

# 3D Computer Vision

Radim Šára    Martin Matoušek

Center for Machine Perception  
Department of Cybernetics  
Faculty of Electrical Engineering  
Czech Technical University in Prague

<https://cw.fel.cvut.cz/wiki/courses/tdv/start>

<http://cmp.felk.cvut.cz>

<mailto:sara@cmp.felk.cvut.cz>

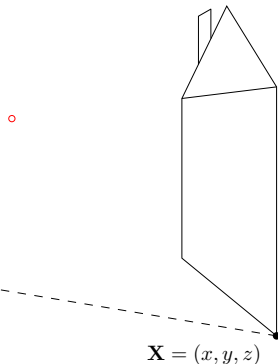
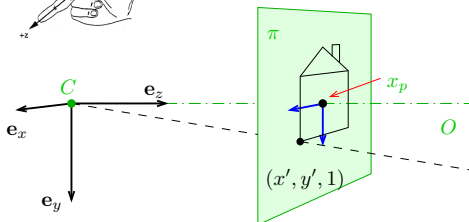
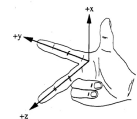
phone ext. 7203

rev. October 3, 2023

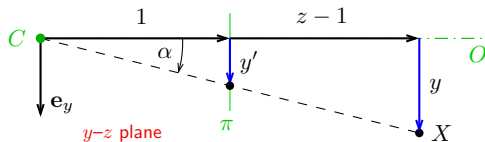


Open Informatics Master's Course

## ► Canonical Perspective Camera (Pinhole Camera, Camera Obscura)



1. in this picture we are looking 'down the street'
2. right-handed canonical coordinate system  $(x, y, z)$  with unit vectors  $e_x, e_y, e_z$
3. origin = center of projection  $C$
4. image plane  $\pi$  at unit distance from  $C$
5. optical axis  $O$  is perpendicular to  $\pi$
6. principal point  $x_p$ : intersection of  $O$  and  $\pi$
7. perspective camera is given by  $C$  and  $\pi$



projected point in the natural image coordinate system:

$$\tan \alpha = \frac{y'}{1} = y' = \frac{y}{1 + z - 1} = \frac{y}{z}, \quad x' = \frac{x}{z}$$

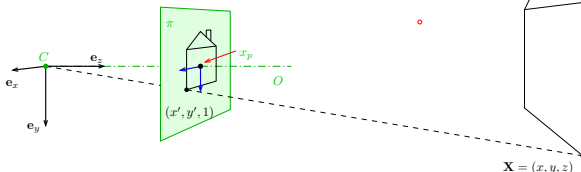
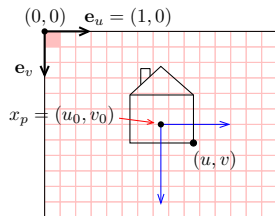
## ► Natural and Canonical Image Coordinate Systems

projected point **in canonical camera** ( $z \neq 0$ )

$$(x', y', 1) = \left( \frac{x}{z}, \frac{y}{z}, 1 \right) = \frac{1}{z}(x, y, z) \simeq (x, y, z) \equiv \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\mathbf{P}_0 = [\mathbf{I} \quad \mathbf{0}]} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \mathbf{P}_0 \underline{\mathbf{X}}$$

projected point **in scanned image**

scale by  $f$  and translate origin to image corner



$$u = f \frac{x}{z} + u_0$$

$$v = f \frac{y}{z} + v_0$$

$$\frac{1}{z} \begin{bmatrix} f x + z u_0 \\ f y + z v_0 \\ z \end{bmatrix} \simeq \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \mathbf{K} \mathbf{P}_0 \underline{\mathbf{X}} = \mathbf{P} \underline{\mathbf{X}}$$

- 'calibration' matrix  $\mathbf{K}$  transforms canonical  $\mathbf{P}_0$  to standard perspective camera  $\mathbf{P}$

## ► Computing with Perspective Camera Projection Matrix

Projection from world to image in standard camera  $\mathbf{P}$ :

$$\underbrace{\begin{bmatrix} f & 0 & u_0 & 0 \\ 0 & f & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\mathbf{P}} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} fx + u_0z \\ fy + v_0z \\ z \end{bmatrix} \simeq \underbrace{\begin{bmatrix} x + \frac{z}{f}u_0 \\ y + \frac{z}{f}v_0 \\ \frac{z}{f} \end{bmatrix}}_{(a)} \simeq \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \underline{\mathbf{m}}$$

cross-check:  $\frac{m_1}{m_3} = \frac{fx}{z} + u_0 = u$ ,  $\frac{m_2}{m_3} = \frac{fy}{z} + v_0 = v$  when  $m_3 \neq 0$

