Image acquisition, camera geometry

Contributor: Prof. Ioannis Pitas

Presenter: Prof. Ioannis Pitas
Aristotle University of Thessaloniki
pitas@aiia.csd.auth.gr
www.multidrone.eu
Presentation version 1.1

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Image acquisition

- A still image visualizes a still object or scene, using a still picture camera.
- A video sequence (moving image) is the visualization of an object or scene illuminated by a light source, using a video camera.
- The captured object, the light source and the video camera can all be either moving or still.
- Thus, moving images are the projection of moving 3D objects on the camera image plane, as a function of time.
- Digital video corresponds to their spatiotemporal sampling.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Light reflection

- Objects reflect or emit light.
- Reflection can be decomposed in two components:
  - *Diffuse reflection* (distributes light energy equally along any spatial direction, allows perceiving object color).
  - *Specular reflection* (strongest along the direction of the incident light, incident light color is perceived).
- *Lambertian surfaces* perform only diffuse reflection, thus being dull and matte (e.g., cement surface).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

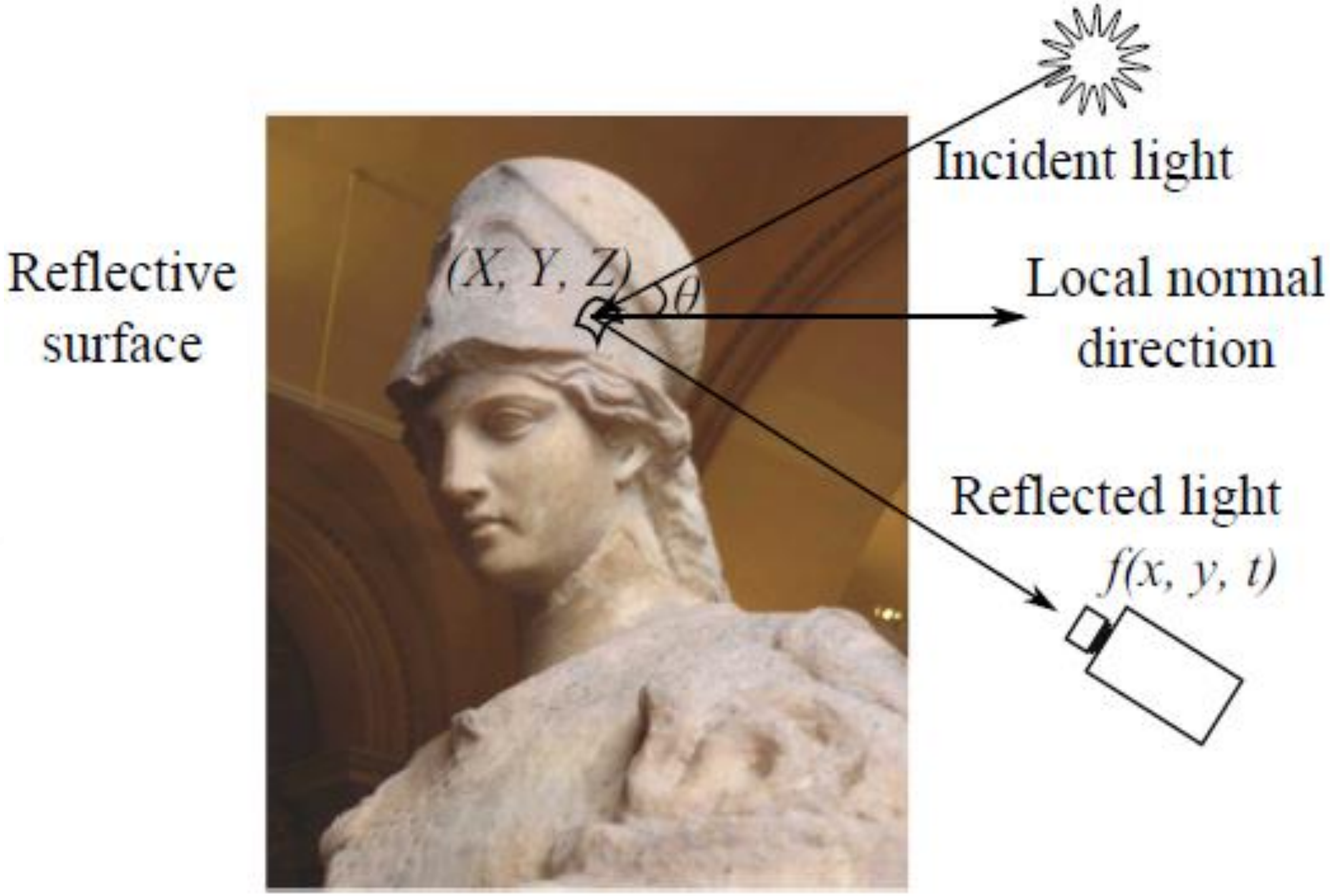
Light reflection

- *Ambient illumination* sources emit the same light energy in all directions (e.g., a cloudy sky).

- *Point illumination* sources emit light energy isotropically or anisotropically (e.g., ordinary light bulbs) along various directions.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Light reflection



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Light reflection

- Reflected irradiance when object surface produces diffuse reflectance and incident light source comes from:
  - Ambient illumination:

$$f_r(X,Y,Z,t,\lambda) = r(X,Y,Z,t,\lambda) \cdot E_\alpha(t,\lambda)$$

  - Point light source:

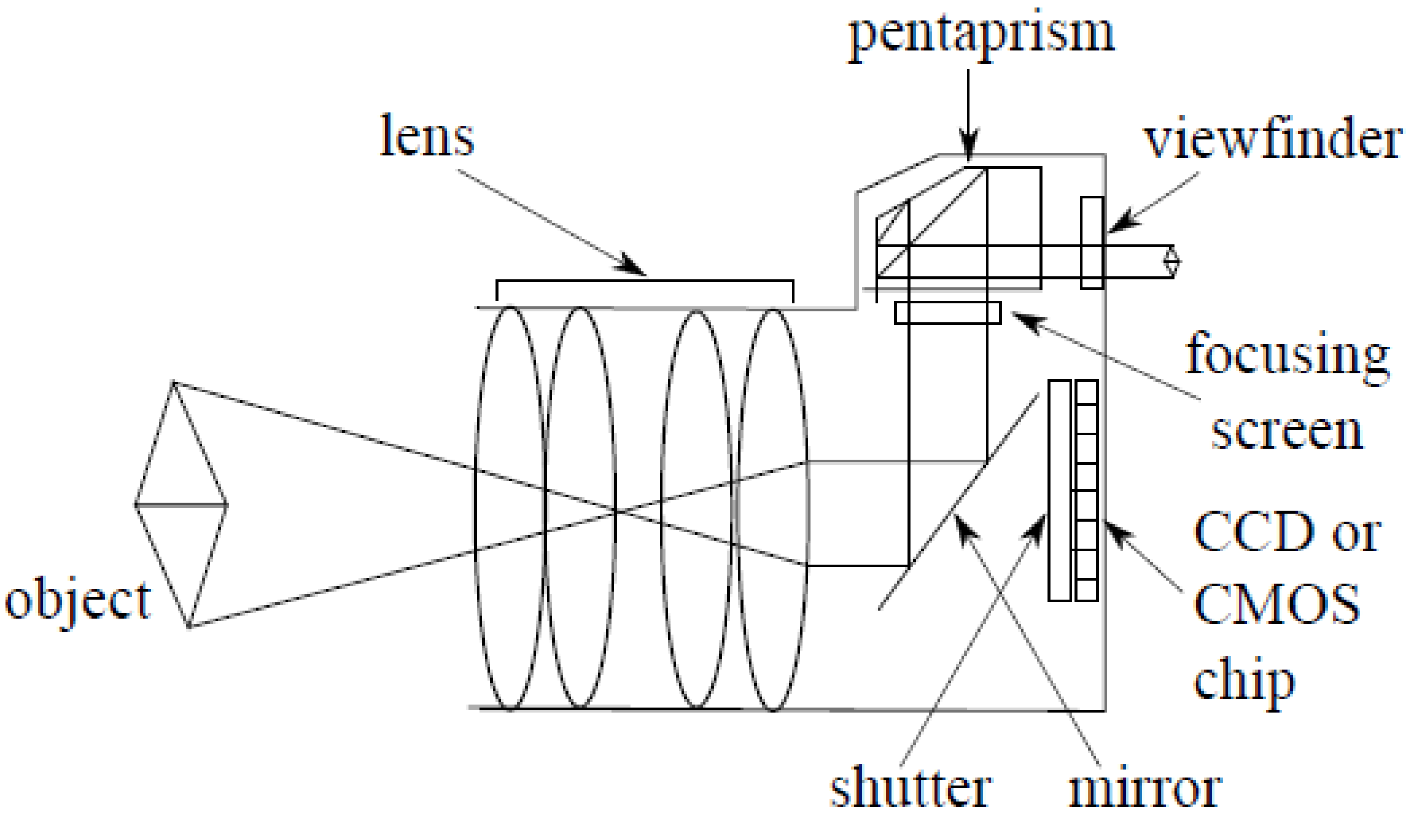
$$f_r(X,Y,Z,t,\lambda) = r(X,Y,Z,t,\lambda) \cdot E_p(t,\lambda) \cdot \cos\theta$$

  - Distant point source and ambient illumination:

$$E(t,\lambda) = E_a(t,\lambda) + E_p(t,\lambda) \cdot \cos\theta$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera structure



This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera structure

- The *lens* is the most important part of the camera.
- Incident light rays pass through a lens (or a group of lenses) and get focused on the semiconductor chip.
- The distance between the lens center (*optical center*, $\mathbf{O}$) and the point of convergence of the light rays inside the camera (*focal point*, $\mathbf{F}$) is called *focal length.*
- Focal length characterizes the lens and determines the scene part to be captured as well as scene object sizes (*magnification*).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera structure

- Two kinds of lenses:
  - Fixed (e.g., *prime*) and
  - Zoom (e.g., *telephoto*)

- Based on their focal length, lenses are categorized in wide-angle, normal and telephoto:
  - Wide-angle lenses have smaller focal length than normal, thus capturing wider parts of the scene and exaggerating differences in the relative distance and size between foreground and background objects.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera structure

- The *shutter* opens and closes to control the time interval during which light rays can hit the CCD or CMOS chip.

