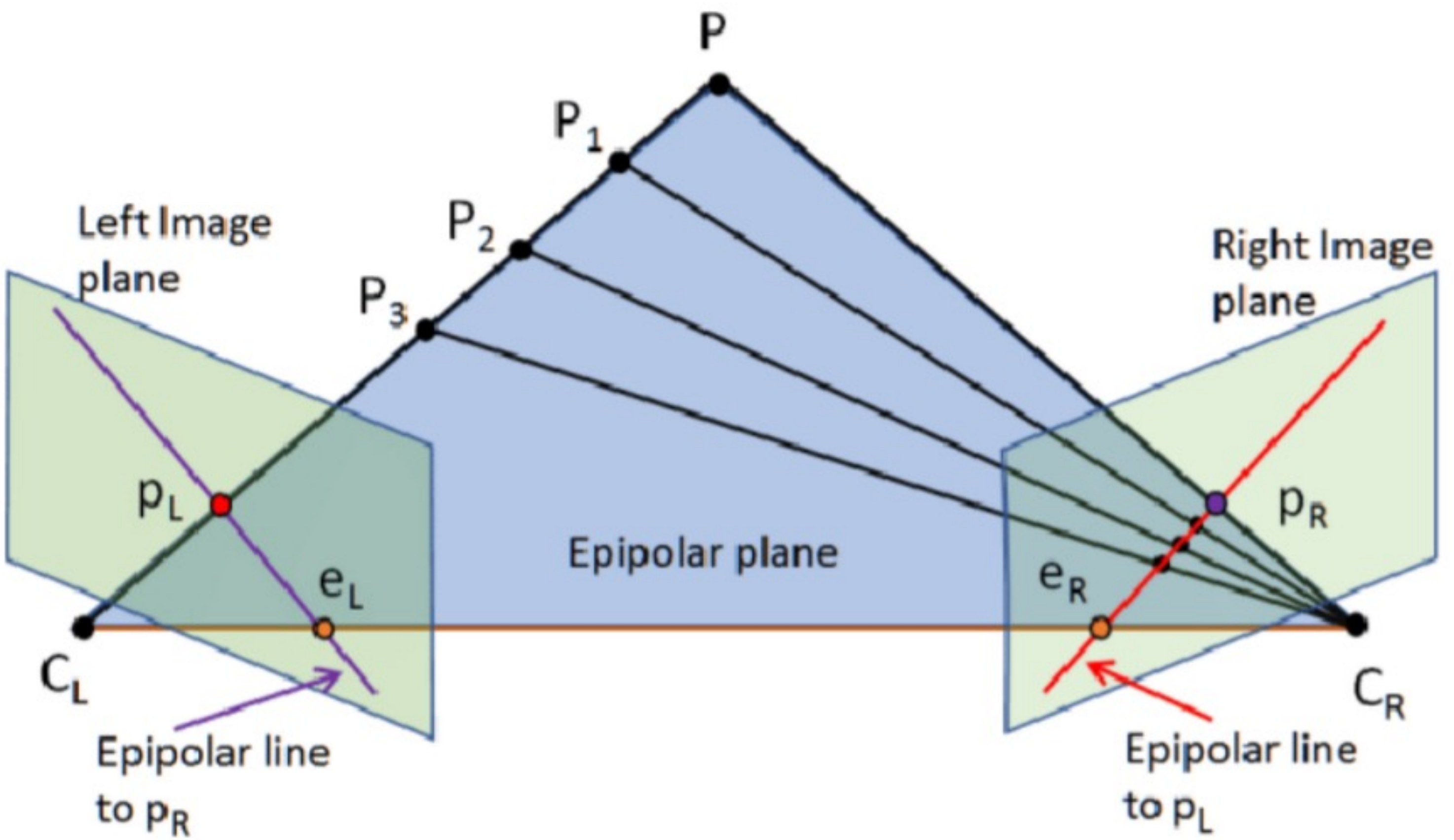
Camera Calibration 2

2024.08.06
ysjeong@rcv.sejong.ac.kr

Stereo Vision

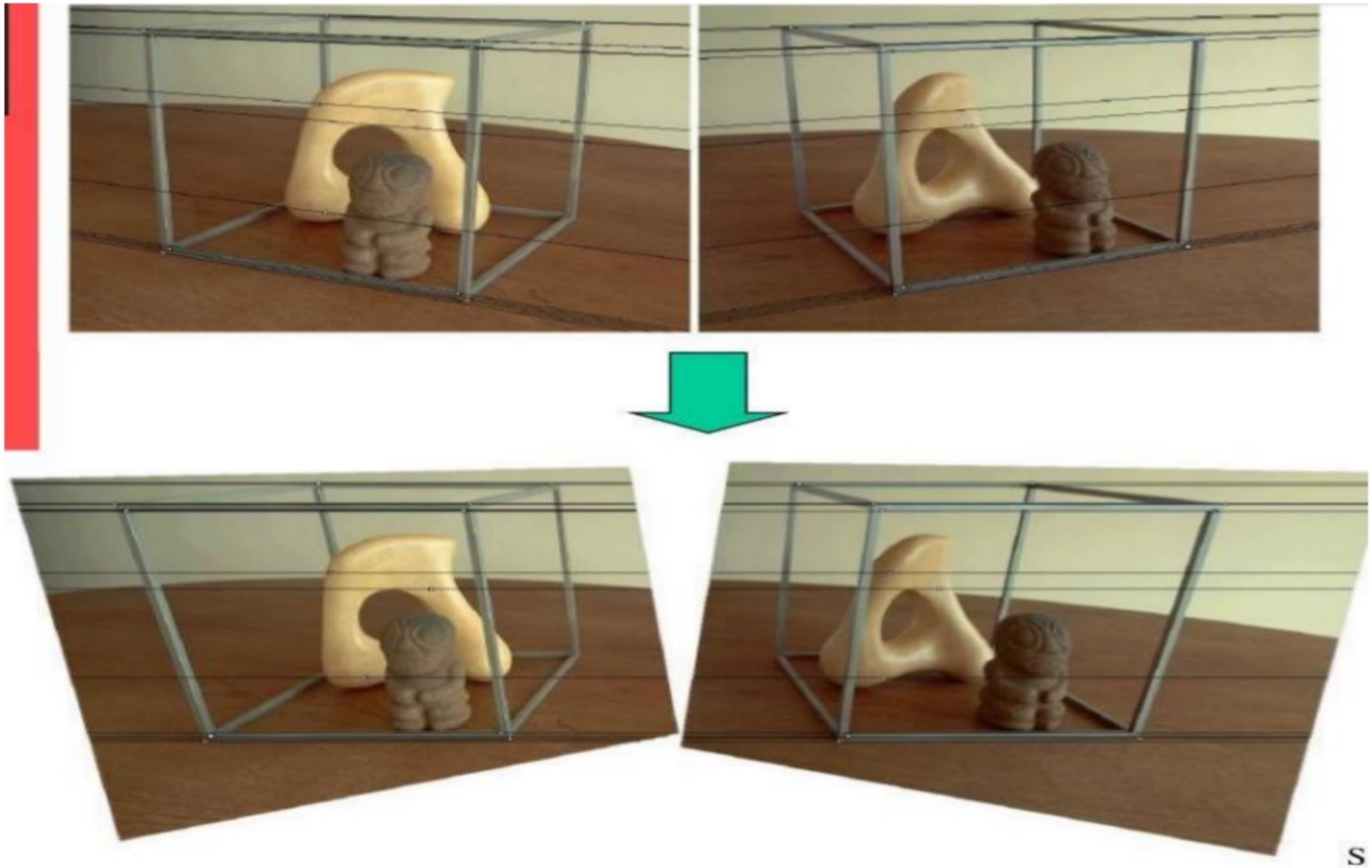
Epipolar Geometry

- A영상 속 한 점에 대하여 깊이를 추정할 수는 없지만, B 영상을 통해 한 점이 어느 직선 위에 존재하는지는 알 수 있다.



