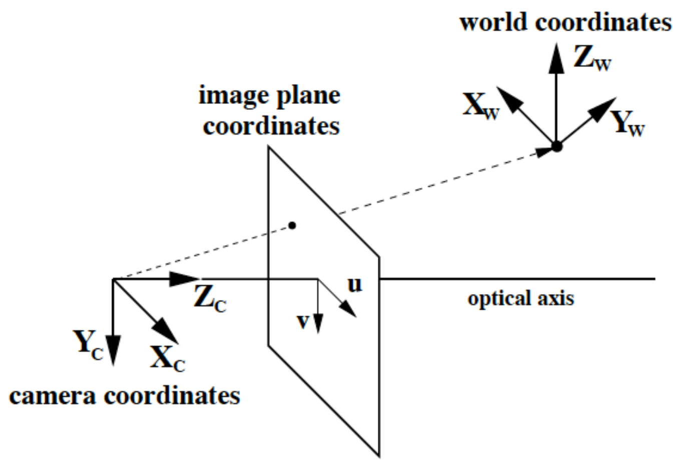
Camera Calibration

Kostas Daniilidis

Camera model



world coordinate system coordinates (X_w, Y_w, Z_w),

camera coordinate system coordinates (X_c, Y_c, Z_c).

image homogeneous coordinates (u, v). The optical axis is the z-axis and the image plane is perpendicular to the optical axis. Intersection of the image plane with the optical axis is the point (u_o, v_o).

f is the distance of the image plane from the origin (effective focal length f in pixels).

$$u - u_o = f\frac{X_c}{Z_c} \quad v - v_o = f\frac{Y_c}{Z_c}.$$

We use s_x and s_y instead of f to capture different sizes of each pixel in mm (and not only): $s_x = f/d_x, s_y = f/d_y$.

In terms of projective geometry

Projection from world to the image plane is a singular tranformation from $\mathbb{P}^3$ to $\mathbb{P}^2$:

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} \sim \begin{pmatrix} s_x & 0 & u_o \\ 0 & s_y & v_o \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} R & t \end{pmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ W_w \end{pmatrix} = \begin{pmatrix} \tilde{M} & \tilde{m} \end{pmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ W_w \end{pmatrix}$$

Image point homogeneous coordinates

Intrinsic parameters

Extrinsic parameters

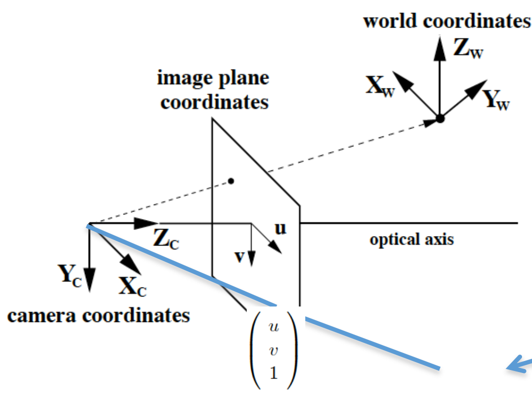
3x4 Camera Matrix

World point homogeneous coordinates

Intrinsic calibration means finding intrinsic parameters
Extrinsic calibration means finding transformation between world coordinate frame and camera coordinate frame. World is defines on a reference pattern like a checkerboard.

What does calibration really mean?

It means that we can transform PIXELS to RAYS !



$$\lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \tilde{M} \begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} + \tilde{m}$$

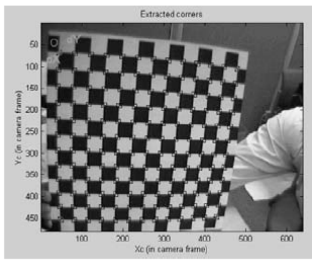
hence

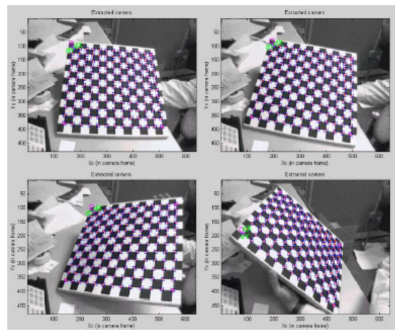
$$\begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} = \lambda \tilde{M}^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} - \tilde{M}^{-1}\tilde{m}.$$