- *Shutter speed* is the speed at which the shutter opens and closes and determines the amount of incoming light.
- Higher speed is required for capturing unblurred, fast moving scenes, while lower speed is used in night shooting, along with bigger *aperture* size.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera structure

MultiDrone

- Aperture size is usually expressed in *f-numbers.* The bigger the f-number the smaller the aperture size.

- It controls the *depth of field* (DOF), the distance between the nearest and farthest focused objects in the image

- The smaller the aperture size is, the longer the depth of field, since less light rays are captured on the image for each visible 3D scene point.

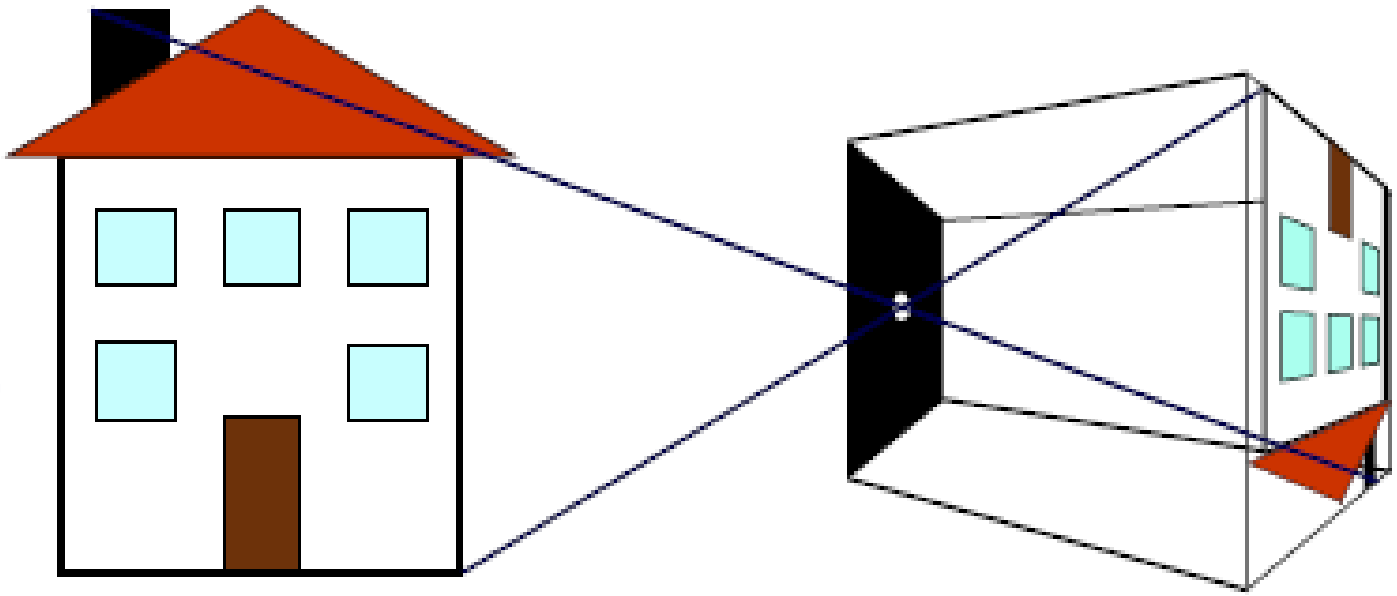
This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Pinhole Camera and Perspective Projection

- The rather naïve *pinhole camera system* can be used to accurately model the geometric and optical aspects of most modern cameras through the *pinhole perspective projection model* or *central perspective projection model.*
  - A very small aperture size is considered
  - Camera pinhole coincides with the *optical center*, or *center of projection* or *camera center.*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Pinhole Camera and Perspective Projection



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

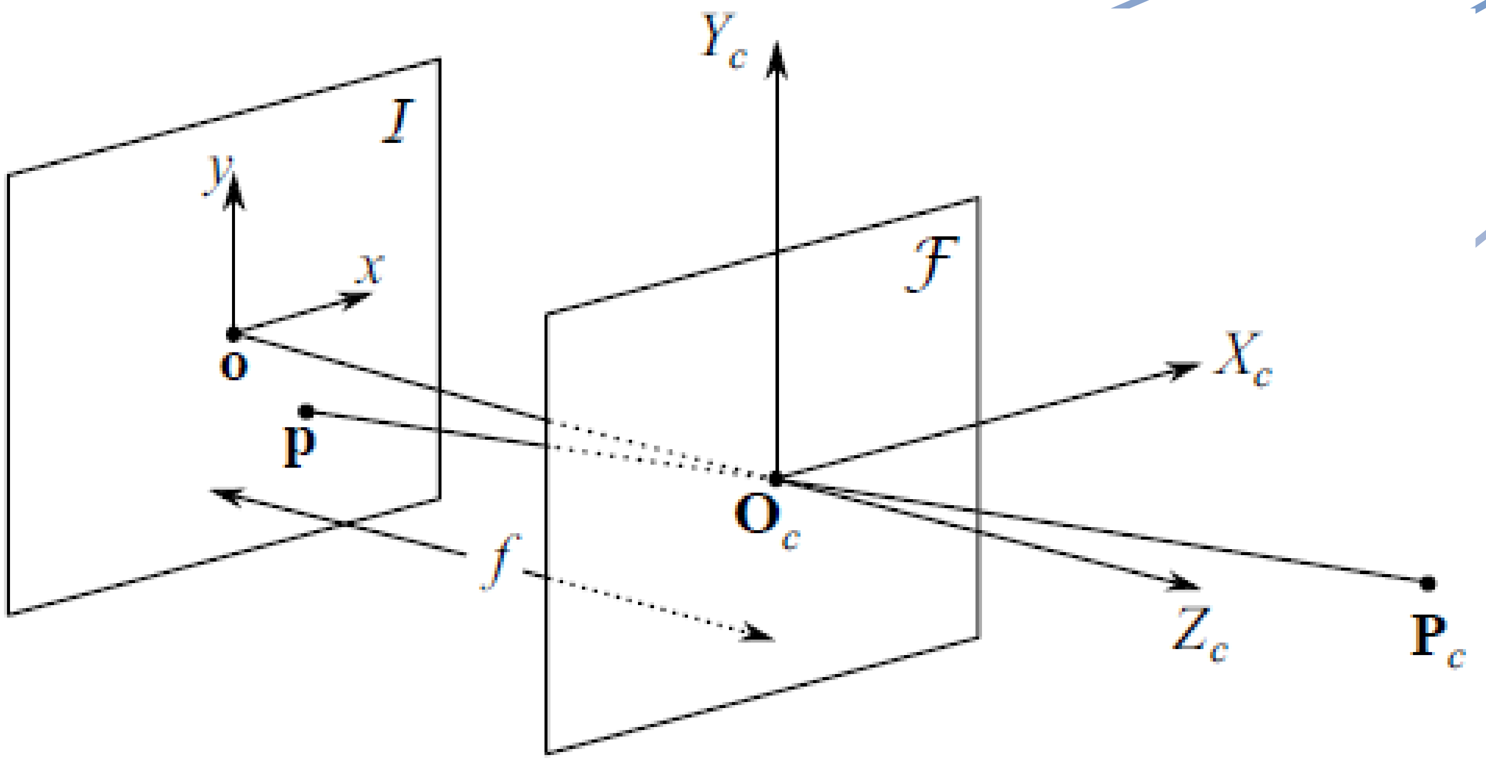
Pinhole Camera and Perspective Projection

- Let us consider 2 coordinate systems:
  - the *camera* (or *standard*) *coordinate system* $(\mathbf{O}_c, X_c, Y_c, Z_c)$ and
  - the image coordinate system $(\mathbf{o}, x, y)$.

- X_c, *and* Y_c define the plane $\mathcal{F}$ that is parallel to the camera image plane $\mathcal{I}$, lying at a focal length f behind the optical center $\mathbf{O}_c$ along the optical axis Z_c.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Pinhole Camera and Perspective Projection



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

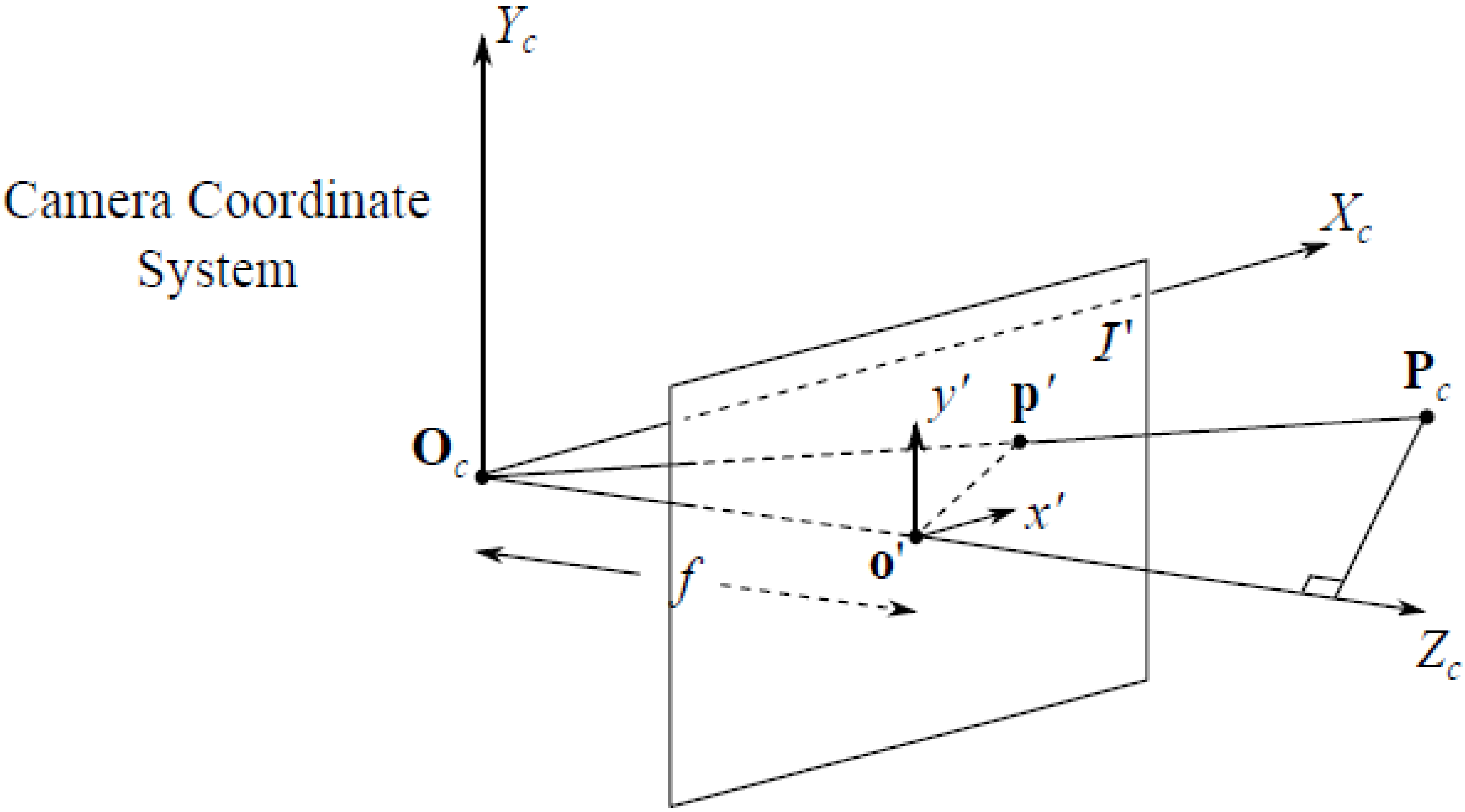
Pinhole Camera and Perspective Projection

- Points projected on the image plane are assigned camera coordinates of opposite sign:
  - images are inverted.

