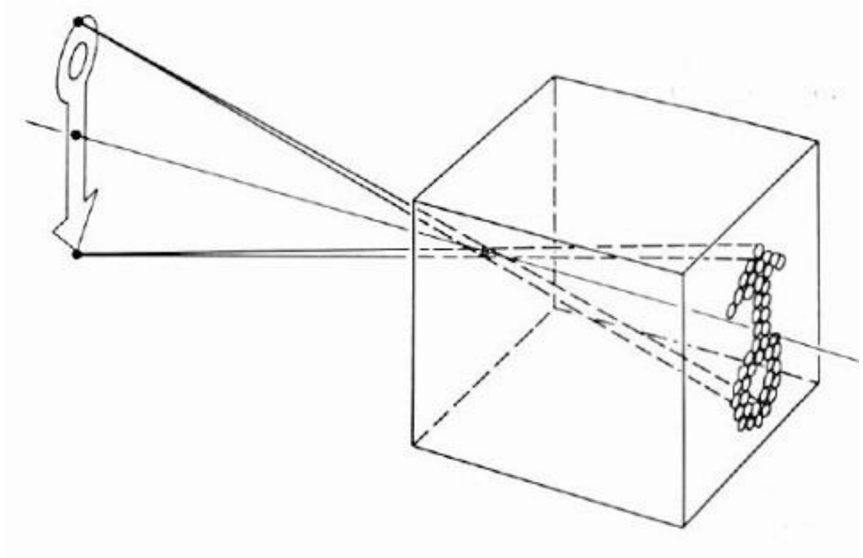
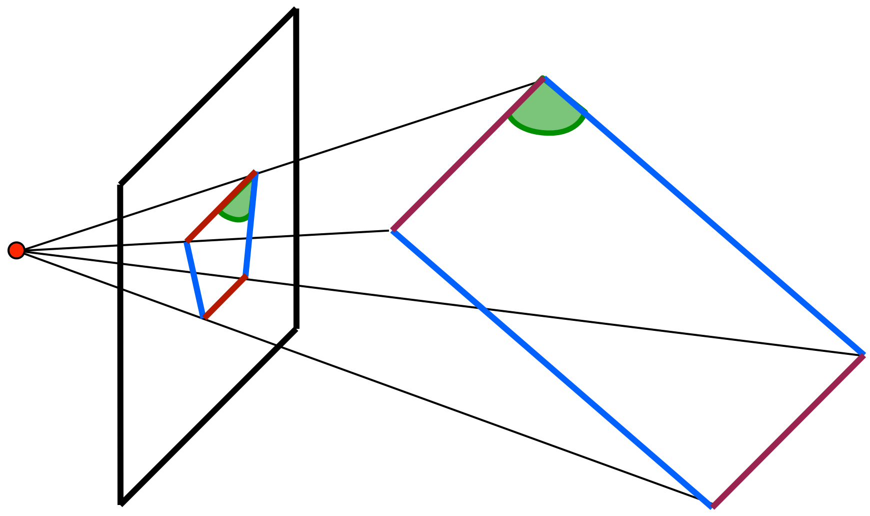
Camera model

Pinhole camera



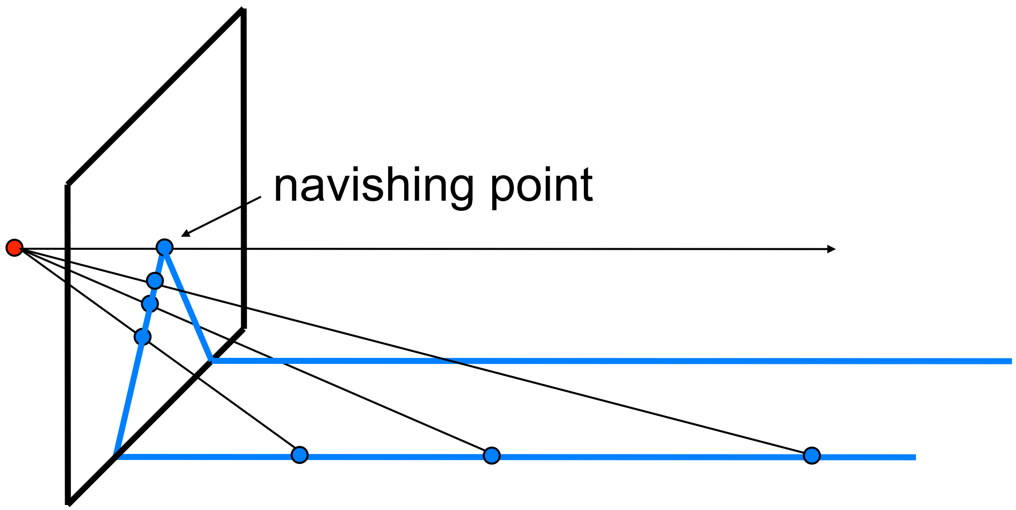
- closed box with a small hole
  - ideal camera: hole is infinitely small
  - incoming rays are a pencil of lines
  - image is created on the back wall, the ***image plane***
  - ***camera center***: (ideal) intersection point of rays
  - ***camera constant***: distance between camera center and image plane

