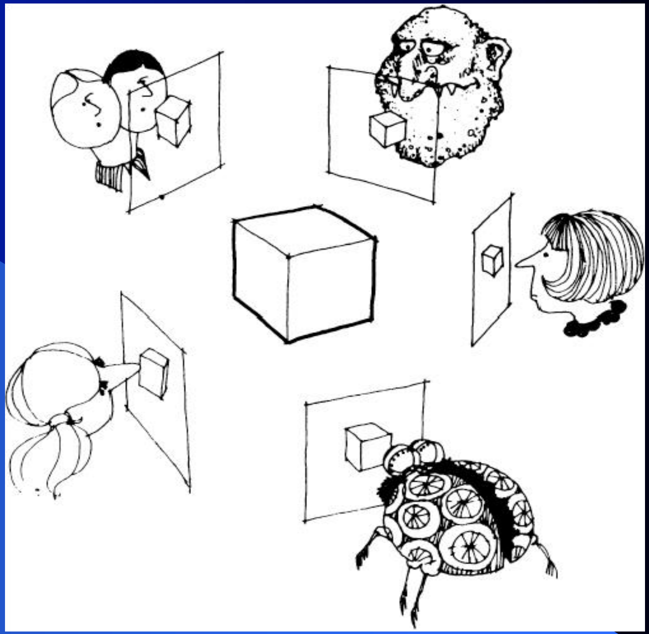
Computer Vision

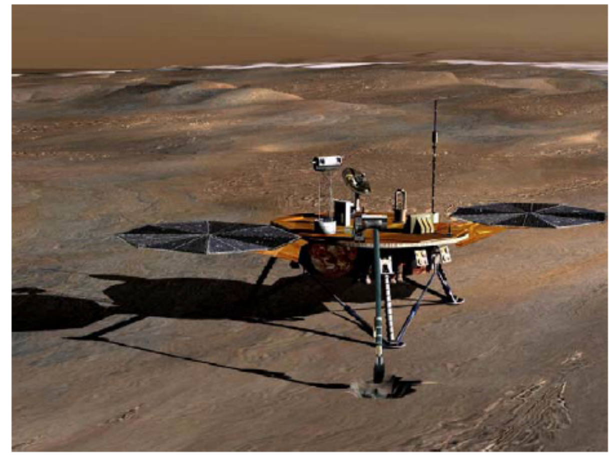
Lecture X

Multiview Stereo



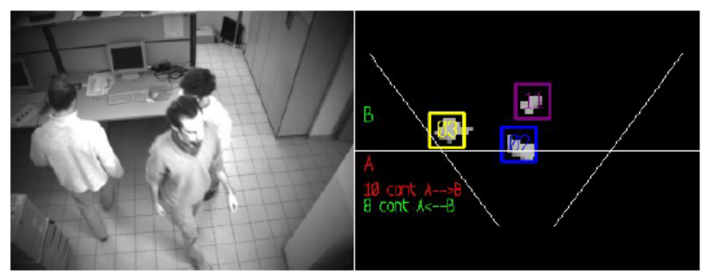
Applications

www.nasa.gov



www.nasa.gov

www.vislab.it



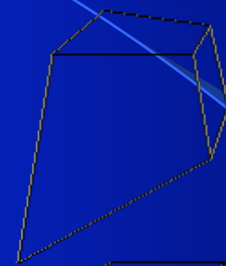
www.vision.deis.unibo.it/smatt/stereo

Stefano Mattoccia

3D Transformations

Projective
15dof

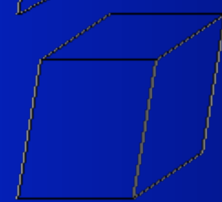
$$\begin{bmatrix} A & t \\ v^T & v \end{bmatrix}$$



Intersection and tangency

Affine
12dof

$$\begin{bmatrix} A & t \\ 0^T & 1 \end{bmatrix}$$



Parallelism of planes,
Volume ratios, centroids,
The plane at infinity π_∞

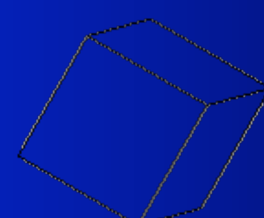
Similarity
7dof

$$\begin{bmatrix} sR & t \\ 0^T & 1 \end{bmatrix}$$

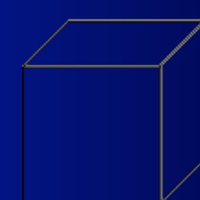
Angles, ratios of length
The absolute conic Ω_∞

Euclidean
6dof

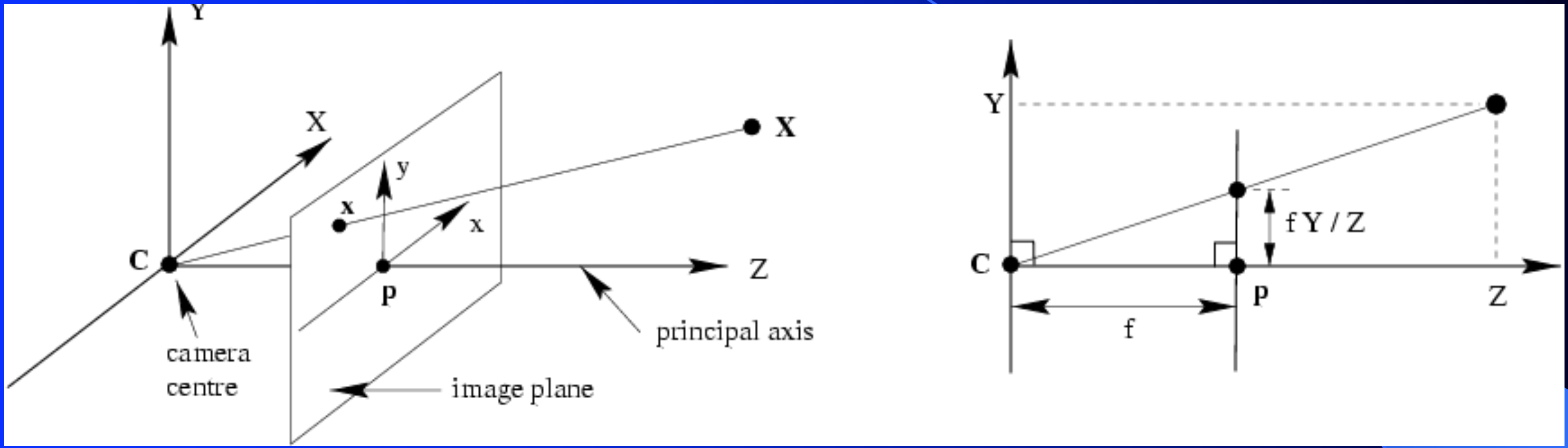
$$\begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix}$$