- In order to facilitate the mathematical treatment, we can define a virtual image plane $\mathcal{I}'$, in front of $\mathcal{F}$ at a positive distance f.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Pinhole Camera and Perspective Projection



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Pinhole Camera and Perspective Projection

MultiDrone

- We want to derive the equations that connect a 3D point (3D vector) $\mathbf{P}_c = [X_c, Y_c, Z_c]^T$ referenced in the camera coordinate system with its projection point (2D vector) $\mathbf{p}' = [x', y']^T$ on the virtual image plane.
- By employing the similarity of triangles $\mathbf{O}_c\mathbf{o}'\mathbf{p}'$ and $\mathbf{O}_c\mathbf{Z}_c\mathbf{P}_c$:

$$\frac{x'}{X_c} = \frac{y'}{Y_c} = \frac{f}{Z_c}, \qquad x' = f\frac{X_c}{Z_c}, \qquad y' = f\frac{Y_c}{Z_c}$$

- Coordinates on the real image plane are given by the same equations, differing only by a minus sign.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

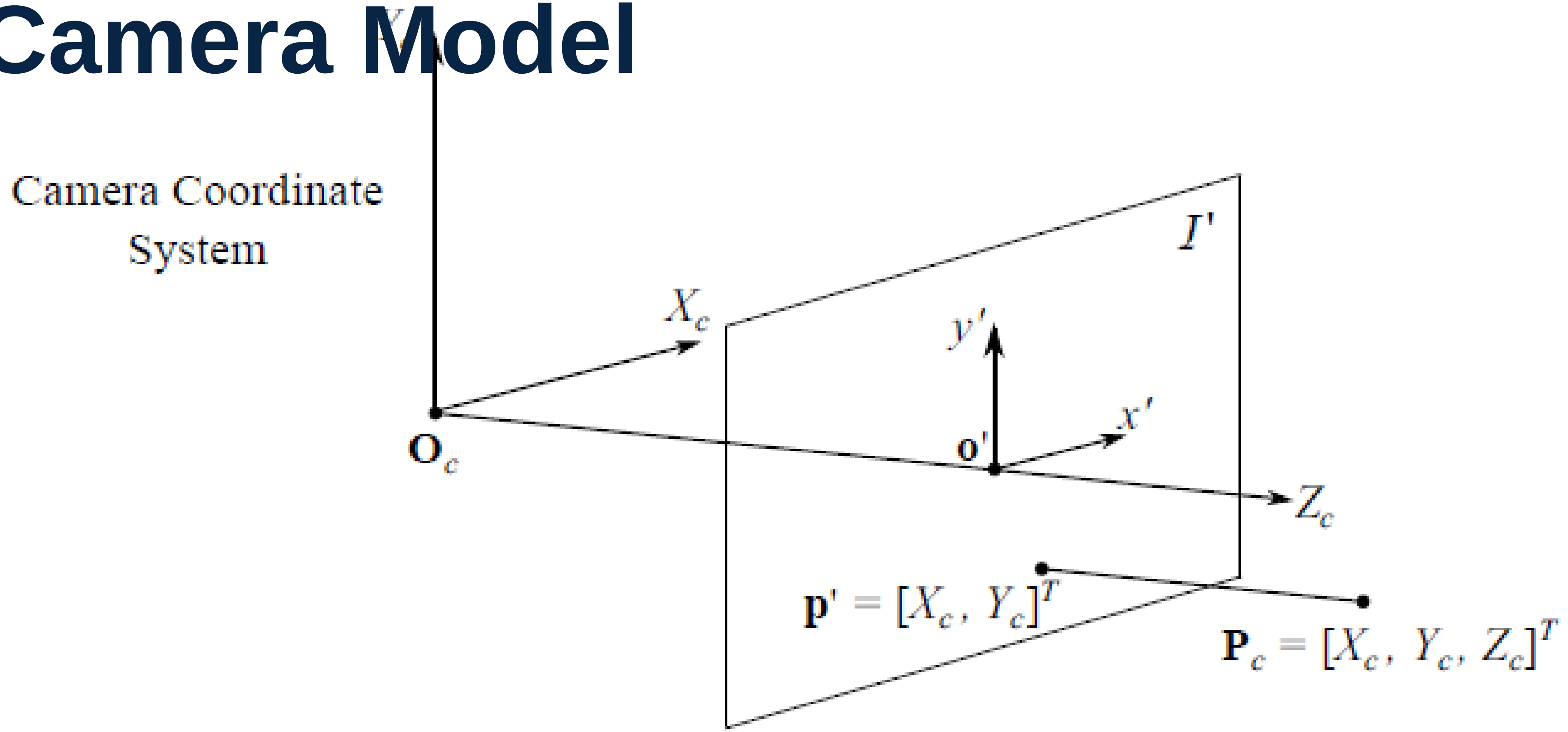
The Weak-Perspective Camera Model

- Perspective projection equations are rational, rather than linear. Thus:
  - straight lines are mapped to straight lines but
  - distances between points and the angles between straight lines are not preserved after projection.

- We can linearise them applying two transformations:
  - *orthographic projection* $x' = X_c$ and $y' = Y_c$ and
  - *isotropic scaling* $f/\bar{Z}$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

The Weak-Perspective Camera Model



Orthographic projection

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

The Weak-Perspective Camera Model

- Isotropic scaling transformation leads to linearly approximate perspective projection equations, defining the so called *weak-perspective* camera model:

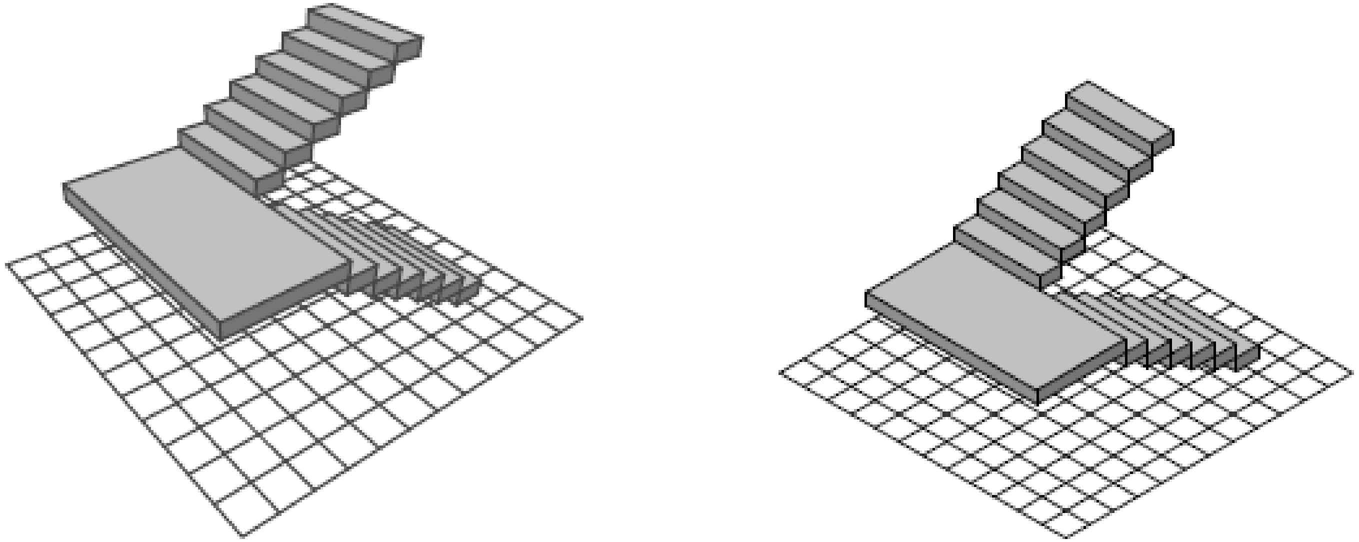
$$x' = f\frac{X_c}{Z_c} \approx \frac{f}{\bar{Z}}X_c, \qquad y' = f\frac{Y_c}{Z_c} \approx \frac{f}{\bar{Z}}Y_c$$

- The weak-perspective camera model is only an approximation of the pinhole camera imaging. It holds if the *relative distance* dZ_c for any pair of scene points along the optical axis is much smaller than $\bar{Z}$.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

The Weak-Perspective Camera Model

- While a weak-perspective camera preserves parallelism in the projected lines, as orthographic projection does (b), perspective projection (a) does not.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Central Projection and Homogeneous Coordinates

- A way to linearize the perspective projection equations is by using the so-called *homogeneous coordinates.*
  - A 2D image point is mapped to a 3D point:

    $\mathbf{p} = [x, y]^T \in \mathbb{R}^2 \longrightarrow \mathbf{p}_H = [x, y, 1]^T \in \mathbb{P}^2$.
  - A 3D scene point is mapped to a 4D point:

    $\mathbf{P} = [X, Y, Z]^T \in \mathbb{R}^3 \longrightarrow \mathbf{P}_H = [X, Y, Z, 1]^T \in \mathbb{P}^3$.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Central Projection and Homogeneous Coordinates

- The linear relationship that connects $\mathbf{p}_H$ and $\mathbf{P}_H$ is:

$$Z\mathbf{p}_H = \begin{bmatrix} Zx \\ Zy \\ Z \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \mathcal{P}\mathbf{P}_H$$

where $\mathcal{P}$ is the so-called camera *perspective projection matrix,* a 3×4 full row rank homogeneous matrix with 11 *degrees of freedom.*