Pinhole camera



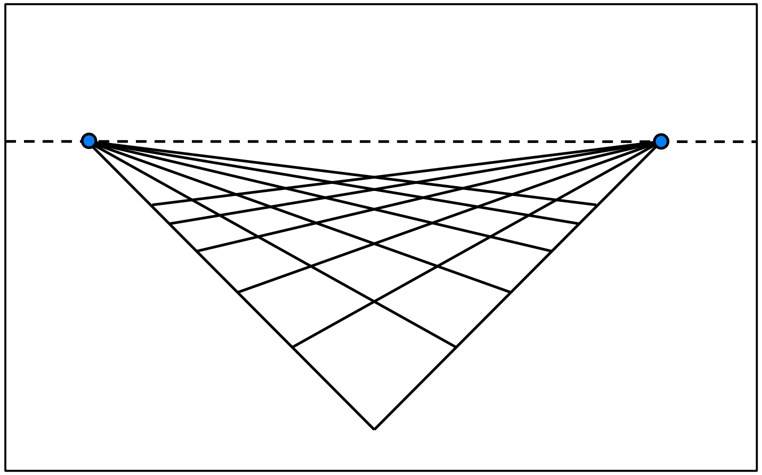
- Properties of the mapping
  - **line-preserving**: straight lines are mapped to straight lines
    rays through all points of a line form a plane
  - **not length-preserving**: lengths (and length ratios) are lost
    scale of mapping is inversely proportional to object distance
  - **not angle-preserving**: angles between lines change
    note: follows from the two previous ones - why?

Pinhole camera



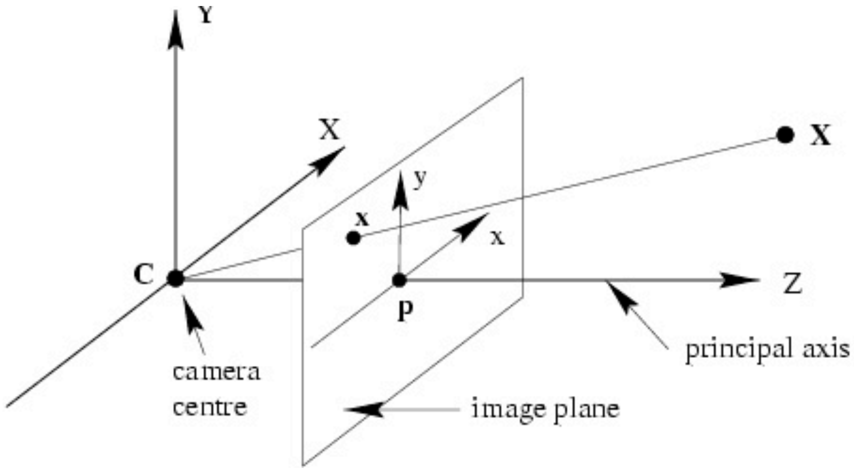
- Vanishing points
  - parallel lines are mapped to non-parallel ones
  - images of all lines that are parallel to eathother intersect in a vanishing point
  - vanishing point is the image of their (common) point at infinity
  - every direction in space has exactly one corresponding vanishing point