P : 3차원 물체의 좌표
C_L, C_R: 좌,우 카메라의 중심
p_L, p_R: 좌,우 이미지 평면에 P를 투영시킨 점.
e_L, e_R: 좌,우 카메라의 epipole

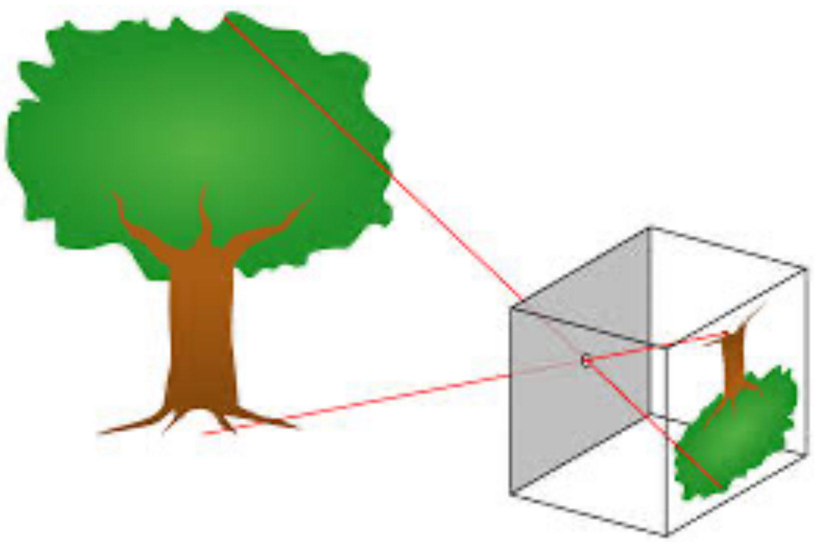
Image Rectification

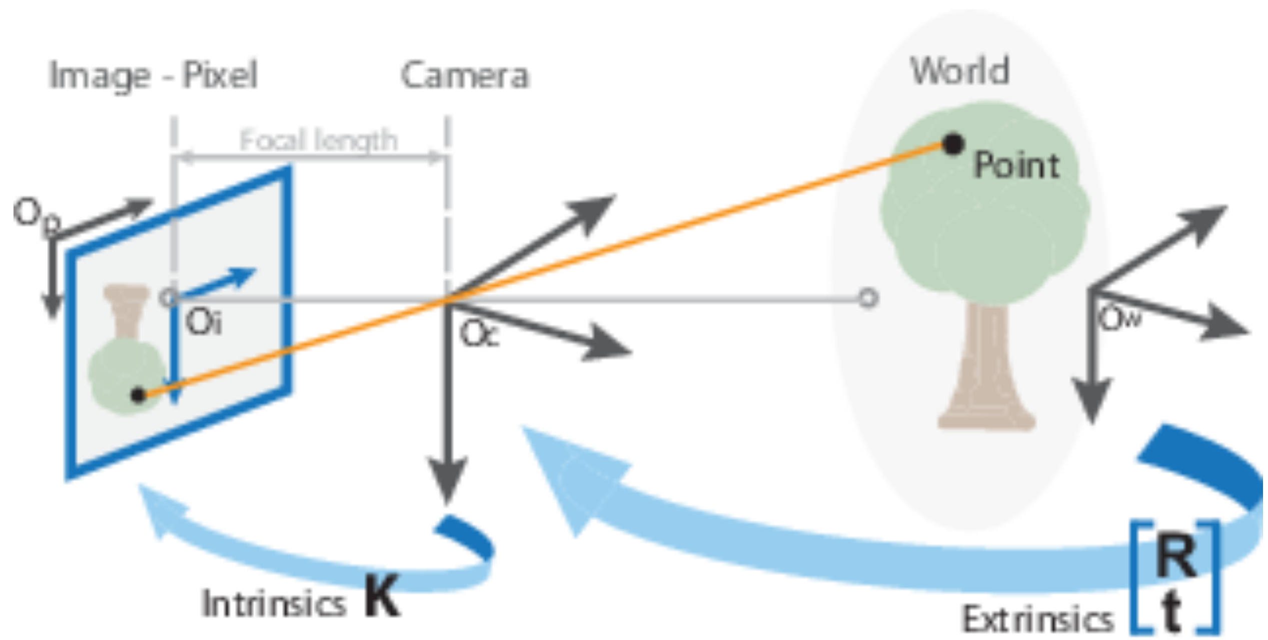


Coordinate System

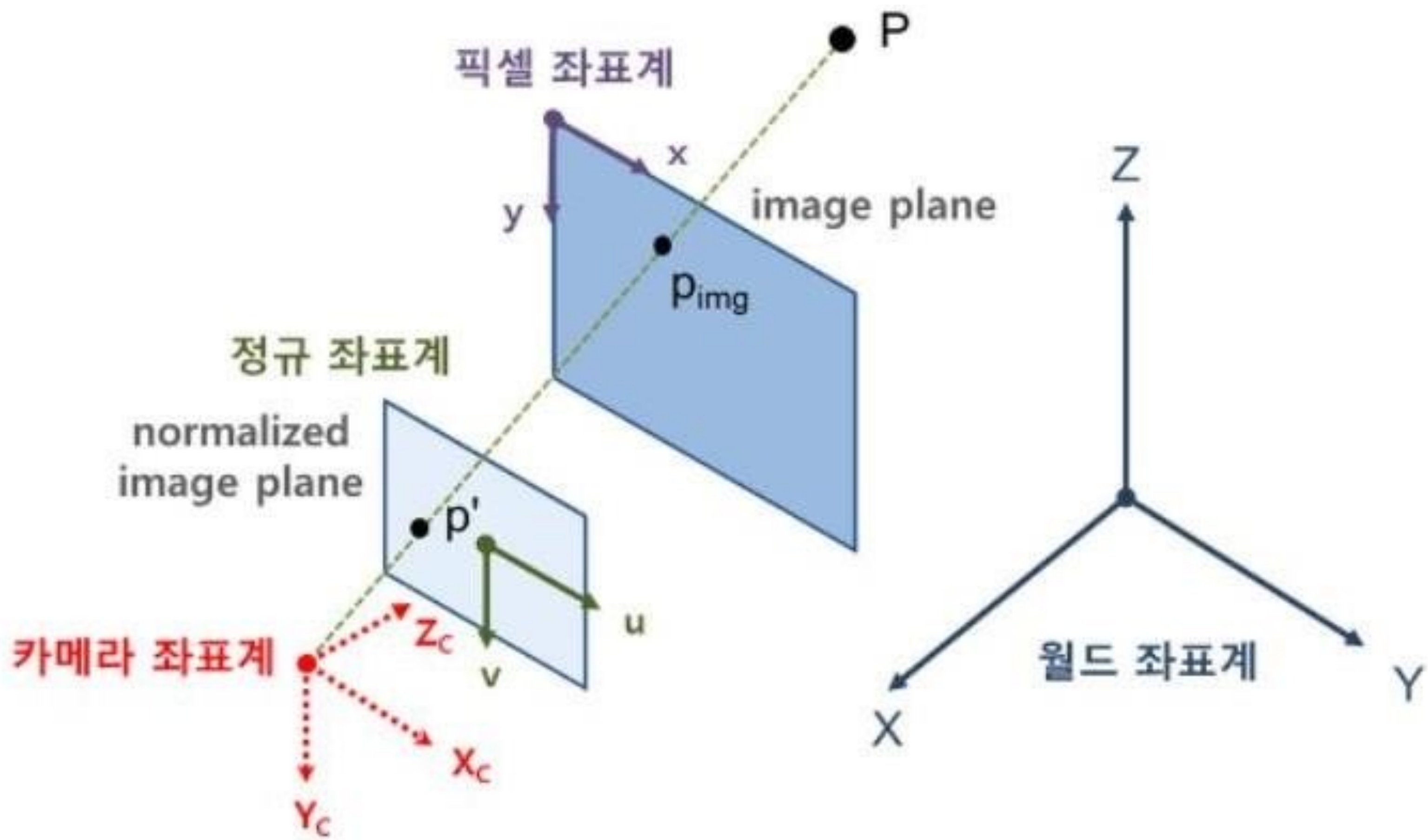
Camera Geometry

- 우리가 다루는 입력은 2D, 실제 동작 환경은 3D이므로 현실세계의 차원과 영상 속 차원의 관계를 알아야 함.





Coordinate System