Volume



Pinhole Camera Model



non-homogeneous coordinates

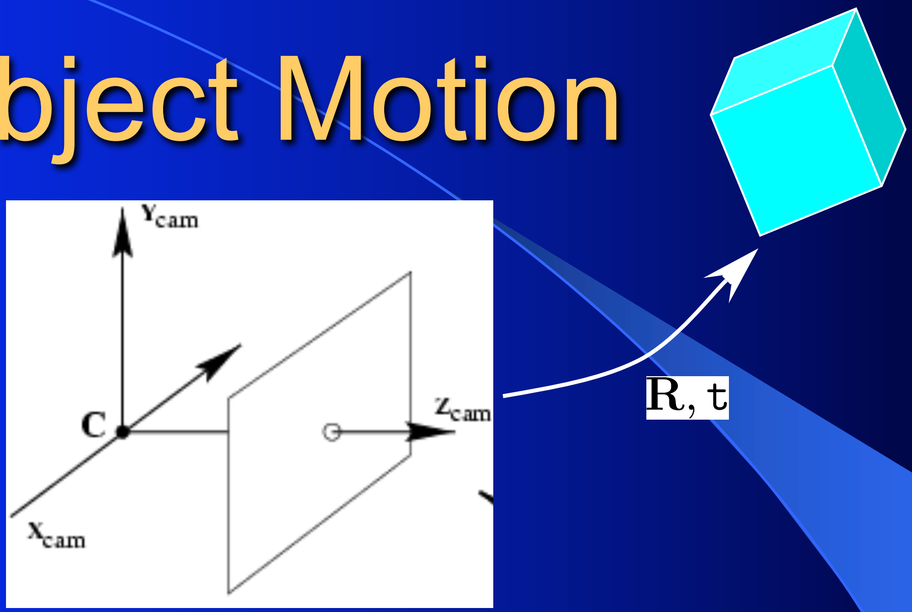
$$(X, Y, Z)^T \mapsto (fX/Z, fY/Z)^T$$

homogeneous coordinates

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

Object Motion

$$K = \begin{bmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

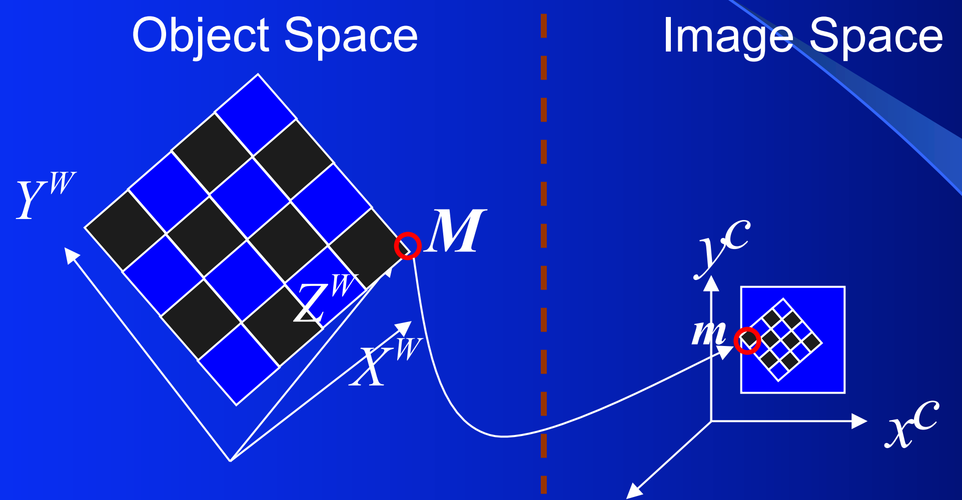


$$\mathtt{x} = \mathbf{K}\begin{bmatrix} \mathbf{I} & \mathtt{0} \end{bmatrix} \mathtt{X}_{\mathtt{cam}} \qquad \mathtt{X}_{\mathtt{cam}} = \begin{bmatrix} \mathbf{R} & \mathtt{t} \\ \mathtt{0} & \mathtt{1} \end{bmatrix} \mathtt{X}_{\mathtt{obj}}$$

$$\mathtt{x} = \mathbf{K}\begin{bmatrix} \mathbf{I} & \mathtt{0} \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathtt{t} \\ \mathtt{0} & \mathtt{1} \end{bmatrix} \mathtt{X}_{\mathtt{obj}}$$

$$= \mathbf{K}\begin{bmatrix} \mathbf{R} & \mathtt{t} \end{bmatrix} \mathtt{X}_{\mathtt{obj}}$$

Camera Calibration



$$\boldsymbol{m} = [\text{Camera Projection Matrix}]\ \boldsymbol{M}$$

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

A = [$\underbrace{\textbf{R t}}$]

intrinsic camera parameters

extrinsic camera parameters

Binocular Stereo

- Given a calibrated binocular stereo pair, fuse it to produce a depth image

image 1

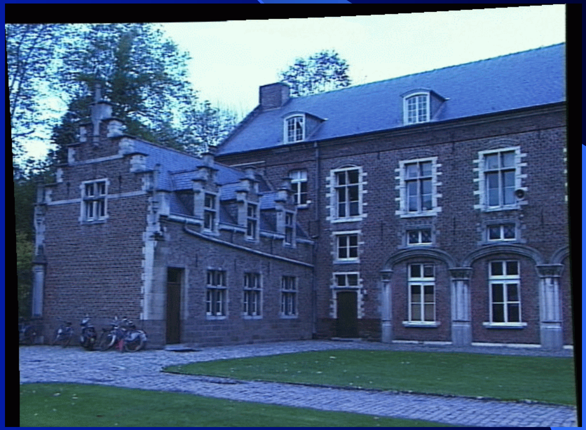
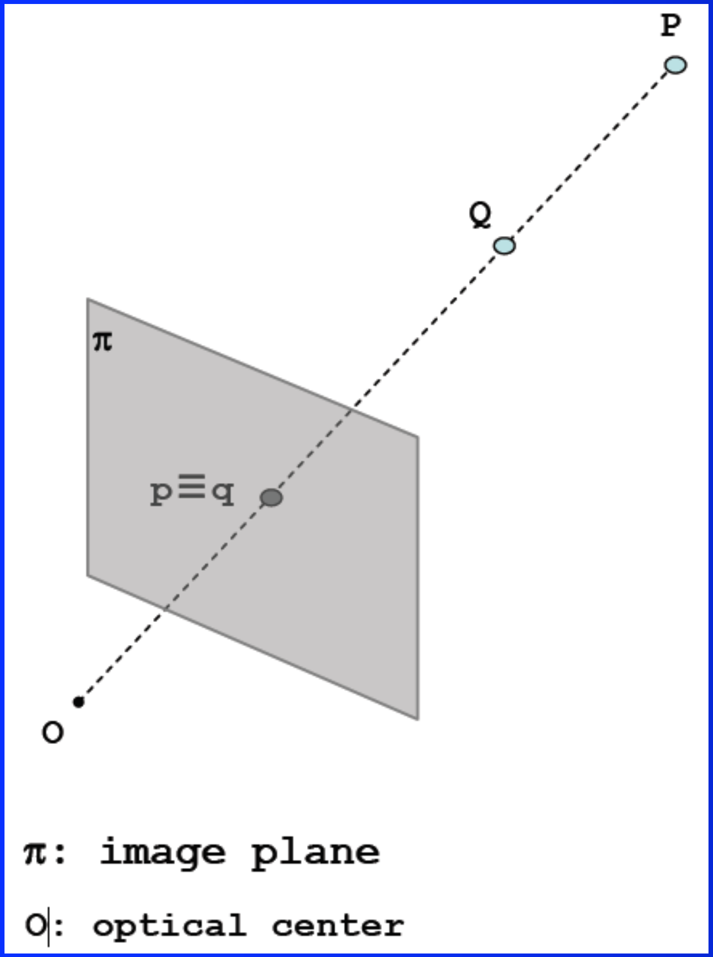