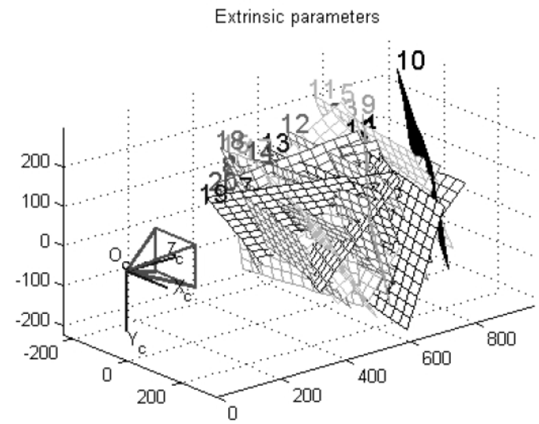
Ray equation

We observe that if we know camera matrix $\begin{pmatrix} \tilde{M} & \tilde{m} \end{pmatrix}$ then we can find the equation of the ray through origin and pixel point (u,v)

Camera Calibration Toolbox





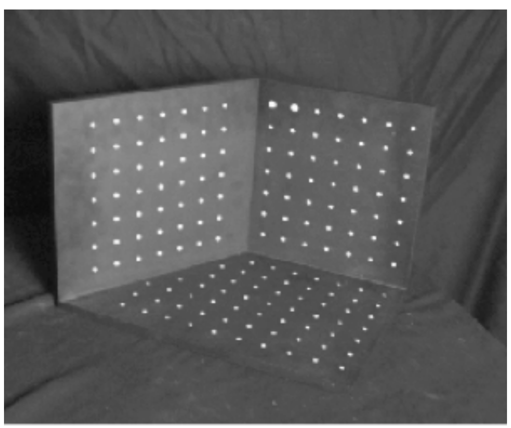


(Bouguet 1997) MATLAB toolbox (and OpenCV similar):
Waiving a checkerboard in front of the camera. Origin of world is on the first pose of the checkerboard.

Principle (Zhang 1993-5): Extract collineations between image plane and checkerboard at each pose (see next page)

In the old days…

(Faugeras' book 93) researchers were using 3D reference cubes to compute M matrix from one shot. At least six point correspondences are needed to compute M:



Estimation of M is linear but M has to be decomposed in K, R, T

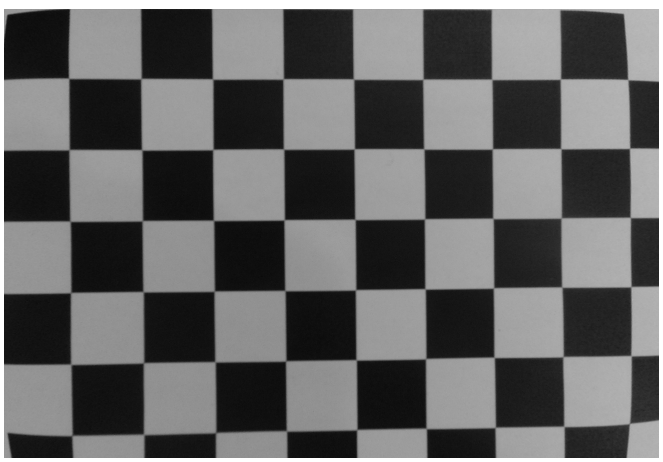
$$\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} \cong \begin{bmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix}$$