https://darkpgmr.tistory.com/77?category=460965

월드 좌표계 : $P=(X,Y,Z)$
카메라 좌표계 : $Pc = (X_C,Yc,Zc)$
픽셀(영상) 좌표계 : $P_{img} = (x,y)$
정규 좌표계 : $p' = (u,v)$

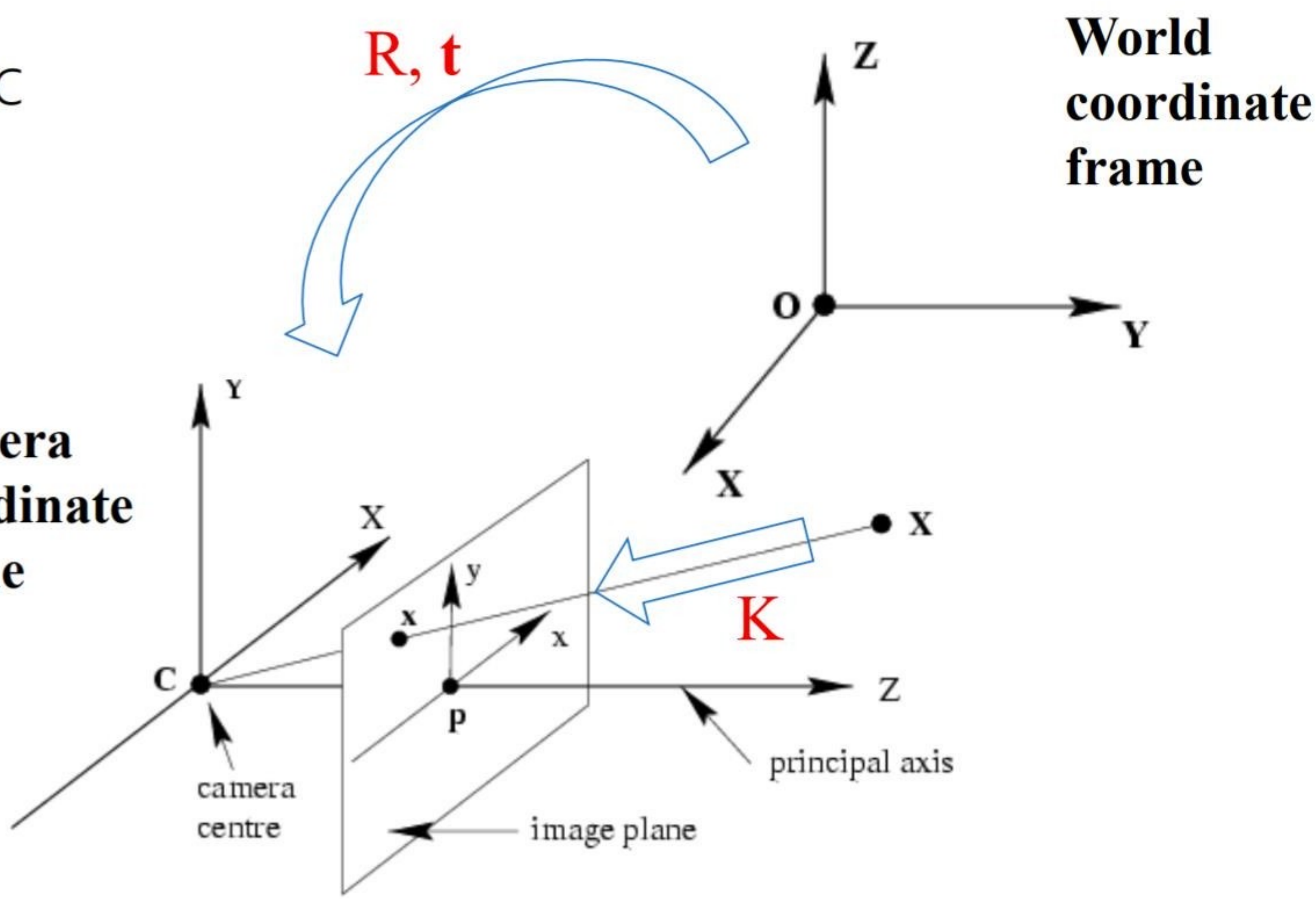
$$p_{img} = Kp'$$

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

Coordinate System

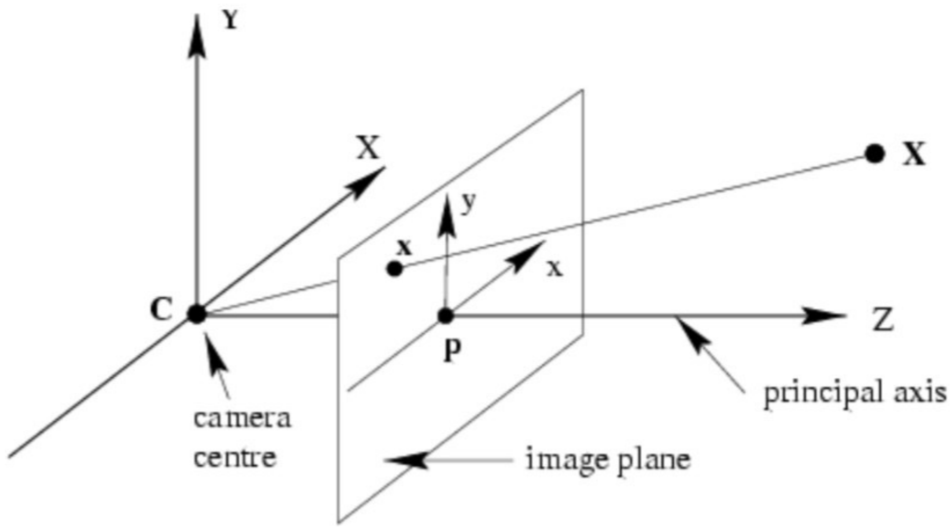
$$\mathbf{x} = K\left[R \mid \mathbf{t}\right]\mathbf{X}$$