image 2

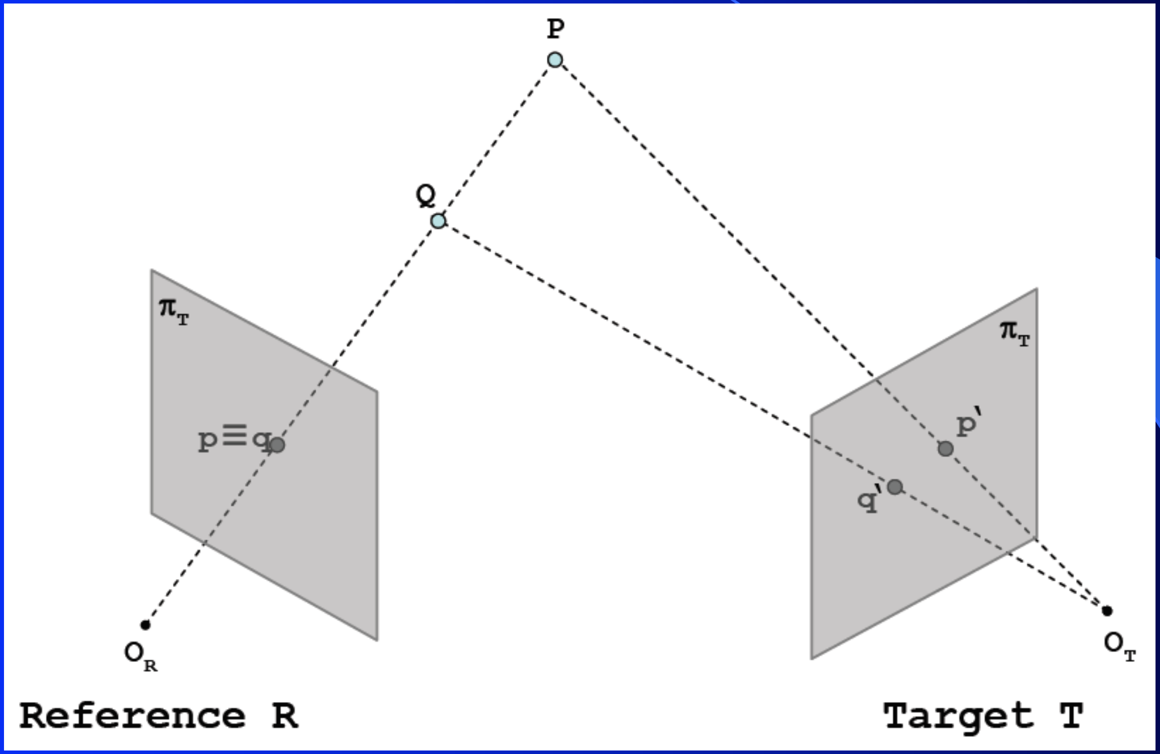
dense depth map

Single Camera



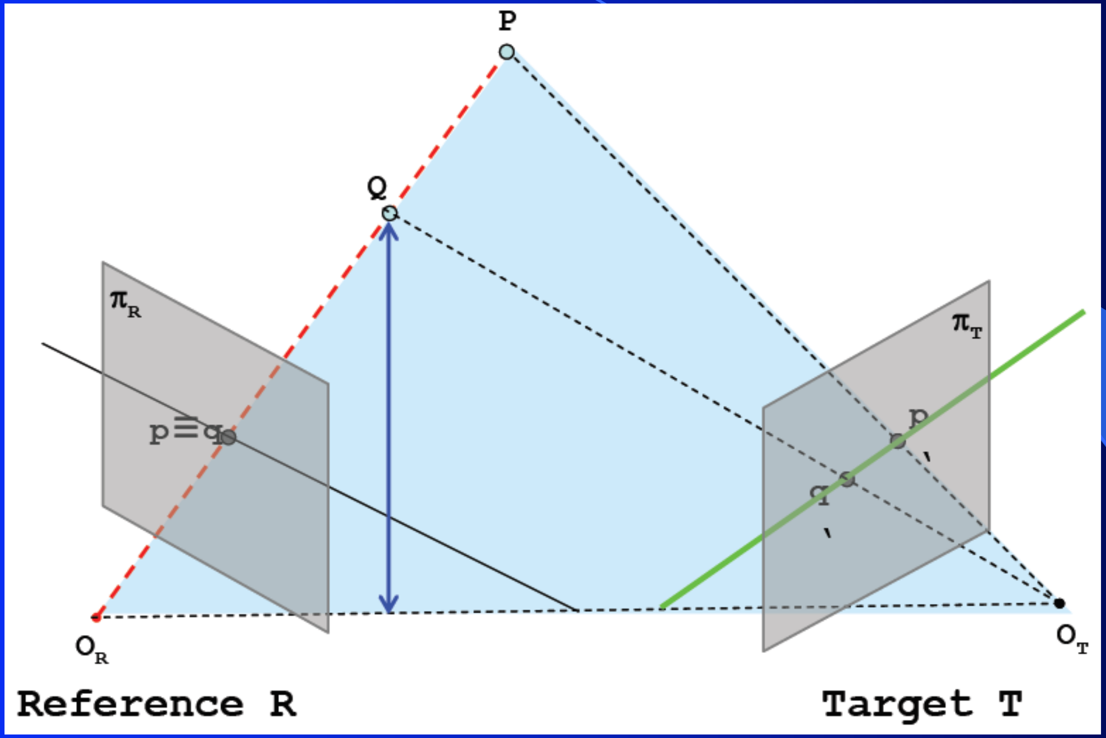
- Both (real) points (P and Q) project into the same image point (p ≡q)
- This occurs for each point along the same line of sight
- Useful for optical illusions…

