3D VISION

Monocular POSE

Homography between Projective Planes

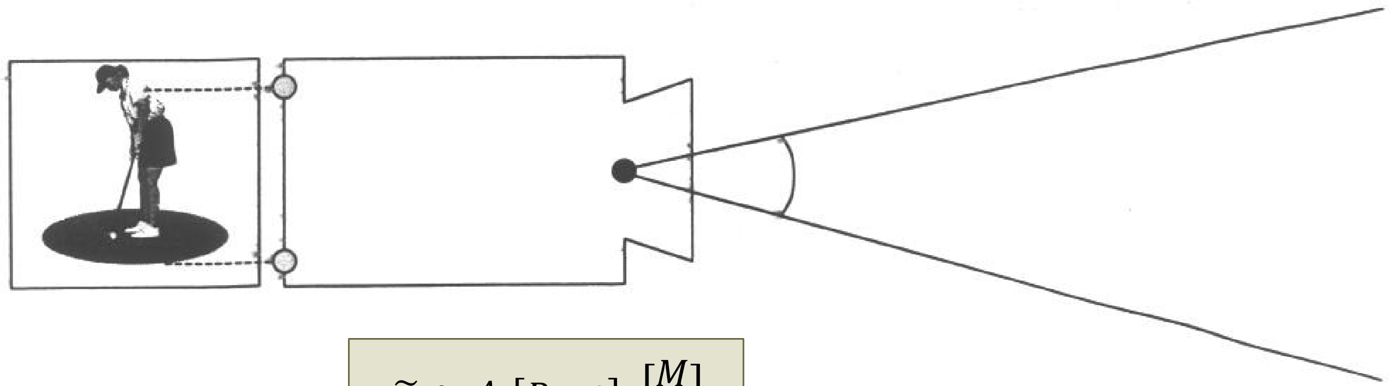
Systems Engineering and Automation

Miguel Hernández University

Table of Contents

Single-Image POSE Estimation

- Objectives
  - To what extent we can get information about the location of objects in the scene from a single view.
  - A camera measures angles, not distances
- Options:
  - Calibrated Camera
  - Not-Calibrated Camera



$$\tilde{m} \cong A\,[R \quad t]\begin{bmatrix} M \\ 1 \end{bmatrix}$$

POSE Estimation

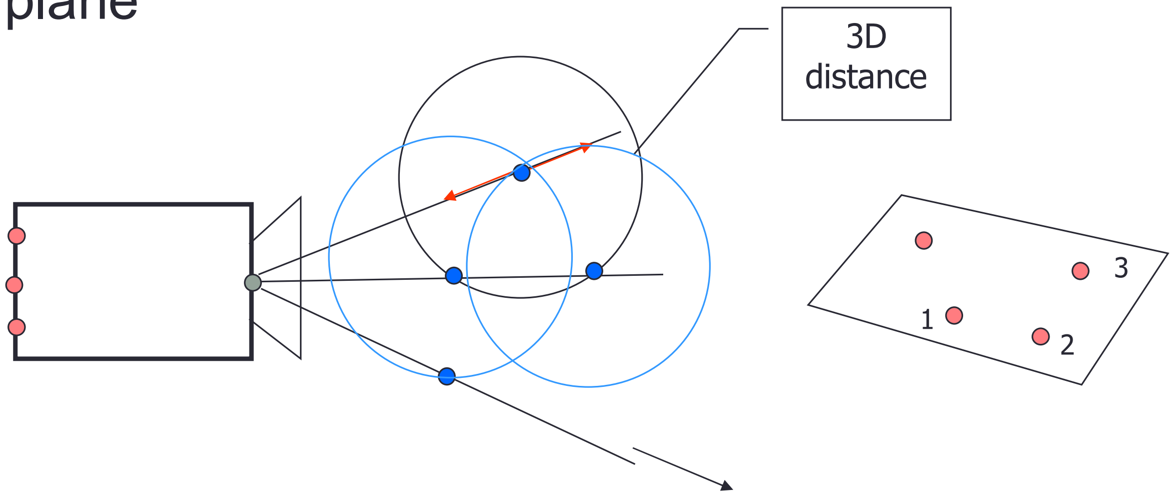
- Definition:
  - '**Pose**' (position and orientation) of an object from:
    - Knowledge of the structure of the object (e.g. Distances between marks on the object)
    - An image taken with a calibrated camera in which the marks can be identified
  - There are many algorithms depending on the information and structure of the object.

POSE Estimation

- ‘**Pose**’ estimation of an object using artificial marks (points):
  - We know the camera matrix (intrinsic parameters) ***A*** ([***R,t***] allow to set an absolute reference)
  - 6 dof (3 Position, 3 Orientation)
  - Each mark (point) provides two restrictions
  - At least 3 marks are required (distances between them)

POSE Estimation: 3 points on a plane

- 3 points on a plane



Optical Center

Direction: point at infinite

$$\tilde{m} \cong A\ [R \quad t] \begin{bmatrix} M \\ 1 \end{bmatrix}$$

$$\tilde{C} \cong \begin{bmatrix} -R^T t \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} C \\ 1 \end{bmatrix}$$

$$\tilde{M}_\infty \cong \begin{bmatrix} R^T A^{-1} \tilde{m} \\ 0 \end{bmatrix}$$

$$\tilde{m} \cong A\ [R \quad t] \begin{bmatrix} M_\infty \\ 0 \end{bmatrix}$$

POSE Estimation: 3 points on a plane

- Analytical Solution:
  - Equation of the projection rays:

$$\tilde{Q}_i \cong \begin{bmatrix} -R^T t \\ 1 \end{bmatrix} + \lambda_i \begin{bmatrix} R^T A^{-1} \tilde{m}_i \\ 0 \end{bmatrix} = \begin{bmatrix} -R^T t + \lambda_i R^T A^{-1} \tilde{m}_i \\ 1 \end{bmatrix} = \begin{bmatrix} Q_i \\ 1 \end{bmatrix}$$

$$\begin{aligned} d^2{}_{12} &= \|\lambda_1 R^T A^{-1} \tilde{m}_1 - \lambda_2 R^T A^{-1} \tilde{m}_2\|^2 \\ d^2{}_{13} &= \|\lambda_1 R^T A^{-1} \tilde{m}_1 - \lambda_3 R^T A^{-1} \tilde{m}_3\|^2 \\ d^2{}_{23} &= \|\lambda_2 R^T A^{-1} \tilde{m}_2 - \lambda_3 R^T A^{-1} \tilde{m}_3\|^2 \end{aligned}$$

- 3 quadratic equations with three unknowns
  - 2^3=8 solutions if the projections are non-coplanar (the plane does not pass through the optical center)
  - 4 in front of the camera
  - Unique solution: Use a 4th point