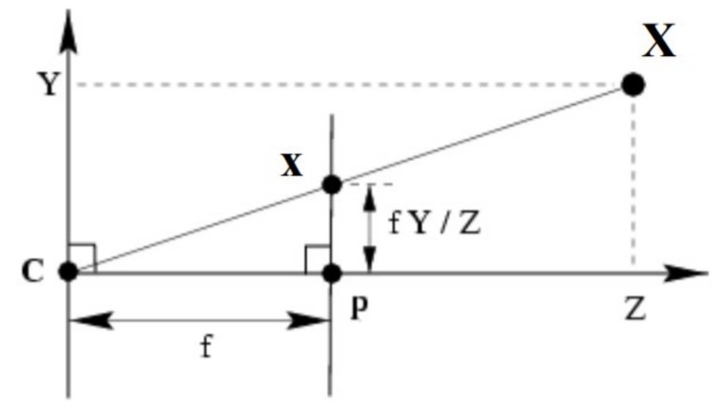
C2I W2C



Coordinate System

- By similar triangles, the point $(X, Y, Z)^T$ is mapped to the point $(fX/Z, fY/Z, f)^T$ on the image plane.
- Ignoring the final image coordinate, we see that
$$(X, Y, Z)^T \mapsto (fX/Z, fY/Z)^T$$
- This is a mapping from Euclidean 3-space IR^3 to Euclidean 2-space IR^2.





Principal Point

- The expression below assumes that the origin of coordinates in the image plane is at the principal point.

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

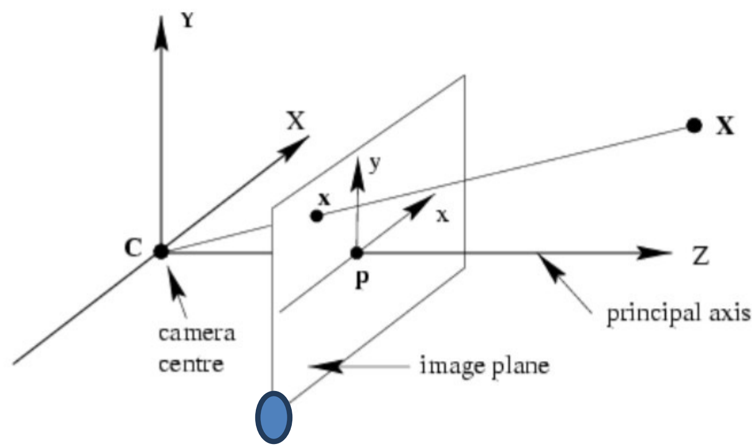
- In practice, it may not be and there is a more general expression as

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

  - $(p_x, p_y)^T$ are the coordinates of the principal point.

- This equation may be expressed conveniently in homogeneous coordinates as

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



Principal Point

$$\begin{pmatrix} fX + Zp_x \\ fY + Zp_x \\ Z \end{pmatrix} = \begin{bmatrix} f & & p_x & 0 \\ & f & p_y & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \begin{bmatrix} f & & p_x \\ & f & p_y \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & 0 \\ & 1 & & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

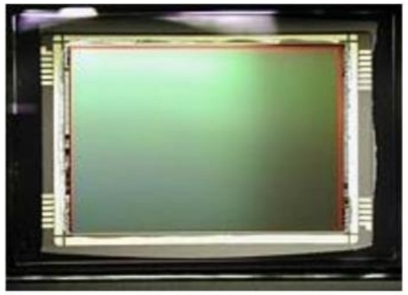
⇨ $\mathbf{x} = \mathrm{K}[\mathrm{I} \mid 0]\mathbf{X}_{\mathrm{cam}}$ $\mathbf{X}_{\mathrm{cam}}$ is expressed in the camera coordinate frame.

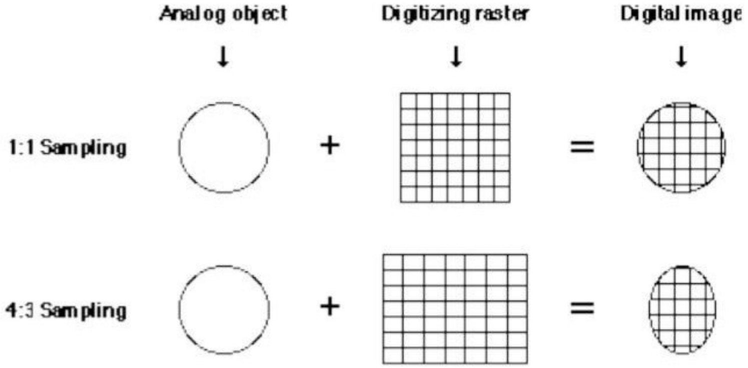
$$\mathrm{K} = \begin{bmatrix} f & & p_x \\ & f & p_y \\ & & 1 \end{bmatrix}$$

Calibration matrix or
Intrinsic parameters matrix

Intrinsic Parameter

- The pinhole camera model assumes that the image coordinates are Euclidean coordinates having equal scales in both axial directions.
- In the case of CCD cameras, there is the additional possibility of having non-square pixels.
- If image coordinates are measured in pixels, then this has the extra effect of introducing unequal scale factors in each direction.
- If the number of pixels per unit distance in image coordinates are m_x and m_y in the x and y directions,
  - then the transformation from world coordinates to pixel coordinates is obtained by multiplying K on the left by an extra factor diag(m_x, m_y, 1).





Intrinsic Parameter

❖ The general form of the calibration matrix of a CCD camera is as follows:

$$\mathrm{K} = \begin{bmatrix} m_x & & \\ & m_y & \\ & & 1 \end{bmatrix} \begin{bmatrix} f & & p_x \\ & f & p_y \\ & & 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix}$$

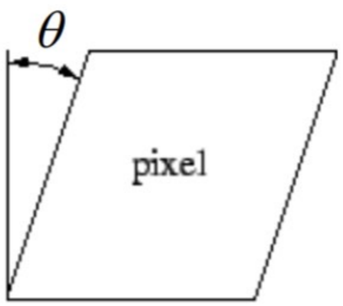
- $\alpha_x = fm_x$ and $\alpha_y = fm_y$ represent the focal length of the camera in terms of pixel dimensions in the x and y directions, respectively.
- (x_0, y_0) is the principal point in terms of pixel dimensions, with coordinates $x_0 = p_x m_x$ and $y_0 = p_y m_y$.