Stereo Camera



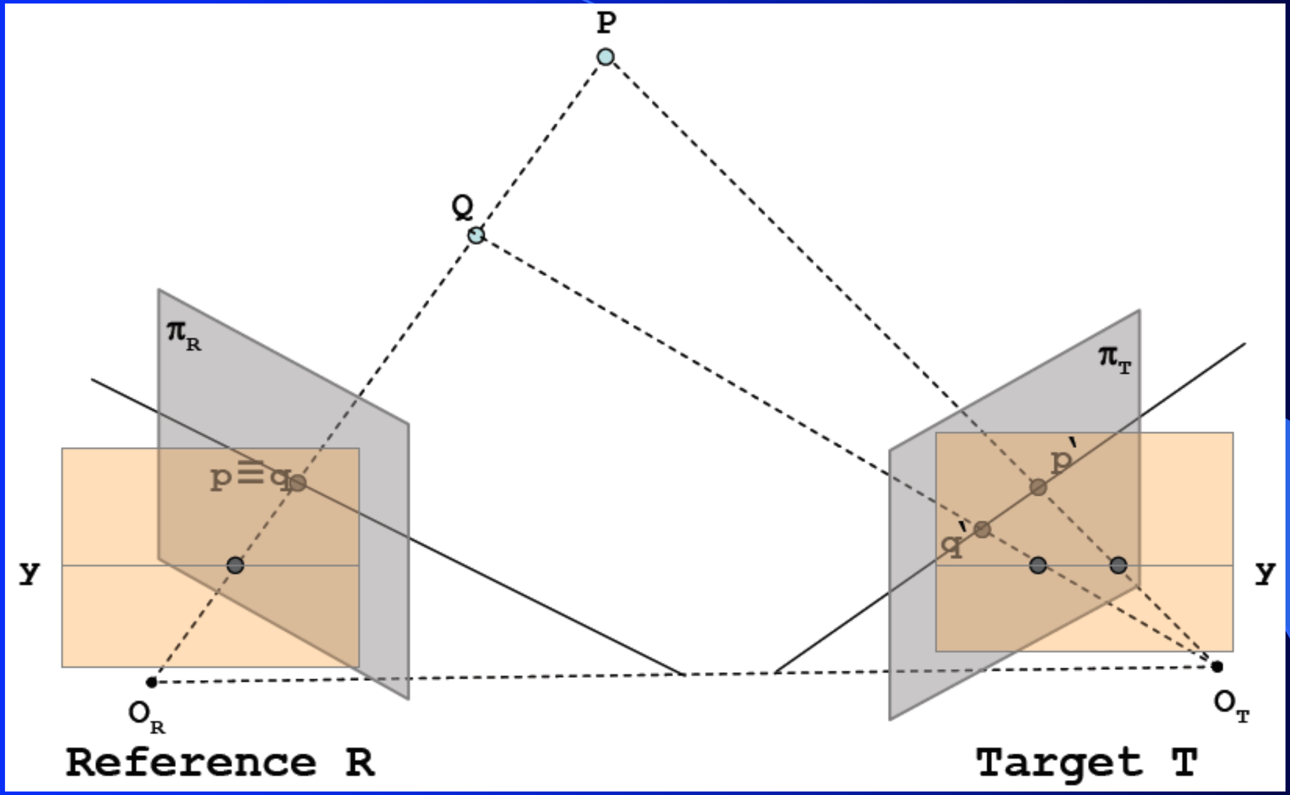
With two (or more) cameras we can infer depth, by means of triangulation, if we are able to find corresponding (homologous) points in the two images

Epipolar Constraint



- Consider two points P and Q on the same **line of sight of the reference image R** (both points project into the same image point p≡q on image plane π_R of the reference image)
- The epipolar constraint states that the correspondence for a point belonging to the (red) line of sight lies on the **green line on image plane π_T of target image**

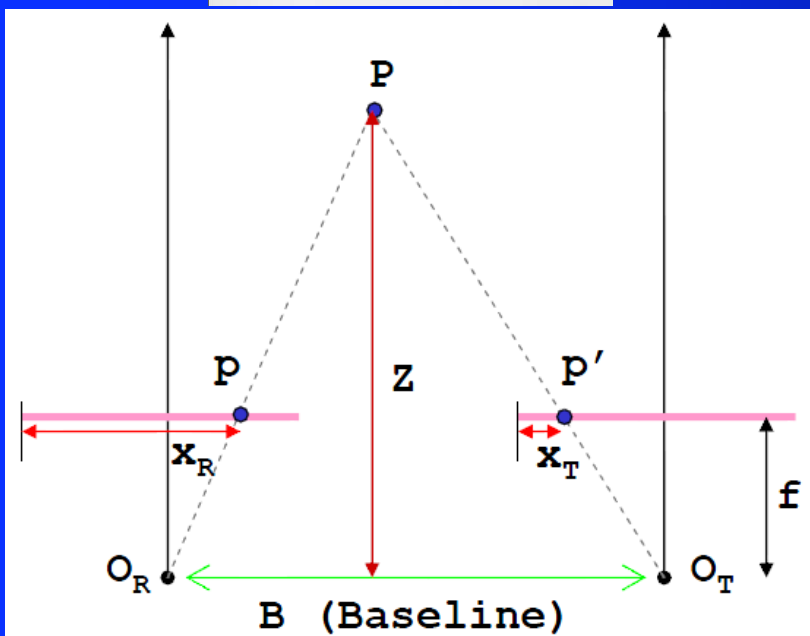
Stereo Camera in Standard Form



Once we know that the search space for corresponding points can be narrowed from 2D to 1D, we can put (virtually) the stereo rig in a more convenient configuration (standard form) - corresponding points are constrained on the same image scanline

Binocular Stereo Cameras

Disparity and Depth



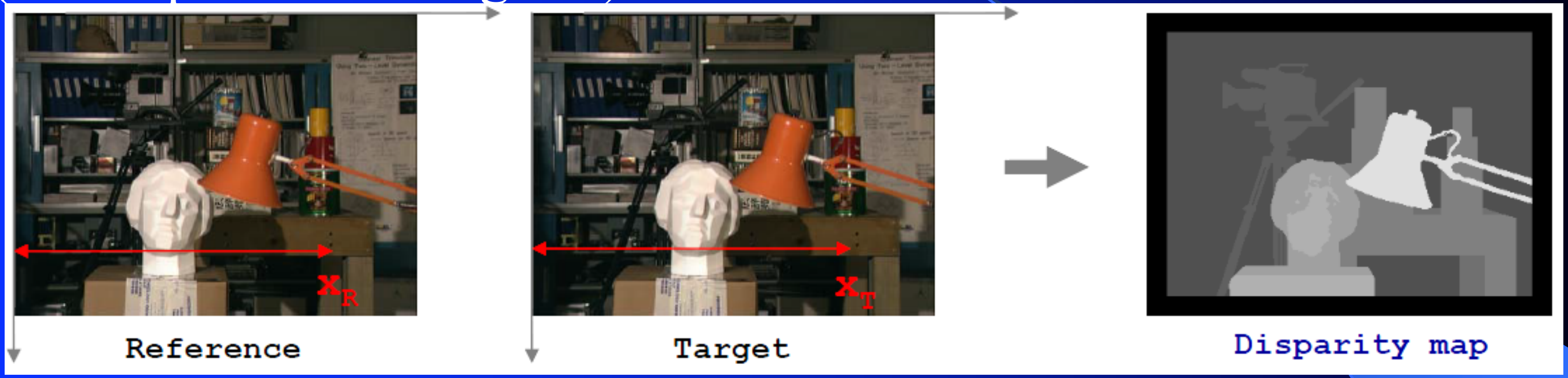
With the stereo rig in standard form and by considering similar triangles (PO_RO_T and Ppp’):