Perspective-n-Point (PnP): Plane Homography

- POSE calculation the from a set of 2D-3D matched points on a plane (linear method):

$$\tilde{m} \cong A[R \quad t]\tilde{M} = \tilde{P}\ \tilde{M} \qquad \text{Si} \quad \tilde{M} \cong \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} \in \prod \;\Rightarrow\; \tilde{m} \cong A \cdot \begin{array}{c} \begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ r_{31} & r_{32} & t_z \end{bmatrix} \\ \begin{array}{ccc} \bar{\mathbf{r}}_1 & \bar{\mathbf{r}}_2 & \bar{\mathbf{t}} \end{array} \end{array} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \qquad \bar{\mathbf{r}}_3 = \bar{\mathbf{r}}_1 \times \bar{\mathbf{r}}_2$$

$$\begin{bmatrix} \rho \cdot u \\ \rho \cdot v \\ \rho \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad \Rightarrow \quad \begin{array}{l} (h_{31}X + h_{32}Y + h_{33}) \cdot u = h_{11}X + h_{12}Y + h_{13} \\ (h_{31}X + h_{32}Y + h_{33}) \cdot v = h_{21}X + h_{22}Y + h_{23} \end{array}$$

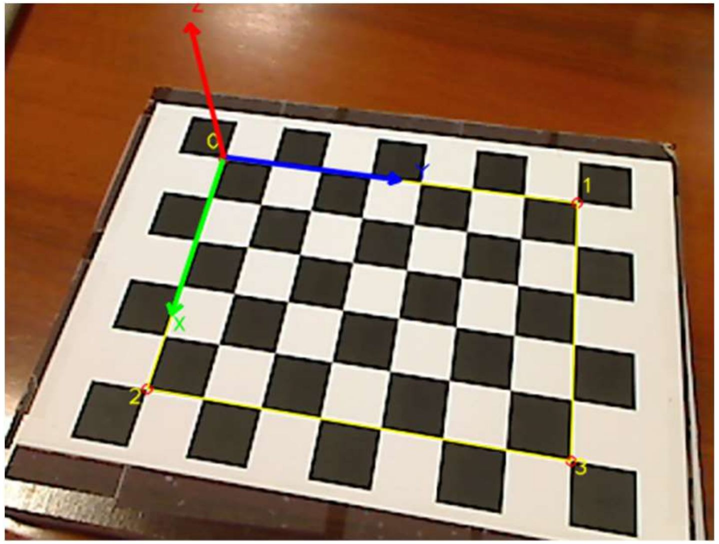
At least 4 points are needed:

$$\mathbf{A}_{xy} = \begin{bmatrix} -X^{(1)} & -Y^{(1)} & -1 & 0 & 0 & 0 & X^{(1)}u^{(1)} & Y^{(1)}u^{(1)} & u^{(1)} \\ 0 & 0 & 0 & -X^{(1)} & -Y^{(1)} & -1 & X^{(1)}v^{(1)} & Y^{(1)}v^{(1)} & v^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

$$\mathbf{h} = [h_{11}, h_{12}, h_{13}, h_{21}, h_{22}, h_{23}, h_{31}, h_{32}, h_{33}]^T$$

$$\mathbf{A}_{xy} \cdot \mathbf{h} = 0 \qquad \min_{\mathbf{h}} \left[\left(\mathbf{A}_{xy} \mathbf{h} \right)^2 \right]$$

$$\boxed{SVD(A_{xy}) \rightarrow A_{xy} = UDV^T \Rightarrow \mathbf{h} = v_9}$$



Perspective-n-Point (PnP): Plane Homography

- Homography Calculation:
  - Non-homogeneous solution (Euclidean): we assume $(h_{33} \neq 0)$ $H = \lambda A[\bar{\mathbf{r}}_1 \quad \bar{\mathbf{r}}_2 \quad \bar{\mathbf{t}}]$

$$\begin{bmatrix} \rho \cdot u \\ \rho \cdot v \\ \rho \end{bmatrix} = \begin{bmatrix} h_{11}/h_{33} & h_{12}/h_{33} & h_{13}/h_{33} \\ h_{21}/h_{33} & h_{22}/h_{33} & h_{23}/h_{33} \\ h_{31}/h_{33} & h_{32}/h_{33} & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

$h_{33} \cong t_z$ Distance between planes

$$\begin{bmatrix} \rho \cdot u \\ \rho \cdot v \\ \rho \end{bmatrix} = \begin{bmatrix} h'_{11} & h'_{12} & h'_{13} \\ h'_{21} & h'_{22} & h'_{23} \\ h'_{31} & h'_{32} & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \Longrightarrow \begin{matrix} (h'_{31}X + h'_{32}Y + 1) \cdot u = h'_{11}X + h'_{12}Y + h'_{13} \\ (h'_{31}X + h'_{32}Y + 1) \cdot v = h'_{21}X + h'_{22}Y + h'_{23} \end{matrix}$$

$$\mathbf{A}_{xy} = \begin{bmatrix} -x^{(1)} & -y^{(1)} & -1 & 0 & 0 & 0 & x^{(1)}x'^{(1)} & y^{(1)}x'^{(1)} \\ 0 & 0 & 0 & -x^{(1)} & -y^{(1)} & -1 & x^{(1)}y'^{(1)} & y^{(1)}y'^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} -u^{(1)} \\ -v^{(1)} \\ \vdots \end{bmatrix}^T$$

$$\mathbf{h} = [h'_{11}, h'_{12}, h'_{13}, h'_{21}, h'_{22}, h'_{23}, h'_{31}, h'_{32}]^T$$

- Least Squares:

$$\mathbf{A}_{xy} \cdot \mathbf{h} = \mathbf{b}$$

$$\min_{\mathbf{h}} \left[\left(\mathbf{A}_{xy}\mathbf{h} - \mathbf{b} \right)^2 \right] \Longrightarrow \boxed{\mathbf{h} = \left(A_{xy}^T A_{xy} \right)^{-1} \cdot A_{xy}^T \cdot \mathbf{b}}$$

Data Normalization

- More robust estimation against noise
  - Equation System better conditioned
- Specify a transformation
  - The center of gravity of all points is the origin
  - The mean distance to origin is $\sqrt{2}$
  - Applies to image and space points

$$\left.\begin{aligned}\tilde{\tilde{m}}_w &= T \cdot \tilde{m}_w \\ \tilde{\tilde{m}} &= T' \cdot \tilde{m}\end{aligned}\right\} \rightarrow \boxed{\tilde{\tilde{m}} \cong \hat{H} \cdot \tilde{\tilde{m}}_w} \qquad \boxed{H \cong T'^{-1} \cdot \hat{H} \cdot T}$$