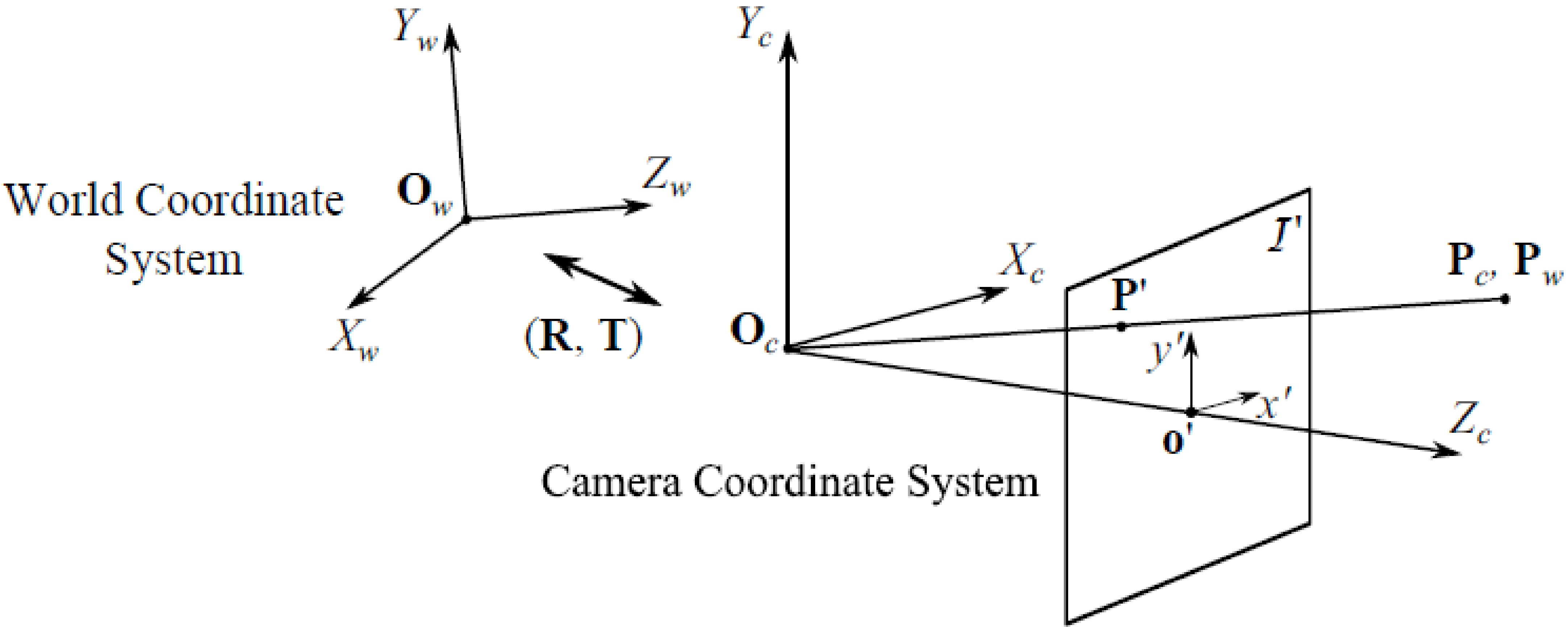
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera Parameters and Projection Matrix

- Object coordinates **P** in the camera coordinate system are, in most cases, unknown, whereas in the *world coordinate system* they may be known.
- The required transformation from the world to the camera coordinate system involves a translation followed by a rotation, based on the *extrinsic* camera parameters.
- Projection on the image plane requires the *intrinsic* camera parameters.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera Parameters and Projection Matrix



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera Parameters and Projection Matrix

- Extrinsic camera parameters:
  - *Translation vector* $\mathbf{T} \in \mathbb{R}^3$ (3 degrees of freedom)
  - *Orthonormal rotation matrix* $\mathbf{R} \in \mathbb{R}^{3\times3}$ (3 degrees of freedom: only 3 of the 9 rotation matrix entries are independent from each other).
- The relationship between a point $\mathbf{P}_w \in \mathbb{R}^3$ in world coordinates and its camera coordinate counterpart $\mathbf{P}_c \in \mathbb{R}^3$ is:

$$\mathbf{P}_c = \mathbf{R}(\mathbf{P}_w - \mathbf{T})$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera Parameters and Projection Matrix

- If the three rows of the rotation matrix and focal length *f* are known, the image coordinates *x'*, *y'* on the virtual image plane are given by:

$$x' = f\frac{\mathbf{R}_1^T(\mathbf{P}_w - \mathbf{T})}{\mathbf{R}_3^T(\mathbf{P}_w - \mathbf{T})}$$

$$y' = f\frac{\mathbf{R}_2^T(\mathbf{P}_w - \mathbf{T})}{\mathbf{R}_3^T(\mathbf{P}_w - \mathbf{T})}$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera Parameters and Projection Matrix

- In reality, the virtual image plane $\mathcal{I}'$ does not exist. The real 2D image plane (image sensor surface) is digitized.
- The transformation of an image point $\mathbf{p} = [x, y]^T$ on the image plane coordinates to the corresponding discrete point $\mathbf{p}_d = [x_d, y_d]^T$ in pixel coordinates, is given by:

$$x = -(x_d - o_x)s_x, \qquad y = -(y_d - o_y)s_y$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

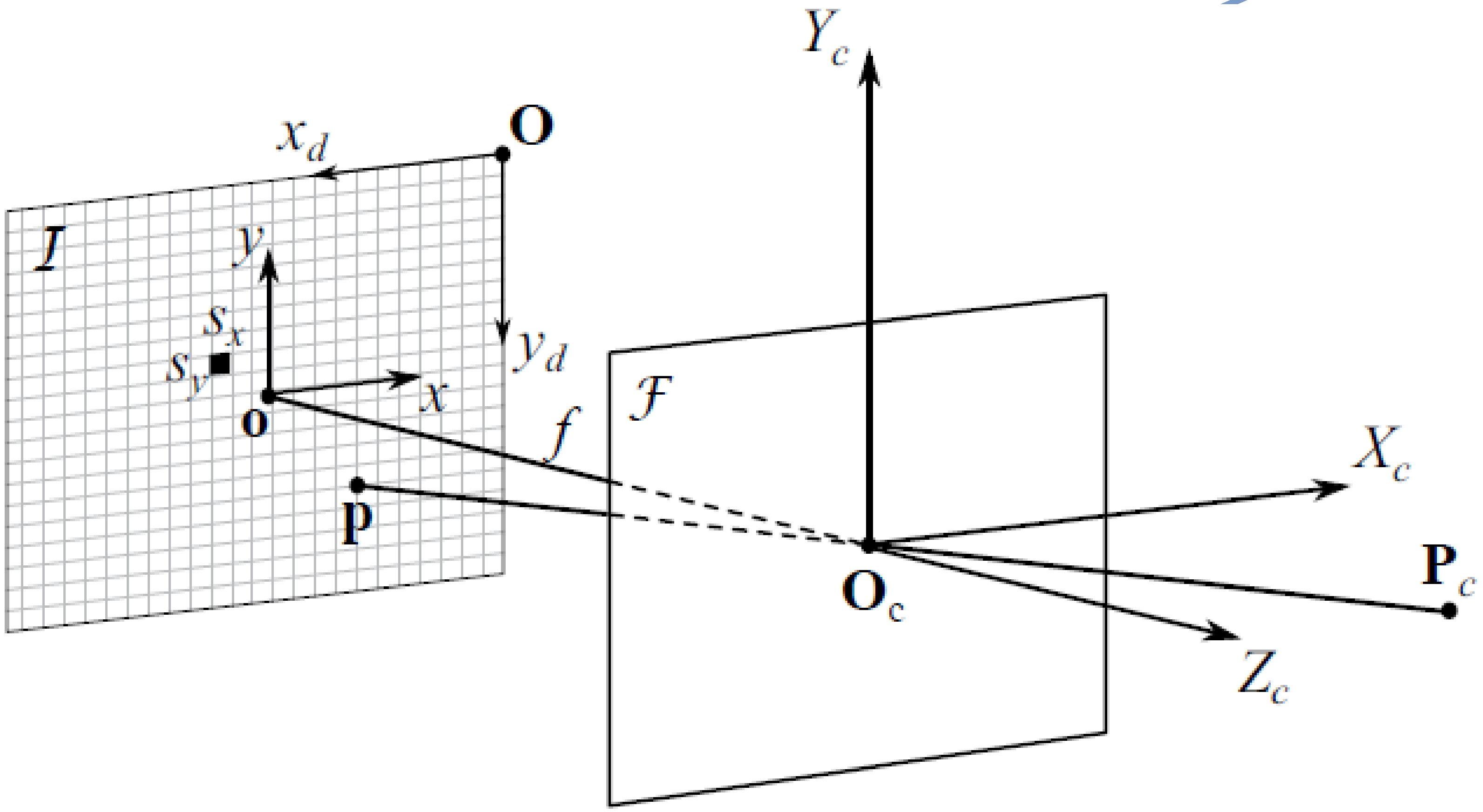
Camera Parameters and Projection Matrix

where:

- $[o_x, o_y]^{\mathrm{T}}$: the location of the principal camera point $\mathbf{o}$ in pixel coordinates.
- s_x, s_y: the effective pixel sizes in millimeters
- Coordinate system origin: at the top left corner of the image, not at the image center.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera Parameters and Projection Matrix



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera Parameters and Projection Matrix

- The transformation relating the image pixel coordinates with the world coordinates is:

$$x_d = o_x - \frac{f}{s_x} \frac{\mathrm{R}_1^T(\mathrm{P}_w - \mathrm{T})}{\mathrm{R}_3^T(\mathrm{P}_w - \mathrm{T})} \qquad y_d = o_y - \frac{f}{s_y} \frac{\mathrm{R}_2^T(\mathrm{P}_w - \mathrm{T})}{\mathrm{R}_3^T(\mathrm{P}_w - \mathrm{T})}$$

- It can be linearized in homogeneous coordinates, by decomposing the transformation into a sequence of two transformations:
  - Map a world coordinate point to camera coordinates.
  - Map the came coordinate point to homogeneous image pixel coordinates.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera Parameters and Projection Matrix