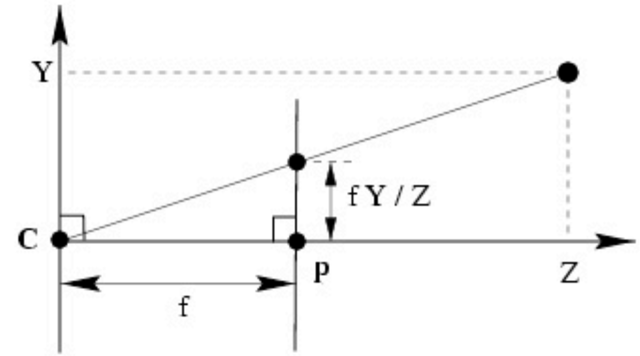
Pinhole camera



- Vanishing lines
  - by the same argument, all parallel planes have a common vanishing line
  - rays through the line at infinity form one of these planes and map it to a vanishing line
  - if a line is parallel to a plane, its vanishing point will lie on the plane's vanishing line
  - vanishing line of horizontal planes is often called the "horizon"

Perspective projection

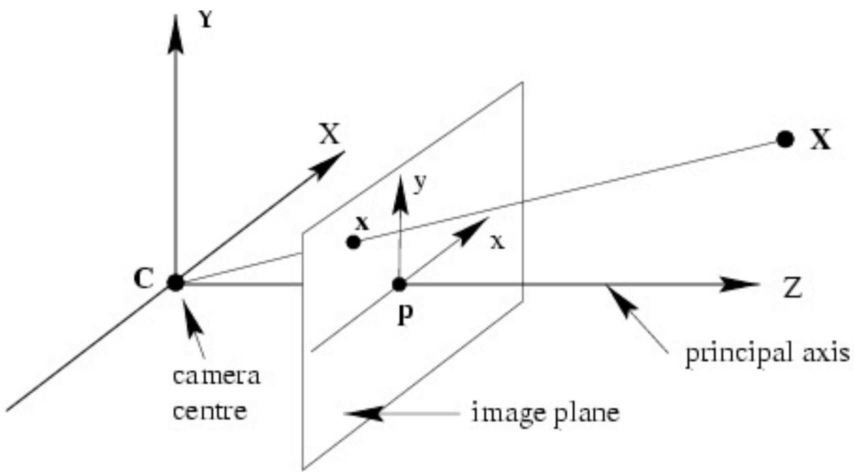


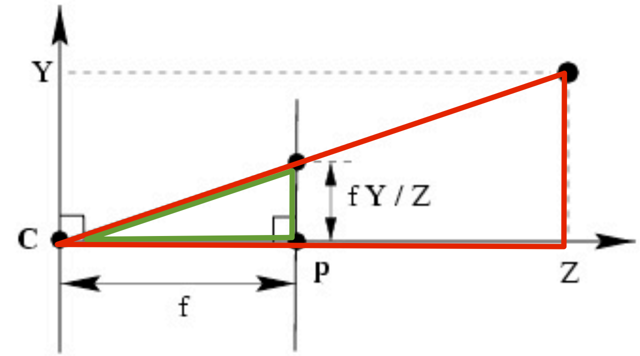


- Choice of camera coordinate system
  - Hole (projection center) is the origin
  - image plane orthogonal to z-axis (z -axis = principal axis)
  - image plane **in front of** projection center
  - physically correct location is equivalent

    (image plane behind camera center, camera constant $-f$)

Perspective projection





- Mapping
  - ray through object point and camera center intersects image plane → image point
  - from equal triangles

$$(X, Y, Z) \rightarrow (f\frac{X}{Z}, f\frac{Y}{Z}, f)$$

  - 2D coordinates in image plane (z same for all image points)

$$\mathbf{x} = (f\frac{X}{Z}, f\frac{Y}{Z})$$

Projection matrix

- projective representation
  - in euclidean coordinates projection is non-linear (division by Z)
  - in homogeneous coordinates

$$\mathbf{x} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} fX \\ fY \\ Z \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} f & & \\ & f & \\ & & 1 \end{bmatrix} \begin{bmatrix} \mathsf{I} & \mathbf{0} \end{bmatrix} \mathbf{X}$$

  - Normalisation

$$\begin{bmatrix} fX \\ fY \\ Z \end{bmatrix} \rightarrow \begin{bmatrix} fX/Z \\ fY/Z \end{bmatrix}$$

