{\mathbf{0}} & 1 \end{bmatrix} \equiv \boldsymbol{A}[\boldsymbol{R}|\mathbf{t}]$$

1.3 Lens distortion

The PPM is based on the pinhole model and is unable to capture distortions that a lens introduces.

Radial distortion Deviation from the ideal pinhole caused by the lens curvature.

Radial distortion

Barrel distortion Defect associated with wide-angle lenses that causes straight lines to bend outwards.

Barrel distortion

Pincushion distortion Defect associated with telephoto lenses that causes straight lines to bend inwards.

Pincushion distortion

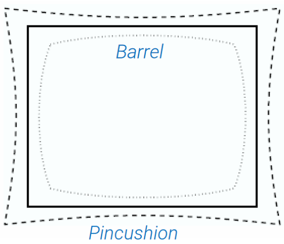


Figure 1.6: Example of distortions w.r.t. a perfect rectangle

Tangental distortion Second-order effects caused by misalignment or defects of the lens (i.e. capture distortions that are not considered in radial distortion).

1.3.1 Modeling lens distortion

Lens distortion can be modeled using a non-linear transformation that maps ideal (undistorted) image coordinates $(x_{\text{undist}}, y_{\text{undist}})$ into the observed (distorted) coordinates (x, y):

Modeling lens distortion

$$\begin{bmatrix} x \\ y \end{bmatrix} = \underbrace{L(r)\begin{bmatrix} x_{\text{undist}} \\ y_{\text{undist}} \end{bmatrix}}_{\text{Radial distortion}} + \underbrace{\begin{bmatrix} dx(x_{\text{undist}}, y_{\text{undist}}, r) \\ dy(x_{\text{undist}}, y_{\text{undist}}, r) \end{bmatrix}}_{\text{Tangential distortion}}$$

where:

- r is the distance from the distortion center which is usually assumed to be the piercing point $c = (0, 0)$. Therefore, $r = \sqrt{(x_\text{undist})^2 + (y_\text{undist})^2}$.
- $L(r)$ is the radial distortion function which is a linear operator defined for positive r only and is approximated using the Taylor series:

$$L(0) = 1 \qquad L(r) = 1 + k_1 r^2 + k_2 r^4 + k_3 r^6 + \dots$$

  where k_i are additional intrinsic parameters.
- The tangential distortion is approximated as:

$$\begin{bmatrix} dx(x_\text{undist}, y_\text{undist}, r) \\ dy(x_\text{undist}, y_\text{undist}, r) \end{bmatrix} = \begin{bmatrix} 2p_1 x_\text{undist} y_\text{undist} + p_2(r^2 + 2(x_\text{undist})^2) \\ 2p_1 y_\text{undist} x_\text{undist} + p_2(r^2 + 2(y_\text{undist})^2) \end{bmatrix}$$

  where p_1 and p_2 are additional intrinsic parameters.

  Remark. This approximation has empirically been shown to work.

Remark. The additivity of the two distortions in an assumption. Other models might add arbitrary complexity.

1.3.2 Image formation with lens distortion

Lens distortion is applied after the canonical perspective projection. Therefore, the complete workflow for image formation becomes the following: Image formation with lens distortion

1. Transform points from WRF to CRF:

$$\boldsymbol{G}\tilde{\mathbf{M}}_W \equiv \begin{bmatrix} x_C & y_C & z_C & 1 \end{bmatrix}^T$$

2. Apply the canonical perspective projection:

$$\begin{bmatrix} \frac{x_C}{z_C} & \frac{y_C}{z_C} \end{bmatrix}^T = \begin{bmatrix} x_\text{undist} & y_\text{undist} \end{bmatrix}^T$$

3. Apply the lens distortion non-linear mapping:

$$L(r) \begin{bmatrix} x_\text{undist} \\ y_\text{undist} \end{bmatrix} + \begin{bmatrix} dx(x_\text{undist}, y_\text{undist}, r) \\ dy(x_\text{undist}, y_\text{undist}, r) \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

4. Transform points from CRF to IRF:

$$\boldsymbol{A} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \equiv \begin{bmatrix} ku \\ kv \\ k \end{bmatrix} \mapsto \begin{bmatrix} u \\ v \end{bmatrix}$$

1.4 Zhang's method

Calibration patterns There are two approaches to camera calibration: Calibration patterns

- Use a single image of a 3D calibration object (i.e. image with at least 2 planes with a known pattern).
- Use multiple (at least 3) images of the same planar pattern (e.g. a chessboard).

