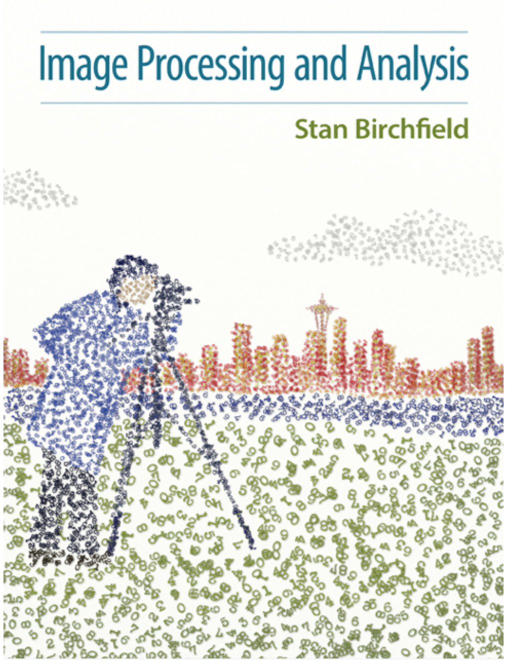
Prof. Kjersti Engan

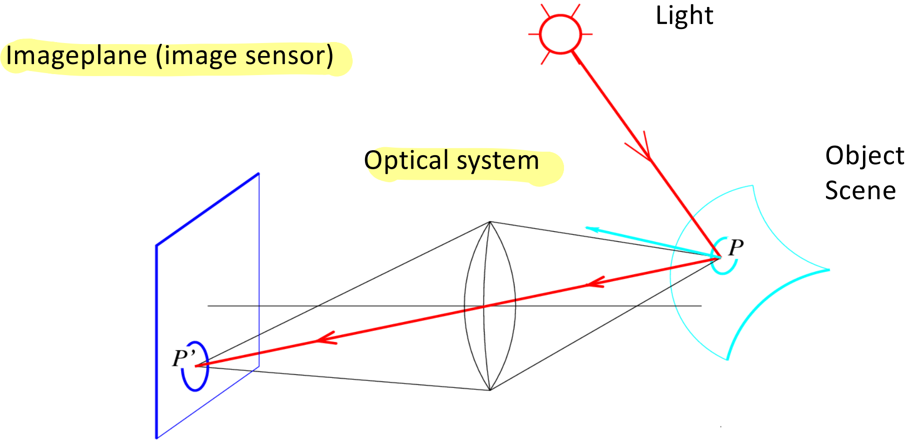
ELE510 Image processing and computer vision

Camera and calibration, (chap 13.4, 13.5 Birchfield) 2020



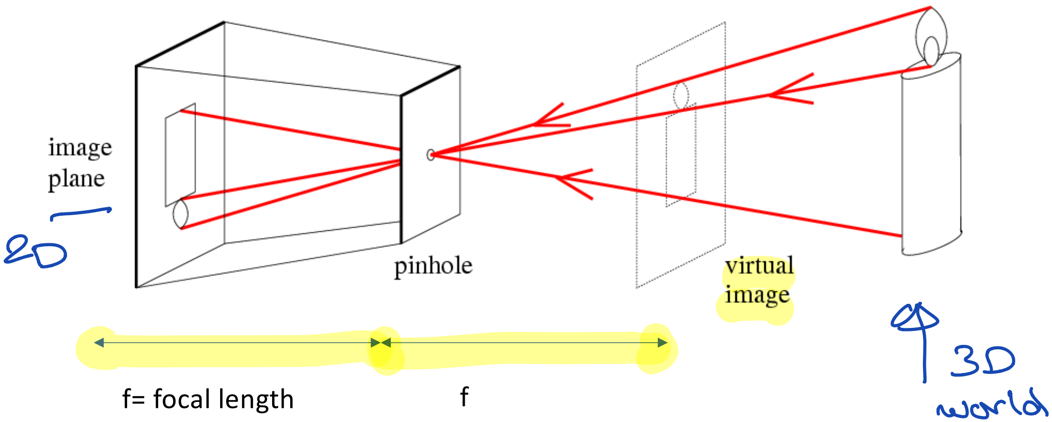
Parts in course presentations is material from Cengage learning. It can be used for teaching, and it can be share it with students on access controlled web-sites (Canvas) for use in THIS course. **But it should not be copied or distributed further in any way** (by you or anybody).

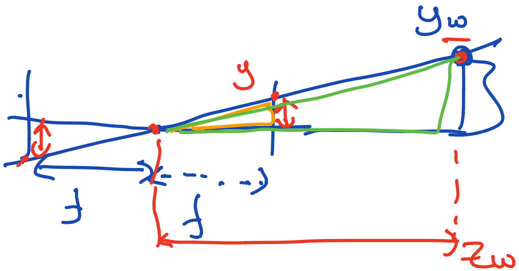
Image Formation



Geometric **Optical** **Photometric**

Pinhole camera - revisited





$$\frac{y}{f} = \frac{y_w}{z_w} \Rightarrow y = f \cdot \frac{y_w}{z_w}$$

$$x = f \cdot \frac{x_w}{z_w}$$

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{f}{z_w} & 0 & 0 & 0 \\ 0 & \frac{f}{z_w} & 0 & 0 \\ 0 & 0 & \frac{1}{z_w} & 0 \end{bmatrix} \cdot \overbrace{\begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}}^{P^w}$$

we dont know z_w!