Projection matrix

- Interior Orientation (transition to image coordinates)
  - scaling from world coordinate units to pixels (image coordinate units)
  - pixel size in world units

$$\frac{1}{m_x} \times \frac{1}{m_y}$$

  - pixels ar eusually square, $m_x = m_y$

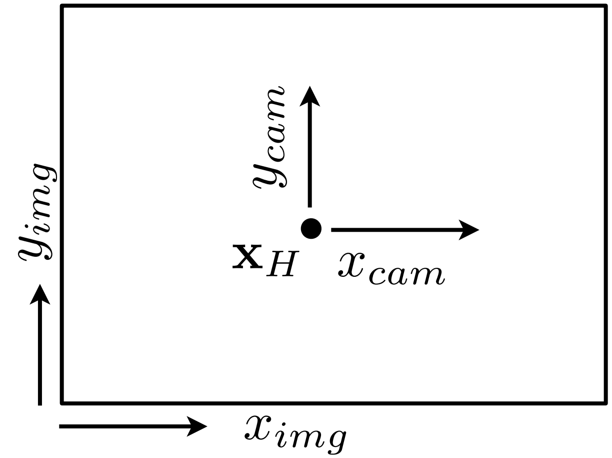
$$\mathbf{x} = \begin{bmatrix} m_x & & \\ & m_y & \\ & & 1 \end{bmatrix} \begin{bmatrix} f & & \\ & f & \\ & & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} \mathbf{X} = \begin{bmatrix} c_x & & \\ & c_y & \\ & & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} \mathbf{X}$$

pixel   pixel / meter   meter   meter

Projection matrix

- Interior Orientation (transition to image coordinates)
  - Translation (shift) from principal point to origin of image coordinate system (usually in the image corner)

$$\mathbf{x} = \begin{bmatrix} c_x & & x_H \\ & c_y & y_H \\ & & 1 \end{bmatrix} \begin{bmatrix} \mathsf{I} & \mathbf{0} \end{bmatrix} \mathbf{X}$$



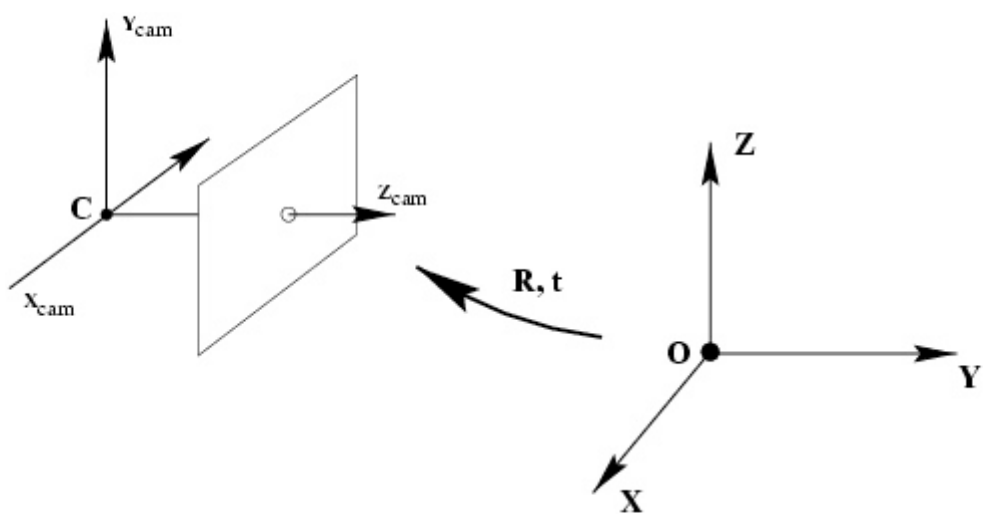
  - Shear of image coordinate system (usually no shear in real cameras)

$$\mathbf{x} = \begin{bmatrix} c_x & s & x_H \\ & c_y & y_H \\ & & 1 \end{bmatrix} \begin{bmatrix} \mathsf{I} & \mathbf{0} \end{bmatrix} \mathbf{X} = \mathsf{K} \begin{bmatrix} \mathsf{I} & \mathbf{0} \end{bmatrix} \mathbf{X}$$

$$(X, Y, Z) \rightarrow \left(c_x \frac{X}{Z} + s \frac{Y}{Z} + x_H \ , \ c_y \frac{Y}{Z} + y_H \right)$$

Projection matrix

- Exterior Orientation
  - assumption thus far: camera at the origin of the 3D world coordinate system, looking along z-axis
  - normally not the case → 3D point must first be shifted and rotated accordingly



