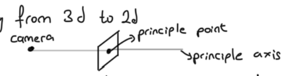
12- sensing depth

camera parameters

(içsel)
intrinsic parameters = mapping from 3d to 2d

↳ principle point coordinates =



↳ scaling factors = for mapping from word to pixel coordinates (focal length, pixel magnification factor)

↳ $\begin{bmatrix} \alpha_x & 0 & \beta_x \\ 0 & \alpha_y & \beta_y \\ 0 & 0 & 1 \end{bmatrix}$ calibration matrix

↳ skew = non-regtengular pixels (ignored mostly)

↳ radial distortions = straight lines appearing curved in the image → mostly in wide angles

extrinsic parameter = how the camera is located in 3d space (how a scene captured from a specific viewpoint)

↳ translation (position) = camera coordinates in the space (x, y, z)

↳ rotation (orientation) = the direction that the camera is pointing

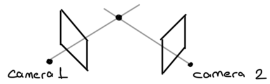
camera calibration

- word coordinate system ⟶ image coordinate system

$$\begin{bmatrix} \times \\ \text{2D point} \\ (3\times1) \end{bmatrix} = \underbrace{\begin{bmatrix} \text{camera to pixel} \\ \text{coord. trans matrix} \\ K\ (3\times3) \end{bmatrix}}_{\text{instrinsic camera param}} \underbrace{\begin{bmatrix} \text{canonical} \\ \text{projection matrix} \\ [I|0]\ (3\times4) \end{bmatrix} \begin{bmatrix} \text{word to camera coord} \\ \text{translation matrix} \\ \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix}\ (4\times4) \end{bmatrix}}_{\text{extrinsic camera param}} \begin{bmatrix} \times \\ \text{3D point} \\ 4\times1 \end{bmatrix}$$

$P = K[R|t]$ = general camera projection matrix

triangulation = determining the location of a point in 3d space by measuring its position from multiple cameras/points

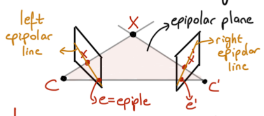


stereo reconstruction = two dimensional images are used to reconstruct a 3d scene

① for each point in the first image determine the corresponding point in the second image → search

↳ epipolar constraint = simplify the process from 2d to 1d problem

- a point in one image must lie on the corresponding epipolar line in the other image



coplanarity = points are all located on the same plane

epipole = where the line connecting two camera centers intersets the image plane (image of the centre of the other camera) $e = PC'$ $e' = P'C$

↳ epipolar geometry depends only the relative pose (position and orientation) and internel params of the two cameras (position of the cameras and image planes)

↳ it does not depend on the scene structure (3d points external to the camera)

↳ buttom-up algorithms

- dense = compute a correspondance at every pixel
- sparse = compute correspondances only for features (eg SIFT features)

② for each pair of matched points, determine the 3d point by triangulation → estimate

↳ vector solution = compute a midpoint of the shortest line between CI and $C'I'$

↳ line triangulation = algebric solution, solve X using $x = PX$ $x' = P'X$

↳ minimize a geometric/statistical error = $\hat{x} = P\hat{X}$ $\hat{x}' = P'\hat{X}$ estimate $\hat{X}$

$$\min(\hat{X}) \to C(x, x') = d(x, \hat{x})^2 + d(x', \hat{x}')^2$$

↳ laser scanning or active stereo with structured light patterns can be used to determine the 3d point too