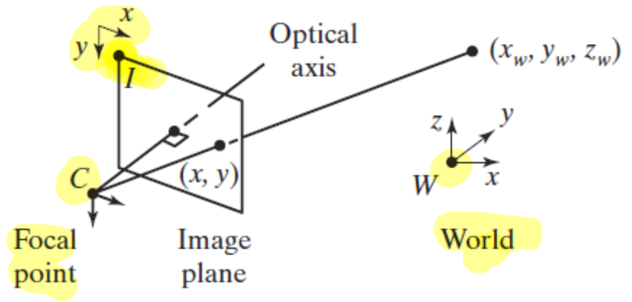
$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \underbrace{\frac{1}{z_w}}_{\text{dont know!}} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}}_{m.} \underbrace{\propto}_{\text{proportional to}} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \underbrace{\begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}}_{m}$$

Perspective Imaging

- Three coordinate systems are involved in the process of imaging:
  - one attached to the world
  - one attached to the camera
  - one associated with the image

- Points in the world are described in the *world coordinate system*, in which lengths are measured in meters.



- The camera coordinate system is centered at the focal point, x and y axes aligned with the rightward horizontal and downward vertical directions, respectively, of the image plane, and the positive z axis points along the optical axis toward the world.

- The *image coordinate system* is centered at the top left corner of the image, with the positive x and y axes pointing along the rows and columns, respectively, of the imaging sensor. In the image coordinate system, measurements are made in pixels.

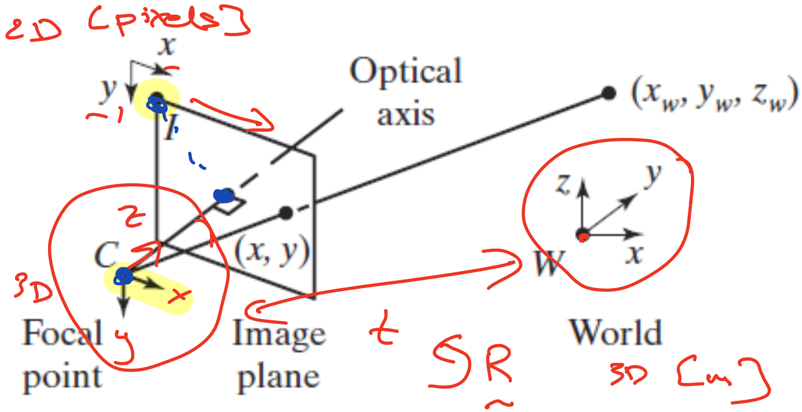


Figure 13.24 The projection of a world point (x_w, y_w, z_w) onto an image plane at point (x, y), assuming a pinhole camera model with no diffraction. The three coordinate systems are the world coordinate system (W), the camera coordinate system (C), and the image coordinate system (I).

Perspective Imaging

- Using homogeneous coordinates, the imaging process is captured mathematically as:

$$\underset{\substack{2D\\ [\text{pixels}]}}{\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}} \propto \underbrace{\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix}}_{\mathbf{P}_{\{3\times 4\}}} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad 3D \quad [m] \quad \boxed{z_w}$$

- Where **P** is a 3 X 4 projection matrix that itself is composed of two parts:

$$\mathbf{P}_{\{3\times 4\}} = \underbrace{\begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{K}_{\{3\times 3\}}} [\mathbf{R}_{\{3\times 3\}} \ \ \mathbf{t}_{\{3\times 1\}}]$$

- **Intrinsic** (camera dependent) parameters are found in **K.** Converts meters to pixels, moves the origo in the image plane etc.
- **Extrinsic** (how do we describe the world relative to the camera) parameters in **R** and **t**

Intrinsic Camera parameters

$$\underbrace{\begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{K}_{\{3\times 3\}}}$$

- u0 and v0 specifies the principal point, i.e. The intersection of the optical axis and the image plane.
- α, β, ɣ are related to the focal length in x direction (fx), in y direction (fy) and the skew θ between x and y axis in the image plane.

x

y

$\frac{\pi}{2}$

$$\boxed{\begin{aligned} \alpha &= f_x \\ \beta &= f_y/\sin(\theta) \\ \gamma &= -f_x/\tan(\theta) \end{aligned}}$$

$f_x = f/\Delta x$, $f_y = f/\Delta y$ Often: $f_x \approx f_y$ (tolerance of 5%)

$$\mathbf{K} = \begin{pmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

Often we can assume θ≈(π/2) :
you would need to account for axis skew when calibrating unusual cameras or cameras taking photographs of photographs, else you can usually ignore the skew parameter.

Lens Distortion

- Real cameras have lenses. Light bends due to curvature in the lens.
- The dominant distortion of a typical lens is **radial distortion**

It is a function only of the radial distance from the center of the image.

- Let (x_u, y_u) be the undistorted coordinates of a pixel in the image:

$$x_u = x_d + \bar{x}_d f(r_d)$$

$$y_u = y_d + \bar{y}_d f(r_d)$$

Where (r is the radial distance between the pixel and the undistorted point):

$$f(r_d) = k_1 r_d^2 + k_2 r_d^4 + k_3 r_d^6 + \cdots$$

- If a more accurate model is desired, **tangential distortion** (or *decentering distortion*) can be included

- The coefficients have to be estimated, thereafter the image can be unwarped

Extrinsic parameters

- Rotation and translation of the World coordinates with respect to the Camera coordinates:

$$\mathbf{R} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \mathbf{T} = \begin{pmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

3x3 3x1

$$\mathbf{TR} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \mathcal{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{pmatrix} \qquad \boxed{{}^{C}\mathbf{P} = \mathbf{TR}\,{}^{W}\mathbf{P}}$$

$${}^{w}P \quad \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$