Projection matrix

- Exterior Orientation
  - Application of translation and rotation

$$\mathbf{x} = \mathsf{K}\begin{bmatrix}\mathsf{I} & \mathbf{0}\end{bmatrix}\mathsf{R}\mathsf{T}\mathbf{X} = \mathsf{K}\begin{bmatrix}\mathsf{R} & -\mathsf{R}\mathbf{X}_0\end{bmatrix}\mathbf{X}$$

$$\mathbf{x} = \mathsf{P}\mathbf{X}$$

"projective equality": same geometric entity, i.e.equal up to a constant factor

$$\lambda\mathbf{x} = \mathsf{P}\mathbf{X}$$

$$\tilde{\mathsf{R}} = \begin{bmatrix}\mathsf{R} & \mathbf{0}\\ \mathbf{0}^\top & 1\end{bmatrix} \qquad \mathsf{T} = \begin{bmatrix}1 & 0 & 0 & -X_0\\ 0 & 1 & 0 & -Y_0\\ 0 & 0 & 1 & -Z_0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

- ***Note***: camera position in world coordinates is the null space of the projection matrix

$$\mathsf{P}\mathbf{X}_0 = \mathbf{0}$$

Projection matrix

- Summary
  - collinearity (mapping with pinhole camera)

    $$\mathbf{x} = \mathsf{P}\mathbf{X}$$

  - 11 degrees of freedom

    12 matrix elements - 1 arbitrary projective scale

    5 interior orientation + 6 exterior orientation

- Perspective projection is not invertible
  - one dimension (depth) is lost
  - reconstruction of the ray in 3D space

$$\lambda\mathbf{x} = \begin{bmatrix}\mathsf{KR} & -\mathsf{KR}\mathbf{X}_0\end{bmatrix}\begin{bmatrix}\mathbf{X}_e \\ 1\end{bmatrix}$$

$$\mathbf{X}^e = \mathbf{X}_0 + \underline{\lambda}(\mathsf{KR})^{-1}\mathbf{x}$$

Projection matrix

- Summary
  - collinearity equations in euclidean coordinates

$$x = \frac{p_{11}X + p_{12}Y + p_{13}Z + p_{14}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}}$$

$$y = \frac{p_{21}X + p_{22}Y + p_{23}Z + p_{24}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}}$$

  - without shear or scale difference in image coordinate system
    this notation is used in older textbooks

$$\mathsf{K} = \begin{bmatrix} c & 0 & x_H \\ 0 & c & y_H \\ 0 & 0 & 1 \end{bmatrix}$$

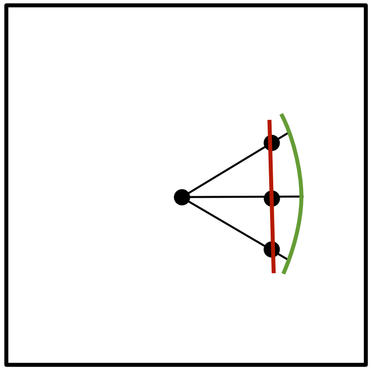
$$x = c\frac{r_{11}(X - X_0) + r_{12}(Y - Y_0) + r_{13}(Z - Z_0)}{r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)} + x_H$$

$$y = c\frac{r_{21}(X - X_0) + r_{22}(Y - Y_0) + r_{23}(Z - Z_0)}{r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)} + y_H$$

Interior Orientation

- Modeling real cameras
  - Manufacturing precision of camera lower than measurement precision
    → principal point *not* exactly in image center
  - Lens system not an ideal lens (analog film also not planar)
    → non-linear image distortions

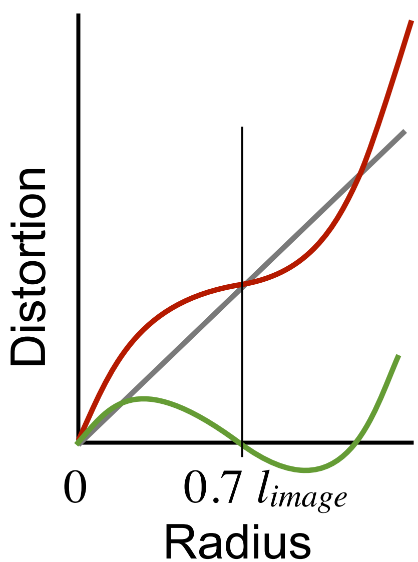
- Lens distortion
  - ray path deviates from a straight line
  - deviation depends on position of the image point
    → mapping is *no longer line-preserving*
  - modeled through a correction term:
    perspective projection + non-linear correction