$f$  – ‘focal length’ – converts length ratios to pixels,  $[f] = \text{px}$ ,  $f > 0$

$(u_0, v_0)$  – principal point in pixels

### Perspective Camera:

1. dimension reduction

since  $\mathbf{P} \in \mathbb{R}^{3,4}$

2. nonlinear unit change  $\mathbf{1} \mapsto \mathbf{1} \cdot z/f$ , see (a)

for convenience we use  $P_{11} = P_{22} = f$  rather than  $P_{33} = 1/f$  and the  $u_0, v_0$  in relative units

3.  $(m_1, m_2, 0)$  represents points at infinity in image plane  $\pi$

i.e. points with  $z = 0$

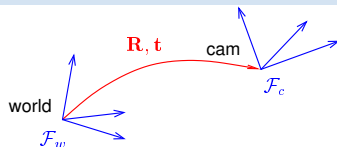
## ► Changing The Outer (World) Reference Frame

A transformation of a point from the world to camera coordinate system:

$$\mathbf{X}_c = \mathbf{R} \mathbf{X}_w + \mathbf{t}$$

$\mathbf{R}$  – rotation matrix    world orientation in the camera coordinate frame  $\mathcal{F}_c$

$\mathbf{t}$  – translation vector    world origin in the camera coordinate frame  $\mathcal{F}_c$



$$\mathbf{P} \underline{\mathbf{X}}_c = \mathbf{K} \mathbf{P}_0 \begin{bmatrix} \mathbf{X}_c \\ 1 \end{bmatrix} = \mathbf{K} \mathbf{P}_0 \begin{bmatrix} \mathbf{R} \mathbf{X}_w + \mathbf{t} \\ 1 \end{bmatrix} = \mathbf{K} \underbrace{\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{bmatrix}}_{\mathbf{P}_0} \underbrace{\begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix}}_{\mathbf{T}} \begin{bmatrix} \mathbf{X}_w \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix} \underline{\mathbf{X}}_w \quad \mathbf{T}^{-1} = ?$$

$\mathbf{P}_0$  (a  $3 \times 4$  mtx) discards the last row of  $\mathbf{T}$

- $\mathbf{R}$  is rotation,  $\mathbf{R}^\top \mathbf{R} = \mathbf{I}$ ,  $\det \mathbf{R} = +1$
- **6 extrinsic parameters**: 3 rotation angles (Euler theorem), 3 translation components
- alternative, often used, camera representations

$\mathbf{I} \in \mathbb{R}^{3,3}$  identity matrix

$$\mathbf{P} = \mathbf{K} \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix} = \mathbf{K} \mathbf{R} \begin{bmatrix} \mathbf{I} & -\mathbf{C} \end{bmatrix} \quad \text{i.e. } \mathbf{C} = -\mathbf{R}^\top \mathbf{t}$$

$\mathbf{C}$  – camera position in the world reference frame  $\mathcal{F}_w$

$\mathbf{r}_3^\top$  – optical axis in the world reference frame  $\mathcal{F}_w$     cam:  $\mathbf{o}_c = (1, 0, 0)$ ,    world:  $\mathbf{o}_w = -\mathbf{R}^\top \mathbf{o}_c = \mathbf{r}_3^\top$

$\mathbf{t} = -\mathbf{R} \mathbf{C}$   
third row of  $\mathbf{R}$

- we can save some conversion and computation by noting that  $\mathbf{K} \mathbf{R} \begin{bmatrix} \mathbf{I} & -\mathbf{C} \end{bmatrix} \underline{\mathbf{X}} = \mathbf{K} \mathbf{R} (\underline{\mathbf{X}} - \mathbf{C})$

## ► Changing the Inner (Image) Reference Frame

### The general form of calibration matrix $\mathbf{K}$ includes

- skew angle  $\theta$  of the digitization raster
- pixel aspect ratio  $a$

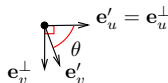
$$\mathbf{K} = \begin{bmatrix} a f & -a f \cot \theta & u_0 \\ 0 & f / \sin \theta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{units: } [f] = \text{px}, [u_0] = \text{px}, [v_0] = \text{px}, [a] = 1$$

⊗ H1; 2pt: Give the parameters  $f, a, \theta, u_0, v_0$  a precise meaning by decomposing  $\mathbf{K}$  to simple maps; deadline LD+2 wk