Remark. In practice, it is easier to get multiple images of the same pattern.

Zhang's method Algorithm to determine the intrinsic and extrinsic parameters of a camera setup given multiple images of a pattern.

Image acquisition Acquire n images of a planar pattern with c internal corners.

Consider a chessboard for which we have prior knowledge of:

- The number of internal corners,
- The size of the squares.

Remark. To avoid ambiguity, the number of internal corners should be odd along one axis and even along the other (otherwise, a 180° rotation of the board would be indistinguishable).

The WRF can be defined such that:

- The origin is always at the same corner of the chessboard.
- The z-axis is at the same level of the pattern so that $z = 0$ when referring to points of the chessboard.
- The x and y axes are aligned to the grid of the chessboard. x is aligned along the short axis and y to the long axis.

Remark. As each image has its own extrinsic parameters, during the execution of the algorithm, for each image i will be computed an estimate of its own extrinsic parameters $\boldsymbol{R}_i$ and $\mathbf{t}_i$.

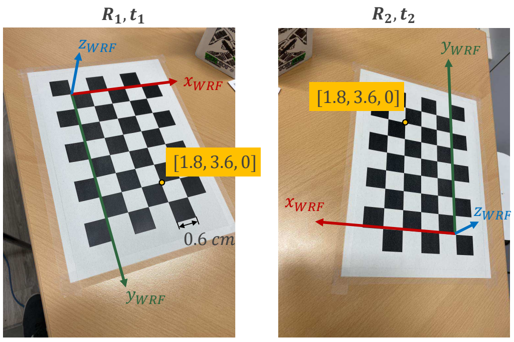


Figure 1.7: Example of two acquired images

Initial homographies guess For each image i, compute an initial guess of its homography H_i.

Due to the choice of the z-axis position, the perspective projection matrix and the WRF points can be simplified:

$$k\tilde{\mathbf{m}} \equiv k \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \equiv \boldsymbol{P}\tilde{\mathbf{M}}_W \equiv \begin{bmatrix} p_{1,1} & p_{1,2} & \cancel{p_{1,3}} & p_{1,4} \\ p_{2,1} & p_{2,2} & \cancel{p_{2,3}} & p_{2,4} \\ p_{3,1} & p_{3,2} & \cancel{p_{3,3}} & p_{3,4} \end{bmatrix} \begin{bmatrix} x \\ y \\ \cancel{0} \\ 1 \end{bmatrix}$$

$$\equiv \begin{bmatrix} p_{1,1} & p_{1,2} & p_{1,4} \\ p_{2,1} & p_{2,2} & p_{2,4} \\ p_{3,1} & p_{3,2} & p_{3,4} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \equiv \boldsymbol{H}\tilde{\mathbf{w}}$$

where $\boldsymbol{H}$ is a homography and represents a general transformation between projective planes.

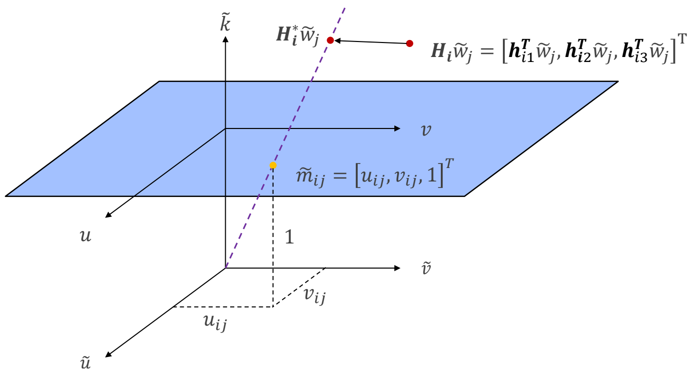
DLT algorithm Consider the i-th image with its c corners. For each corner j, we have prior knowledge of:

- Its 3D coordinates in the WRF.
- Its 2D coordinates in the IRF.