Intrinsic Parameter

❖ Sometimes, the skew parameter (*s*) is added to the calibration matrix.

$$K = \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix}$$

- The skew parameter will be zero for most normal cameras.
- However, in certain unusual instances, it can take non-zero values.
- If s≠0, then this can be interpreted as a skewing of the pixel elements in the CCD array so that the x- and y-axes are not perpendicular.



$$s = \alpha_x \tan\theta$$

Extrinsic Parameter

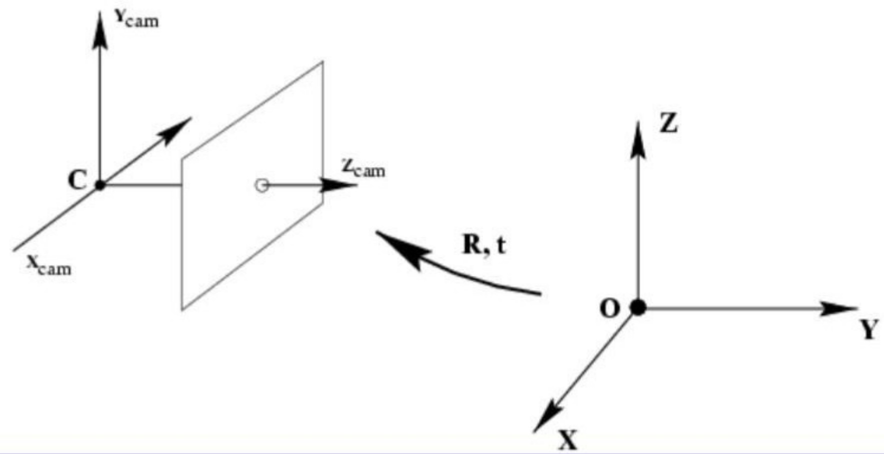
- $\tilde{\mathbf{X}}$ represents a 3D point in the world coordinate frame.
- $\tilde{\mathbf{X}}_{cam}$ represents the same point in the camera coordinate frame.

  $$\mathbf{x} = \mathrm{K}\left[\mathrm{I} \mid 0\right]\mathbf{X}_{\mathrm{cam}}$$

- Two points are related via a rotation and a translation.

  $$\tilde{\mathbf{X}}_{\mathrm{cam}} = \mathrm{R}\left(\tilde{\mathbf{X}}\text{-}\tilde{\mathbf{C}}\right)$$ → inhomogeneous coordinates

  - $\tilde{\mathbf{C}}$ represents the coordinates of the camera.
  - R is a 3x3 rotation matrix representing the orientation of the camera coordinate frame.



Extrinsic Parameter

$$\begin{aligned}\tilde{\mathbf{X}}_{\mathrm{cam}} &= \mathrm{R}\left(\tilde{\mathbf{X}}\text{-}\tilde{\mathbf{C}}\right) \\ &= \mathrm{R}\tilde{\mathbf{X}} - \mathrm{R}\tilde{\mathbf{C}}\end{aligned}$$

→ inhomogeneous coordinates

❖ This equation may be written in homogeneous coordinates as

$$\mathbf{X}_{\mathrm{cam}} = \begin{bmatrix} \mathrm{R} & -\mathrm{R}\tilde{\mathbf{C}} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \begin{bmatrix} \mathrm{R} & -\mathrm{R}\tilde{\mathbf{C}} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix} \mathbf{X}$$

- $\mathbf{X}_{cam}$ and $\mathbf{X}$ are all in homogeneous coordinate system.
- $\mathbf{X}$ represents a 3D point in the world coordinate frame.
- $\mathbf{X}_{cam}$ represents the same point in the camera coordinate frame.

Extrinsic Parameter

$$\mathbf{X}_{\text{cam}} = \begin{bmatrix} \mathrm{R} & -\mathrm{R}\tilde{\mathbf{C}} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \begin{bmatrix} \mathrm{R} & -\mathrm{R}\tilde{\mathbf{C}} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix} \mathbf{X}$$

$$\mathbf{x} = \mathrm{K}[\mathrm{I} \mid \mathbf{0}]\mathbf{X}_{\text{cam}} = \mathrm{K}[\mathrm{I} \mid \mathbf{0}] \begin{bmatrix} \mathrm{R} & -\mathrm{R}\tilde{\mathbf{C}} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix} \mathbf{X} = \mathrm{K}\left[\mathrm{R} \mid -\mathrm{R}\tilde{\mathbf{C}}\right]\mathbf{X} = \mathrm{KR}\left[\mathrm{I} \mid -\tilde{\mathbf{C}}\right]\mathbf{X}$$

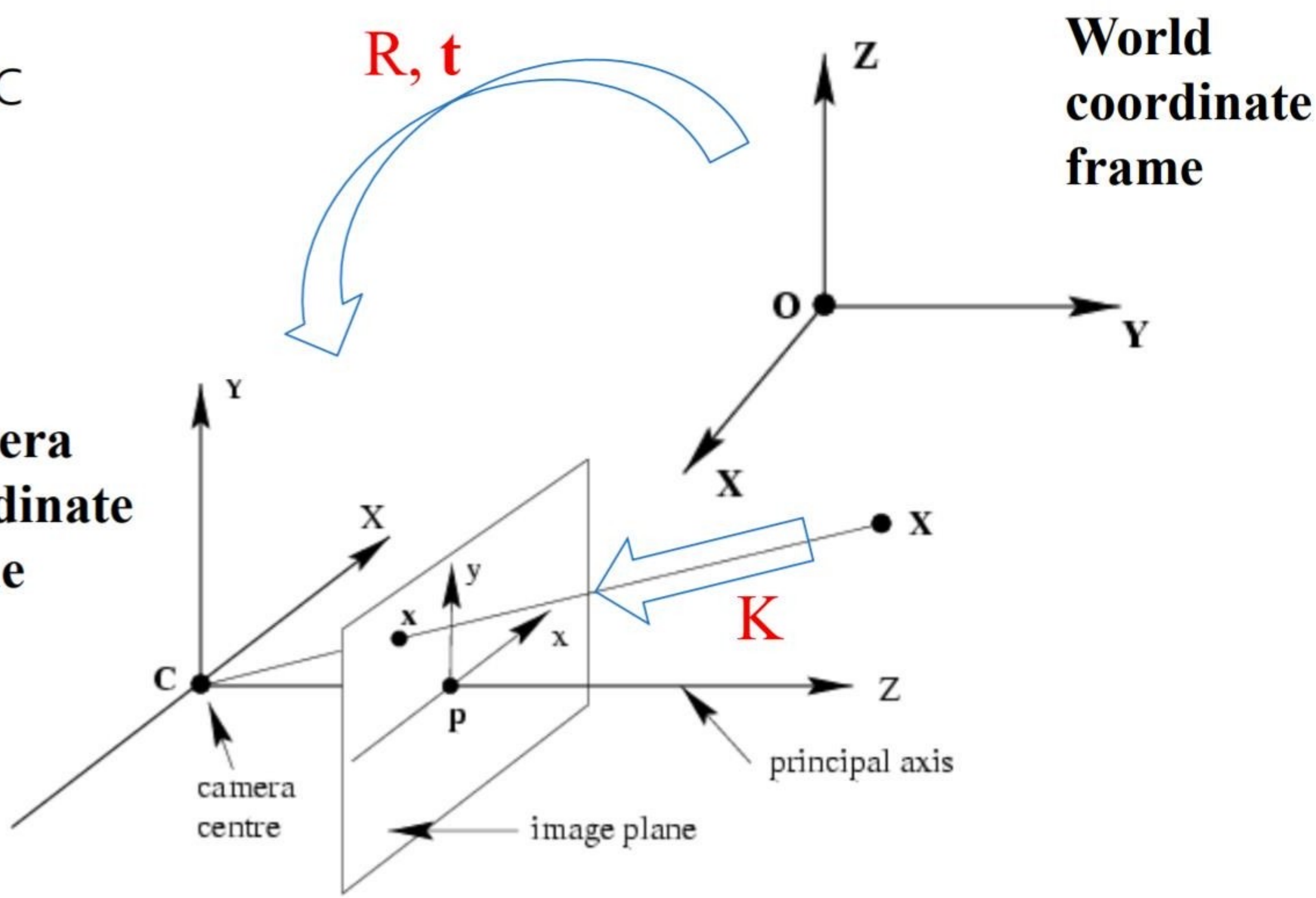
$$\mathbf{x} = \mathrm{KR}\left[\mathrm{I} \mid -\tilde{\mathbf{C}}\right]\mathbf{X} = \mathrm{K}\left[\mathrm{R} \mid -\mathrm{R}\tilde{\mathbf{C}}\right]\mathbf{X} = \mathrm{K}[\mathrm{R} \mid \mathbf{t}]\mathbf{X}$$