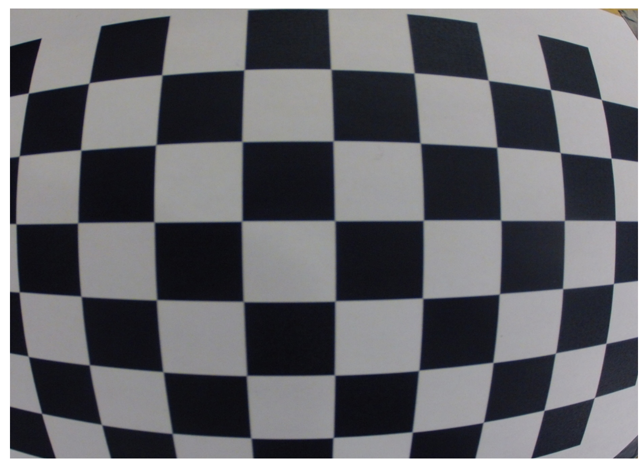
$$\begin{aligned} u_i &= \frac{m_{00}X_i + m_{01}Y_i + m_{02}Z_i + m_{03}}{m_{20}X_i + m_{21}Y_i + m_{22}Z_i + m_{23}} \\ v_i &= \frac{m_{10}X_i + m_{11}Y_i + m_{12}Z_i + m_{13}}{m_{20}X_i + m_{21}Y_i + m_{22}Z_i + m_{23}} \end{aligned}$$

$$\begin{bmatrix} X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & 0 & -u_iX_i & -u_iY_i & -u_iZ_i & -u_i \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & -v_iX_i & -v_iY_i & -v_iZ_i & -v_i \end{bmatrix} \begin{bmatrix} m_{00} \\ m_{01} \\ m_{02} \\ m_{03} \\ m_{10} \\ m_{11} \\ m_{12} \\ m_{13} \\ m_{20} \\ m_{21} \\ m_{22} \\ m_{23} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Cameras with large field of view have radial distortions

$$
\begin{aligned}
u^{dist} &= u(1 + k_1 r + k_2 r^2 + k_3 r^3 + \ldots) \\
v^{dist} &= v(1 + k_1 r + k_2 r^2 + k_3 r^3 + \ldots) \\
\text{where} \quad r^2 &= u^2 + v^2
\end{aligned}
$$