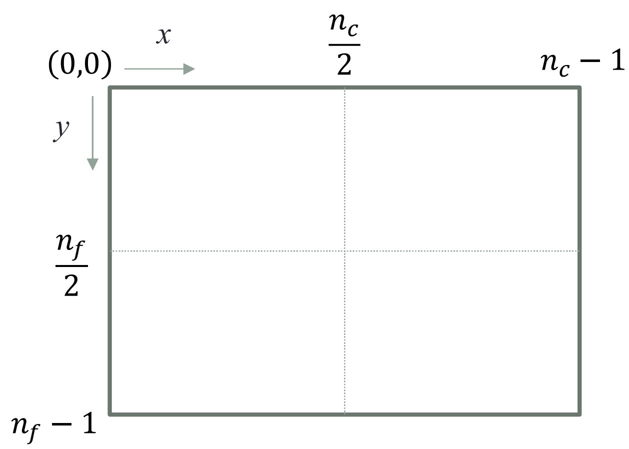
$$\boldsymbol{T} = \begin{bmatrix} 1/\sigma_x & 0 & -\mu_x/\sigma_x \\ 0 & 1/\sigma_y & -\mu_y/\sigma_y \\ 0 & 0 & 1 \end{bmatrix}$$

- Using simpler transformations
  - All the points between -1 y 1

$$\boldsymbol{T} = \begin{bmatrix} 2/n_c & 0 & -1 \\ 0 & 2/n_f & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

  - All the points between 0 y 2

$$\boldsymbol{T} = \begin{bmatrix} 2/n_c & 0 & 0 \\ 0 & 2/n_f & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



Perspective-n-Point (PnP): Plane Homography

- POSE calculation the from a set of 2D-3D matched points on a plane (linear method):
  - Due to the noise in the data we cannot directly assign the rotation vectors from the homography (they are not orthonormal)

$$A \cdot \begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ r_{31} & r_{32} & t_z \end{bmatrix} = \lambda \cdot \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$$

$$[\bar{\mathbf{r}}_1 \quad \bar{\mathbf{r}}_2 \quad \bar{\mathbf{t}}] = \lambda \cdot A^{-1}[\bar{\mathbf{h}}_1 \quad \bar{\mathbf{h}}_2 \quad \bar{\mathbf{h}}_3]$$

$$\|\bar{\mathbf{r}}_1\| = 1 \Rightarrow \quad \lambda = 1/\left\|A^{-1}\bar{\mathbf{h}}_1\right\|$$

$$\bar{\mathbf{r}}_1 = A^{-1}\bar{\mathbf{h}}_1/\left\|A^{-1}\bar{\mathbf{h}}_1\right\|; \quad \bar{\mathbf{r}}_2 = A^{-1}\bar{\mathbf{h}}_2/\left\|A^{-1}\bar{\mathbf{h}}_1\right\|; \quad \bar{\mathbf{r}}_3 = \bar{\mathbf{r}}_1 \times \bar{\mathbf{r}}_2; \quad \bar{\mathbf{t}} = A^{-1}\bar{\mathbf{h}}_3/\left\|A^{-1}\bar{\mathbf{h}}_1\right\|$$

- Robust procedure:
  1. We calculate **H** from 2D-3D correspondences
  2. A first estimation of **R*** is calculated considering $\lambda=1$
  3. We adjust the rotation matrix so that its vectors are orthoormal → **R**
  4. We calculate the scale factor λ as the inverse of the biggest singular value $1/\sigma_1$
  5. We calculate the translation vector with the scale factor λ y $\mathbf{h}_3$

$$\boxed{R^* = \left[A^{-1}\bar{\mathbf{h}}_1 \quad A^{-1}\bar{\mathbf{h}}_2 \quad \left(A^{-1}\bar{\mathbf{h}}_1 \times A^{-1}\bar{\mathbf{h}}_2\right)\right]}$$

$$\boldsymbol{SVD}(R^*) \rightarrow R^* = U \cdot D \cdot V^T$$

$$D = diag(\sigma_1, \sigma_2, \sigma_3)$$

$$\boxed{R = U \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \underbrace{|U \cdot V^T|}_{\pm 1} \end{bmatrix} \cdot V^T}$$

$$\lambda = 1/\sigma_1$$

$$\boxed{\bar{\mathbf{t}} = A^{-1}\bar{\mathbf{h}}_3/\sigma_1}$$

Perspective-n-Point (PnP): General Disposition

- POSE calculation the from a set of 2D-3D matched points in a general disposition (linear method):
  - Problem similar to calibration but in this case we know **A**

$$\tilde{m} \cong A[R \quad t]\tilde{M} = \tilde{P}\ \tilde{M}$$

$$\begin{cases} p_{31}X_w u + p_{32}Y_w u + p_{33}Z_w u + p_{34}u - p_{11}X_w - p_{12}Y_w - p_{13}Z_w - p_{14} = 0 \\ p_{31}X_w v + p_{32}Y_w v + p_{33}Z_w v + p_{34}v - p_{21}X_w - p_{22}Y_w - p_{23}Z_w - p_{24} = 0 \end{cases}$$

$$\begin{bmatrix} \rho u \\ \rho v \\ \rho \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

$$\mathbf{A}_{xy} = \begin{bmatrix} -X_w^{(1)} & -Y_w^{(1)} & -Z_w^{(1)} & -1 & 0 & 0 & 0 & 0 & X_w^{(1)}u^{(1)} & Y_w^{(1)}u^{(1)} & Z_w^{(1)}u^{(1)} & u^{(1)} \\ 0 & 0 & 0 & 0 & -X_w^{(1)} & -Y_w^{(1)} & -Z_w^{(1)} & -1 & X_w^{(1)}v^{(1)} & Y_w^{(1)}v^{(1)} & Z_w^{(1)}v^{(1)} & v^{(1)} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \end{bmatrix}$$

$$\bar{\mathbf{p}} = [p_{11}, p_{12}, p_{13}, p_{14}, p_{21}, p_{22}, p_{23}, p_{24}, p_{31}, p_{32}, p_{33}, p_{34}]$$

At least 6 points are needed:

$$\mathbf{A}_{xy} \cdot \mathbf{p} = 0$$

$$\boxed{SVD(A_{xy}) \rightarrow A_{xy} = UDV^T \Rightarrow \mathbf{p} = v_{12}}$$