- Definition of the 3×4 matrix of extrinsic parameters $\mathbf{P}_E$:

$$\mathrm{P}_E = \begin{bmatrix} r_{11} & r_{12} & r_{13} & -\mathrm{R}_1^T\mathrm{T} \\ r_{21} & r_{22} & r_{23} & -\mathrm{R}_2^T\mathrm{T} \\ r_{31} & r_{32} & r_{33} & -\mathrm{R}_3^T\mathrm{T} \end{bmatrix}$$

- Definition of the 3×3 matrix of intrinsic parameters $\mathbf{P}_I$:

$$\mathrm{P}_I = \begin{bmatrix} -\frac{f}{s_x} & 0 & o_x \\ 0 & -\frac{f}{s_y} & o_y \\ 0 & 0 & 1 \end{bmatrix}$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera Parameters and Projection Matrix

- The transformation of a point $\mathbf{P} \in \mathbb{P}^3$ to $\mathbf{p} \in \mathbb{P}^2$ is given by:

$$\begin{bmatrix} Zx_d \\ Zy_d \\ Z \end{bmatrix} = \mathrm{P}_I \mathrm{P}_E \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \qquad \mathrm{p} = \mathrm{P}_I \mathrm{P}_E \mathrm{P} = \mathcal{P}\mathrm{P}$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera Parameters and Projection Matrix

- Where $\mathcal{P} = \mathbf{P}_I \; \mathbf{P}_E$ is the 3×4 *camera projection matrix*, also called *camera calibration matrix*

$$\mathcal{P} = \begin{bmatrix} -\frac{f}{s_x} & 0 & o_x \\ 0 & -\frac{f}{s_y} & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & -\mathbf{R}_1^T\mathbf{T} \\ r_{21} & r_{22} & r_{23} & -\mathbf{R}_2^T\mathbf{T} \\ r_{31} & r_{32} & r_{33} & -\mathbf{R}_3^T\mathbf{T} \end{bmatrix}$$

- $\mathbf{P}_E$ has the form $\mathbf{P}_E = [\mathbf{R}_1 | \mathbf{R}_2 | \mathbf{R}_3 | -\mathbf{RT}]$, since we assume that the camera coordinate system is first translated and then rotated. Otherwise it would be $\mathbf{P}_E = [\mathbf{R}_1 | \mathbf{R}_2 | \mathbf{R}_3 | \mathbf{T}]$.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera Parameters and Projection Matrix

- For reasons of simplicity, it is common to assume that:
  - the origins of both the pixel coordinate system and the image plane coordinate system coincide with the principal point, $o_x = o_y = 0$ and
  - pixels are square having unit edge length $s_x = s_y = 1$.

- The projection matrix, can thus be rewritten as:

$$\mathcal{P} = \begin{bmatrix} -fr_{11} & -fr_{12} & -fr_{13} & f\mathrm{R}_1^T\mathrm{T} \\ -fr_{21} & -fr_{22} & -fr_{23} & f\mathrm{R}_2^T\mathrm{T} \\ r_{31} & r_{32} & r_{33} & -\mathrm{R}_3^T\mathrm{T} \end{bmatrix}$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera Parameters and Projection Matrix

- If the two axes of the coordinate system (x_d, y_d) are not exactly perpendicular (non-rectangular pixels), the projection matrix takes the form:

$$\mathcal{P} = \begin{bmatrix} -\frac{f}{s_x} & s_\theta & o_x \\ 0 & -\frac{f}{s_y} & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & -\mathbf{RT} \\ \mathbf{0} & 1 \end{bmatrix}$$

- s_θ: *skew factor,* proportional to $\frac{1}{\tan\theta}$,
- θ: the angle between the pixel coordinate system axes.
- Typically $\theta = 90^o$, and hence $s_\theta = 0$.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera Parameters and Projection Matrix

- Assuming $s_\theta = 0$ and treating the ratios $a_x = -\frac{f}{s_x}$ and $a_y = -\frac{f}{s_y}$ as single quantities, by expressing the focal length in terms of pixel dimensions along the horizontal and vertical dimension, $\mathbf{P}_I$ can be rewritten as:

$$\mathbf{P}_I = \begin{bmatrix} a_x & 0 & o_x \\ 0 & a_y & o_y \\ 0 & 0 & 1 \end{bmatrix}$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

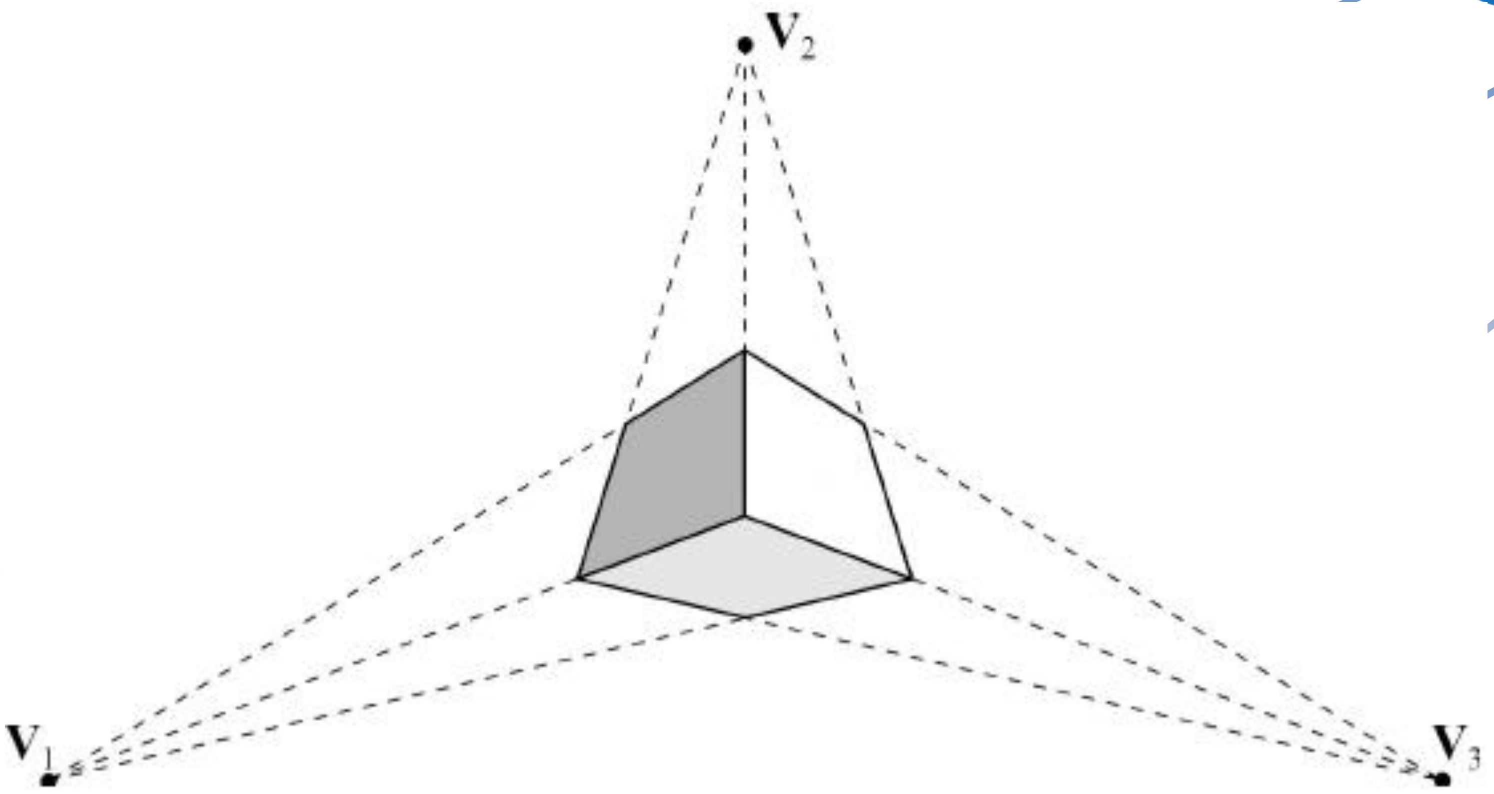
Properties of the Projective Transformation

MultiDrone

- The projective transformation does not change the cross-ratio C_r of four collinear points.
- *Line at infinity*: the set of all points at infinity in $\mathbb{P}^2$.
- *Plane at infinity*: the set of all points at infinity in $\mathbb{P}^3$, formed by an infinite number of lines at infinity corresponding to different plane directions.
- *Vanishing points:* the points of intersection of the projected lines formed by parallel lines in the 3D Euclidean space.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Properties of the Projective Transformation

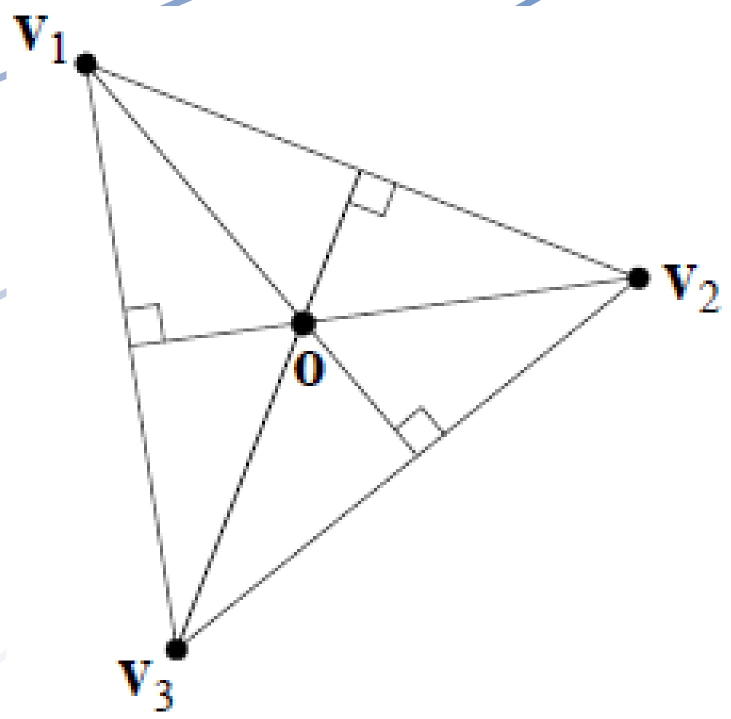


Vanishing points

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Properties of the Projective Transformation

- The orthocenter of the triangle formed by the vanishing points $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, corresponding to three perpendicular parallel line directions in the world reference system, is the principal point $\mathbf{o}$ on the image plane.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Properties of the Projective Transformation

MultiDrone

- *Cross-ratio (or anharmonic ratio)* C_r: ratio of ratios of distances between collinear points.
  - It is a geometric property invariant under a projective transformation.