Hints:

1. image projects to orthogonal system  $F^\perp$ , then it maps by skew to  $F'$ , then by scale  $a f$ ,  $f$  to  $F''$ , then by translation by  $u_0, v_0$  to  $F'''$
2. Skew: Do not confuse it with the **shear mapping**. Express point  $\mathbf{x}$  as

$$\mathbf{x} = u' \mathbf{e}_{u'} + v' \mathbf{e}_{v'} = u^\perp \mathbf{e}_u^\perp + v^\perp \mathbf{e}_v^\perp, \quad u, v \in \mathbb{R}$$



$\mathbf{e}_u^\perp$  are unit-length basis vectors  $\mathbf{e}_u^\perp = \mathbf{e}'_u = (1, 0)$ ,  $\mathbf{e}_v^\perp = (0, 1), \dots$

consider their four pairwise dot-products  $(\mathbf{e}'_u)^\top \mathbf{e}_u^\perp = 0$ ,  $(\mathbf{e}'_u)^\top \mathbf{e}'_v = \cos(\theta)$ ,  $\dots$

3.  $\mathbf{K}$  maps from  $F^\perp$  to  $F'''$  as

$$w''' [u''', v''', 1]^\top = \mathbf{K} [u^\perp, v^\perp, 1]^\top$$

## ► Summary: Projection Matrix of a General Finite Perspective Camera

$$\underline{\mathbf{m}} \simeq \mathbf{P}\underline{\mathbf{X}}, \quad \mathbf{P} = [\mathbf{Q} \quad \mathbf{q}] \simeq \mathbf{K} [\mathbf{R} \quad \mathbf{t}] = \mathbf{K}\mathbf{R}[\mathbf{I} \quad -\mathbf{C}]$$

a recipe for filling  $\mathbf{P}$

**general finite perspective camera has 11 parameters:**

- 5 intrinsic parameters:  $f, u_0, v_0, a, \theta$
- 6 extrinsic parameters:  $\mathbf{t}, \mathbf{R}(\alpha, \beta, \gamma)$

finite camera:  $\det \mathbf{K} \neq 0$

**Representation Theorem:** The set of projection matrices  $\mathbf{P}$  of finite perspective cameras is isomorphic to the set of homogeneous  $3 \times 4$  matrices with the left  $3 \times 3$  submatrix  $\mathbf{Q}$  non-singular.

random finite camera: `Q = rand(3,3); while det(Q)==0, Q = rand(3,3); end, P = [Q, rand(3,1)];`

## ► Projection Matrix Decomposition

$$\mathbf{P} = [\mathbf{Q} \quad \mathbf{q}] \rightarrow \mathbf{K} [\mathbf{R} \quad \mathbf{t}]$$

$$\mathbf{Q} \in \mathbb{R}^{3,3}$$

$$\mathbf{K} \in \mathbb{R}^{3,3}$$

$$\mathbf{R} \in \mathbb{R}^{3,3}$$

full rank

(if finite perspective camera; see [H&Z, Sec. 6.3] for cameras at infinity)

upper triangular with positive diagonal elements

rotation mtx:  $\mathbf{R}^T \mathbf{R} = \mathbf{I}$  and  $\det \mathbf{R} = +1$

1.  $[\mathbf{Q} \quad \mathbf{q}] = \mathbf{K} [\mathbf{R} \quad \mathbf{t}] = [\mathbf{KR} \quad \mathbf{Kt}]$

also →35

2. RQ decomposition of  $\mathbf{Q} = \mathbf{KR}$  using three Givens rotations

[H&Z, p. 579]

$$\mathbf{K} = \mathbf{Q} \underbrace{\mathbf{R}_{32} \mathbf{R}_{31} \mathbf{R}_{21}}_{\mathbf{R}^{-1}} \quad \mathbf{Q} \mathbf{R}_{32} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & 0 & \cdot \end{bmatrix}, \quad \mathbf{Q} \mathbf{R}_{32} \mathbf{R}_{31} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 0 & 0 & \cdot \end{bmatrix}, \quad \mathbf{Q} \mathbf{R}_{32} \mathbf{R}_{31} \mathbf{R}_{21} = \begin{bmatrix} \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot \\ 0 & 0 & \cdot \end{bmatrix}$$

$\mathbf{R}_{ij}$  zeroes element  $ij$  in  $\mathbf{Q}$  affecting only columns  $i$  and  $j$  and the sequence preserves previously zeroed elements, e.g.