$$\mathbf{x} = \mathrm{P}\mathbf{X} \qquad \mathrm{P} = \mathrm{K}[\mathrm{R} \mid \mathbf{t}] \qquad \mathbf{t} = -\mathrm{R}\tilde{\mathbf{C}}$$

Extrinsic Parameter

$$\mathbf{x} = \mathrm{K}\left[\mathrm{R} \mid \mathbf{t}\right]\mathbf{X}$$

C2I W2C



Camera Calibration

$$s\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & \text{skew\_c}f_x & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$= A[R \mid t]\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad (1)$$

Solve Calibration matrix

$\mathrm{x}_i = \mathrm{PX}_i$ → Two vector differ in magnitude but have the same direction.

$\mathrm{x}_i \times \mathrm{PX}_i = 0 \qquad [\mathrm{x}_i]_\times \mathrm{PX}_i = 0$

Cross products

Of particular interest are 3×3 skew-symmetric matrices. If $\mathbf{a} = (a_1, a_2, a_3)^\mathsf{T}$ is a 3-vector, then one defines a corresponding skew-symmetric matrix as follows:

$$[\mathbf{a}]_\times = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}. \tag{A4.5}$$

Note that any skew-symmetric 3×3 matrix may be written in the form $[\mathbf{a}]_\times$ for a suitable vector $\mathbf{a}$. Matrix $[\mathbf{a}]_\times$ is singular, and $\mathbf{a}$ is its null-vector (right or left). Hence, a 3×3 skew-symmetric matrix is defined up to scale by its null-vector.

The cross product (or vector product) of two 3-vectors $\mathbf{a} \times \mathbf{b}$ (sometimes written $\mathbf{a} \wedge \mathbf{b}$) is the vector $(a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1)^\mathsf{T}$. The cross product is related to skew-symmetric matrices according to

$$\mathbf{a} \times \mathbf{b} = [\mathbf{a}]_\times \mathbf{b} = \left(\mathbf{a}^\mathsf{T}[\mathbf{b}]_\times\right)^\mathsf{T}. \tag{A4.6}$$

Solve Calibration matrix

$$\mathrm{x}_i = \mathrm{P}\mathrm{X}_i$$

$$\mathrm{x}_i \times \mathrm{P}\mathrm{X}_i = 0 \qquad [\mathrm{x}_i]_\times \mathrm{P}\mathrm{X}_i = 0$$

$$\begin{bmatrix} \mathbf{0}^\top & -w_i\mathbf{X}_i{}^\top & y_i\mathbf{X}_i{}^\top \\ w_i\mathbf{X}_i{}^\top & \mathbf{0}^\top & -x_i\mathbf{X}_i{}^\top \\ -y_i\mathbf{X}_i{}^\top & x_i\mathbf{X}_i{}^\top & \mathbf{0}^\top \end{bmatrix} \begin{pmatrix} \mathbf{P}^1 \\ \mathbf{P}^2 \\ \mathbf{P}^3 \end{pmatrix} = \mathbf{0}$$

$\mathbf{P}^{i\mathrm{T}}$ → i-th row of $\mathbf{P}$

Because only two rows are linearly independent, the last row is deleted.

$$\begin{bmatrix} \mathbf{0}^\top & -w_i\mathbf{X}_i{}^\top & y_i\mathbf{X}_i{}^\top \\ w_i\mathbf{X}_i{}^\top & \mathbf{0}^\top & -x_i\mathbf{X}_i{}^\top \end{bmatrix} \begin{pmatrix} \mathbf{P}^1 \\ \mathbf{P}^2 \\ \mathbf{P}^3 \end{pmatrix} = \mathbf{0}$$