- For four points $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4$ in $\mathbb{P}^2$:

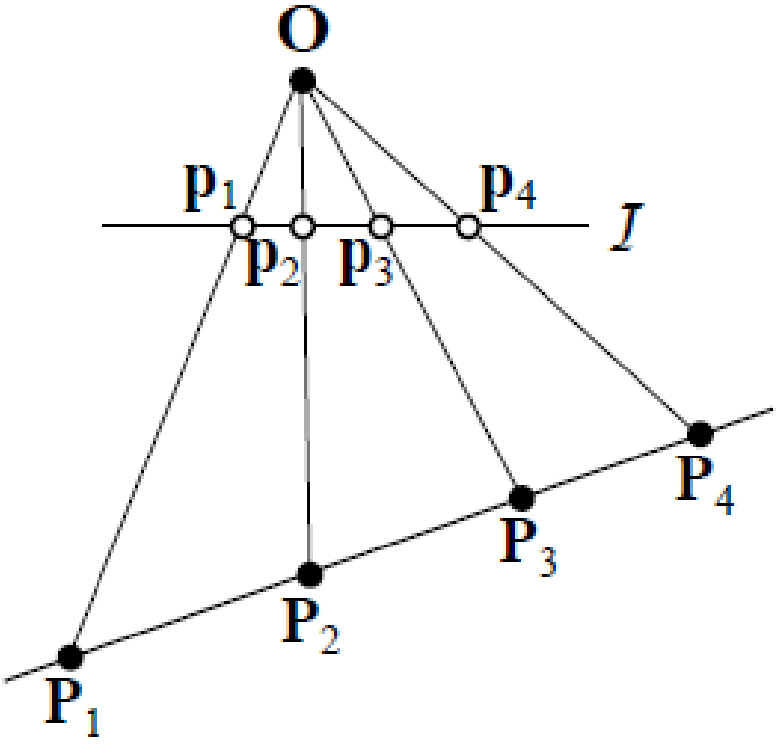
$$C_r(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4) = \frac{\Delta_{13}\Delta_{24}}{\Delta_{14}\Delta_{23}}$$

Δ_{ij}: the Euclidean distance between points $\mathbf{p}_i$ and $\mathbf{p}_j$.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Properties of the Projective Transformation

- The cross-ratio of four collinear points remains invariant under a projective transformation: $C_r(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4) = C_r(\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3, \mathbf{P}_4)$.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

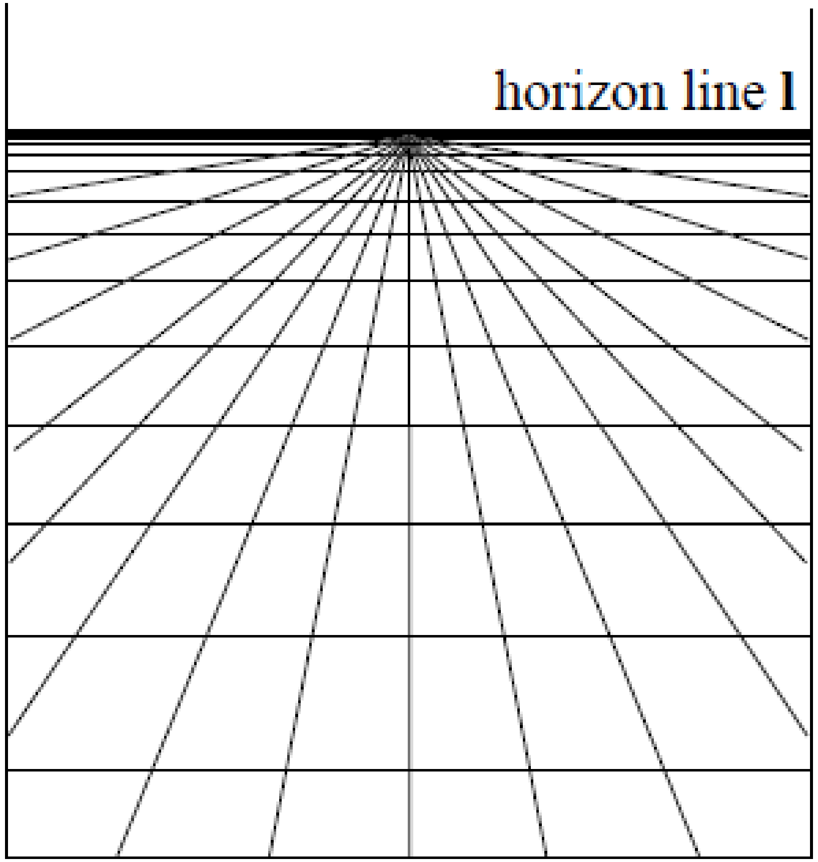
Properties of the Projective Transformation

- *Vanishing lines*: the projections on the image of lines at infinity in $\mathbb{P}^3$, where parallel planes in the 3D Euclidean space intersect.
- *Horizon line*: the vanishing line of the ground plane and its parallel planes.
  - The projections of parallel lines lying on a plane that forms an angle θ with the ground, intersect either above or below the horizon line, depending on the sign of $\cos\theta$.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Properties of the Projective Transformation

- *Chirp effect*: the increase in local image spatial frequency proportionally to the distance of the projected scene area from the camera.
- It is evident in 2D image regions where distant and close-up scene parts are projected.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Cameras at Infinity

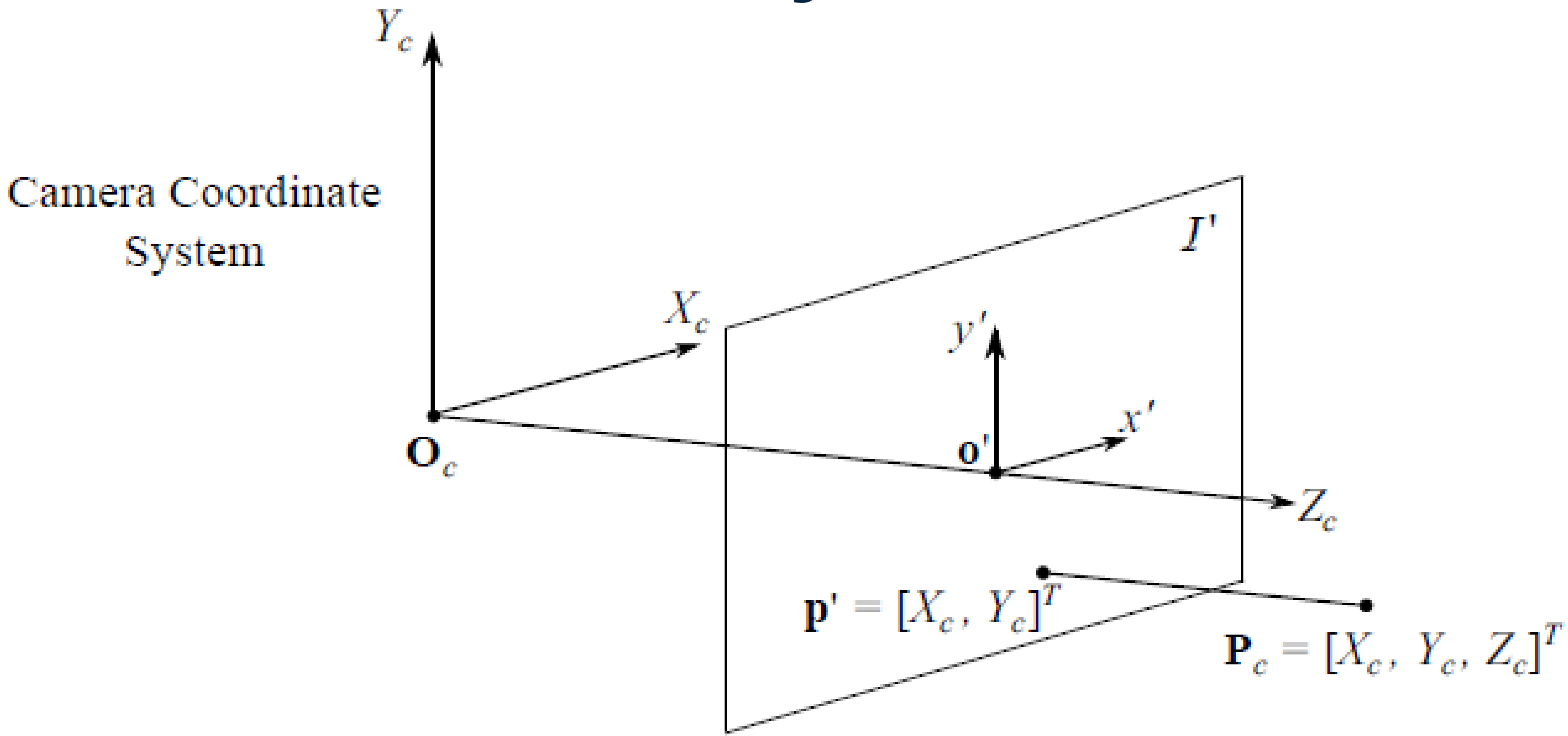
- When the center of projection lies at infinity, instead of perspective projection, the orthographic camera and the weak-perspective camera model are used.
- Weak-perspective camera model projection matrix $\mathcal{P}_{wp}$:

$$\mathcal{P}_{wp} = \begin{bmatrix} -fr_{11} & -fr_{12} & -fr_{13} & f\mathrm{R}_1^T\mathrm{T} \\ -fr_{21} & -fr_{22} & -fr_{23} & f\mathrm{R}_2^T\mathrm{T} \\ 0 & 0 & 0 & \mathrm{R}_3^T(\overline{\mathrm{P}} - \mathrm{T}) \end{bmatrix}$$

$\overline{\mathbf{P}}$: centroid of the viewed object.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Cameras at infinity



Orthographic projection

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Cameras at Infinity

- The weak-perspective projection is valid, if the depth variations amongst the points of a viewed object are small, in comparison with its average distance from the camera (object depth), represented by $\mathbf{R}_3{}^T(\bar{\mathbf{P}} - \mathbf{T})$.
- Another camera-at-infinity model is the *affine camera model,* with projection matrix:

$$\mathcal{P}_{af} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ 0 & 0 & 0 & a_{34} \end{bmatrix}$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Cameras at Infinity

- The left 3×3 sub-matrix of $\mathcal{P}_{af}$ is singular.
- $\mathcal{P}_{af}$ has 9 independent entries.