(see the next slide for derivation details)

$$\mathbf{R}_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{bmatrix} \text{ gives } 0 = k_{32} = c q_{32} + s q_{33} \Rightarrow c = \frac{q_{33}}{\sqrt{q_{32}^2 + q_{33}^2}} \quad s = \frac{-q_{32}}{\sqrt{q_{32}^2 + q_{33}^2}}$$

⊛ P1; 1pt: Multiply known matrices  $\mathbf{K}$ ,  $\mathbf{R}$  and then decompose back; discuss numerical errors

- RQ decomposition nonuniqueness:  $\mathbf{KR} = \mathbf{KT}^{-1}\mathbf{TR}$ , where  $\mathbf{T} = \text{diag}(-1, -1, 1)$  is also a rotation, we must correct the result so that the diagonal elements of  $\mathbf{K}$  are all positive 'thin' RQ decomposition
- care must be taken to avoid overflow, see [Golub & van Loan 2013, sec. 5.2]



## RQ Decomposition Step

```
Q = Array [qm1,m2 &, {3, 3}];  
R32 = {{1, 0, 0}, {0, c, -s}, {0, s, c}}; R32 // MatrixForm
```

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{pmatrix}$$

```
Q1 = Q . R32 ; Q1 // MatrixForm
```

$$\begin{pmatrix} q_{1,1} & c q_{1,2} + s q_{1,3} & -s q_{1,2} + c q_{1,3} \\ q_{2,1} & c q_{2,2} + s q_{2,3} & -s q_{2,2} + c q_{2,3} \\ q_{3,1} & c q_{3,2} + s q_{3,3} & -s q_{3,2} + c q_{3,3} \end{pmatrix}$$

```
s1 = Solve [{Q1[[3]][[2]] == 0, c^2 + s^2 == 1}, {c, s}][[2]]
```

$$\left\{ c \rightarrow \frac{q_{3,3}}{\sqrt{q_{3,2}^2 + q_{3,3}^2}}, s \rightarrow -\frac{q_{3,2}}{\sqrt{q_{3,2}^2 + q_{3,3}^2}} \right\}$$

```
Q1 /. s1 // Simplify // MatrixForm
```

$$\begin{pmatrix} q_{1,1} & \frac{-q_{1,3} q_{3,2} + q_{1,2} q_{3,3}}{\sqrt{q_{3,2}^2 + q_{3,3}^2}} & \frac{q_{1,2} q_{3,2} + q_{1,3} q_{3,3}}{\sqrt{q_{3,2}^2 + q_{3,3}^2}} \\ q_{2,1} & \frac{-q_{2,3} q_{3,2} + q_{2,2} q_{3,3}}{\sqrt{q_{3,2}^2 + q_{3,3}^2}} & \frac{q_{2,2} q_{3,2} + q_{2,3} q_{3,3}}{\sqrt{q_{3,2}^2 + q_{3,3}^2}} \\ q_{3,1} & 0 & \sqrt{q_{3,2}^2 + q_{3,3}^2} \end{pmatrix}$$

## ► Center of Projection (Optical Center)

**Observation:** finite  $\mathbf{P}$  has a non-trivial right null-space

rank 3 but 4 columns

### Theorem

Let  $\mathbf{P}$  be a camera and let there be  $\underline{\mathbf{B}} \neq \mathbf{0}$  s.t.  $\mathbf{P}\underline{\mathbf{B}} = \mathbf{0}$ . Then  $\underline{\mathbf{B}}$  is equivalent to the projection center  $\underline{\mathbf{C}}$  (homogeneous, in world coordinate frame).

### Proof.

1. Let  $AB$  be a spatial line ( $B$  given from  $\mathbf{P}\underline{\mathbf{B}} = \mathbf{0}$ ,  $A \neq B$ ). Then

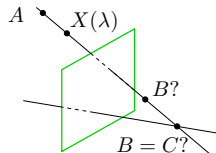
$$\underline{\mathbf{X}}(\lambda) \simeq \lambda \underline{\mathbf{A}} + (1 - \lambda) \underline{\mathbf{B}}, \quad \lambda \in \mathbb{R} \quad (\text{world frame})$$

2. It projects to

$$\mathbf{P}\underline{\mathbf{X}}(\lambda) \simeq \lambda \mathbf{P}\underline{\mathbf{A}} + (1 - \lambda) \mathbf{P}\underline{\mathbf{B}} \simeq \mathbf{P}\underline{\mathbf{A}}$$