$$\frac{B}{Z} = \frac{(B + x_T) - x_R}{Z - f} \Rightarrow Z = \frac{B \cdot f}{x_R - x_T}$$

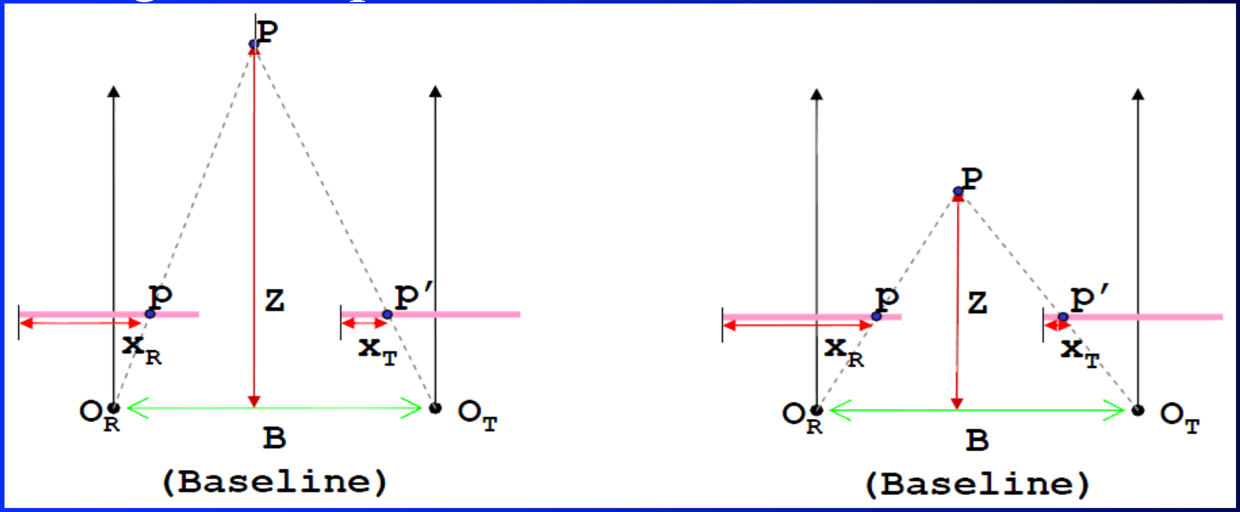
$x_R - x_T$ is the disparity

Disparity and Depth

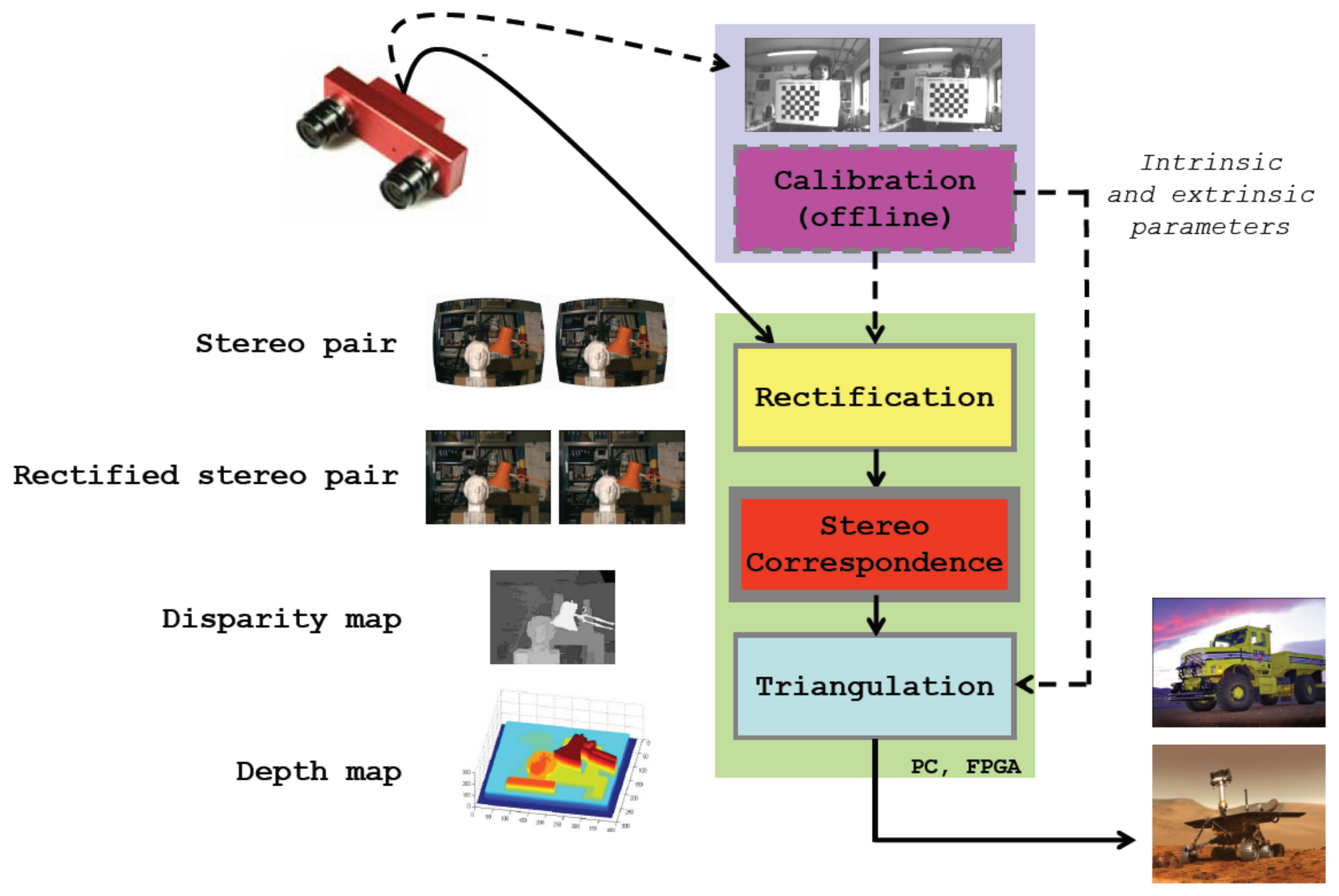
The disparity is the difference between the x coordinate of two corresponding points; it is typically encoded with grayscale image (closer points are brighter).