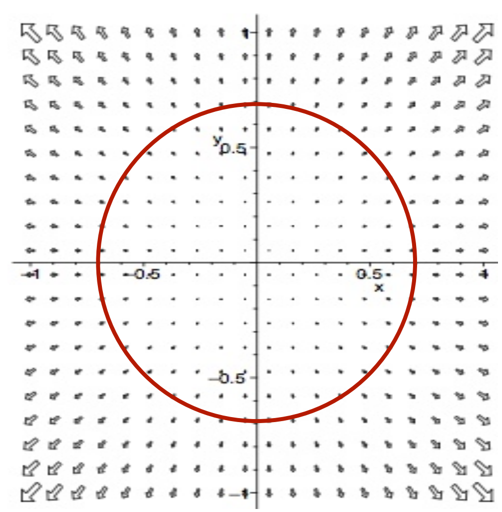


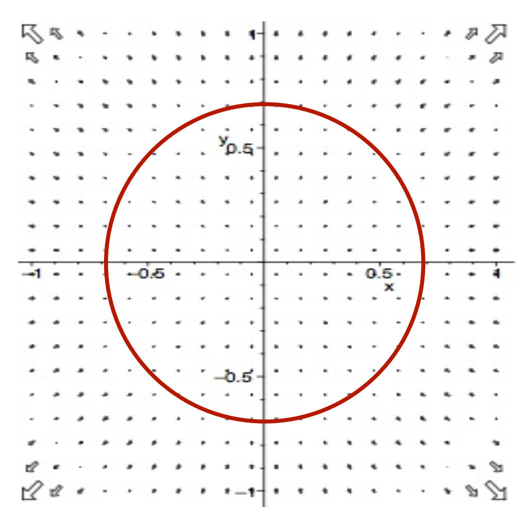
Interior Orientation

- Note: linear lens distortion
  - linear component of distortion already contained in matrix K (e.g. different scale in x and y)
  - also, a linear error in radial direction cannot be distinguished from a change of focal length
  - to determine the “focal length” we need an arbitrary definition at which radius the radiale distortion should be 0

    → camera constant is a virtual quantity







Interior Orientation

- Taking distortion into account
  - the mapping is no longer perspective, but more general

$$\begin{bmatrix} x^a \\ y^a \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \Delta x(x, y, \mathbf{q}) \\ \Delta y(x, y, \mathbf{q}) \end{bmatrix}$$

  - in homogeneous coordinates

$$\mathbf{x}^a = \begin{bmatrix} 1 & 0 & \Delta x(\mathbf{x}, \mathbf{q}) \\ 0 & 1 & \Delta y(\mathbf{x}, \mathbf{q}) \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x} = \mathsf{H}^a(\mathbf{x})\, \mathbf{x}$$

$$\mathbf{x}^a = \mathsf{H}^a(\mathbf{x})\mathsf{P}\mathbf{X} = \mathsf{P}^a(\mathbf{x})\mathbf{X}$$

Interior Orientation

- Working with general imaging equations
  - from 3D object space to 2D image space:
    distortion depends on ideal (distortion-free) image point, thus two-step computation

$$\text{Schritt 1:} \quad \mathbf{x} = \mathsf{P}\mathbf{X}$$

$$\text{Schritt 2:} \quad \mathbf{x}^a = \mathsf{H}^a(\mathbf{x})\,\mathbf{x}$$

  - from 2D image space to 3D object space:
    distortion-free image point is not accessible, thus iterative computation

$$\mathbf{x}_{(1)} = \left(\mathsf{H}^a(\mathbf{x}^a)\right)^{-1}\mathbf{x}^a$$

$$\mathbf{x}_{(t+1)} = \left(\mathsf{H}^a(\mathbf{x}_{(t)})\right)^{-1}\mathbf{x}^a$$

Interior Orientation

- Modeling lens distortion
  - physical model

    goal: describe physical process

    advantage: describes actual cause

    disadvantage: effects are complex and difficult to model, often multiple physical effects overlap