- Once calculated matrix **P** we can identify **R, t**

$$[R \quad t] = \lambda \cdot A^{-1}\tilde{P} \;\Rightarrow\; [\bar{\mathbf{r}}_1 \quad \bar{\mathbf{r}}_2 \quad \bar{\mathbf{r}}_3 \quad \bar{\mathbf{t}}] = \lambda \cdot A^{-1}[\bar{\mathbf{p}}_1 \quad \bar{\mathbf{p}}_2 \quad \bar{\mathbf{p}}_3 \quad \bar{\mathbf{p}}_4]$$

$$\boxed{R^* = A^{-1}[\bar{\mathbf{p}}_1 \quad \bar{\mathbf{p}}_2 \quad \bar{\mathbf{p}}_3]}$$

$$SVD(R^*) \rightarrow R *= U \cdot D \cdot V^T$$

$$D = diag(\sigma_1, \sigma_2, \sigma_3)$$

$$\boxed{R = U \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & |U \cdot V^T| \end{bmatrix} \cdot V^T}$$

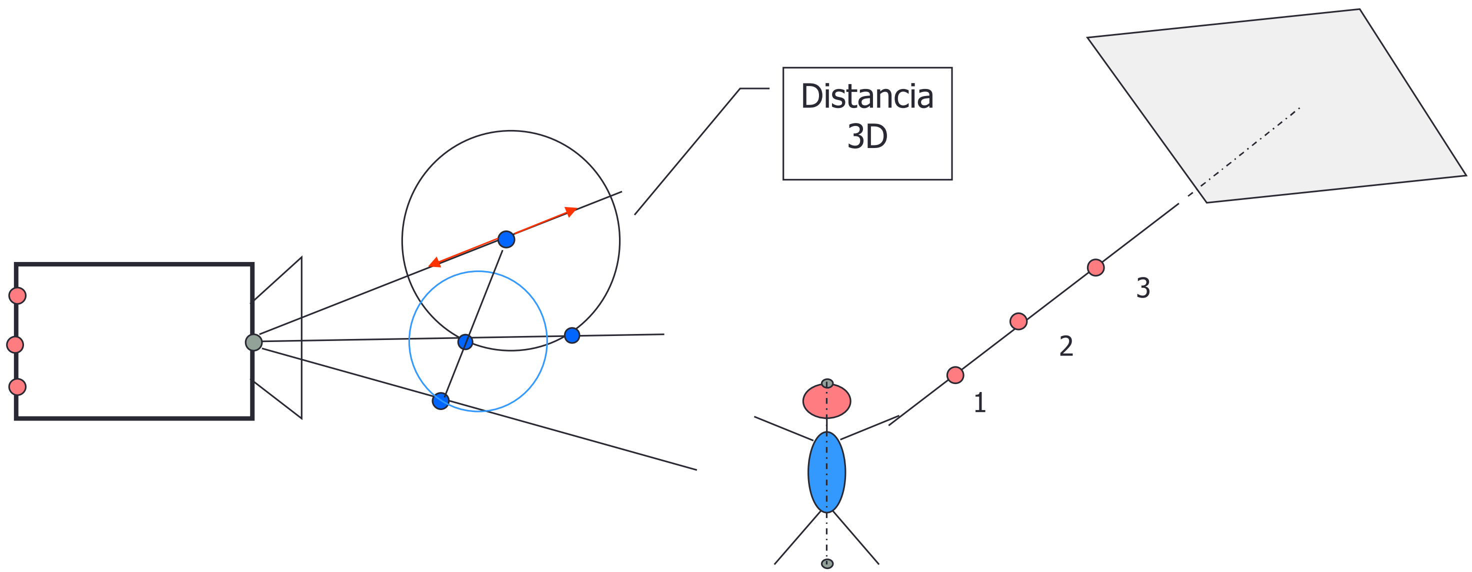
$$\pm 1$$

$$\lambda = 1/\sigma_1$$

$$\boxed{\bar{\mathbf{t}} = A^{-1}\bar{\mathbf{p}}_4/\sigma_1}$$

POSE Estimation: 3 aligned points

- Example: POSE Estimation of a bar with 3 marks:
  - We know the distance between marks and the calibration matrix
  - 5 dof (3 Position of the 3 marks + 2 Rotation)



POSE Estimation: 3 aligned points

- Analytical Solution:
  - Equation of the projection rays:

$$\tilde{Q}_i \cong \begin{bmatrix} -R^T t \\ 1 \end{bmatrix} + \lambda_i \begin{bmatrix} R^T A^{-1} \tilde{m}_i \\ 0 \end{bmatrix} = \begin{bmatrix} -R^T t + \lambda_i R^T A^{-1} \tilde{m}_i \\ 1 \end{bmatrix} = \begin{bmatrix} Q_i \\ 1 \end{bmatrix}$$

distances

$$d^2{}_{12} = \|\lambda_1 R^T A^{-1} \tilde{m}_1 - \lambda_2 R^T A^{-1} \tilde{m}_2\|^2$$
$$d^2{}_{13} = \|\lambda_1 R^T A^{-1} \tilde{m}_1 - \lambda_3 R^T A^{-1} \tilde{m}_3\|^2$$

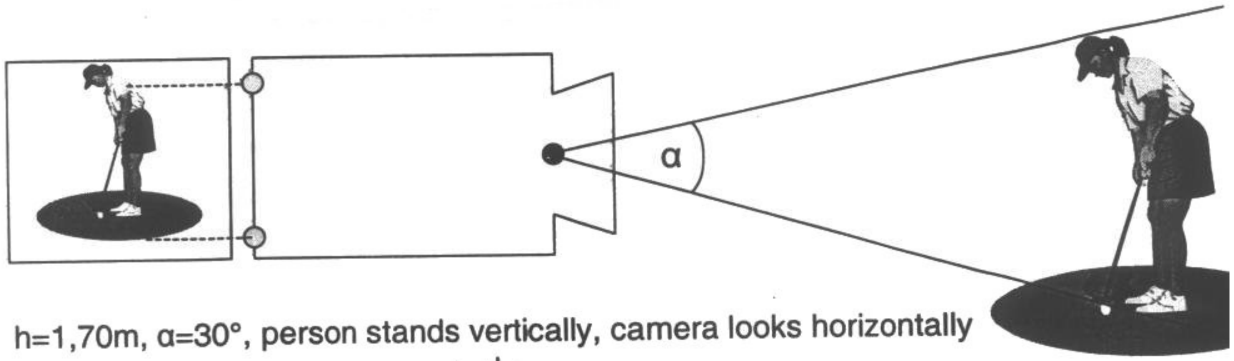
3 aligned points

$$Q_3 \cong Q_1 + \lambda\,(Q_2 - Q_1)$$
$$R^T A^{-1}(\lambda_3 \tilde{m}_3 - \lambda_1 \tilde{m}_1) = R^T A^{-1} \lambda\,(\lambda_2 \tilde{m}_2 - \lambda_1 \tilde{m}_1)$$
$$\lambda_3 \tilde{m}_3 - \lambda_1 \tilde{m}_1 = \lambda\,(\lambda_2 \tilde{m}_2 - \lambda_1 \tilde{m}_1)$$