$$\mathbf{A}\mathbf{p} = \mathbf{0}$$

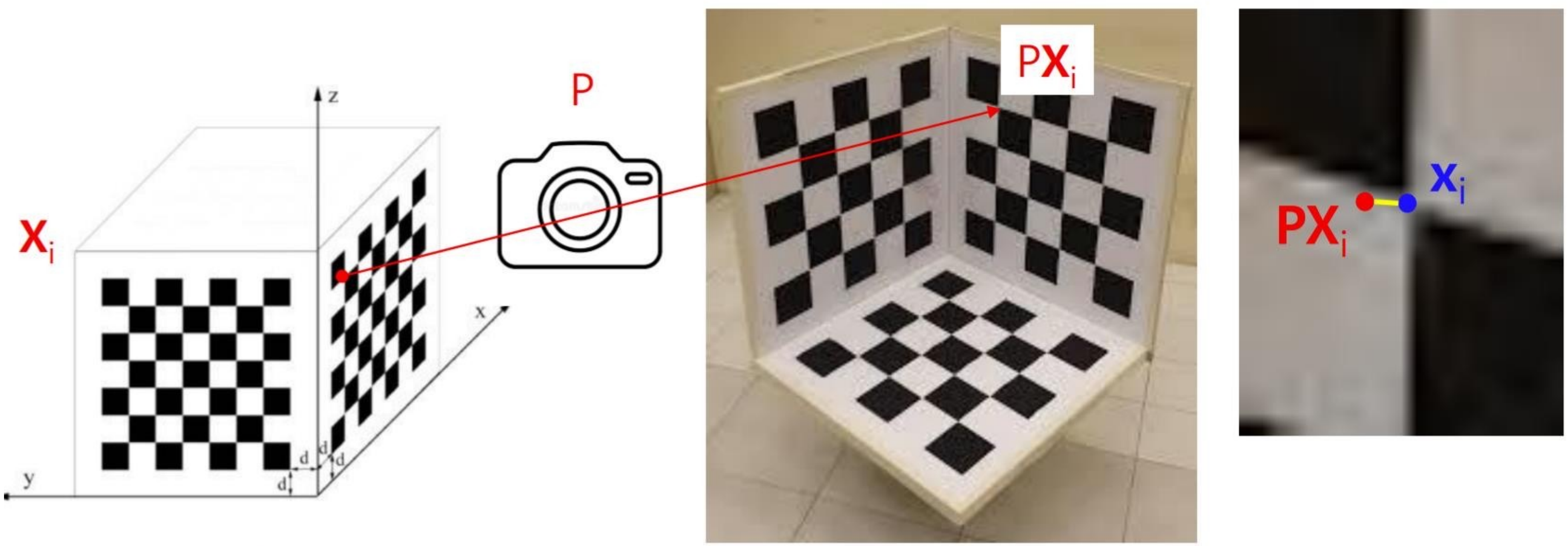
SVD can be used to find the least squares solution.

Solve Calibration matrix

- The geometric error in the image can be define as

$$\sum_i d(\mathrm{x}_i, \mathrm{PX}_i)^2$$

  - $\mathbf{x}_i$ is the measured point and P$\mathbf{X}_i$ is the image of $\mathbf{X}_i$.



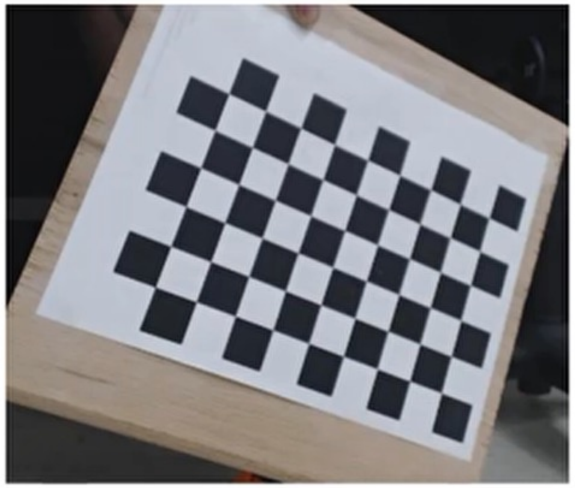
3D object

Image of 3D object

Zhang's Method

- Z. Zhang, "Flexible Camera Calibration By Viewing a Plane From Unknown Orientations," ICCV 1999.
  - Most widely used camera calibration method
  - Implemented in Matlab and OpenCV

- This method is convenient because it requires a simple calibration object formed by a single plane.



Zhang's Method

- Homography between the model plane and its image
  - Without loss of generality, we assume the model plane is on Z=0 of the world coordinate system.

$$s\begin{bmatrix}u\\v\\1\end{bmatrix} = \mathbf{A}\begin{bmatrix}\mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & \mathbf{t}\end{bmatrix}\begin{bmatrix}X\\Y\\0\\1\end{bmatrix} = \mathbf{A}\begin{bmatrix}\mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t}\end{bmatrix}\begin{bmatrix}X\\Y\\1\end{bmatrix}$$

  - A model point $\widetilde{M}$ and its image $\widetilde{\boldsymbol{m}}$ is related by a homography $\mathbf{H}$:

$$s\widetilde{\mathrm{m}} = \mathbf{H}\widetilde{\mathrm{M}} \qquad \text{with} \quad \mathbf{H} = \mathbf{A}\begin{bmatrix}\mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t}\end{bmatrix} \qquad \widetilde{\mathrm{M}} = [X, Y, 1]^T$$

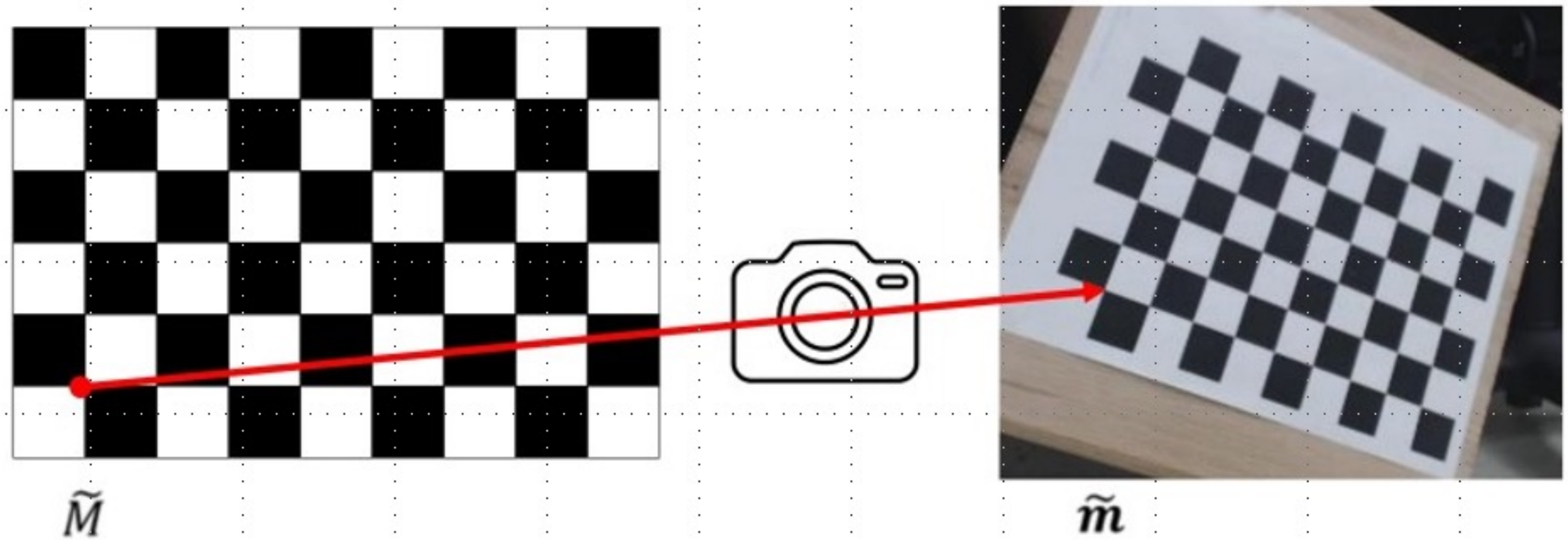
    - 3x3 matrix $\mathbf{H}$ is defined up to a scale factor.