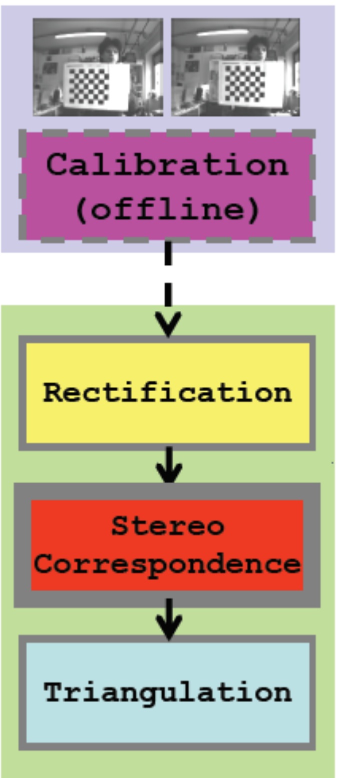
Disparity is higher for points closer to the camera



Stereo Vision System

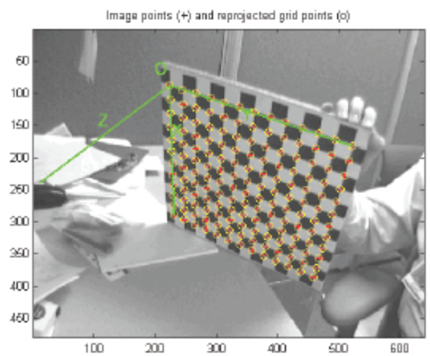


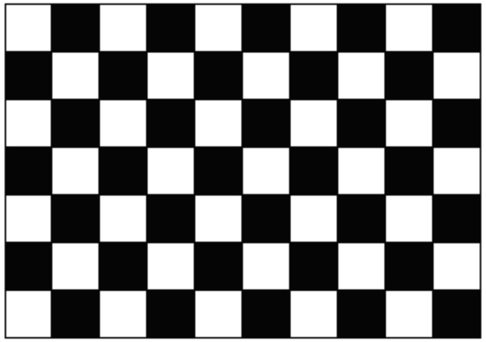
Calibration



Offline procedure aimed at finding:

- Intrinsic parameters of the two cameras (focal length, image center, parameters of lenses distortion, etc)
- Extrinsic parameters (R and T that aligns the two cameras)





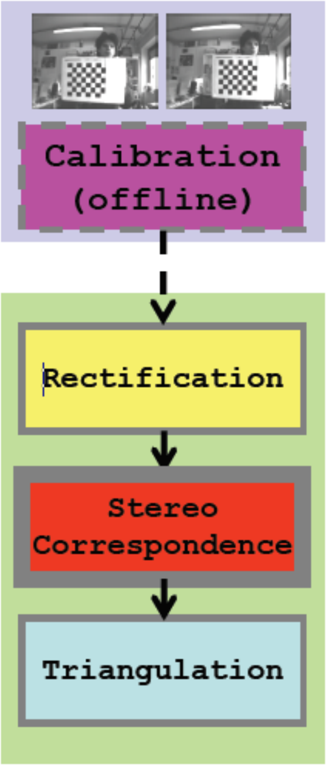
Calibration is carried out acquiring and processing 10+ stereo pairs of a known pattern (typically a checkerboard)

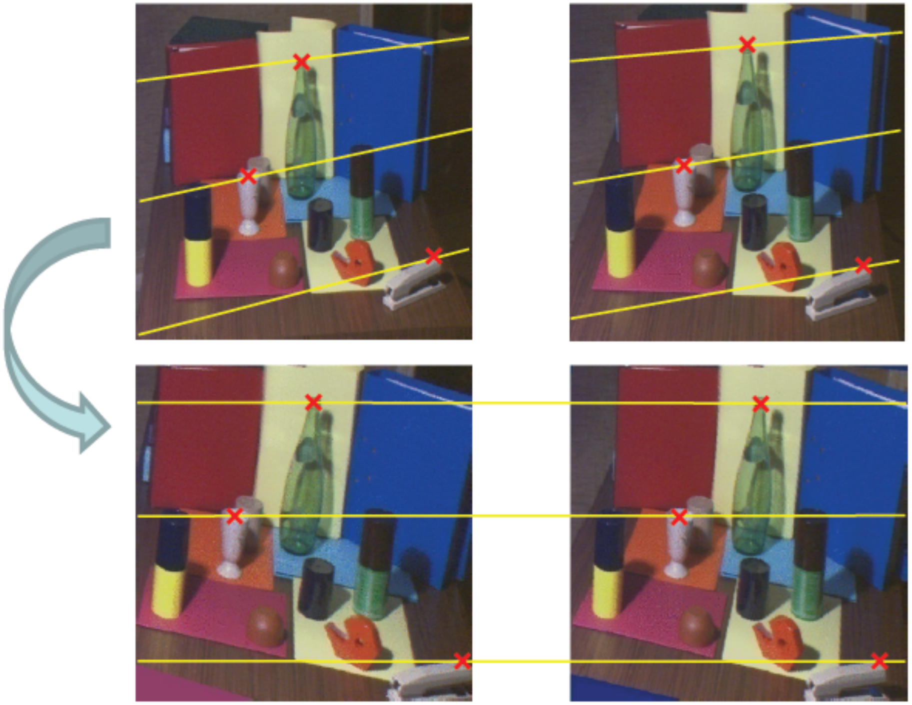
Rectification

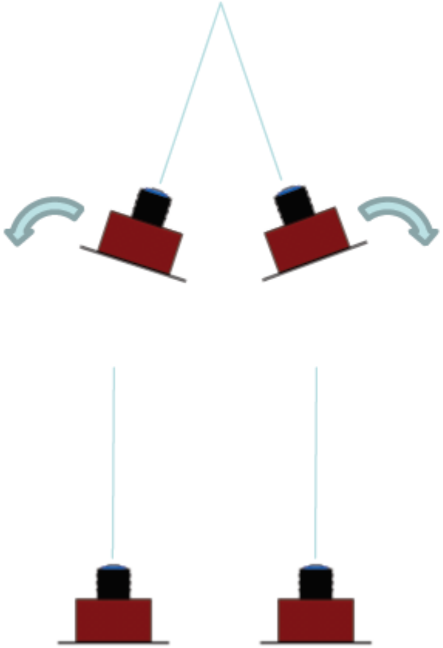
Using the information from the calibration step:

a) removes lens distortions

b) turns the stereo pair in standard form



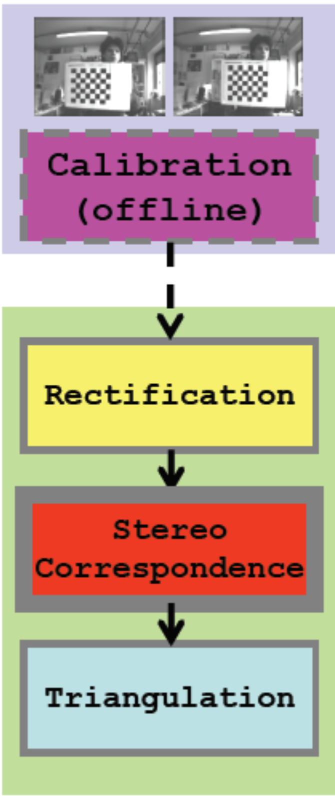


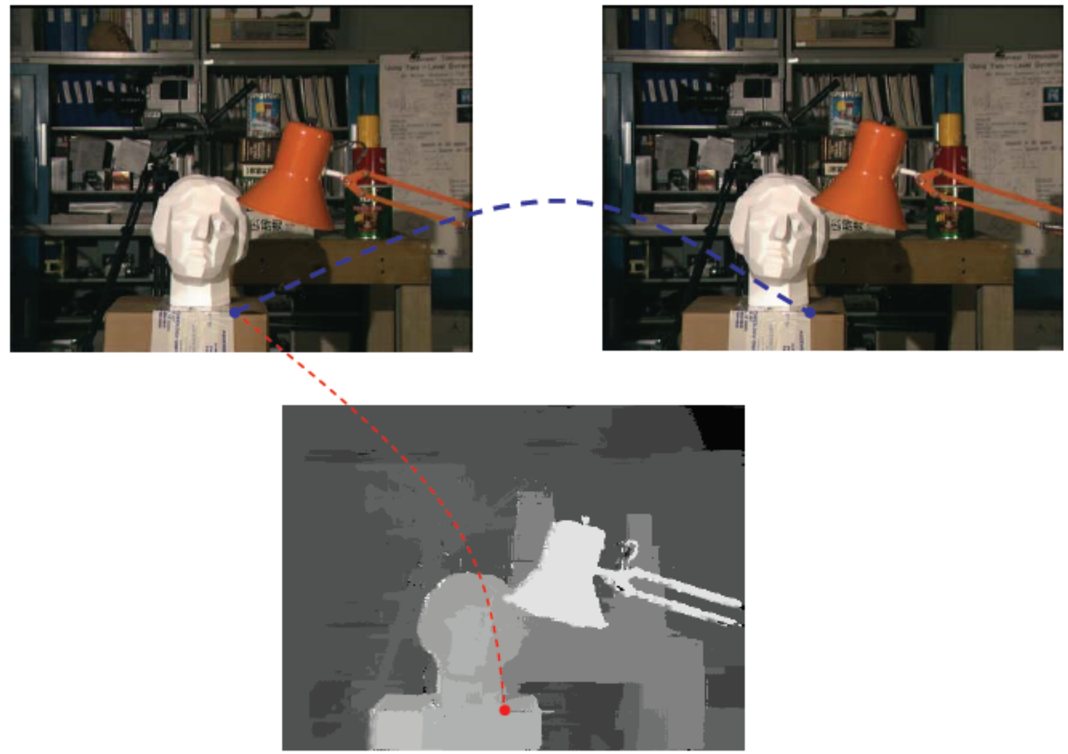


Stereo camera in standard form

Stereo Correspondence

Aims at finding homologous points in the stereo pair.



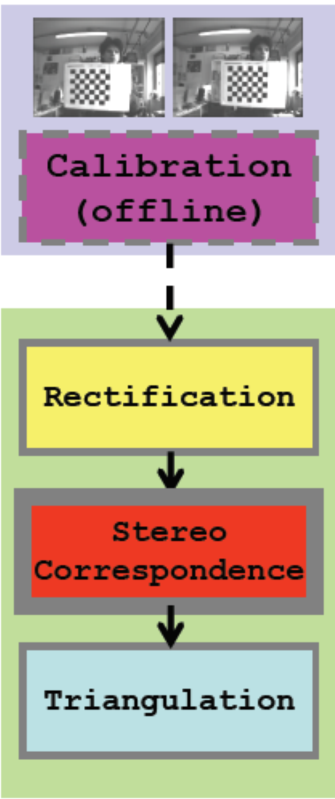


disparity map

This topic will be extensively analyzed in the next slides...

Triangulation

Given the disparity map, the baseline and the Focal length (calibration): triangulation computes the position of the correspondence in the 3D space

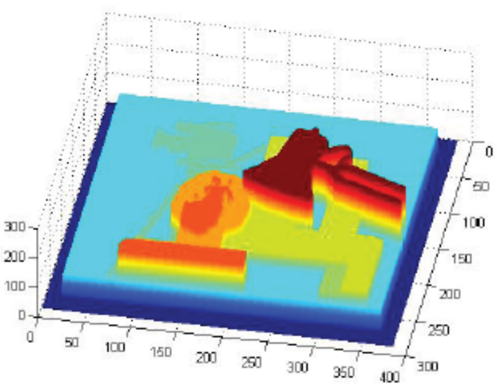


disparity map

$$Z = \frac{b \cdot f}{d}$$

$$X = Z\frac{x_R}{f}$$

$$Y = Z\frac{y_R}{f}$$



depth map

Undistort Image (OpenCV)