- Unique solution in front of the camera

POSE Estimation: height on a plane

- Height of an object sitting on a plane (floor)

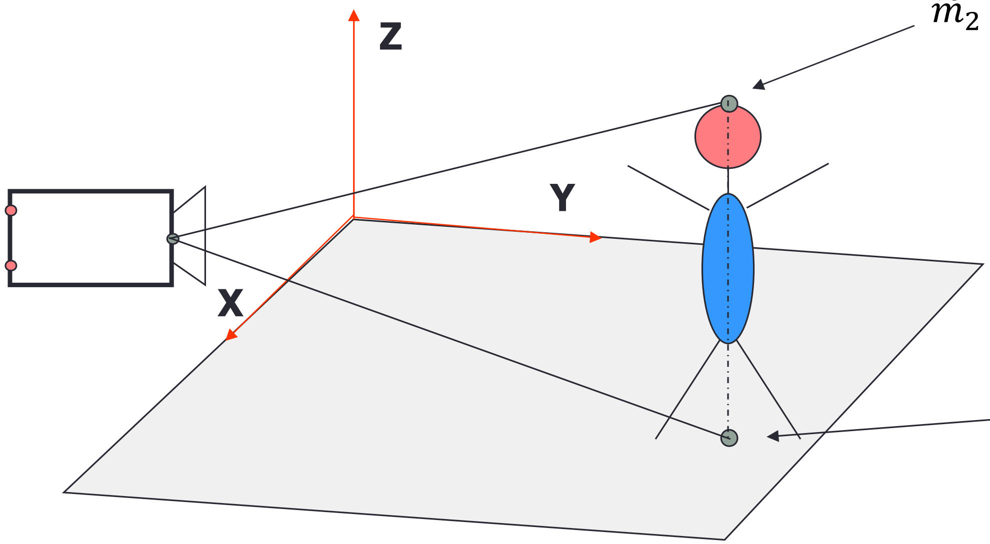


h=1,70m, α=30°, person stands vertically, camera looks horizontally
→ d=...

POSE Estimation: height on a plane

- Example:

$$\tilde{m}_2 \cong A[R\ \ t]\begin{bmatrix} X_1 \\ Y_1 \\ h \\ 1 \end{bmatrix} = AR\left(\begin{bmatrix} X_1 \\ Y_1 \\ h \end{bmatrix} + R^T t\right)$$



Calculate X1,Y1

$$\tilde{m}_1 \cong A[R\ \ t]\begin{bmatrix} X_1 \\ Y_1 \\ 0 \\ 1 \end{bmatrix}$$

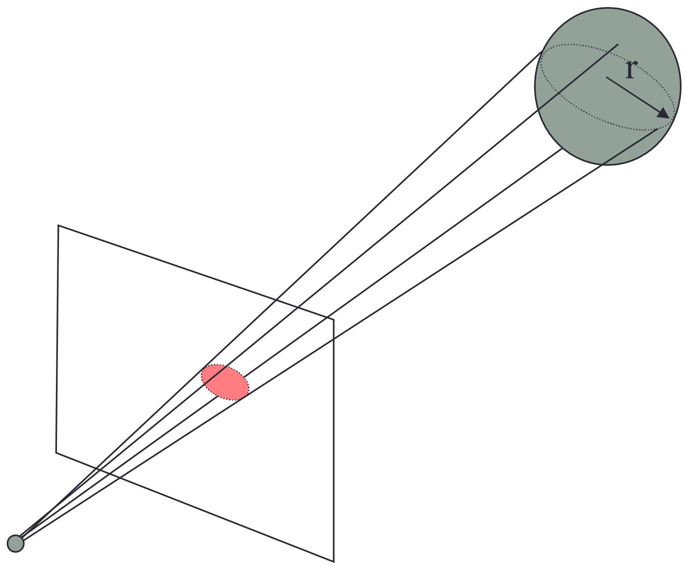
$$[\tilde{m}_2]_\times AR\left(\begin{bmatrix} X_1 \\ Y_1 \\ h \end{bmatrix} + R^T t\right) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$A = \begin{bmatrix} 1000 & 0 & 200 \\ 0 & 1000 & 200 \\ 0 & 0 & 1 \end{bmatrix},\ R = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix},\ t = \begin{bmatrix} -3.5355 \\ -5.0000 \\ -42.4264 \end{bmatrix}$$

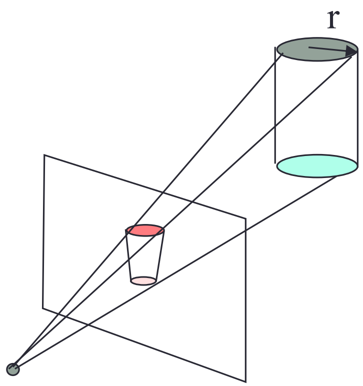
$$m_1 = [181.48 \quad 121.43]^T,\ \ m_2 = [219.23 \quad 118.41]^T$$

POSE Estimation: other examples

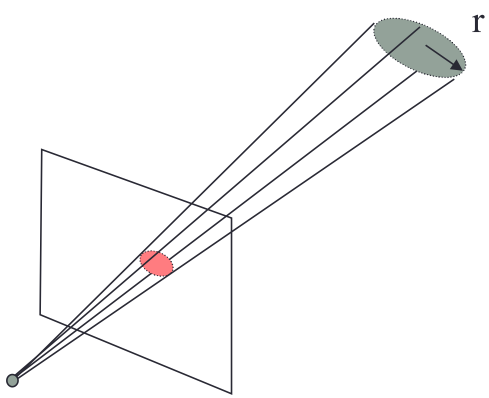
- Pose knowing the radius of a sphere



- Pose knowing the radius of a cylinder



- Pose knowing the radius of a disc

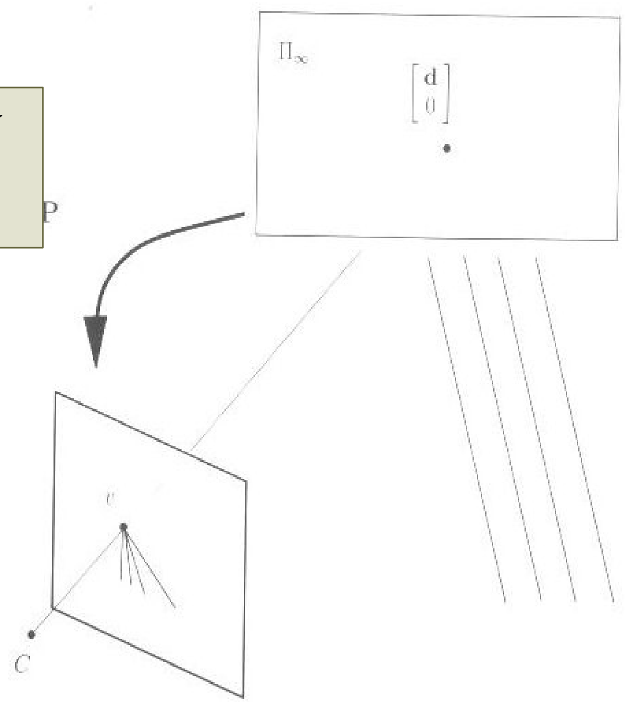


POSE Estimation: Vanishing Points

- An Homography can be calculated between the plane at infinity and the image plane (Allows to calculate the orientation **R**)
  - Vanishing Points: projection of infinity points (intersection point of two parallel lines)