Zhang's Method

❖ Given an image of the model plane, an homography can be estimated using the method we have already studied.

$$s\widetilde{\mathrm{m}} = \mathrm{H}\widetilde{\mathrm{M}}$$

$$\mathbf{A}\mathbf{h} = \mathbf{0} \qquad \begin{bmatrix} \mathrm{A}_1 \\ \mathrm{A}_2 \\ \vdots \\ \mathrm{A}_n \end{bmatrix} \mathbf{h} = 0 \qquad \begin{bmatrix} \mathbf{0}^\mathsf{T} & -w'_i\mathbf{x}_i^\mathsf{T} & y'_i\mathbf{x}_i^\mathsf{T} \\ w'_i\mathbf{x}_i^\mathsf{T} & \mathbf{0}^\mathsf{T} & -x'_i\mathbf{x}_i^\mathsf{T} \end{bmatrix} \begin{pmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \\ \mathbf{h}^3 \end{pmatrix} = \mathbf{0}$$



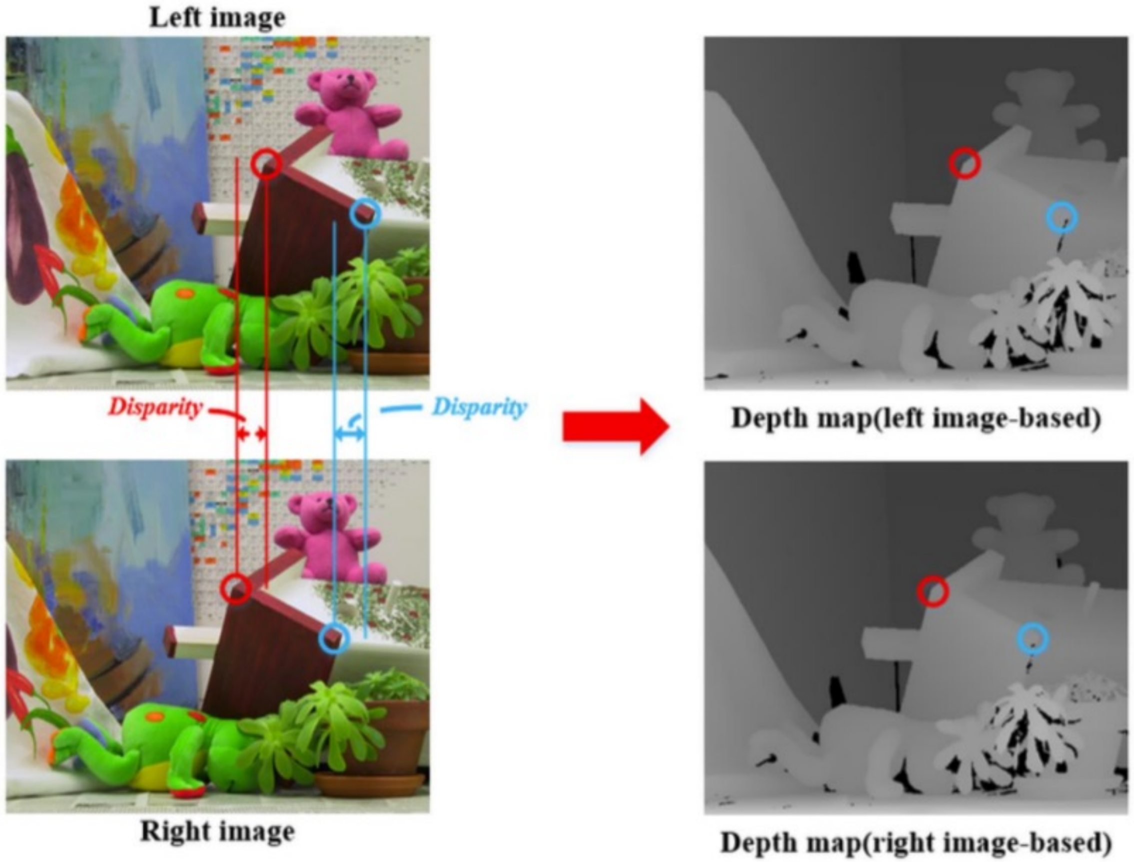
Depth - Disparity

깊이 영상(Depth map)

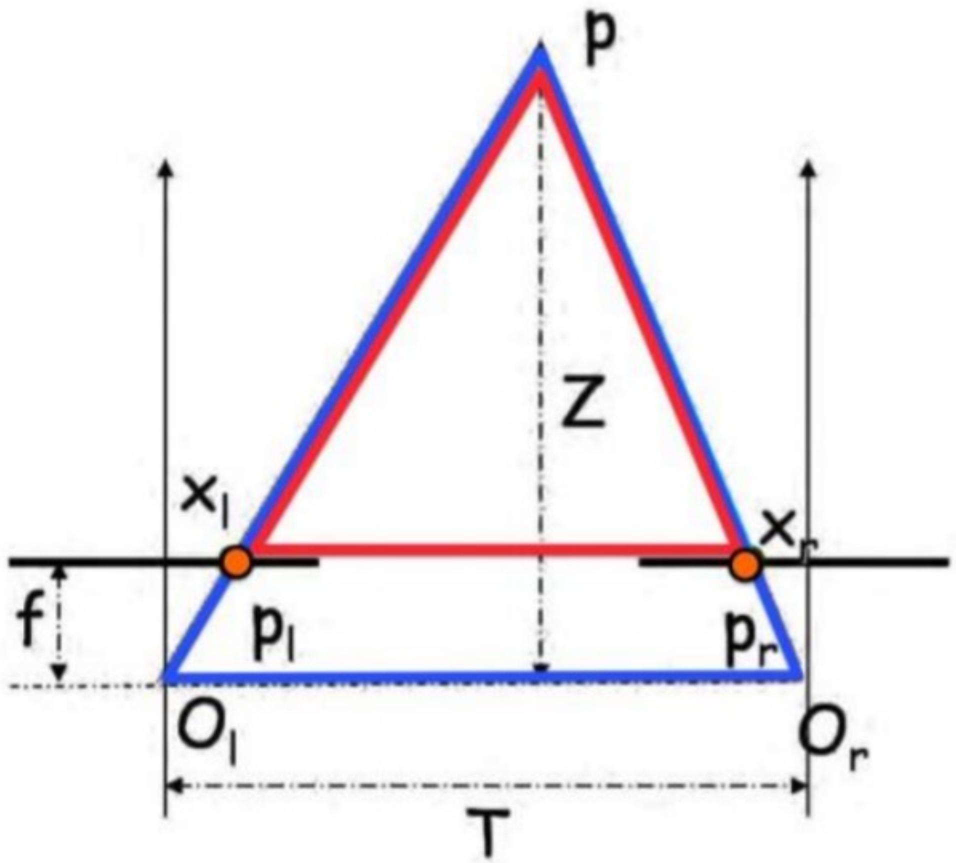
(a) Real-image

(b) Depth-map

시차(Disparity)



깊이와 시차의 관계식



Similar triangles:

$$\frac{T}{Z} = \frac{T + x_r - x_l}{Z - f}$$

$$\boxed{Z = \frac{f \cdot T}{x_l - x_r}}$$

baseline

focal length

disparity

이미지 출처 : CSC420 : Intro to Image Understanding (Sanja Fidler)

3D detection