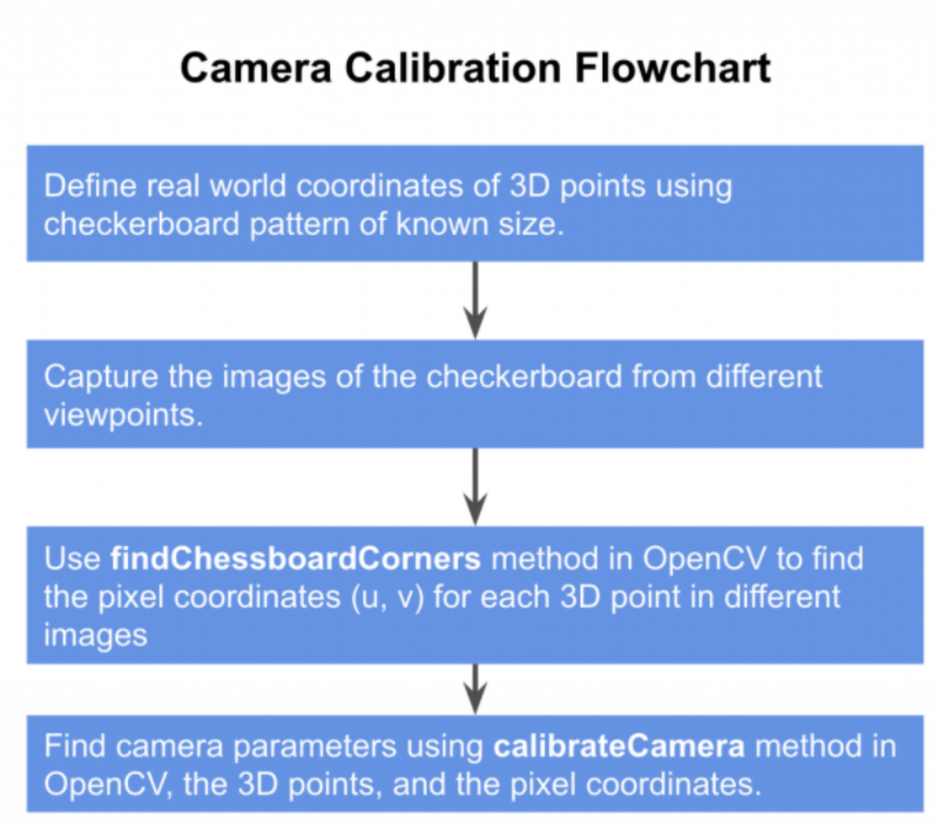




A procedure called calibration

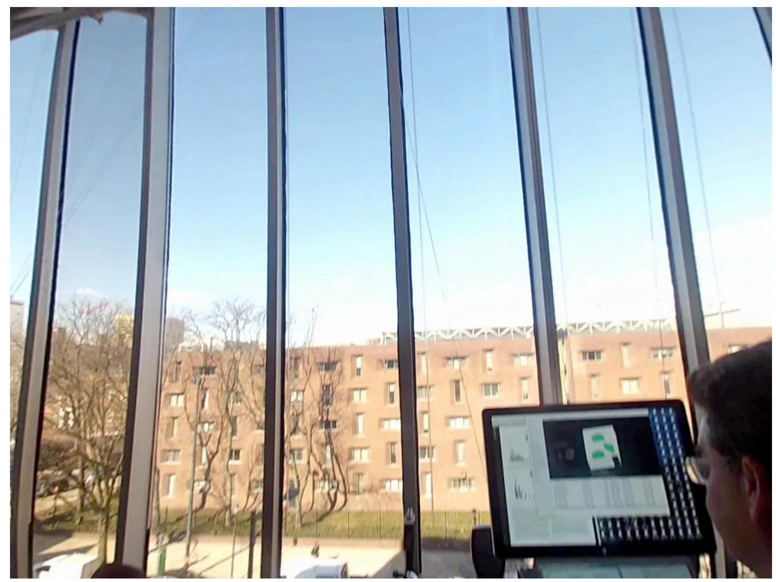
Estimates the *intrinsic parameters*

- f focal length
- (u_o, v_o) image center
- $k_1, k_2, \ldots$ radial distortion parameters

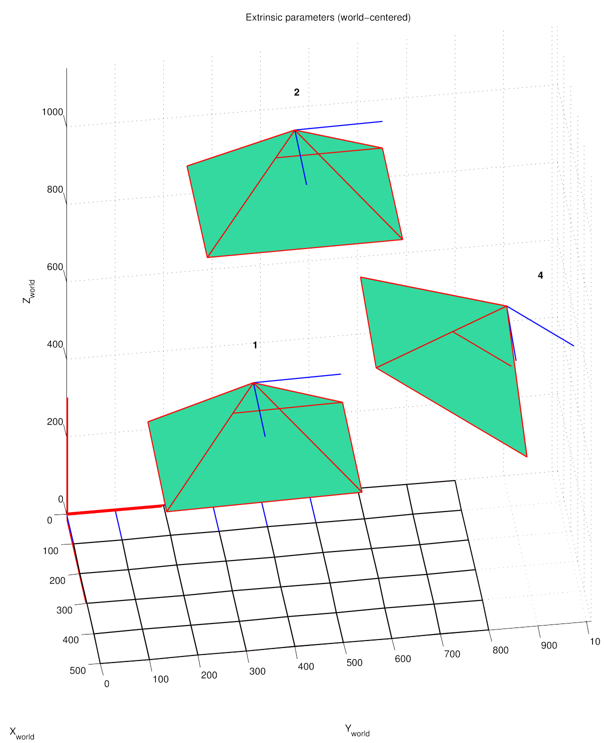


https://learnopencv.com/camera-calibration-using-opencv/

As a result of the calibration we have undistorted images and video



..as well as the poses of the camera and the projection rays in world coordinates



$$\begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} = \boxed{-R^T T} + \lambda \boxed{R^T K^{-1}} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

known

We will return later on the specifics of calibration….