$$\tilde{m} \cong A[R \quad t]\,\tilde{M}$$

$$\text{if} \quad \tilde{M} \cong \begin{bmatrix} M \\ 0 \end{bmatrix} \in \Pi_\infty \Rightarrow \begin{cases} \tilde{m} \cong AR\ M = \tilde{H}_\infty M \\ M \cong \tilde{H}_\infty{}^{-1}\tilde{m} \end{cases}$$



POSE Estimation: Vanishing Points

- Calculation of Orientation **R**:
  - Find the projection of points in infinity (directions)
    - Vanishing Points: intersection point of two parallel lines to the axis Z,X

$$\text{Si} \quad \tilde{M} \cong \begin{bmatrix} M \\ 0 \end{bmatrix} \in \prod_\infty \quad \Rightarrow \quad \tilde{m} \cong A[R \quad t] \begin{bmatrix} M \\ 0 \end{bmatrix} = ARM = \tilde{H}_\infty M$$

$$\tilde{m}_\infty^{(x)} \cong A[\bar{\mathbf{r}}_1 \quad \bar{\mathbf{r}}_2 \quad \bar{\mathbf{r}}_3 \quad \bar{\mathbf{t}}] \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = A \cdot \bar{\mathbf{r}}_1 \qquad \tilde{m}_\infty^{(z)} \cong A[\bar{\mathbf{r}}_1 \quad \bar{\mathbf{r}}_2 \quad \bar{\mathbf{r}}_3 \quad \bar{\mathbf{t}}] \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = A \cdot \bar{\mathbf{r}}_3$$

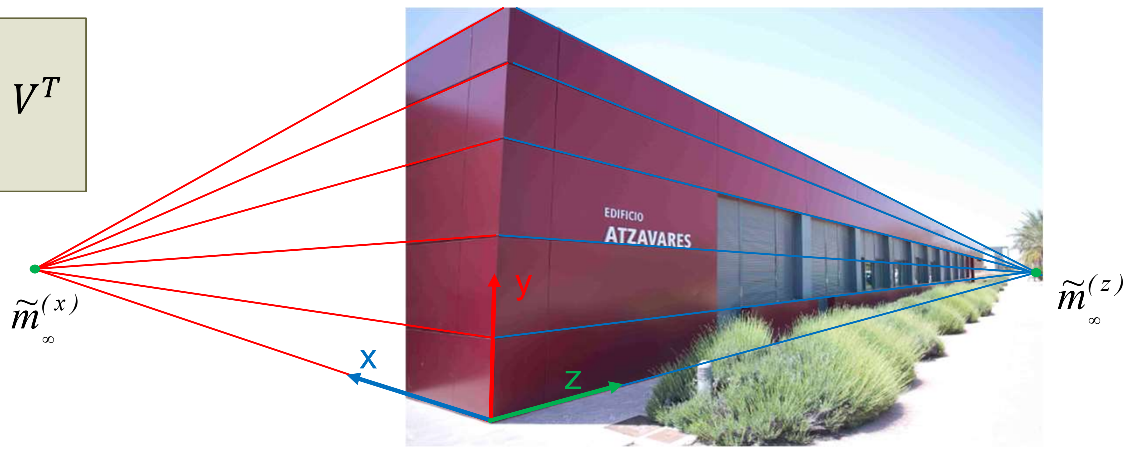
$$\bar{\mathbf{r}}_1 = A^{-1}\tilde{m}_\infty^{(x)} / \left\| A^{-1}\tilde{m}_\infty^{(x)} \right\|$$

$$\bar{\mathbf{r}}_3 = A^{-1}\tilde{m}_\infty^{(z)} / \left\| A^{-1}\tilde{m}_\infty^{(z)} \right\|$$

$$\bar{\mathbf{r}}_2 = \bar{\mathbf{r}}_3 \times \bar{\mathbf{r}}_1$$

$$R^* = [\bar{\mathbf{r}}_1 \quad \bar{\mathbf{r}}_2 \quad \bar{\mathbf{r}}_3] \qquad SVD(R^*) \rightarrow R^* = U \cdot D \cdot V^T$$

$$\boxed{R = U \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & |U \cdot V^T| \end{bmatrix} \cdot V^T}$$



POSE Estimation: Vanishing Points

- Application:
  - Tracking road lines
    - Vanishing Point: intersection point of two parallel lines (Z-axis)
    - Considering that the vehicle moves on a plane we only have an angular degree of freedom

$$\tilde{m}_{\infty}^{(z)} \cong A \begin{bmatrix} \cos\theta & 0 & \sin\theta & t_x \\ 0 & 1 & 0 & t_y \\ -\sin\theta & 0 & \cos\theta & t_z \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

Vanishing Point

Rotation around the axis Y

$$A^{-1} \cdot \tilde{m}_{\infty}^{(z)} = \lambda \cdot \begin{bmatrix} \sin\theta \\ 0 \\ \cos\theta \end{bmatrix}$$