- Main difference between the affine camera model and the projective framework:
  - In the projective framework, no distinction exists between points at infinity and the usual finite points of the affine space.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Cameras at Infinity

- The affine camera model does **not** really describe any physical camera, but is used due to its simplicity.

- Main difference between the affine camera model and the weak-perspective model:
  - It preserves straight line parallelism, but **not** angles, since it may entail anisotropic scaling.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera Calibration

- Camera calibration deals with determining the extrinsic and intrinsic camera parameters.

- To this end, a *calibration pattern*, also called *calibration grid* is employed, so that the projection matrix $\mathcal{P}$ of the camera can be computed from a single view by mapping known points $\mathbf{P}_w = [X_w, Y_w, Z_w]^T$ in world coordinates to their projections $\mathbf{p} = [x, y]^T$ on the image plane.

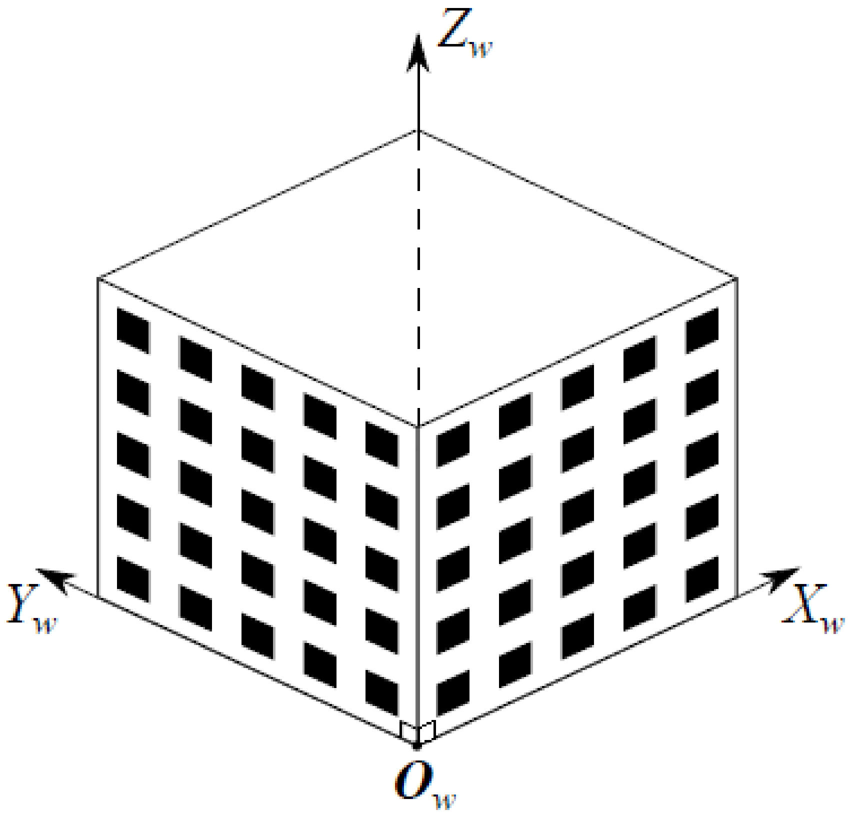
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

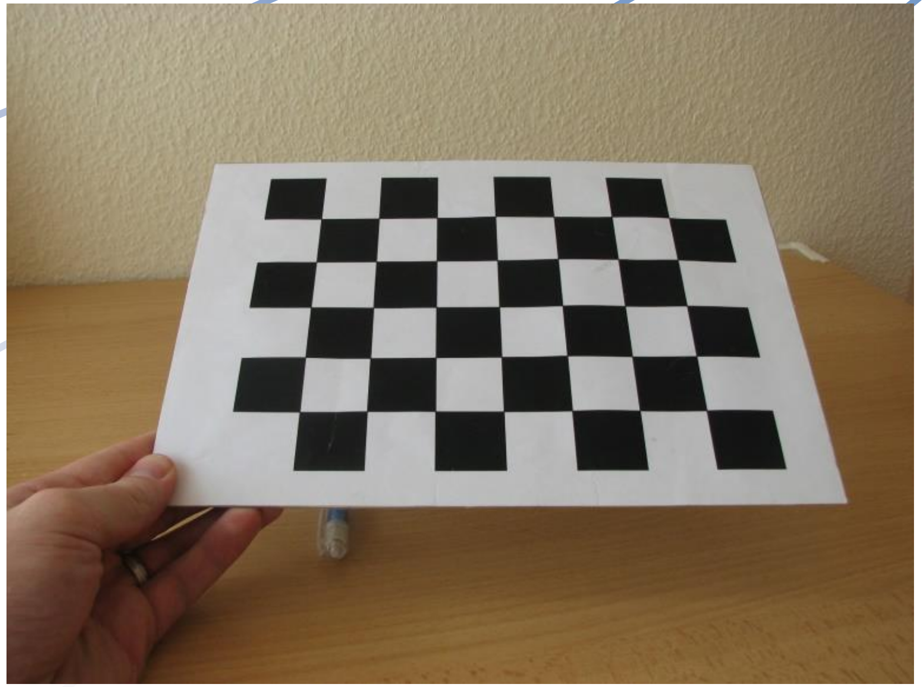
Camera Calibration

- Taking as many such mappings as needed, a system of equations is formed.
- The solution of the system leads to the determination of the unknown extrinsic and intrinsic camera parameters.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera Calibration





Calibration patterns.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera Calibration

MultiDrone

- The calibration pattern is a 3D object of common, known dimensions and positioning, with a checkerboard pattern clearly visible on each side.
- Pattern dimensions must be known, in an accuracy much greater than the desired calibration accuracy.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Camera Calibration

- The most popular calibration techniques utilizing solely a planar calibration pattern are:
  - *Direct camera parameter estimation* and
  - *Zhang's calibration method.*

- Other calibration methods do not require a calibration object and are jointly referenced by the term *self-calibration* or *autocalibration.*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Direct camera parameter estimation

- $\mathbf{P}_w = [X_w, Y_w, Z_w]^T$: a known point in world coordinates.
- $\mathbf{P}_c = [X_c, Y_c, Z_c]^T$: the same point in camera coordinates.
- $\mathbf{p}_d = [x_d, y_d]^T$: its image point in pixel coordinates.

- The transformation between the world and camera coordinate systems involves an orthonormal 3×3 rotation matrix $\mathbf{R}$ and a 3×1 translation vector $\mathbf{T}$ (equivalent to determining the extrinsic camera parameters).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Direct camera parameter estimation

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \mathrm{R} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + \mathrm{T} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{33} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}$$

- It can be decomposed into:

$$X_c = r_{11} X_w + r_{12} Y_w + r_{13} Z_w + T_x$$
$$Y_c = r_{21} X_w + r_{22} Y_w + r_{23} Z_w + T_y$$
$$Z_c = r_{31} X_w + r_{32} Y_w + r_{33} Z_w + T_z$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Direct camera parameter estimation

- Assuming $o_x = o_y = 0$, the point $\mathbf{P}_c$ in the camera coordinate system is related to the pixel coordinates of point $\mathbf{p}_d$ by:

$$x_d = -\frac{f}{s_x}\frac{X_c}{Z_c} \qquad y_d = -\frac{f}{s_y}\frac{Y_c}{Z_c}$$

and finally:

$$x_d = -\frac{f}{s_x}\frac{r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

$$y_d = -\frac{f}{s_y}\frac{r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Direct camera parameter estimation

MultiDrone

- Using a sufficient number of world coordinate and pixel coordinate point pairs, equations can be formulated and solved for the unknown camera parameters:
  - $r_{11}, r_{12}, \dots, r_{33}, T_x, T_y, T_z, s_x, s_y, f$.

- It should be noted that knowledge of the ratios f/s_x , f/s_y , rather than of all internal camera parameters suffices for camera calibration.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Direct camera parameter estimation

- Internal camera parameters:
  - f: focal length in pixel length.
  - s_x, s_y: pixel size.
  - o_x, o_y: camera center coordinates.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Direct camera parameter estimation

$$x_d = -\frac{f}{s_x}\frac{r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

$$y_d = -\frac{f}{s_y}\frac{r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

- Since the two equations have the same denominator, for each pair of 3D points $\mathbf{P}_{wi} = [X_{wi}, Y_{wi}, Z_{wi}]^T$ and their image points $\mathbf{p}_{di} = [x_{di}, y_{di}]^T$, $i = 1, \ldots, N$, we have:

$$x_{di}\frac{f}{s_y}(r_{21}X_{wi} + r_{22}Y_{wi} + r_{23}Z_{wi} + T_y) = y_{di}\frac{f}{s_x}(r_{11}X_{wi} + r_{12}Y_{wi} + r_{13}Z_{wi} + T_x)$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Direct camera parameter estimation

- By using the pixel aspect ration $a = {}^{s_x}/_{s_y}$, putting all the equation terms on the left side and employing N point pairs, we get:

$$
\begin{aligned}
x_{d1} r_{21} X_{w1} &+ \cdots + x_{d1} T_y - y_{d1} \alpha r_{11} X_{w1} - \cdots - y_{d1} \alpha T_x &= 0\\
x_{d2} r_{21} X_{w2} &+ \cdots + x_{d2} T_y - y_{d2} \alpha r_{11} X_{w2} - \cdots - y_{d2} \alpha T_x &= 0\\
\vdots \quad & \quad \vdots \quad \vdots \quad \quad \vdots \quad \quad \vdots \quad \vdots & \\
x_{dN} r_{21} X_{wN} &+ \cdots + x_{dN} T_y - y_{dN} \alpha r_{11} X_{wN} - \cdots - y_{dN} \alpha T_N &= 0.
\end{aligned}
$$

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Direct camera parameter estimation

- Each of the linear homogeneous equations has 8 unknown parameters:
  - $\mathbf{u} = \left[ar_{11}, ar_{12}, ar_{13}, r_{21}, r_{22}, r_{23}, T_x, T_y\right]^T \triangleq [u_1, u_2, \dots, u_8]^T$
- Expressing the homogeneous system of equation as a product of the matrix $\mathbf{X}$

$$\mathbf{X} \triangleq \begin{bmatrix} x_{d1}X_{w1} & x_{d1}Y_{w1} & x_{d1}Z_{w1} & x_{d1} & -y_{d1}X_{w1} & -y_{d1}Y_{w1} & -y_{d1}Z_{w1} & -y_{d1} \\ x_{d2}X_{w2} & x_{d2}Y_{w2} & x_{d2}Z_{w2} & x_{d2} & -y_{d2}X_{w2} & -y_{d2}Y_{w2} & -y_{d2}Z_{w2} & -y_{d2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{dN}X_{wN} & x_{dN}Y_{wN} & x_{dN}Z_{wN} & x_{dN} & -y_{dN}X_{wN} & -y_{dN}Y_{wN} & -y_{dN}Z_{wN} & -y_{dN} \end{bmatrix}$$

$$\mathbf{X}\mathbf{u} = 0$$

and the vector $\mathbf{u}$:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Direct camera parameter estimation