- the entire line projects to a single point  $\Rightarrow$  it must pass through the projection center of  $\mathbf{P}$
- this holds for any choice of  $A \neq B \Rightarrow$  the only common point of the lines is the  $C$ , i.e.  $\underline{\mathbf{B}} \simeq \underline{\mathbf{C}}$

□



Hence

$$\mathbf{0} = \mathbf{P}\underline{\mathbf{C}} = [\mathbf{Q} \quad \mathbf{q}] \begin{bmatrix} \underline{\mathbf{C}} \\ 1 \end{bmatrix} = \mathbf{Q}\underline{\mathbf{C}} + \mathbf{q} \Rightarrow \underline{\mathbf{C}} = -\mathbf{Q}^{-1}\mathbf{q}$$

⊗ verify from →30

$\underline{\mathbf{C}} = (c_j)$ , where  $c_j = (-1)^j \det \mathbf{P}^{(j)}$ , in which  $\mathbf{P}^{(j)}$  is  $\mathbf{P}$  with column  $j$  dropped

Matlab: `C_homo = null(P);` or `C = -Q\q;`

## ► Optical Ray

Optical ray: Spatial line that projects to a single image point.

1. Consider the following spatial line (world coordinate frame)

$\mathbf{d} \in \mathbb{R}^3$  line direction vector,  $\|\mathbf{d}\| = 1$ ,  $\lambda \in \mathbb{R}$ , Cartesian representation

$$\mathbf{X}(\lambda) = \mathbf{C} + \lambda \mathbf{d}$$

2. The projection of the (finite) point  $X(\lambda)$  is

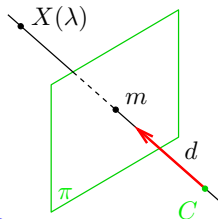
$$\begin{aligned} \underline{\mathbf{m}} &\simeq [\mathbf{Q} \quad \mathbf{q}] \begin{bmatrix} \mathbf{X}(\lambda) \\ 1 \end{bmatrix} = \mathbf{Q}(\mathbf{C} + \lambda \mathbf{d}) + \mathbf{q} = \lambda \mathbf{Q} \mathbf{d} = \\ &= \lambda [\mathbf{Q} \quad \mathbf{q}] \begin{bmatrix} \mathbf{d} \\ 0 \end{bmatrix} \end{aligned}$$

... which is also the image of a point at infinity in  $\mathbb{P}^3$

- optical ray line corresponding to image point  $m$  is the set

$$\mathbf{X}(\mu) = \mathbf{C} + \mu \mathbf{Q}^{-1} \underline{\mathbf{m}}, \quad \mu \in \mathbb{R} \quad (\mu = 1/\lambda)$$

- optical ray direction may be represented by a point at infinity  $(\mathbf{d}, 0)$  in  $\mathbb{P}^3$
- optical ray is expressed in the world coordinate frame



## ► Optical Axis

Optical axis: Optical ray that is perpendicular to image plane  $\pi$

1. points  $X$  on a given line  $N$  parallel to  $\pi$  project to a point at infinity  $(u, v, 0)$  in  $\pi$ :

$$\begin{bmatrix} u \\ v \\ 0 \end{bmatrix} \simeq \mathbf{P}\mathbf{X} = \begin{bmatrix} \mathbf{q}_1^\top & q_{14} \\ \mathbf{q}_2^\top & q_{24} \\ \mathbf{q}_3^\top & q_{34} \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ 1 \end{bmatrix}$$

2. therefore the set of points  $X$  is parallel to  $\pi$  iff

$$\mathbf{q}_3^\top \mathbf{X} + q_{34} = 0$$

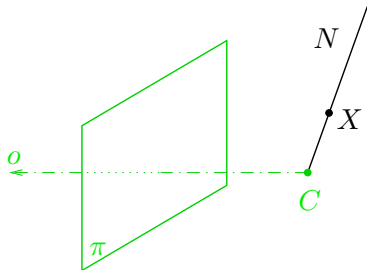
3. this is a plane equation with  $\pm \mathbf{q}_3$  as the normal vector
4. optical axis direction: substitution  $\mathbf{P} \mapsto \lambda \mathbf{P}$  must not change the direction
5. we select (assuming  $\det(\mathbf{R}) > 0$ )

$$\mathbf{o} = \det(\mathbf{Q}) \mathbf{q}_3$$

if  $\mathbf{P} \mapsto \lambda \mathbf{P}$  then  $\det(\mathbf{Q}) \mapsto \lambda^3 \det(\mathbf{Q})$  and  $\mathbf{q}_3 \mapsto \lambda \mathbf{q}_3$ , hence  $\mathbf{o} \mapsto \mathbf{o}$