$$\lambda = \left\| A^{-1} \cdot \tilde{m}_{\infty}^{(z)} \right\|$$

$\tilde{m}_{\infty}^{(z)}$

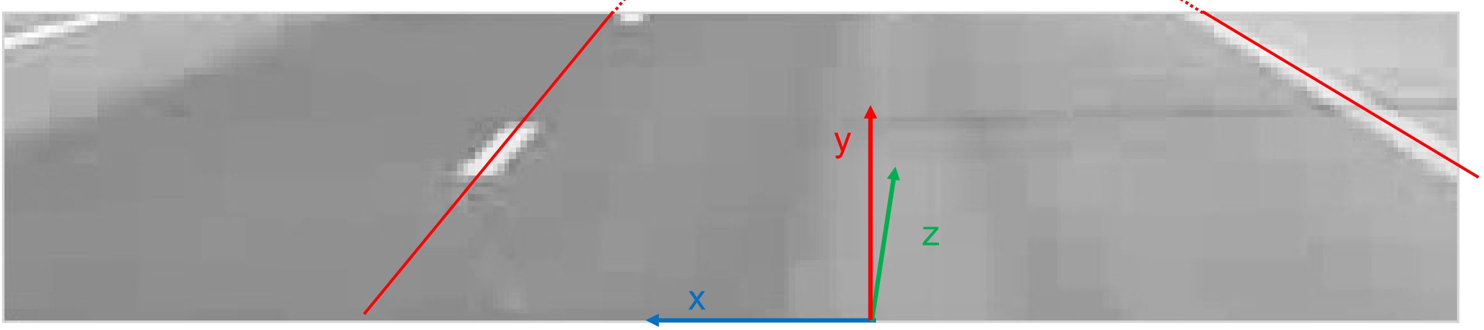


Table of Contents

- Monocular POSE
- Homography between Projective Planes

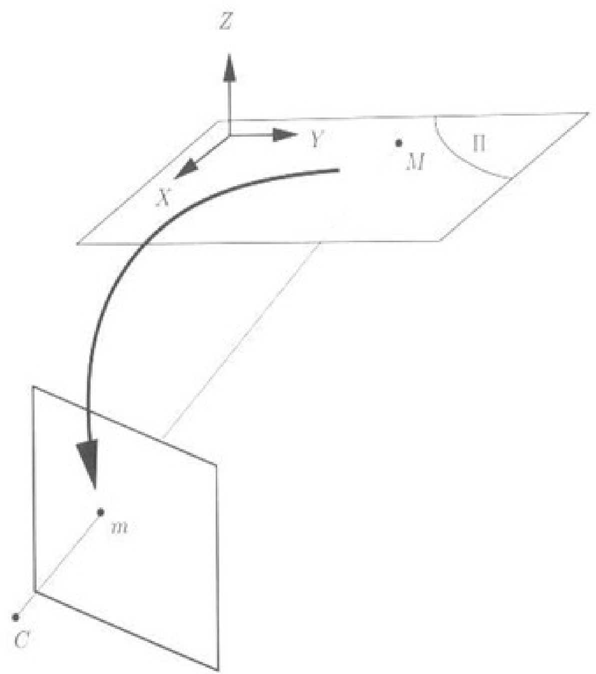
Homographies between Projective Planes

- Pane – Image
  - NOT-calibrated camare
  - The Plane Π does not contains the optical center ***C***
  - *Points on the plane:*

$$\tilde{m} \cong \tilde{P}\ \tilde{M}\ = [P \quad p]\ \tilde{M}$$

$$\text{if } \tilde{M} \cong \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} \in \prod \Rightarrow \tilde{m} \cong \begin{bmatrix} P_{11} & P_{12} & P_{14} \\ P_{21} & P_{22} & P_{24} \\ P_{31} & P_{32} & P_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

$$\boxed{\tilde{m} \cong H_{3x3}\ \tilde{M}_{\Pi}, \quad \tilde{M}_{\Pi} \cong {H^{-1}}_{3x3}\tilde{m}}$$



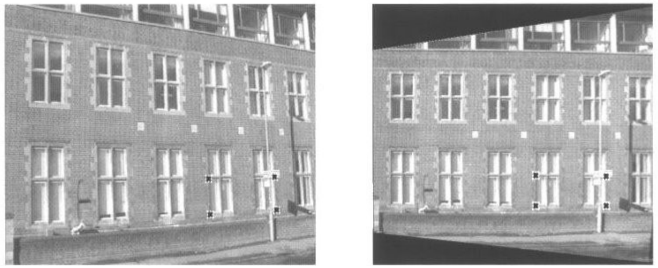
- *Calculation of **H**: 4 points of the plane projected into the image*
  - *Non collinear taken in threes*

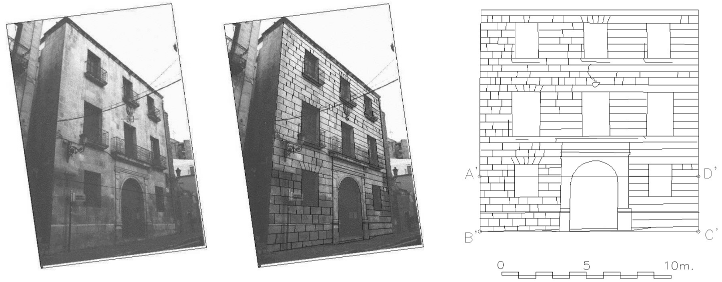
Homographies between Projective Planes

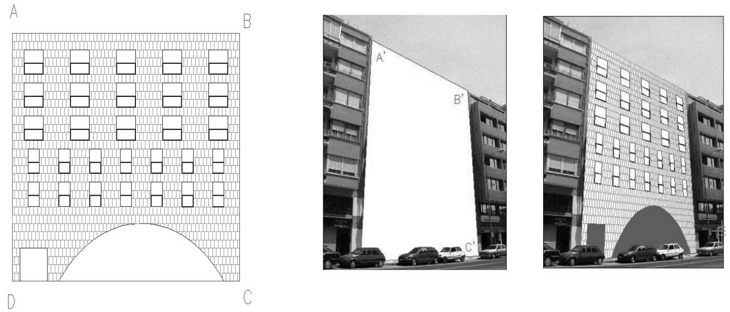
- Examples:
  - Projective insertion of one projective plane into another
    - Camera NOT Calibrated
  - Recognition of projective patterns
    - Camera NOT Calibrated

Homographies between Projective Planes

- Allows:
  - To know the projection of any point contained in the plane (outside the boundaries of the image)
  - To check if a point belongs to a plane
  - To know the 3D position on the plane from the projection in a single camera (2D)
  - Rectification: remove projective distortion in 2D vision
  - Photogrammetry:

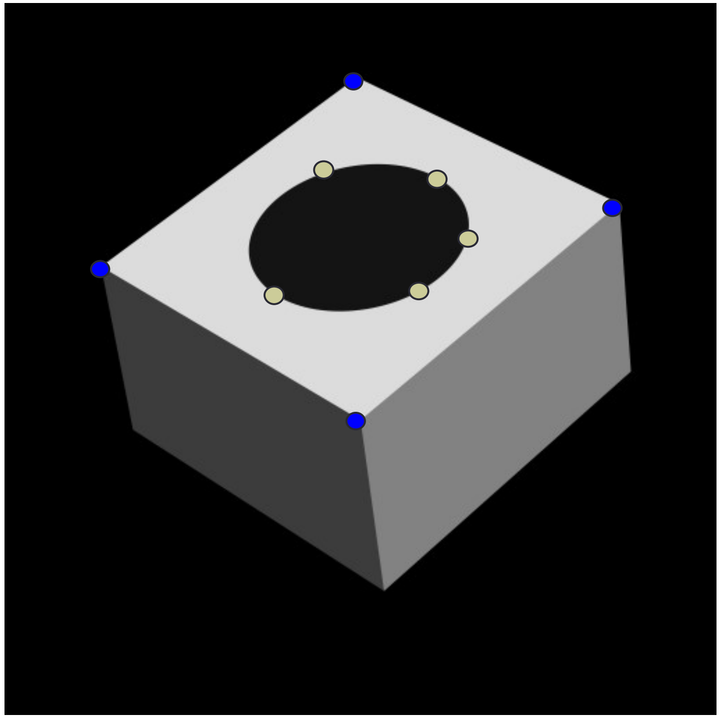




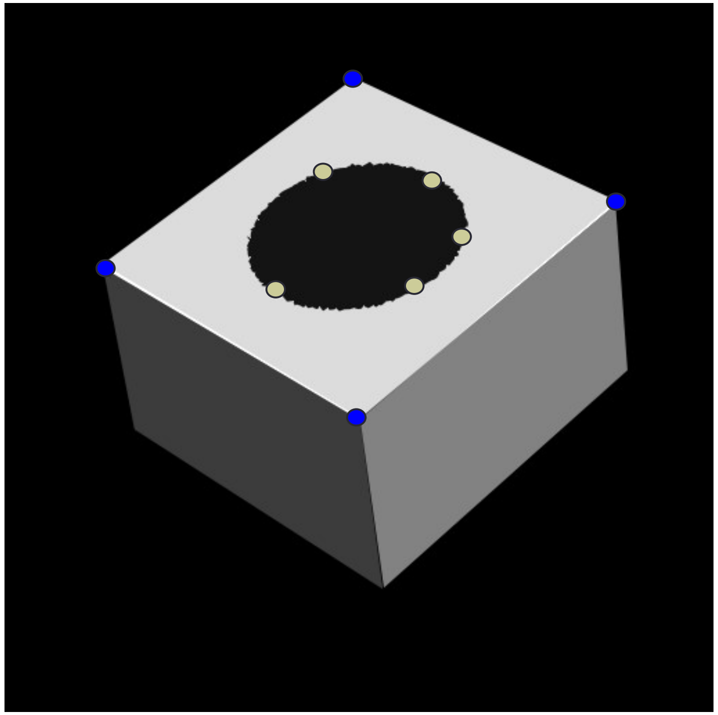


Homographies between Projective Planes

- Quality control in projective views



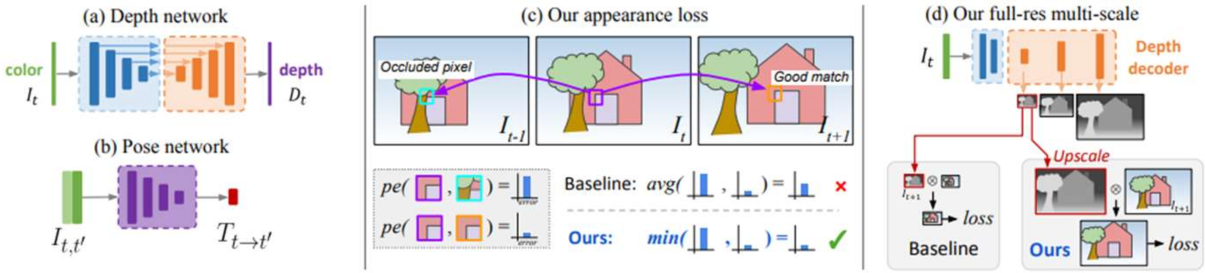
Correct Drill

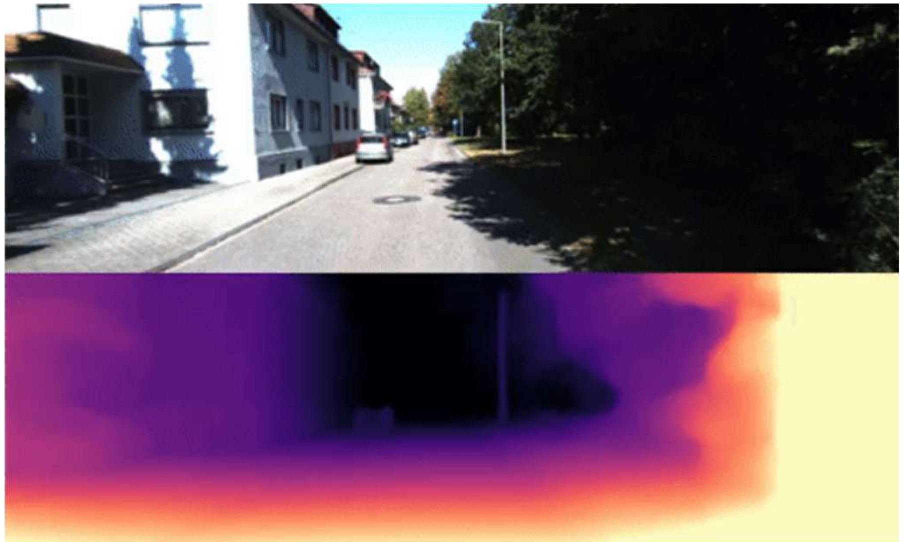


Defective Drill

Advanced Monocular Techniques: CNN

- Monodepth2: Self-Supervised Monocular Depth Estimation
  - U-Net CNN





"Digging into Self-Supervised Monocular Depth Prediction" Clément Godard, Oisin Mac Aodha, Michael Firman and Gabriel J. Brostow, (2019) The International Conference on Computer Vision (ICCV), www.github.com/nianticlabs/monodepth2

Advanced Monocular Techniques: CNN

Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data

Lihe Yang[1] Bingyi Kang[2†] Zilong Huang[2] Xiaogang Xu[3,4] Jiashi Feng[2] Hengshuang Zhao[1†]
[1]The University of Hong Kong [2]TikTok [3]Zhejiang Lab [4]Zhejiang University
† corresponding authors

https://depth-anything.github.io

Figure 1. Our model exhibits impressive generalization ability across extensive unseen scenes. **Left two columns:** COCO [35]. **Middle two:** SA-1B [27] (a hold-out unseen set). **Right two:** photos captured by ourselves. Our model works robustly in low-light environments (1st and 3rd column), complex scenes (2nd and 5th column), foggy weather (5th column), and ultra-remote distance (5th and 6th column), *etc*.

"Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data" Lihe Yang1, Bingyi Kang, Zilong Huang, Xiaogang Xu, Jiashi Feng, Hengshuang Zhao1 (2024) CVPR Computer Vision and Pattern Recognition https://github.com/LiheYoung/Depth-Anything

Summary

- Monocular POSE
- Homography between Projective Planes

- Next step is the use of more than one point of view: Stereo Vision
  - 3D Reconstruction