- Thus, the desired solution is the *null space* of the matrix $\mathbf{X}$.
- Provided that $N \geq 7$ pairs are not coplanar:
  - Matrix $\mathbf{X}$ will have rank 7.
  - The system of equations will have one non-trivial solution $\mathbf{u} \neq \mathbf{0}$, obtained via the *singular value decomposition* (SVD) of matrix $\mathbf{X}$:
  $$\mathbf{X} = \mathbf{U\Sigma V}^T$$
  - $\mathbf{\Sigma}$ is a diagonal matrix containing the singular values. The solution $\mathbf{u}$ is the column of matrix $\mathbf{V}$ corresponding to the zero singular value of $\boldsymbol{\Sigma}$.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Zhang's calibration method

- It uses a planar calibration pattern, posing in N different orientations $(N \geq 2)$, by moving either the pattern of the camera.
  - Precise knowledge of this motion is not required.
  - The calibration pattern can simply be printed on a paper and attached to any planar surface.
  - It has educed complexity.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Zhang’s calibration method

- Let $\mathbf{P}_I$ be the intrinsic parameters matrix:

$$\mathbf{P}_I = \begin{bmatrix} \alpha_x & s_\theta & o_x \\ 0 & \alpha_y & o_y \\ 0 & 0 & 1 \end{bmatrix}$$

where $= [o_x, o_y]^T$ the principal point coordinates, s_θ the skew factor and $a_x = -\frac{f}{s_x}, a_y = -\frac{f}{s_y}$.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Zhang's calibration method

- Assuming the calibration pattern lies on the scene plane $Z_w = 0$ (in world coordinates):

$$s \begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix} = \mathbf{P}_I[\mathbf{R}_1|\mathbf{R}_2|\mathbf{R}_3|\mathbf{T}] \begin{bmatrix} X_w \\ Y_w \\ 0 \\ 1 \end{bmatrix} = \mathbf{P}_I[\mathbf{R}_1|\mathbf{R}_2|\mathbf{T}] \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix}$$

where $\mathbf{P}_w = [X_w, Y_w, 0]^T$ a scene point on the calibration pattern, $\mathbf{p} = [x_d, y_d]^T$ its two-dimensional image and s is just a scale factor.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Zhang's calibration method

- A 3×3 homography matrix $\mathbf{H}$ can be defined up to a scale factor, relating $\mathbf{P}$ and $\mathbf{p}$:

$$s\mathbf{p} = \mathbf{HP} = \mathbf{P}_I[\mathbf{R}_1|\mathbf{R}_2|\mathbf{T}]\mathbf{P}$$

Therefore:

$$\mathbf{H} = \lambda\mathbf{P}_I[\mathbf{R}_1|\mathbf{R}_2|\mathbf{T}]$$

where λ is a scale factor and $\mathbf{H} \triangleq [\mathbf{h}_1\mathbf{h}_2\mathbf{h}_3]$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Zhang's calibration method

- From known 3D and 2D correspondences, such a homography can be estimated iteratively and by utilizing

$$s\mathbf{p} = \mathbf{HP} = \mathbf{P}_I[\mathbf{R}_1|\mathbf{R}_2|\mathbf{T}]\mathbf{P}$$

to form an objective function for optimization with the aid of a non-linear optimization algorithm, like Levenberg-Marquardt:

$$\min_{\mathbf{H}_i} E(\mathbf{H}_i) = \sum_{j=1}^{M} \|\mathbf{p}_{ij} - \hat{\mathbf{H}}_i \mathbf{P}_j\|^2, \ i = 1, \ldots, N.$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Zhang's calibration method

- Given a known $\mathbf{H}$, let ω be a symmetric matrix:

$$\boldsymbol{\omega} \triangleq \begin{bmatrix} \omega_{11} & \omega_{12} & \omega_{13} \\ \omega_{21} & \omega_{22} & \omega_{23} \\ \omega_{31} & \omega_{32} & \omega_{33} \end{bmatrix} = \mathbf{P}_I^{-T}\mathbf{P}_I^{-1} =$$

$$= \begin{bmatrix} \frac{1}{\alpha_x^2} & \frac{-s_\theta}{\alpha_x^2\alpha_y} & \frac{o_y s_\theta - o_x\alpha_y}{\alpha_x^2\alpha_y} \\ \frac{-s_\theta}{\alpha_x^2\alpha_y} & \frac{s_\theta^2}{\alpha_x^2\alpha_y^2} + \frac{1}{\alpha_y^2} & \frac{-s_\theta(o_y s_\theta - o_x\alpha_y)}{\alpha_x^2\alpha_y^2} - \frac{o_y}{\alpha_y^2} \\ \frac{o_y s_\theta - o_x\alpha_y}{\alpha_x^2\alpha_y} & \frac{-s_\theta(o_y s_\theta - o_x\alpha_y)}{\alpha_x^2\alpha_y^2} - \frac{o_y}{\alpha_y^2} & \frac{(o_y s_\theta - o_x\alpha_y)^2}{\alpha_x^2\alpha_y^2} + \frac{o_y^2}{\alpha_y^2} + 1 \end{bmatrix}$$

which can be represented by $\mathbf{b} = [\omega_{11}, \omega_{12}, \omega_{22}, \omega_{13}, \omega_{23}, \omega_{33}]^T$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Zhang's calibration method

- Since $\mathbf{R}_1$ and $\mathbf{R}_2$ are orthogonal, $s\mathbf{p} = \mathbf{HP} = \mathbf{P}_I[\mathbf{R}_1|\mathbf{R}_2|\mathbf{T}]\mathbf{P}$ and the definition of ω entail that:

$$\mathbf{h}_1^T \omega \mathbf{h}_2 = 0 \qquad \mathbf{h}_1^T \omega \mathbf{h}_1 = \mathbf{h}_2^T \omega \mathbf{h}_2. \qquad \mathbf{h}_i^T \omega \mathbf{h}_j = \mathbf{v}_{ij}^T \mathbf{b}$$

where:

$$\mathbf{v}_{ij}^T \triangleq [h_{i1}h_{j1}, h_{i1}h_{j2}+h_{i2}h_{j1}, h_{i2}h_{j2}, h_{i3}h_{j1}+h_{i1}h_{j3}, h_{i3}h_{j2}+h_{i2}h_{j3}, h_{i3}h_{j3}]$$

- Based on the way $\mathbf{v}$ is defined,

$$\begin{bmatrix} \mathbf{v}_{12}^T \\ (\mathbf{v}_{11} - \mathbf{v}_{22})^T \end{bmatrix} \mathbf{b} = \mathbf{0}$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Zhang's calibration method

- By taking simultaneously into account N different images of the calibration pattern, and using N different homography matrices $\mathbf{H}_i, i = 1, \ldots, N$, the system of N corresponding equations can be compactly restated as:

$$\mathbf{Ab} = 0,$$

where $\mathbf{A}$ is a $2N \times 6$ matrix.

- This system can be solved for $\mathbf{b}$ applying SVD to $\mathbf{A} = \mathbf{UDV}^T$.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Zhang's calibration method

- The intrinsic camera parameters can be estimated from ω:

$$o_y = \frac{(\omega_{12}\omega_{13} - \omega_{11}\omega_{23})}{\omega_{11}\omega_{22} - \omega_{12}^2}$$

$$\lambda = \omega_{33} - \frac{\omega_{13}^2 + o_y(\omega_{12}\omega_{13} - \omega_{11}\omega_{23})}{\omega_{11}}$$

$$\alpha_x = \sqrt{\frac{\lambda}{\omega_{11}}}$$

$$\alpha_y = \sqrt{\frac{\lambda\omega_{11}}{\omega_{11}\omega_{22} - \omega_{12}^2}}$$

$$s_\theta = -\frac{\omega_{12}\alpha_x^2\alpha_y}{\lambda}$$

$$o_x = \frac{s_\theta o_y}{\alpha_y} - \frac{\omega_{13}\alpha_x^2}{\lambda}.$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Zhang's calibration method

- Having computed the intrinsic camera parameters, the matrix $\mathbf{P}_I$ can also be estimated.
- Thus, by substituting $\mathbf{P}_I$ in the equations:

$$\mathbf{p}_{ij} = \lambda \mathbf{P}_I [\mathbf{R}_{1i} | \mathbf{R}_{2i} | \mathbf{T}_i] \mathbf{P}_j, \quad i = 1, \ldots, N, \quad j = 1, \ldots, M$$

we can solve for the columns of the rotation matrix and the translation vector, in order to obtain the extrinsic camera parameters (rotation matrix $\mathbf{R}_i$, translation vector $\mathbf{T}_i$).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Zhang's calibration method

$$\mathbf{R}_{1i} = \lambda \mathbf{P}_I^{-1} \mathbf{h}_{1i}$$
$$\mathbf{R}_{2i} = \lambda \mathbf{P}_I^{-1} \mathbf{h}_{2i}$$
$$\mathbf{R}_{3i} = \mathbf{R}_{1i} \times \mathbf{R}_{2i}$$
$$\mathbf{T}_i = \lambda \mathbf{P}_I^{-1} \mathbf{h}_{3i}$$

$$\lambda = \frac{1}{\|\mathbf{P}_I^{-1}\mathbf{h}_{1i}\|} = \frac{1}{\|\mathbf{P}_I^{-1}\mathbf{h}_{2i}\|}.$$

- The above estimated results can be used as initializations for some repetitive optimization algorithm, so that refined results ones can be derived.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Self-calibration

- These approaches:
  - do not require a calibration object and recover the camera parameters from image information alone;
  - are flexible but not very robust;
  - exploit properties of the absolute conic of the projective geometry;
  - typically require information equivalent to a partial 3D reconstruction of the scene.
- Self-calibration is strongly related to the geometry of multiple cameras.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)

Q & A

Thank you very much for your attention!

Contact: Prof. I. Pitas
pitas@aiia.csd.auth.gr
www.multidrone.eu

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)