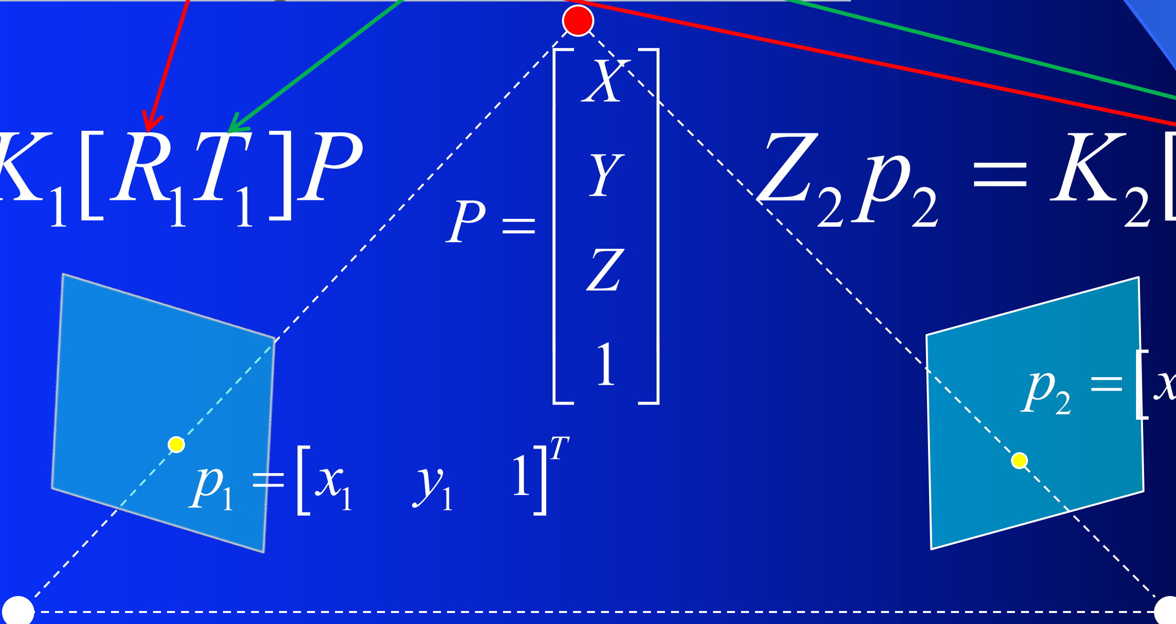
```
void undistort( const Mat& src, Mat& dst, const Mat& cameraMatrix,
          const Mat& distCoeffs, const Mat& newCameraMatrix=Mat() );
```

Solve PnP

```
void solvePnP( const Mat& objectPoints,
               const Mat& imagePoints,
               const Mat& cameraMatrix,
               const Mat& distCoeffs,
               Mat& rvec, Mat& tvec,
               bool useExtrinsicGuess=false );
```

Finds the object pose from the 3D-2D point correspondences

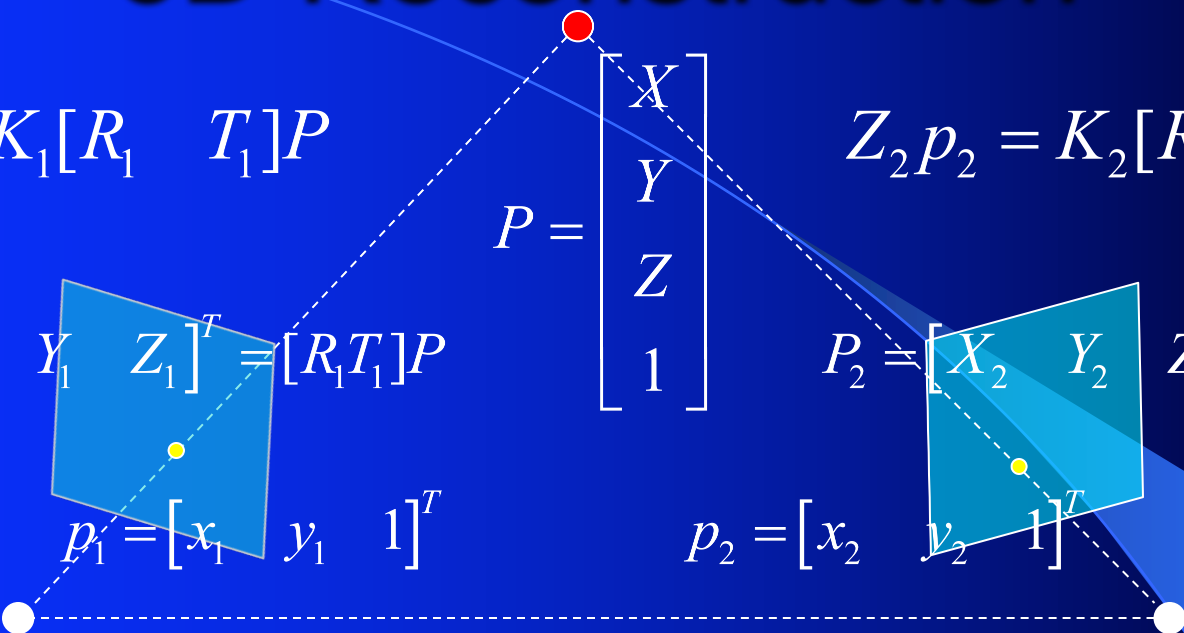
$$Z_1 p_1 = K_1 [R_1 T_1] P \qquad P = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad Z_2 p_2 = K_2 [R_2 T_2] P$$



$$p_1 = \begin{bmatrix} x_1 & y_1 & 1 \end{bmatrix}^T \qquad p_2 = \begin{bmatrix} x_2 & y_2 & 1 \end{bmatrix}^T$$

$$K_1 = \begin{bmatrix} f_{1x} & 0 & c_{1x} \\ 0 & f_{1y} & c_{1y} \\ 0 & 0 & 1 \end{bmatrix} \qquad K_2 = \begin{bmatrix} f_{2x} & 0 & c_{2x} \\ 0 & f_{2y} & c_{2y} \\ 0 & 0 & 1 \end{bmatrix}$$

3D Reconstruction



$$Z_1 p_1 = K_1[R_1 \quad T_1]P \qquad Z_2 p_2 = K_2[R_2 \quad T_2]P$$

$$P = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$P_1 = [X_1 \quad Y_1 \quad Z_1]^T = [R_1 T_1]P \qquad P_2 = [X_2 \quad Y_2 \quad Z_2]^T = [R_2 T_2]P$$

$$K_1 = \begin{bmatrix} f_{1x} & 0 & c_{1x} \\ 0 & f_{1y} & c_{1y} \\ 0 & 0 & 1 \end{bmatrix} \qquad p_1 = [x_1 \quad y_1 \quad 1]^T \qquad p_2 = [x_2 \quad y_2 \quad 1]^T \qquad K_2 = \begin{bmatrix} f_{2x} & 0 & c_{2x} \\ 0 & f_{2y} & c_{2y} \\ 0 & 0 & 1 \end{bmatrix}$$

$$Z_1 a_1 = K_1^{-1} Z_1 p_1 = [R_1 T_1]P := P_1 \qquad Z_2 a_2 = K_2^{-1} Z_2 p_2 = [R_2 T_2]P := P_2$$

$$Z_1 a_1 = [R_1 T_1]P = [R_1 T_1][R_2 T_2]^{-1}[R_2 T_2]P = [R_1 T_1][R_2 T_2]^{-1} Z_2 a_2$$

$$Z_1 a_1 - Z_2 R_1 R_2^{-1} a_2 = T_1 - R_1 R_2^{-1} T_2 \iff \begin{bmatrix} a_1 & -R_1 R_2^{-1} a_2 \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = T_1 - R_1 R_2^{-1} T_2$$

$$N \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = T \Rightarrow N^T N \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = N^T T \Rightarrow \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = (N^T N)^{-1} N^T T$$

Solve It!