Then, for each corner j, we can define 3 linear equations where the homography $\boldsymbol{H}_i$ of the i-th image is the unknown:

$$\tilde{\mathbf{m}}_{i,j} \equiv \begin{bmatrix} u_{i,j} \\ v_{i,j} \\ 1 \end{bmatrix} \equiv \begin{bmatrix} p_{i,1,1} & p_{i,1,2} & p_{i,1,4} \\ p_{i,2,1} & p_{i,2,2} & p_{i,2,4} \\ p_{i,3,1} & p_{i,3,2} & p_{i,3,4} \end{bmatrix} \begin{bmatrix} x_j \\ y_j \\ 1 \end{bmatrix} \equiv \boldsymbol{H}_i \tilde{\mathbf{w}}_j \equiv \underbrace{\begin{bmatrix} \mathbf{h}_{i,1}^T \\ \mathbf{h}_{i,2}^T \\ \mathbf{h}_{i,3}^T \end{bmatrix}}_{\mathbb{R}^{3\times 3}} \tilde{\mathbf{w}}_j \equiv \underbrace{\begin{bmatrix} \mathbf{h}_{i,1}^T \tilde{\mathbf{w}}_j \\ \mathbf{h}_{i,2}^T \tilde{\mathbf{w}}_j \\ \mathbf{h}_{i,3}^T \tilde{\mathbf{w}}_j \end{bmatrix}}_{\mathbb{R}^{3\times 1}}$$

Geometrically, we can interpret $\boldsymbol{H}_i \tilde{\mathbf{w}}_j$ as a point in $\mathbb{P}^2$ that we want to align to the projection of $(u_{i,j}, v_{i,j})$ by tweaking $\boldsymbol{H}_i$ (i.e. find $\boldsymbol{H}_i^*$ such that $\boldsymbol{H}_i^* \tilde{\mathbf{w}}_j \equiv k \begin{bmatrix} u_{i,j} & v_{i,j} & 1 \end{bmatrix}^T$).



It can be shown that two points lay on the same line if their cross product is $\bar{\mathbf{0}}$:

$$\tilde{\mathbf{m}}_{i,j} \equiv \boldsymbol{H}_i \tilde{\mathbf{w}}_j \iff \tilde{\mathbf{m}}_{i,j} \times \boldsymbol{H}_i \tilde{\mathbf{w}}_j = \bar{\mathbf{0}} \iff \begin{bmatrix} u_{i,j} \\ v_{i,j} \\ 1 \end{bmatrix} \times \begin{bmatrix} \mathbf{h}_{i,1}^T \tilde{\mathbf{w}}_j \\ \mathbf{h}_{i,2}^T \tilde{\mathbf{w}}_j \\ \mathbf{h}_{i,3}^T \tilde{\mathbf{w}}_j \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\iff \begin{bmatrix} v_{i,j} \mathbf{h}_{i,3}^T \tilde{\mathbf{w}}_j - \mathbf{h}_{i,2}^T \tilde{\mathbf{w}}_j \\ \mathbf{h}_{i,1}^T \tilde{\mathbf{w}}_j - u_{i,j} \mathbf{h}_{i,3}^T \tilde{\mathbf{w}}_j \\ u_{i,j} \mathbf{h}_{i,2}^T \tilde{\mathbf{w}}_j - v_{i,j} \mathbf{h}_{i,1}^T \tilde{\mathbf{w}}_j \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\iff \underbrace{\begin{bmatrix} \bar{\mathbf{0}}_{1\times 3} & -\mathbf{w}_j^T & v_{i,j} \mathbf{w}_j^T \\ \mathbf{w}_j^T & \bar{\mathbf{0}}_{1\times 3} & -u_{i,j} \mathbf{w}_j^T \\ -v_{i,j} \mathbf{w}_j^T & u_{i,j} \mathbf{w}_j^T & \bar{\mathbf{0}}_{1\times 3} \end{bmatrix}}_{\mathbb{R}^{3\times 9}} \underbrace{\begin{bmatrix} \mathbf{h}_{i,1} \\ \mathbf{h}_{i,2} \\ \mathbf{h}_{i,3} \end{bmatrix}}_{\mathbb{R}^{9\times 1}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\mathbf{h}_*^T \tilde{\mathbf{w}}_j = \tilde{\mathbf{w}}_j^T \mathbf{h}_*$ and factorization

$$\iff \underbrace{\begin{bmatrix} \bar{\mathbf{0}}_{1\times 3} & -\mathbf{w}_j^T & v_{i,j} \mathbf{w}_j^T \\ \mathbf{w}_j^T & \bar{\mathbf{0}}_{1\times 3} & -u_{i,j} \mathbf{w}_j^T \end{bmatrix}}_{\mathbb{R}^{2\times 9}} \underbrace{\begin{bmatrix} \mathbf{h}_{i,1} \\ \mathbf{h}_{i,2} \\ \mathbf{h}_{i,3} \end{bmatrix}}_{\mathbb{R}^{9\times 1}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

only the first two equations are linearly independent

Homographies refinement

Initial intrinsic parameters guess

Initial extrinsic parameters guess

Initial distortion parameters guess

Parameters refinement