  - phenomenological model

    aim: compensate effect of distortion on the image

    advantage: can choose mathematically convenient model

    disadvantage: causes remain unknown

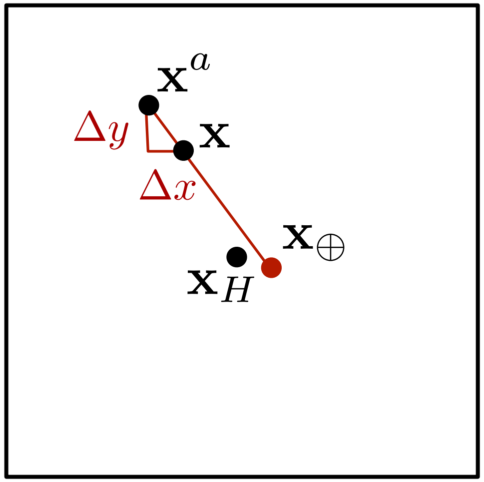

Interior Orientation

- physically motivated distortion model
  - example: radially symmetric distortion

$$\Delta \mathbf{x}^e = \frac{\mathbf{x}^e - \mathbf{x}^e_{\oplus}}{r}(q_2 r^2 + q_4 r^4 + q_6 r^6)$$

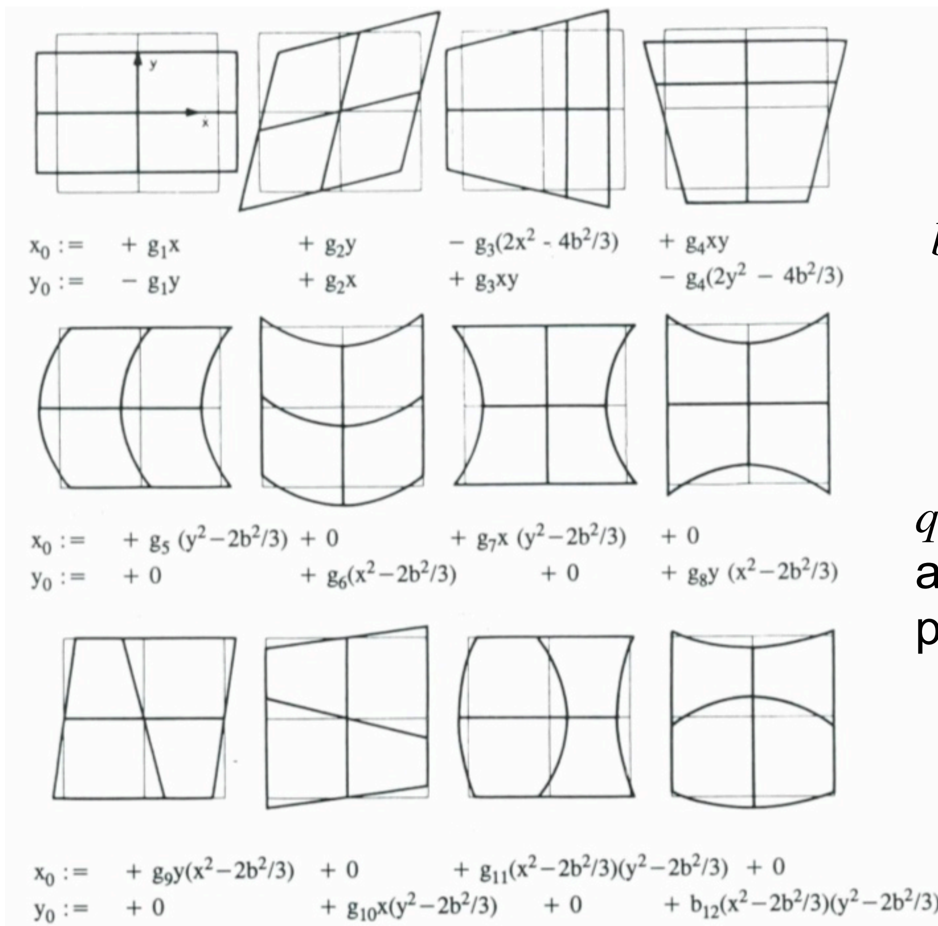
$$r = \sqrt{(x - x_{\oplus})^2 + (y - y_{\oplus})^2}$$

  - $\mathbf{x}_{\oplus}$ is the radial symmetry center of the distortion (near the image center)



Interior Orientation

- phenomenological distortion model
  - example: 10 orthogonal polynomials due to Ebner (originally 12, of which 2 linear ones cover affine distortion of image coordinates)



b: normalisation factor (largest image coordinate)

q_i: (almost) orthogonal mutually and w.r.t. remaining orientation parameters