[H&Z, p. 161]

- the axis is expressed in the world coordinate frame



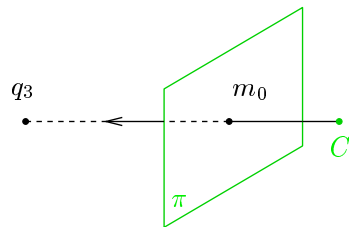
## ► Principal Point

Principal point: The intersection of image plane and the optical axis

1. as we saw,  $\mathbf{q}_3$  is the directional vector of optical axis
2. we take point at infinity on the optical axis that must project to the principal point  $m_0$

3. then

$$\underline{\mathbf{m}}_0 \simeq [\mathbf{Q} \quad \mathbf{q}] \begin{bmatrix} \mathbf{q}_3 \\ 0 \end{bmatrix} = \mathbf{Q} \mathbf{q}_3$$

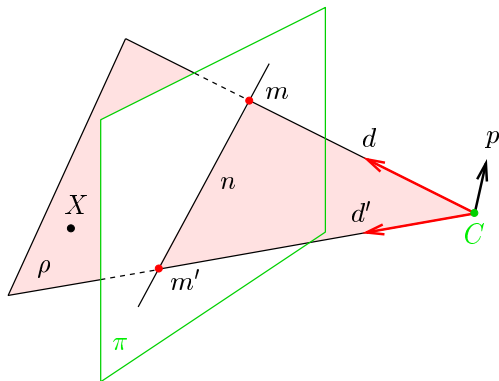


principal point:  $\underline{\mathbf{m}}_0 \simeq \mathbf{Q} \mathbf{q}_3$

- principal point is also the center of radial distortion

## ► Optical Plane

A spatial plane with normal  $p$  containing the projection center  $C$  and a given image line  $n$ .



optical ray given by  $m$        $\underline{d} \simeq \mathbf{Q}^{-1} \underline{m}$   
 optical ray given by  $m'$        $\underline{d}' \simeq \mathbf{Q}^{-1} \underline{m}'$



$$\underline{p} \simeq \underline{d} \times \underline{d}' = (\mathbf{Q}^{-1} \underline{m}) \times (\mathbf{Q}^{-1} \underline{m}') \stackrel{\textcircled{*}}{=} \mathbf{Q}^T (\underline{m} \times \underline{m}') = \mathbf{Q}^T \underline{n}$$

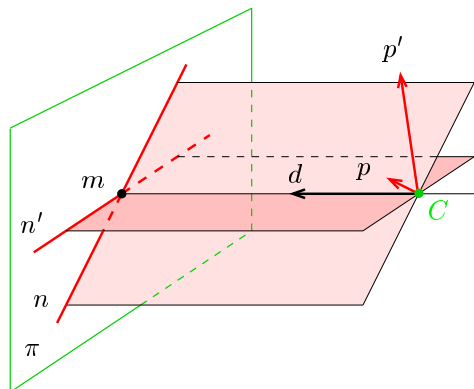
• note the way  $\mathbf{Q}$  factors out!

hence,  $0 = \underline{p}^T (\underline{X} - \underline{C}) = \underline{n}^T \underbrace{\mathbf{Q}(\underline{X} - \underline{C})}_{\rightarrow 30} = \underline{n}^T \mathbf{P} \underline{X} = (\mathbf{P}^T \underline{n})^T \underline{X}$  for every  $X$  in plane  $\rho$

optical plane is given by  $n$ :  $\underline{\rho} \simeq \mathbf{P}^T \underline{n}$

$$\rho_1 x + \rho_2 y + \rho_3 z + \rho_4 = 0$$

## Cross-Check: Optical Ray as Optical Plane Intersection



optical plane normal given by  $\underline{n}$

optical plane normal given by  $\underline{n}'$

$$\underline{d} = \underline{p} \times \underline{p}' = (\mathbf{Q}^T \underline{n}) \times (\mathbf{Q}^T \underline{n}') = \mathbf{Q}^{-1}(\underline{n} \times \underline{n}') = \mathbf{Q}^{-1} \underline{m}$$

$$\underline{p} = \mathbf{Q}^T \underline{n}$$

$$\underline{p}' = \mathbf{Q}^T \underline{n}'$$

Thank You